AGROCLIMATOLOGY:

MAIZE YIELDS IN THE ORANGE FREE STATE

BY

JANE FRANCES GILLOOLY

A Dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg for the Degree of Master of Science.

Johannesburg, 1977.

ABSTRACT

The dissertation opens with a review of current world-wide developments in agroclimatology. This highlights the dearth of objective crop-weather studies on a macroscale in South Africa. Maize yields for 38 magisterial districts of the Orange Free State are examined in relation to mean monthly values of rainfall and temperature for the period 1950 to 1965 inclusive. Initially these districts are grouped into homogeneous maize regions using principal component analysis; after which maize-climate relationships are analysed on a regional basis using multiple regression techniques. Regression models are then extended to enable forecasts of maize yields to be given in January.

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

This dissertation is dedicated to

Barbara and Laurence Gillooly.

PREFATORY NOTE

The world food crisis is a problem which is receiving ever increasing attention from government agencies, scientists and decision makers in general. The current food shortage is a consequence of the surge in world population over the past few decades, and the resultant intensified pressure on areas of arable land. Coupled with the increased human pressure on land is the control exerted by climatic factors and variations thereof, on agricultural production. These climatic changes occur on both short and long term time scales and their effects become more drastic as man penetrates areas which are climatically marginal for agriculture.

The concern over the food shortages has led to the initiation of several international agrometeorological projects for the large scale monitoring of crop production. Until now, predictions of crop yields were limited to small scale studies using detailed climatic data. For large scale regional forecasting of crop yields to be feasible, the data must be commonly used and readily available.

The present dissertation comprises an analysis of the potential monitoring of regional crop yields in South Africa. A single crop, namely maize , has been chosen for analysis within a defined region, the Orange Free State. More specifically, the aims of this dissertation are:

(i) to examine any structural relationships existing between selected climatic parameters and maize yields in the Orange Free State and

(ii) to investigate the plausibility of using stochastic models for forecasting annual maize yields.

Chapter 1 presents an introduction to the world food problem; an historical synopsis of rainfall variations in South Africa, and the scope of agroclimatology in South Africa and the rest of the world. Analytical techniques used in the analysis of maize yields and climatic factors are contained in Chapter 2. Chapter 3 is concerned with the formation of homogeneous maize growing regions; and in Chapter 4, structural maize yield-climate relationships are investigated. In Chapter 5 the models are extended for forecasting purposes. Finally in Chapter 6, conclusions are drawn and the results of the dissertation summarized.

During the period of research for this dissertation, parts of the material presented herein have been published. A synopsis of the work contained in Chapters 3 and 4 has been accepted for publication in *Nature* (1977). Those aspects of Chapter 4 not contained in the first paper have been presented in a paper in *Agricultural Meteorology* (1977) vol.17.

*   *   *

I wish to acknowledge the assistance of Mr. P. Stickler in the final preparation of diagrams. I am most grateful to my parents for the typing and proof reading of the text.

I acknowledge the helpful criticisms of Professor P.D. Tyson, Head of the Department of Geography and Environmental Studies during the final drafting of the dissertation. To my mentor and supervisor Dr. T.G.J. Dyer, I extend my warmest thanks for the stimulating discussions, valuable advice and assistance that he provided throughout the period of research.

The research was funded by grants from the University of the Witwatersrand and the Council for Scientific and Industrial Research.

CONTENTS

CHAPTER I

INTRODUCTION

1.1 BACKGROUND TO THE WORLD FOOD PROBLEM

That the world is suffering from a food crisis due to its burgeoning population is a well known fact. Yet it is only now that this problem is receiving the attention it warrants from international organisations, governments and the public (Baier,1974).

Since the middle of this century, increases in agricultural production have become more and more dependent on the intensified exploitation of existing agricultural land. This was made possible by an advancing technology such as the increased use of agrochemicals, the introduction of high yielding varieties of crops and other cultivation practices, in what became known as the Green Revolution (Revelle, 1974). Although increased yields of major world crops resulted from these measures, the rising demand for food from the expanding world population and greater buying power of the developing countries exceeded the grain supply (Landsberg, 1974). The net result has been a downward trend in world grain reserves (Robertson, 1974a).

The problem of nourishing a world population already too large to support itself, is further complicated by variations in climate (Newman and Pickett, 1974). Such climatic variations occur on short term scales, for example year to year changes and on long term time scales in the form of climatic trends (Thompson, 1975a,b). The increased awareness of such phenomena stems from a number of climatic shocks since 1960. These range from the reappearance of the Arctic ice off the Icelandic coast (Clarke, 1970; Kristjannson, 1969; Kukla and Kukla, 1974) and an increased incidence of monsoon failures in India (Jagannathan and Parthasarathy, 1973; Winstanley, 1973a,b), to alternations of unaccustomed droughts and hot weather in many regions in middle latitudes in both hemispheres (Bryson, 1973; Gribbin, 1974; Jenkinson, 1975). These changes have tended to dislocate world trade and price stability and have raised the spectre of food shortages (Lamb, 1975a).

The recent drought in the Sahelian region of Africa has aroused unprecedented attention. This drought reached a climax in 1973 after three successive bad years in which rainfall averaged approximately two-thirds of the median rainfall value (Winstanley,1973 a). During this period many people and vast numbers of cattle and game died (Bunting *et al*, 1975, 1976). Median rainfall in 1974 and 1975 alleviated conditions in the area. Landsberg (1975) considers the 1970-72 Sahelian drought to be a repetition of similar prolonged droughts in 1913-14 and 1940-45. The greater concern felt about the recent drought reflects the degree to which population pressure in the 1970's and the associated effects on rural activities have triggered greater ecological and socio-economic effects than have previous droughts. (Aust. Acad. of Science, 1976). It is significant that much agricultural production in many parts of the world is undertaken in regions climatically marginal for agriculture. Hence the increasingly severe effects of recent droughts and floods are due more to the increased human pressure on the land, than [illegible] increase in climatic variability.

Considering the socio-economic consequences of climatic changes in conjunction with changes in production policies in many countries (Gribbin, 1974), it is clear that the present food situation is highly vulnerable to the vagaries of climate. The low world food stocks are of the same order of magnitude as the weather-induced range of variability, and are close to the critical limit for annual carry-over, which is eight to ten percent of world food production (National Oceanic and Atmos. Admin., 1973). Hence we are not even in a position to cope with those climatic fluctuations which normally occur from year to year in any epoch.

This new hand-to-mouth food situation necessitates fresh technical approaches to assess current crop conditions and yield expectations (Russell, 1976). The analysis of regional crop yields in relation to climatic factors has important practical applications. It is for instance, most useful to be able to provide information on probable crop yields during the growing season. Yields obtained on a commercial basis are likely to differ somewhat from those produced in field experiments. Data used in microscale and mesoscale studies are detailed and frequently of a phenological nature, whereas macroscale analyses use measurements over a geographical region and represent much larger quantities of production.

1.2 PROGRESS IN AGROCLIMATOLOGY

Agricultural meteorology is the study of those aspects of meteorology of direct relevance to agriculture, and currently concentrates mainly on micrometeorological analyses. On the other hand, agroclimatology is concerned with the broader applications and use of such studies in agricultural planning (Smith, 1976).

There are a number of crop-weather simulation schemes for the assessment of agricultural risk, for example that of Baier (1973a,b). Several crop-weather models have been developed for specific crops and climatic regimes. They are commonly based on simple or multiple regression methods, and include selected climatic parameters within the set of predictor variables. For large scale studies it has been found that precipitation and temperature terms are useful predictors of crop yields, as well as being commonly recorded and readily available (Bourke, 1975). Climatic data are usually expressed as pentades or monthly values but where records are available, daily values are also used. Non-climatic data such as soil type and fertilizer application rates have also been included in the regression models (for example, Bryson, 1974).

Canadian wheat yields have been studied by Baier and Robertson (1967), Baier and Williams (1974), Robertson (1974b) and Williams *et al* (1975); other North American work includes that of Dale and Hodges (1972), Gross and Rust (1972), Hanks (1974), Pochop *et al*

(1975) and Runge (1968). Lewin and Lomas (1974), Lomas and Shashoua (1973) and Lomas *et al* (1974) have concentrated on the influence of weather on agriculture in semi-arid regions. A great deal of research on the effects of climate on maize and rice yields is currently underway in India (Bhargawa *et al*, 1974; Ghildyal and Jana, 1967; Huda *et al*, 1975, 1976; Mane and Dastane, 1971; Sreenivasan, 1974 and Sreenivasan and Banerjee, 1973). Recent work in Japan on rice yields includes that of Murata (1975) and Murata and Togari (1972) amongst others. Seasonal crop yield predictions are currently being made in several wheat producing countries, especially the U.S.S.R. (Degtyareva, 1973; Federov, 1974; Polevoj and Myzina, 1973 and Zheltaya, 1974).

Following a proposal passed at the recent conference on the world food supply concerning the monitoring of crop expectations, the United States is mounting a Large Area Crop Inventory Experiment (L.A.C.I.E.) (Hammond, 1975). The first part of this project is the acquisition of information from satellite photo-images, and the second part involves the use of weather data in regression models.

The cost involved in a project such as L.A.C.I.E. would be prohibitive for many countries. Sufficient agricultural and climatic data is however, available in the government records of most countries. It is now time for the climatic information buried in the meteorological archives of the world to be unearthed and used in agricultural planning. Similarly crop yield data, for example from government censuses, must now be exploited to provide weather-based assessments of the current state of regional food production.

1.3 EFFECTS OF CLIMATIC VARIATIONS ON FOOD PRODUCTION

A problem in crop yield forecasting arises when we come to consider long term predictions of crop production. The choice to be made lies between basing long term agricultural development on climatic normals or on projected trends. Climatic normals are generally based on the periods 1920-1950 and 1930-1960, which were in fact, highly abnormal (Mason, 1976). The alternative is to extrapolate trends observed over a long period of time and over large geographical areas. The latter alternative, provided our concern over climatic variations is real, should be the obvious choice. Indeed, Lamb (1975b) has described the need for the understanding and prediction of long term climatic fluctuations as *a matter of immediate and growing urgency* (p.9).

In an analysis of rainfall records for India, Winstanley (1973a,b) showed that irregular fluctuations are superimposed on longer term trends. Monsoon rainfall during the drought of 1973 was at a low point of an irregular fluctuation, coincident with a longer-term downward trend, which had first become apparent in the late 1920's. By extrapolating past trends, Winstanley showed that the future seasonal rainfall is likely to decrease to a minimum around A.D. 2030. Even if this downward trend does not persist, it is evident that seasonal rainfall receipts in India are not stable. This pattern also holds true for the Sahel region of Africa and the Middle East. Agricultural and economic planners should therefore anticipate periods of above and below average rainfall in the future (Bunting *et al*, 1975; Landsberg, 1975). The logical conclusion of this discussion is that seasonal rainfall forecasting is vital if early warnings of food shortages are to be obtained.

The analysis of Indian rainfall variations clearly shows that instability of climatic patterns is of the utmost importance to agriculture. In South Africa there have been several investigations into variations in rainfall. Bearing in mind the agricultural implications of such knowledge, a brief historical account of this work is given.

1.4 RAINFALL VARIATIONS OVER SOUTH AFRICA: A BRIEF SYNOPSIS

For many years, the question as to whether South Africa is drying up or alternatively whether rainfall is undergoing cyclic variation has been a subject for discussion. Strictly speaking the term *cyclical* should refer only to oscillations of constant amplitude and period. In the early South African literature *cyclical* has been used, rather loosely, to imply the presence of oscillatory patterns in the rainfall series.

Early explorers such as Moffat (1842) and Livingstone (1857) reported that lakes and rivers were abundant in areas that are now semi-arid. Indeed Wilson (1865) was adamant that the Kalahari had become drier, a statement which accords with the conclusion reached by Wallis (1935) after analysing rainfall records for Bechuanaland. The view generally held by the public was that South Africa as a whole was drying up (Agric. Jour. Union S.Afr. 1913; Barber, 1910). These popular beliefs however were refuted by the findings of the Union Parliament Senate Select Committee on Drought, Rainfall and Soil Erosion. In 1918 Schwartz wrote the first of a series of papers dealing with the desiccation theory. His views and plans were later criticized by Cox (1925), Kanthack (1930), Kokot (1948) and Schumann and Thompson (1934). Hutchins (1888) and Tripp (1888) were the first to claim that rainfall variation was *cyclical* in nature. Hutchins was concerned with correlations between sunspot activity and rainfall, and concluded that the 12.6 year sunspot cycle determined rainfall patterns in Natal and the Cape. Tripp, on the other hand, believed that cycles did exist but found no evidence to support the view that they were related to sunspot activity. Nevill (1908) found evidence for a periodic 18-year signal in Natal's rainfall, while Cox (1925) suggested a periodicity, 14 years in length, for Cape Town. Van Reenen (1925) examined rainfall records for South West Africa for the period 1771-1925 and isolated an oscillation with a mean period of 30 years. Peres (1930) examined the record for Lourenco Marques (1909-1929) and suggested an oscillation with a period length of 20 years. This latter finding is questionable because of the short time series for rainfall.

By dividing the country into four climatic regions and considering rainfall over each, separately and in combination, Levinkind (1941) found that rainfall return periods varied from one part of the country to another. More recently, Glover *et al* (1954) considered the perio'''' behaviour of rainfall in terms of return periods, whilst Vo. '957) suggested a general decline in *rainfall in the sou Cape Province between 1881 and* 1950. This find: .owever, was not corroborated by Brook and Mametse (1969) whose work supported the conclusions of Hofmeyr and Schulz (1963) regarding the spatial dependence of rainfall variations.

Investigations of rainfall variations in South Africa culminated in the recent analyses of Tyson (1971, 1972) whose work suggests the reality of an 18-20 year oscillation and delineates its distinctive regional occurrence. Further support for the physical reality of this particular rainfall oscillation lies in the similar but inverse oscillation in temperature that has been observed over the country (Keen, 1971). In an analysis of 157 stations with records extending from 1910 to 1972, Tyson *et al* (1975) dispelled the belief that South Africa is undergoing *progressive desiccation.* Instead it has been shown that rainfall has ordered temporal variations which are spatially dependent. In particular they isolated weak oscillations with wavelengths of 16-20, 10-12 and 3-4 years, and the quasi-biennial oscillation, each having a specific regional occurrence.

The latest work is that of Dyer (1975 a,b,c; 1976 a,b,c,d) who has dealt with several problems related to the statistical analysis of rainfall series over South Africa. Much of his work has arisen from an interest in the possible use of stochastic models in forecasting annual precipitation totals. Dyer's most recent work is concerned with the interrelationships between meridional fluctuations of the Southern hemisphere high pressure belt and precipitation receipts over South Africa (Dyer, 1976d). These findings suggest that the two variables are significantly associated, and provide a means of supplying forecasts for rainfall with small standard errors of the residuals.

With this background in mind it is appropriate to consider the broader applications of this knowledge in the agricultural field. It is within the framework of the world food problem that the recent research into rainfall behaviour patterns must now be viewed. In South Africa, as in the Sahel and India, extended periods of below and above average rainfall are to be expected. Future agricultural and economic development should therefore, take note of these trends so that the effects of adverse climatic fluctuations can be offset, or at least eased. Year-to-year variations of crop yields in South Africa are largely due to climatic factors, of which precipitation and temperature are the more important (Wellington, 1960). The fact that climatic factors limit crop growth in this country, as elsewhere, is sufficient justification for the analysis of crop-weather relationships in South Africa.

1.5 AGROCLIMATOLOGY IN SOUTH AFRICA

The available literature has shown that agroclimatological research in the Northern hemisphere is flourishing. Many of the world's important cereal producers are situated in this hemisphere as illustrated by Table 1. Thus the maintenance of their key positions in a competitive world market is sufficient incentive for their being leaders in the field of agroclimatology.

To date little progress in the field of agroclimatology has been made in the Southern hemisphere. This is probably due to the combined effects of the geographical isolation of the southern continents; their underdeveloped nature and their minor status in world agricultural trade. Excluding the work done (1974), Nix and Fitzpatrick (1969) and Russell (1973) in a general feeling of apathy towards agroclimatology ap o prevail. Nevertheless the expense involved in providing aid to famine-stricken areas of the world is by no means insignificant. An assistance scheme such as that in current use, becomes unviable when there is no longer a grain surplus in developed countries. It is imperative therefore, that all countries should forecast crop expectations so that early warnings of food shortages can be obtained. Progress in agroclimatology in South Africa is no different to the rest of the Southern hemisphere. Numerous popular articles on agriculture and the weather have been published. These papers tend to be descriptions of weather-crop interrelation-

TABLE 1.- AGRICULTURAL STATISTICS FOR 1973 AND 1974 *SHOWING TOP WORLD CEREAL PRODUCERS* (All figures in 1000's of tonnes). (After Fullard, 1976).

Maize Yields	1973	1974
World	310391	292990
U.S.A.	143435	118144
China	30384	31085
Brazil	14059	16065
U.S.S.R.	14216	12142
South Africa	4160	11035
Argentina	9700	9900
Wheat Yields	**1973**	**1974**
World	377272	3[illegible]0231
U.S.S.R.	109784	63800
U.S.A.	46408	48807
China	36002	37002
India	24735	22073
France	17792	18906
Canada	16459	14221
Australia	12094	11200
Rice Yields	**1973**	**1974**
World	324468	323201
China	111954	115275
India	65613	61500
Indonesia	21900	22800
Bangaladesh	19356	17222
Japan	15717	15902
Thailand	14898	13175

ships or advice to farmers on how to overcome climatic hazards; for example Baier*(1962): *Make better use of rainfall* or De Jager (1971): *How to beat the weather*. Other articles in a similar vein include those of de Bruyn (1970), Fabricius (1971 a,b); Haylett (1968); Moll (1969); Tidmarsh (1968); Tredoux (1962) and Whitmore (1948).

Several attempts at assessing the agroclimatological potential of South Africa have been m 'e. For instance, Nuttonson (1961) based his study of winter rainfall areas on official records and the reports of government agencies; Whitmore (1950) discussed agrometeorological problems of farming in the Orange Free State in relation to seasonal fluctuations of temperature and rainfall, optimal climatic conditions for different crops and finally, to land use. At present the Department of Agriculture is running a research project dealing with agroclimatic evaluation in the Republic of South Africa (Buys, 1973). Several workers have adopted a more objective approach to climatic data and crop records; such as Baier*(1963) Crafford and Nott (1970, 1971); de Jager (1974); de Jager and Mallett (1972); Garlipp (1976); Jarrett *et al* (1966); Mallett (1972); Myburgh (1971) and van Jaarsveld (1973).

Garlipp (1976) studied annual maize yields recorded at Potchefstroom from 1953-54 to 1966-67 in relation to rainfall over different parts of the growing season, and thereby isolated that part of the growing season most sensitive to rainfall. Unfortunately his misuse of simple linear regression in the analysis detracts from the reliability of the findings. Mallett (1973) described a mathematical model for maize yield predictions using daily climatic data, especially those of soil moisture, atmospheric evaporation and rainfall in conjunction with soil composition. Such an approach would be of limited value however, for broad scale analyses since these climatic data (excluding rainfall) are not readily available or commonly recorded.

A striking characteristic of the local scientific literature in agroclimatology is that the analyses are restricted to the micro- and mesoscales. This commonly takes the form of working

Baier* - different to previous author.

in small areas under selected field conditions; using detailed climatic data, such as daily values and where possible, crop phenological data. This has obvious disadvantages for the large scale monitoring of crop yields.

There are two distinct approaches to the computer simulation of agricultural production systems; that of the deterministic type and that of the stochastic type of modelling (Newman, 1974). The deterministic approach is based on mathematically formulated relationships with empirical constants when necessary (see Mallett, 1973). This type of model is more applicable to a specific production system and usually requires meteorological data taken over short periods. The stochastic approach depends on fitting statistically the best possible empirical relationships and is more adaptable to relating climatological data to agricultural production in a geographical region. It is the latter approach that is more appropriate for the present study and will be used in this dissertation.

In agroclimatological studies that focus on the possible *prediction of crop yields, it is usual to concentrate on the major* staple crop for the region under study. The climatic parameters included in the statistical model are most commonly, precipitation, temperature, incoming solar radiation and if possible, evaporation (Smith, 1976).

In this dissertation maize, being the major cereal crop produced in South Africa, is the object of attention. It is grown on a large scale in an area known locally as the maize quadrangle, of which the Orange Free State forms a substantial part (Fig. 1). In view of the importance of maize to South African agriculture and the implications of the aforementioned rainfall variations, it would be valuable to examine structural relationships between maize and climate in South Africa.

The Orange Free State in being representative of the maize-producing block, provides a framework for such an analysis. Annual maize yields for the period 1949-50 to 1964-65 will be

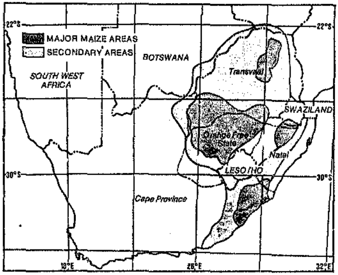


Fig. 1 - Map of Southern Africa showing the maize quadrangle of South Africa, of which the Orange Free State forms a substantial part.

analysed for 38 magisterial districts. The climatic data to be examined are monthly rainfall and temperature values for the same interval (1950-1965). This 16-year interval has been chosen since, after 1965, the maize yield records are incomplete. It was considered better to work with an homogeneous series of shorter length rather than one which included interpolations.

Various statistical techniques will be employed for the analysis of crop-climate relationships. In the first instance, principal component analysis and subsequent vector rotation will be used to form homogeneous maize growing regions. This serves to reduce the complexity of the original data set and the number of variables to be analysed. The use of principal component analysis has not hitherto been used in agroclimatology, although its use in climatology is widespread (Colgate, 1975; Craddock and Flood, 1969; Dyer, 1975c; Hannes, 1976; Kidson, 1975; Kutzbach, 1967; Trenberth, 1975). This particular application of the technique in the formation of maize growing regions is original and offers numerous possibilities for other types of agricultural data.

By fitting a two-way fixed effects linear model to the data within each region it will be possible to determine single regional time series. These single time series will then be submitted to stepwise multiple regression for the analysis of structural maize yield-climate relationships. Finally the feasibility of using the regression models to forecast maize yields as early as possible in the season will be explored.

In the present analysis the use of a shorter record (1949-50-1964-65) offers an advantage that should not be overlooked. It has been shown that South African rainfall series, 60 years in length, have associated variability due to oscillatory behaviour patterns. The reliability of using these longer but unstable series in a crop forecasting model is doubtful. It would be preferable to use short series, their complexities being less than longer ones, and hence would be more appropriate for any forecasting equations that may be developed.

* * *

There are two main factors which contribute to the current world food crisis. Firstly, the expanding world population which is rapidly reaching its limit for the available food supply; secondly, climatic variations, both short term fluctuations and long term trends which tend to dislocate the precarious balance existing between the supply and demand for food. The situation can only deteriorate as agriculture in the marginal regions intensifies, since it is these areas that are most sensitive to the vagaries of climate. The need for weather forecasting so that crop expectations and hence food shortages can be predicted, has been stressed.

In South Africa, rainfall and to a lesser extent temperature limit crop production. Rainfall patterns however, are not stable and exhibit oscillations that are spatially dependent in their occurrence. Hence agricultural planning should not be based on long term means, but should anticipate extended periods of above and below average rainfall in the future.

The present state of agroclimatological research in South Africa has been examined and a change of emphasis from micro- and mesoscale studies to broad scale analyses has been advocated. The object of this dissertation is to examine structural relationships between maize yields and climate in the Orange Free State. Original applications of sophisticated statistical techniques will prove useful in analysing the crop and climate data. Finally the feasibility of predicting maize yields using stochastic models will be explored.

CHAPTER 2

THE THEORETICAL ASPECTS OF THE ANALYSES

2.1 INTRODUCTION

Information on the acqui on and selection of maize yield data and climatic data ıs given. The statistical techniques used in the subsequent analyses of the crop and climatic data are not dealt with in much detail, but sufficient is supplied to prevent ambiguities arising in the discussion of the results. The techniques used are principal component analysis, stepwise multiple linear regression and spectral analysis using the maximum entropy method.

SECTION I: THE DATA

2.2 MAIZE DATA

Figures for maize production (200 lb. bags) and area planted (morgen) were obtained for 38 magisterial districts from the agricultural censuses for the period 1950 to 1965 inclusive. The length of series to be analysed may seem rather short but the data available prevented an extension of the time interval. Records after 1965 were incomplete and it was decided that interpolation of the missing values in the series was undesirable, in what is intended to be an exploratory exercise. In considering data prior to 1950 it was evident that fewer magisterial districts could provide the necessary maize figures. Therefore rather than lose spatial coverage by using data before 1950, and not wishing to interpolate gaps in the maize yield series after 1965, it was decided to limit the analysis to the period

1950-1965 inclusive.

The maize production and acreage figures were converted to yields (200 lb. bags/morgen), providing a (16 x 38) input data matrix. This maize yield matrix will be subjected to principal component analysis for the purpose of reducing the number of variables to be analysed and in order that uncorrelated homogeneous maize growing regions will be formed. It is for these maize regions that rainfall and temperature data were sought.

2.3 RAINFALL DATA

It is appropriate to use rainfall records as published in Part 10 of Climate of South Africa (Weather Bureau,1972), because government establishments and commercial concerns make common use of officially defined regions. The records provide monthly mean rainfall series over relatively small areas for the period 1950-1965 inclusive. Further details relating to the selection of rainfall series for each maize region will be given later, once these maize regions have been defined.

2.4 TEMPERATURE DATA

A limited number of temperature recording stations provide continuous records for the period 1950-1965. Since temperature is a conservative element over large areas (Williams *et al*, 1975), particularly when the topography shows little relief, it was decided that a single temperature station for each maize region would be sufficiently representative. Mean monthly values of temperature were extracted from the annual reports published by the Weather Bureau (Weather Bureau, 1949-1965). Complications arose in the selection of temperature data due to the fact that both 08h00 and 14h00 recording times were available. As with rainfall data, details of the procedure adopted for choosing the best series will be furnished once the maize regions have been defined.

In the case of both rainfall and temperature, the seasonal year (July to June) is used since this is relevant to plant growth and hence to maize production. These monthly values of rainfall and temperature constitute the set of independent variables in multiple regression models where annual yields are the dependent variables. This exercise is carried out for each maize region.

SECTION II: ANALYTICAL TECHNIQUES

2.5 PRINCIPAL COMPONENT ANALYSIS

Principal component analysis is a mathematical technique for the rearrangement of the variance in a data matrix and its representation by a new set of orthogonal variables called principal components. The broad objectives of principal component analysis in the field of meteorology are:

(a) to reduce the dimensionality of the problem by replacing the measured and intercorrelated variables by a smaller number of uncorrelated (orthogonal) variables and

(b) *to interpret these uncorrelated variables in* terms of recognizable physical processes (Craddock, 1973).

In recent years principal component analysis has proved to be a useful technique in climatological investigations. The foundations were laid by the work of Lorenz (1956), Bagrov (1959) and Gilman (1957). Since then significant advances have been made by studies dealing with the broader climatological aspects of the technique (Colgate, 1975; Craddock, 1965; Craddock and Flintoff, 1970; Craddock and Flood, 1969; Grimmer, 1963; Jakovleva *et al*, 1968; Kidson, 1975; Kutzbach, 1967, 1970; Rinne, 1971; Stidd, 1967 and Wallace and Dickenson, 1972). The majority of the abovementioned references deal with the analysis of pressure and temperature fields, often with the aim of long range forecasting. An excellent review of such work has been given by Craddock (1973).

In this dissertation principal component analysis will be used in a quite different application and as recommended by Dyer (1975c). The technique will be used to form uncorrelated homogeneous maize growing regions thereby reducing the dimensions of the original (16 x 38) data matrix, and hence the number of variables to be analysed is decreased.

2.5.1 General Properties of Principal Component Analysis

Principal components, also referred to in the literature as eigenvectors or empirical orthogonal functions, have a number of important properties:

(a) they are uncorrelated linear functions of the original variables that is,

$$\lambda_i = \sum_{j=1}^{p} l_{ij} x_j \quad , \qquad i = 1, .., p \quad (2.1)$$

where λ_i is a new set of uncorrelated variables; p is the number of variables x and l_{ij} are the eigenvector or principal component constants and are estimated from the data. In particular if the original variables are normally distributed, the new eigenvectors are not only uncorrelated but independent (Kendall, 1975).

(b) the first principal component is that linear combination of the original variables which extracts the largest possible proportion of the variance attributable to these variables. The second eigenvector has the largest quota of the remaining variance, subject to being uncorrelated with the first eigenvector and so on (Kendall and Stuart, 1966).

(c) they are derived as the eigenvectors of either the covariance or correlation matrix (Cattell, 1965). When the former is used, the patterns that emerge reflect the absolute departures and thereby highlight small scale features present in the data set. On the other hand data used in the correlation matrix are expressed as standardized anomalies (that is, covariance divided by the standard deviation), which tend to emphasize *relative variations and thus broader scale patterns.*

The contribution of any component to the total variance of the data field is given by the corresponding eigenvalue, and provides a convenient measure of its relative importance. The cutoff value for eigenvalues contributing a significant reduction in the total variance is frequently taken as unity (Kaiser, 1961; Jolliffe, 1972). Nevertheless the choice is somewhat arbitrary, and in practice one chooses the criterion that best suits the analysis (Ratkowsky and Martin, 1974).

(d) they are not of a predetermined form as with spherical harmonics for example (Kupper, 1972), but rather the pattern developed is a unique function of the data sample to which they apply (Mitchell *et al,* 1966).

2.5.2 Principal Component Analysis on Maize Yields

There are several modes of analysis available as shown in Table 2.

TABLE 2 - VARIOUS MODES OF PRINCIPAL COMPONENT ANALYSIS

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

In this dissertation S mode is used: magisterial districts represent the variables (entities), years the observations (occasions), and the elements of the data matrix are annual maize yields (Rummel, 1970).

It is not uncommon for the eigenvectors from the direct analysis to present a somewhat confusing picture. This is because some of the loadings, that is the correlations between the original variables and the new eigenvectors, are such that their interpretation is ambiguous. What is needed is for the eigenvectors to be brought into some sort of accordance with the original variables; for example, by maximizing or reducing to zero as many of the loadings as possible (Child, 1970). Furthermore, since the factors are uncorrelated, it is desirable that a particular variable is not heavily loaded on more than one component (Kendall, 1975). The mathematical procedure for doing this is the rotation of the eigenvector matrix.

The technique of rotation was really designed for factor analysis; its justification in principal component analysis is that it often obliges by removing unwanted confusion, the type of which has been mentioned earlier. In rotating the eigenvector matrix, we find the orthogonal linear transformations to a number of new factors which represent the best fit relationship between a cluster of variables.

Rotation of the eigenvectors has been performed using the varimax method of Kaiser (1958). The IBM Scientific Subroutine VARMX has been used for this operation. The eigenvectors are rotated in such a way as to maximize the sum of the variances of the squared loadings within each column of the rotated loading matrix. The general effect is to produce large,or small loadings and to avoid intermediate values. This serves to remove ambiguities from the eigenvector matrix.

The principal component analysis was therefore, usefully employed in reducing the complexity of the original (16 x 38) maize yield matrix. Rotation of the eigenvector matrix brought about further clarification of maize growing districts. By plotting the factor loading of most importance to each magisterial district, delimitation of an area was made possible. Using this method uncorrelated homogeneous maize growing regions were defined.

2.6 THE COMPONENTS OF ANNUAL MAIZE YIELDS

A two-way fixed unreplicated linear model has been fitted to sets of homogeneous time series for groups of districts, as defined by the principal component analysis. Thus for each maize growing region, the input matrix consisted of s columns, where s is the number of districts in that region; and n rows, where n is the number of years in the series.

This technique makes it possible to investigate the behaviour of the three components of which annual yields are assumed to consist. The model can be defined as

$$x_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \; ; \quad i=1,.., n \; ; \; j=1,.., s \quad (2.2)$$

where x_{ij} is the annual maize yield for the i-th year and the j-th district; μ is the overall mean annual yield for a region containing a set of s districts with records of annual yields for a period of n years; α_i is the i-th level of the regional temporal effect series (A); β_j is the j-th level of the regional temporal spatial effect (B) and ε_{ij} is the micro or residual effect, defined as the N $(0,\sigma^2)$ error variable. The interaction $\alpha_i \beta_j$, between the main effects A and B is assumed to be zero (Afifi and Azen, 1972).

2.7 MULTIPLE LINEAR REGRESSION

Multiple linear regression provides a measure of the inter-relationship between a set of independent (predictor) variables and the dependent variable (predictand). There are 2 types of relationship between independent and dependent variables:

deterministic and structural. The latter will be used because the former infers a perfect relationship which does not apply to the data being analysed. A structural relationship accommodates this lack of perfection by including what is commonly referred to as an error term.

The technique of multiple linear regression is therefore, based on a structural relationship between Y_i some dependent variable and a number of x_{ik}'s the independent variables, that is:

$$Y_i = \alpha_o + \sum_{j=1}^{p} \alpha_k x_{ik} + \varepsilon_i \quad , \quad i = 1, .., n \qquad (2.3)$$

where p is the number of independent variables in the regression, the α's are the weights or regression coefficients to be determined and ε_i is the i-th residual on Y_i . These coefficients are determined by least squares. In the present analysis the dependent variable is annual maize yield and the independent variables include monthly values of rainfall and temperature, and a term for ordered temporal variations. Using this technique the internal dependence structure between yield, rainfall, temperature and time will be investigated.

Conventional multiple regression pulls into the regression the full set of predictor variables made available. Some of these variables may however, be redundant causing the regression to be inefficient. Consequently there are a number of procedures for selecting the *best* regression function. Stepwise regression is one such procedure. As new variables are incorporated into the regression model those variables included at earlier stages are re-examined. A variable which may have been the best single predictor to enter at an early step may later become superfluous due to the relationship between it and other variables now in the regression (Smillie, 1966).

The order of stepwise insertion of variables into the regression equation is determined by their partial correlation coefficients, which take as given those variables already in the model. *The squares of the partial correlation coefficients give* the proportion of the variation in Y unaccounted for by those variables already in the regression. Hence the variation introduced next into the regression function is that one which accounts for the largest proportion of the remaining variance (Draper and Smith, 1966).

The exact sequence of inclusion or deletion of a variable is therefore:

(i) if the variance contribution of the variable in the regression is insignificant at a specified level of significance, this variable is removed from the model. If no variable is to be removed then the alternative is

(ii) if the variance reduction obtained by adding a variable to the regression is significant at a specified level of significance, this variable enters the regression. This process is continued until no more variables are admitted to the model and no more rejected.

Straight forward multiple regression was first used in agroclimatological studies by R.A. Fisher in his classic study of the influence of rainfall on wheat yields at Rothamsted (Fisher, 1924). Since then regression techniques have been widely used in crop forecasting models as has been shown earlier. Few research workers have used the stepwise approach in trying to find the optimal yield predictors. Unfortunately where the technique has been employed the description of the methodology is unclear and generally confusing, for example Baier (1973 a,b) and Williams *et al* (1975).

In the present dissertation stepwise regression is used to examine structural relationships between maize yields and climate in the Orange Free State. The technique produces an economical subset of rainfall and temperature values and time, as the set of predictor variables, based on their combined effects on the predictand, annual maize yields.

2.8 THE EXAMINATION OF RESIDUALS FROM THE REGRESSION

The residuals from the regression are defined as the n differences,

$$\varepsilon_i = Y_i - \hat{Y}_i \quad , \qquad i = 1, .., n \qquad (2.4)$$

where Y_i is an observation and $\hat{Y}_i$ is the corresponding predicted value obtained using a fitted regression equation. From this definition we can see that the residuals ε_i are the differences between what is actually observed and what is predicted by the regression equation: that is, that amount of the variance which the regression equation has not been able to explain (Draper and Smith, 1966). Thus the ε_i can be thought of as the observed errors if the model is correct.

These residuals from the regressions will be analysed using the maximum entropy method devised by Burg (1968, 1970, 1972). *The technique involves the maximization of the energy* content of a time series yielding the equation for a power spectrum,

$$P(f) = \frac{2 \Delta P_N}{\left| 1 + \sum_{k=1}^{N} \gamma_k \exp(-i2\pi f k \Delta) \right|^2} \qquad (2.5)$$

where Δ is the sampling interval, f is *the frequency, P the* power or spectral density; P_N is the mean output power of the series after whitening by a filter of length N and γ_k is the *k-th coefficient of the N - point prediction error filter* (Ulrych *et al*, 1973; Smylie *et al*, 1973).

The structure of Eq.(2.5) eliminates resolution constraints in short series because the frequency response of a digital filter can be computed for arbitrary values of frequency (Currie, 1973 a,b ; 1974 a,b,c). The maximum entropy method also provides an *unbiased estimate of spectrum shape because no fixed smoothing* windows are applied (Lacoss, 1971; Currie, 1974 a). Finally in contrast to an unsmoothed fast Fourier transform spectrum, maximum entropy spectra are smoother and more stable.

Limitations of the maximum entropy method are that it is computationally expensive compared to a fast Fourier transform, and that no objective criterion is available for truncation of filter length N (Currie, 1974 c). However, the same weakness is also true of conventional spectral techniques (Jenkins and Watts, 1968). In addition, practice shows that amplitude estimates tend to be biased (Currie, 1974 c).

The purpose of examining the residuals is to investigate that fraction of the variance in maize yields that has not been accounted for by regressions. The spectra produced from the analysis will indicate the presence or absence of any oscillatory patterns present in the yield series. For practical purposes, a trigonometric term would need to be included into the regression model, should such oscillatory patterns be present and significant.

* * *

The techniques outlined above provide an objective and a fresh approach to the analysis of maize yields in the Orange Free State, when records of annual totals for 38 magisterial districts are considered as a set of time series. Furthermore the applicability of the techniques is not spatially restricted, and should prove of value over other areas of the earth's surface; nor is their suitability limited to the analysis of maize yields, but also many types of agricultural data would respond well to these techniques. The results of the analyses of maize yields and climate are next presented.

CHAPTER 3

THE FORMATION OF HOMOGENEOUS MAIZE REGIONS

3.1 INTRODUCTION

Principal component analysis has been used to replace the individual maize yield series with a new set of uncorrelated variables, which account for most of the variance in the original yield data. It provides an objective method for delimiting maize growing regions within the Orange Free State, and has the added advantage of reducing the number of variables to be analysed. Vector rotation brought about further clarification of the newly derived orthogonal variables.

By fitting linear two-way unreplicated models to the results obtained from the principal component analysis, these maize regions could be represented by single uncorrelated time series. These models are effective in decomposing the individual series into three effects, namely, a regional temporal effect, a regional spatial effect and a residual or micro-effect. Subsequent analyses were carried out on the regional temporal effect for each maize growing region.

3.2 PRINCIPAL COMPONENT ANALYSIS

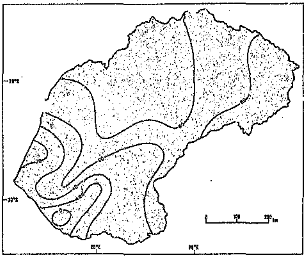
A principal component analysis has been carried out using the correlation matrix on the maize yield series of the 38 magisterial districts; the input matrix over the period 1950-1965 was thus of dimension (16 x 38). The criterion frequently used for the truncation of eigenvectors is that of Kaiser (1961) which rejects eigenvalues less than unity as a cutoff. Using this criterion we find that the first 8 eigenvalues account for 93.25 per cent of the total variance in maize yields (Table 3.1). Kendall (1975) advises strongly against employing rule-of-thumb methods for cutoff values. Rather an intuitive value should be taken after investigating a number of possible values.

TABLE 3.1 VARIANCE EXTRACTED BY EACH OF THE FIRST 8 EIGENVALUES TOGETHER WITH THE ASSOCIATED CUMJLATIVE PERCENTAGE OF THE TOTAL VARIATION IN MAIZE YIELDS.

Eigenvalue number	Variance extracted	Cumulative percentage of total variance
1	17.70	46.57
2	7.09	65.24
3	2.94	72.97
4	2.22	78.81
5	1.92	83.85
6	1.36	87.43
7	1.20	90.59
8	1.01	93.25

Since the object of this exercise is to form homogeneous and uncorrelated maize growing regions, it seems fit to explore the various possible groupings, before deciding on the truncation point for eigenvalues. Thus it would be useful to be able to depict spatial patterns in maize yields that may occur in the Orange Free State. The easiest way of doing this is to map the coefficients or loadings of the eigenvectors for each district. This will show the level of correlation between each district's maize yields and the newly determined orthogonal variables. Using the Burt-Banks formula (Child, 1970), loadings greater than 0.50 are significant.

The contour field of the first eigenvector shows little relief over most of the province with loadings decreasing towards the south and east (Fig. 3.1 upper). This pattern is similar to that



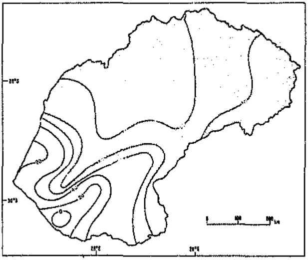


Fig.3.1 - Contours of the first eigenvector show little relief over most of the Orange Free State (upper); but for areas in the southwest, this component shows significant association with maize records over most of the province (lower).

found in the contour pattern of mean yields in the Orange Free State (Fig. 3.2), a relationship suggested before by Craddock and Flood (1969). The first eigenvector relates to mean yields which are significantly common to the northern and central parts of the province (Fig. 3.1 lower). The northeastern half of the Orange Free State is positively correlated with the second eigenvector; negative correlation occurs in the southwest (Fig. 3.3 upper). Areas with significant loadings on this component are shown in Fig. 3.3 lower. The third principal component shows spatial compensation, similar in extent to that of the second eigenvector, between areas of positive association in the northern and southern regions, and areas of negative association in the east and west (Fig. 3.4 upper). On a broadscale this pattern is similar to that of the overall variance field; zones of greater deviation in the north and south being separated by an east-west corridor of lesser variance (Fig. 3.2). Areas bordering Lesotho and southern districts show a particular affinity for the third eigenvector (Fig. 3.4 lower). The pattern of the fourth principal component exhibits more localized features than lower order eigenvectors (Fig. 3.5 upper), yet significant peaks and troughs are still present (Fig. 3.5.lower).

Eigenvectors above the fourth revealed more localized, random patterns rather than depicting regional ones. The purpose of the principal component analysis was to remove redundancies from the data matrix, thereby reducing the number of variables to be analysed, and to form groups of districts based on their similar behaviour patterns. Hence it is clear that localized features must be avoided. The provisional eigenvalue cutoff of 1.0 that was used earlier, is thus no longer appropriate; the number of eigenvectors to be used will be limited to 4, accounting for 78.8 per cent of the total variation in maize yields.

The above figures have proved helpful in analysing maize yields. However, because the loadings are not always clearly defined, that is either high or low, it was felt that eigenvector rotation might further clarify the delimitation.

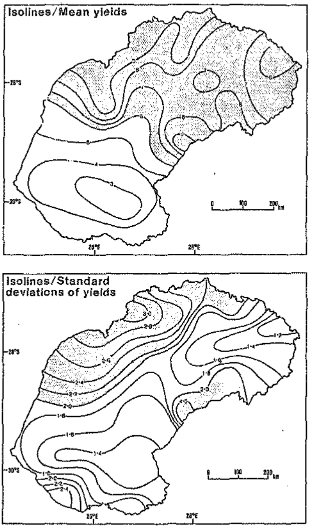


Fig.3.2 - Contour patterns where isolines are mean yields (upper) and variance yields (lower) of the Orange Free State. Areas of strong association ░░░ reveal the broad similarities to the first and third eigenvectors respectively.

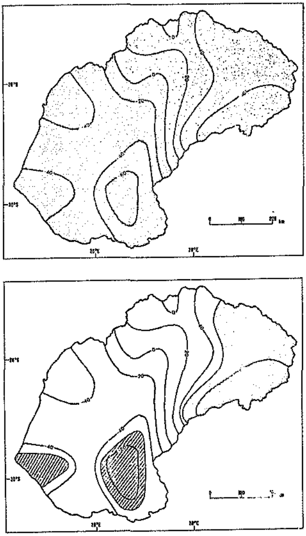


Fig.3.3 - The northeastern districts a.e positively correlated with the second principal component , but negatively correlated in the southwest (upper). A significant ridge of high loadings occurs in the northeast and troughs are found in the southeast and southwest ///// (lower).

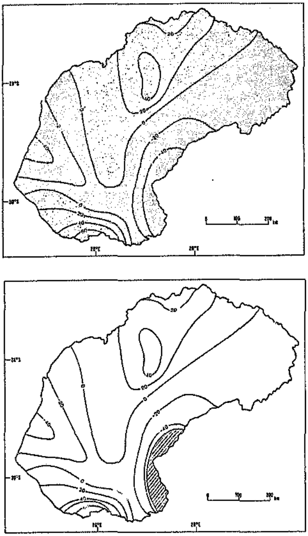


Fig.3.4 - On the third principal component, zones of positive association in the north and south are compensated for by areas of poor correlation in the east and west (upper). Significant coefficients are found bordering Lesotho ///// and *in the south* (*lower*).

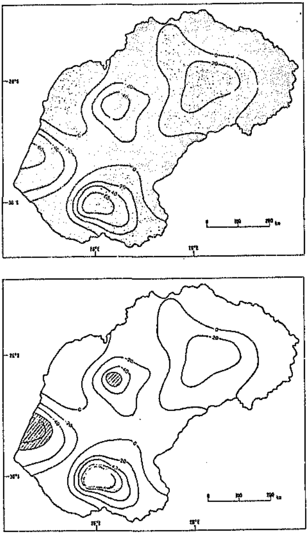


Fig.3.5 - The pattern exhibited by the fourth component between areas of positive [pattern] and negative association [pattern] is fairly localized (upper). Significant troughs ///// and peaks [pattern] however are still found (lower).

3.3 ROTATION OF THE EIGENVECTORS

The main purpose in rotation is to remove, or at least minimize, ambiguities that may have arisen in the straight-forward analysis. In other words, eigenvectors with mediocre loadings will be transferred into either a high or a low coefficient category. An important property of rotation is that the total percentage of variance extracted remains the same on rotation; it is merely redistributed amongst the eigenvectors to enable clarification of the contour field. In this study 4 eigenvectors have been rotated.

Mapping the loadings of the first rotated eigenvector reveals greater spatial contrast (Fig. 3.6 upper), especially the north-eastern area which is positively correlated with this component, again at a level where coefficients greater than 0.5 are significant (Fig. 3.7 lower). Rotation of the second eigenvector produces a contour field, shown in Fig. 3.7 upper, which seems to be less coherent than that produced in the direct analysis (Fig. 3.3 upper). The zones of significant association however, have undergone a pleasing transformation, with the southwestern half of the province now having a strong affinity for this component (Fig. 3.7 lower). The third rotated eigenvector reveals a marked increase in the area of positive correlation (Fig. 3.8 upper) with especially high loadings restricted to the south and a lobe to the north (Fig. 3.8 lower). Rotating the fourth component clarifies areas of localized opposed correlation (Fig. 3.9 upper). Areas warranting specific consideration occur in the south and west (Fig. 3.9 lower).

The results shown in Fig. 3.6 to Fig. 3.9 inclusive suggest that groupings of magisterial districts on a basis of maize yields is possible. This suggestion hinges on the alignment of districts from different parts of the Orange Free State with certain specific eigenvectors. Each yield series, that is for every magisterial district, has 4 rotated eigenvector loadings. For each district, the eigenvector with the highest coefficient has been noted and mapped (Fig. 3.10). The result was extremely enlightening and, on a basis of their similar affinity for a particular eigenvector, the formation of homogeneous groups of districts becomes clear.

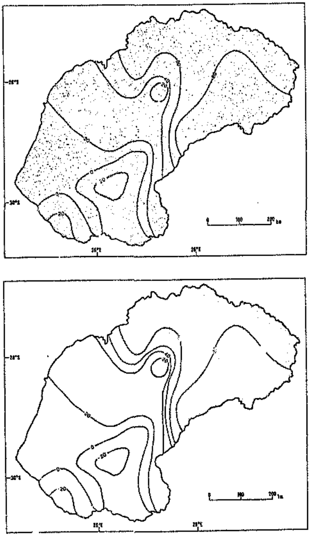
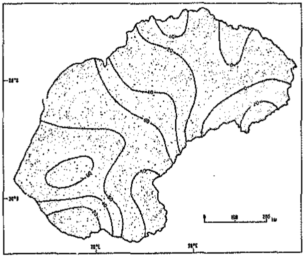


Fig.3.6 - The first rotated eigenvector shows improved spatial contrast between areas of high and low correlation (upper). Significant association is limited to the northeast (lower).



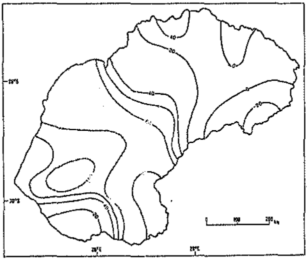


Fig.3.7 - The second factor shows little relief [pattern] but for pockets of negative correlation in the northeast [pattern] (upper). Apart from a zone in *the south, high loadings are ubiquitous in the* southwestern half of the province [pattern] (lower).

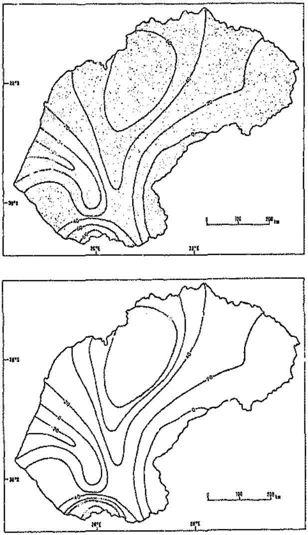


Fig.3.8 - Rotation of the third eigenvector has increased the area of high correlation relative to zones of weak association (upper), and has further clarified significant loadings in the north and south (lower).

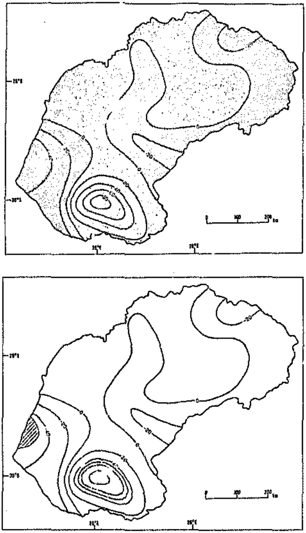


Fig.3.9 - Areas of high [stippled] and low association [shaded] have been damped by the fourth factor (upper). Localized areas of significantly high loadings are still present (lower). [hatched] /////.

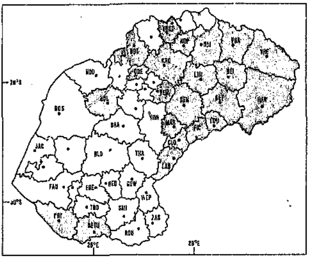


Fig. 3.10 - Showing uncorrelated homogeneous maize growing regions based on their strongest affinity for certain rotated eigenvectors.

- First rotated eigenvector
- Second rotated eigenvector
- Third rotated eigenvector
- Fourth rotated eigenvector
- Magisterial districts with incomplete maize yield records

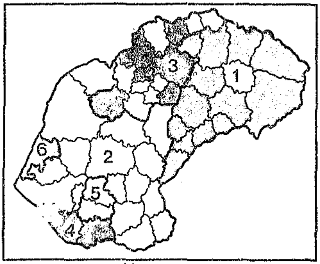


Fig. 3.11 - Showing homogeneous maize growing regions; numbering of regions is based on their unique geographical situations rather than the dominant rotated eigenvectors within a group of districts.
Blank areas depict districts for which complete records were unavailable.

Unique geographical groupings were found for components 1 and 2. For each of the third and fourth rotated eigenvectors, there were two spatially distinct groupings. This made it necessary to refer to 5 spatially distinct maize growing regions. Where possible these maize regions are referred to by the corresponding factor number, but where this is not appropriate, additional numbers have been included (Fig. 3.11).

Having formed homogeneous groups of maize regions it would be advantageous if each group could be represented by a single time series. This would greatly simplify subsequent analyses. By fitting a linear two-way unreplicated model to the maize yield series for each district within a region, it is possible to isolate the components of the annual yield series. These components are: a *regional temporal effect*, a *regional spatial effect* and a micro or residual effect for each region. It is the temporal effect that is of interest to our analyses.

3.4 FORMATION OF REGIONAL TIME SERIES FOR YIELDS

A two-way analysis of variance was performed on the records for each district within the 5 regions. Region 6 has been excluded as this is a single district and therefore has a unique time series. The analysis of variance table for Regions 4 and 5 is not presented here because of problems arising from too few degrees of freedom.

The linear model used in this analysis has been defined earlier as

$$x_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \quad , \quad i=1,..,n \ ; \ j=1,..,s$$

where x_{ij} is the annual maize yield for the i-th year and the j-th district; μ is the overall mean annual yield for each *region*; α_i is the i-th level of the regional temporal effect series; β_j is the j-th level of the regional spatial effect and ε_{ij} is the micro-effect. The analysis of variance table for the temporal effect is shown in Table 3.2; the null hypothesis being H_0 : all $\alpha_i=0$.

TABLE 3.2 - SUMMARY OF THE ANALYSIS OF VARIANCE OF YIELDS FOR AN UNREPLICATED FIXED TWO-WAY MODEL ON REGIONS 1 TO 3; H_0 : ALL $\alpha_i = 0$ (F is Fisher's F-statistic; ν_1 are degrees of freedom of numerator; ν_2 are degrees of freedom of denominator).

Region	F	ν_1	ν_2
1	21.72	15	195
2	19.86	15	180
3	24.45	15	75

The results demonstrate that the regional temporal effect is significantly different from zero in all regions, and the null hypothesis, H_0 : all $\alpha_i = 0$, can be rejected at the 0.01 level of significance. The mean temporal series for Regions 1 to 5 are shown in Table 3.3.

TABLE 3.3 - MEAN TEMPORAL SERIES FOR REGIONS 1 TO 5.

Year	Region				
	1	2	3	4	5
1949-50	-0.64	-1.23	-2.23	1.29	1.72
1950-51	-1.11	0.46	-0.62	-0.30	1.71
1951-52	-1.59	-2.51	-4.79	-1.46	-0.44
1952-53	0.06	-0.50	-0.77	2.21	-1.06
1953-54	-0.27	0.48	0.56	0.37	-1.02
1954-55	-1.08	-1.00	-0.30	-1.03	1.99
1955-56	-0.97	0.33	-0.87	0.75	0.12
1956-57	-0.88	0.89	1.18	3.12	0.67
1957-58	-3.05	-0.16	-0.71	0.64	-1.22
1958-59	-1.54	1.03	-1.80	-0.99	-0.13
1959-60	0.47	0.42	-0.45	-2.37	1.40
1960-61	1.51	-0.38	1.34	-2.52	-0.31
1961-62	2.61	1.91	4.29	-0.12	-0.63
1962-63	3.57	3.06	4.35	0.86	1.25
1963-64	0.97	-1.63	-1.35	-2.63	-1.98
1964-65	1.94	-1.17	2.19	2.18	-0.72

The rejection of the null hypothesis implies that year-to-year variations are not due to chance but that samples have been taken from different populations of annual maize yields.

* * *

Principal component analysis has been employed efficiently in reducing the complexity of the data matrix to be analysed. By mapping the strongest affinity of each district for a certain factor (that is, the highest factor loading), a number of uncorrelated homogeneous maize growing regions evolved. A two-way fixed effects unreplicated linear model was then fitted to the time series within a region, thereby determining regional time series representative of the districts within that group. These new series are used in subsequent analyses and are referred to as the regional temporal series.

CHAPTER 4

MAIZE-CLIMATE RELATIONSHIPS

4.1 INTRODUCTION

Regional temporal series have been determined for each homogeneous maize region using a two-way unreplicated linear model. These new yield series are used with the appropriate rainfall and temperature data in a study of maize-climate interrelationships. Stepwise multiple regression is a useful technique for elucidating such interrelationships, and in the present study shows that a relatively high percentage of the variation in maize yields is attributable to climatic factors and ordered temporal variations.

4.2 DATA USED IN THE MULTIPLE REGRESSIONS

4.2.1 Rainfall

It was mentioned earlier that rainfall records for officially defined regions (Weather Bureau, 1972) would be used, and that these rainfall series were available for relatively small areas of the Orange Free State (Fig. 4.1). From Fig. 4.1 it is evident that a number of rainfall series could be used, either as individual series or as the mean of several individual series, as possible predictors of maize yields within each region. By exploring in the regression all possible combinations of rainfall series for a particular region, this problem was overcome: that rainfall series accounting for the largest amount of variance in maize yields in each region was chosen. Where there was little to choose between having an individual or combined series, the former was chosen for ease in the practical application of forecasting

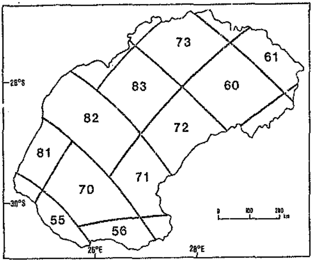


Fig. 4.1- Regional rainfall districts of the Orange Free State, as defined by the South African Weather Bureau (1972).

models to be developed later.

4.2.2 Temperature

Earlier it was stated that single temperature stations were chosen for each maize region and that monthly mean values were available for both 08h00 and 14h00. This latter difficulty was overcome by adopting an approach similar to that used in choosing rainfall series. Both time series of temperature were run through the regression for each region, and once again that series accounting for the largest proportion of the variance in yield was employed in the final regressions. The rainfall and temperature that were finally chosen for each region are given in Table 4.1

TABLE 4.1 - RAINFALL AND TEMPERATURE SERIES USED IN THE FINAL REGRESSIONS. The series were selected on a basis of their contribution to the total variance of regional maize yields (S.A.S.T. is South African Standard Time).

Region	Rainfall Series	Temperature Series
1	Mean 73/60	Bethlehem 08h00 SAST
2	70	Bloemfontein 08h00 SAST
3	83	Kroonstad 08h00 SAST
4	Mean 55/56	Fauresmith 14h00 SAST
5	70	Fauresmith 08h00 SAST
6	81	Fauresmith 14h00 SAST

4.2.3 Temporal Variations

An additional term has been added which encompasses organized miscellaneous factors contributing to yield variations; for example, the use of fertilizers and/or high yielding varieties of maize and other technological advances. For conciseness throughout the dissertation, these factors will be referred to as temporal variations.

Ideally the annual sets of monthly rainfall and temperature values, together with the term for temporal change should be combined to form a set of 25 independent variables. This was not possible however, since the number of years in the yield series (N =16) must be greater than the number of predictor variables (p = 25) if sufficient degrees of freedom are to be available. In the first instance rainfall and temperature would be run through the regression separately. Thus in each case the number of predictor variables would be 13. This exploratory run was useful in providing a guide for choosing the critical rainfall and temperature months for the final regression.

The criterion assumed for entry of a variable into a regression is that the proportion of the variance in yields accounted for by that particular predictor variable must be greater than or equal to 0.01. Using this criterion a maximum of 5 months for each of rainfall and temperature were selected for addition into the final regression. Considering climatic and time dependent factors together, there was no region which contributed more than 6 variables at a 5 per cent level of significance, to the final regression functions.

The validity of the decision to regress yield on each set of predictors separately was checked using a different system of regression. Th[illegible] is based on an iterative method (Veldman, 1967) [illegible] which was not justified because of greater use of compu[illegible]e. This iterative technique pulled into the model the same significant predictor variables as had the earlier method based on the solution of normal equations. The only differences encountered were in those variables that were not significant at the 5 per cent level which relate to random relationships between maize yields and the predictor variables.

Having justified the decision to regress yields on each set of predictor variables separately, it is appropriate to present the results of the individual runs on rainfall and temperature. The results of the final regressions once rainfall and temperature months have been combined will be finally presented.

4.3 REGRESSION WITH RAINFALL AND TIME

In Region 1 maize yields are determined largely by September rainfall, which accounts for 26.6 per cent of the total variance in yields (Fig. 4.2). Further south in Region 2 it is the midseason months from November to January that are critical to maize yields. The northern districts comprising Region 3 show that no single month's rainfall accounts for a marked portion of the variance in yields, but that months from January onwards contribute to yield variations. Region 4 exhibits the seemingly anomalous result where the month accounting for the largest percentage of the variance is not the best predictor; October rain accounts for 22.8 per cent of the variance and is pulled into the regression fifth while May and July rainfall, accounting for 14.9 and 16.5 per cent respectively, are the best predictors of maize yields. This is an advantage of using stepwise rather than conventional multiple regression. Region 5 is similar to Region 4 in that October rain again accounts for a substantial amount of the variation in yields, namely 27.1 per cent; it differs from Region 4 in being the most valuable predictor month. The south western part (Region 6) is unique in that May and June's rainfall are most significant. Since the final yields of maize are determined in part by the extent of moisture retention or loss from the kernel, it is possible that by harvesting later in the season in Region 6 the full effect of rain during May and June would be felt. A summary of the variance accounted for by significant rainfall months and time is provided in Table 4.2.

4.4 REGRESSION WITH TEMPERATURE AND TIME

When considering temperature and time as the set of predictor variables in the regression, it is seen that yields of Region 1 show some degree of dependence on both early (August) and late season (April) temperatures (Fig. 4.3). The temperature pattern for Region 2 is rather similar to that of rainfall; in both a single month is critical for determining yields. November's rainfall accounted for 35.2 per cent of the variance in yields; here temperatures during March account for 38.6 per cent of the variation

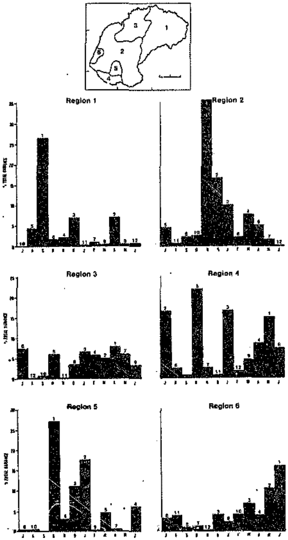


Fig.4.2 - Contribution of different rainfall months to the total variance in maize yields for each region. Numerals denote the order in which predictor months were pulled into the regression.

TABLE 4.2 - SUMMARY OF THE CONTRIBUTION OF RAINFALL MONTHS AND TIME IN ACCOUNTING FOR THE VARIANCE IN MAIZE YIELDS FOR ALL 6 MAIZE REGIONS. PREDICTORS LISTED IN THE ORDER IN WHICH THEY WERE DRAWN INTO THE REGRESSIONS.

Region	Variable	MULTIPLE CORRELATION	
		Coefficient (R)	R^2 (%),variance
1	Time	0.65	42.2
	September	0.83	68.8
	April	0.87	75.7
	December	0.91	82.9
	November	0.92	85.3
2	November	0.59	35.2
	December	0.73	52.9
	January	0.79	62.1
	March	0.85	72.6
	July	0.88	77.2
3	Time	0.61	37.3
	April	0.67	45.1
	March	0.71	50.1
	January	0.75	56.5
	February	0.79	62.2
4	May	0.39	14.9
	July	0.56	31.4
	January	0.69	47.9
	April	0.75	56.4
	October	0.89	79.2
5	October	0.52	27.1
	January	0.67	44.6
	Time	0.77	59.0
	December	0.84	70.2
	June	0.87	76.4
6	June	0.40	16.3
	May	0.52	26.8
	March	0.58	33.6
	April	0.62	38.3
	Time	0.66	42.9

in yields. Whereas rainfall in Region 3 was not critical to yields this pattern is counterbalanced by the importance of temperature. December temperatures are a striking limiting factor to maize yields in this region and account for 46.8 per cent of the variance. Region 4 reveals patterns for temperature similar to those obtained for rainfall. While there is no single month that accounts for a sizeable amount of the variance in yields, September temperature accounts for 25.9 per cent of the variance and late season months (April and May combined) account for a similar percentage of yield variations. Region 5 reflects certain of the properties described for Region 2, in that both mid and late season temperatures are effective in limiting maize yields; in particular, May and December temperatures account for 25.? and 16.2 per cent respectively, of the variation in yields. As with rainfall patterns, temperatures of the late season months are dominant in Region 6. The contributions of temperature months and time to the total variance in maize yields are shown in Table 4.3.

4.5 FINAL REGRESSION OF YIELDS ON RAINFALL, TEMPERATURE AND TIME

This set of predictor variables was based on the rainfall and temperature months selected from the exploratory series of regressions. The temporal term was included in the final regression, only if it played a significant role in the individual regression runs using rainfall or temperature. As mentioned earlier the maximum number of predictor variables pulled into the final regression was 5. The contribution of each predictor to the coefficient of determination can also be expressed as the percentage contribution to the total variance in yields that is accounted for. The latter notation will be used in the discussion which follows. Using Fisher's F-statistic to assess significance we find that variables entering the regression are significant at the 5 per cent level.

Yields from Region 1 show a definite response to the climatic behaviour in the early part of the growing season. It is evident from Table 4.4 that time accounts for 42.2 per cent of the variance in maize yields.

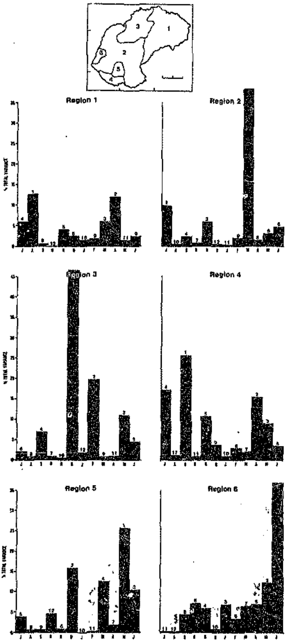


Fig.4.3 - Contribution of different temperature months to the total variance in maize yields for each region. Numerals denote the order in which predictor months were drawn into the regression.

TABLE 4.3 - SUMMARY OF CONTRIBUTION OF TEMPERATURE MONTHS AND TIME TO THE TOTAL VARIANCE IN MAIZE YIELDS FOR REGIONS 1 TO 6. PREDICTOR VARIABLES ARE LISTED IN THE ORDER IN WHICH THEY WERE DRAWN INTO THE REGRESSION.

Region	Variable	MULTIPLE CORRELATION	
		Coefficient (R)	R^2 (%), variance
1	Time	0.65	42.2
	August	0.74	55.0
	April	0.82	66.8
	March	0.85	72.9
	July	0.89	79.0
2	March	0.62	38.6
	July	0.70	48.8
	Time	0.84	71.2
	November	0.88	77.4
	September	0.89	79.8
3	December	0.68	46.8
	February	0.82	66.5
	May	0.88	77.5
	September	0.92	64.5
	June	0.94	89.1
4	September	0.51	25.[illegible]
	April	0.65	41.6
	May	0.70	49.6
	July	0.82	66.9
	November	0.88	77.5
5	May	0.51	25.8
	December	0.65	42.0
	June	0.74	55.0
	March	0.82	67.6
	July	0.85	71.5
6	June	0.61	37.0
	May	0.70	49.4
	January	0.75	56.2
	November	0.79	62.1
	April	0.83	69.1

TABLE 4.4 - CONTRIBUTION OF RAINFALL, TEMPERATURE AND TIME TO THE TOTAL VARIANCE IN MAIZE YIELDS OF REGION 1. R_n refers to rainfall; T to temperature, (R^2 is cumulative).

Variable Entered	Multiple Correlation		Increase in R^2 (%)	F. Value to enter
	Coefficient(R)	R^2(%),variance		
Time	0.65	42.2	42.2	10.24
Sept.(R_n)	0.83	68.8	26.5	11.04
Aug. (T)	0.89	78.4	9.6	5.35
Nov.(R_n)	0.93	87.2	8.8	7.50
Dec.(R_n)	0.96	92.5	5.4	7.19

September rainfall and August temperature account for 26.5 and 9.6 per cent of the variation in yields respectively. Fig.4.4 provides a pictorial representation of these relationships.

The criterion by which variables are entered stepwise into the regression was said earlier to be the maximum amount of extra variance accounted for. For this we look at the partial correlation coefficients. In the case of Region 1 time was pulled into the regression first because its simple correlation coefficient (r) is greatest. Using Table 4.5 as a guide it is seen that September rainfall has the highest partial correlation coefficient at Step 2 (r_p = -0.678) and therefore enters the regression second. At Step 3, August temperature has the largest r_p and therefore is pulled into the equation third and so on. The order in which variables are pulled into the regression denotes their importance as predictor variables of maize yields, when combined with those already in the regression.

A perusal of the simple correlation matrix for Region 1 shows that only simple correlations for time and April temperature with yields are significant at the 95 per cent level (Table 4.6). Had we not examined the partial correlation coefficients, the predictive qualities of September, November and December rainfall

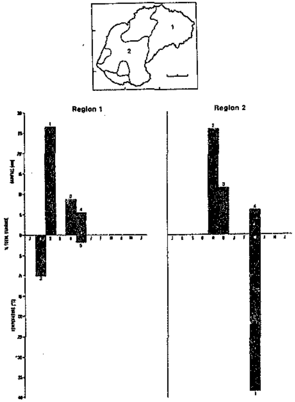


Fig.4.4 - Relative contribution of rainfall and temperature months to the total variance in maize yields for Regions 1 and 2. Numerals refer to order in which predictor months were pulled into the regression.

TABLE 4.5-PARTIAL CORRELATION COEFFICIENTS FOR 4 STAGES OF STEPWISE REGRESSION FOR REGION 1. Time was pulled into the regression first because its simple r was greatest (R_n refers to rainfall, T to temperature).

Variable	Partial correlation coefficients (r_p)			
	Step 2	Step 3	Step 4	Step 5
Aug. (R_n)	-0.1841	0.1605	-0.4230	0.2027
Sep. (R_n)	-0.6777			
Oct. (R_n)	-0.1872	-0.1307	-0.2879	0.1554
Nov. (R_n)	0.3249	0.4233	0.6367	
Dec. (R_n)	-0.1324	-0.2802	-0.6313	-0.6467
Apr. (R_n)	0.1850	0.4698	0.1729	0.0764
Jul. (T)	-0.1223	0.1147	0.0435	-0.0405
Aug. (T)	0.4699	0.5553		
Nov. (T)	-0.1737	-0.0573	0.0280	-0.0168
Dec. (T)	-0.0380	-0.0315	0.0155	0.2621
Mar. (T)	0.0965	-0.1645	-0.2176	-0.3120
Apr. (T)	-0.2807	-0.1168	-0.4076	-0.4915

and August temperatures would have been missed. These months account for a total of 50.3 per cent of the variance in maize yields.

TABLE 4.6- SIMPLE LINEAR CORRELATION COEFFICIENTS (r) OF COMBINED RAINFALL AND TEMPERATURE MONTHS WITH TIME, ON ANNUAL YIELDS FOR REGION 1. (* denotes significance at the 95% level).

Variable	r
Time	-0.65 *
Aug. (R_n)	-0.12
Sep. (R_n)	-0.49
Oct. (R_n)	-0.06
Nov. (R_n)	0.34
Dec. (R_n)	-0.28
Apr. (R_n)	0.35
Jul. (T)	-0.28
Aug. (T)	-0.17
Nov. (T)	-0.22
Dec. (T)	0.10
Mar. (T)	-0.11
Apr. (T)	-0.58 *

In the final model it is interesting to note that April temperature does not emerge as a predictor. This is because the simple correlation coefficients (r) refer to variables in isolation and not in combination with several others. The partial correlation (r_p) on the other hand, takes as given those variables already in the regression and hence describes a combined effect between variables.

The final model for Region 1 is given in Eq.4.1

$$Y = 1.47 \ -0.35 \text{ Time } -0.02 \text{ Sep (R) } + 0.01 \text{ Nov (R) } -0.01 \text{ Dec (R) } + 0.70 \text{ Aug (T)} \qquad (4.1)$$

where Y is the predicted annual maize yield. This regression model accounts for 92.5 per cent of the total variance of maize yields. The overall fit of the equation is significant at the 99 per cent level.

The multiple regression on maize yields of Region 2 shows that climatic factors during the latter part of the season exert an important influence on maize yields. In particular mid and late season temperature months are critical; March temperature accounting for 38.6 per cent of the variance in yields (Table 4.7).

TABLE 4.7 - CONTRIBUTION OF ALL PREDICTOR VARIABLES TO THE TOTAL VARIANCE IN YIELDS FOR REGION 2. R_n refers to rainfall; T to temperature (R^2 is cumulative)

Variable Entered	Multiple Correlation		Increase in R^2(%)	F value to enter
	Coefficient(R)	R^2(%),variance		
Mar.(T)	0.62	38.6	38.6	8.80
Nov.(R_n)	0.81	64.9	26.3	9.73
Dec.(R_n)	0.87	76.5	11.7	5.96
Mar.(R_n)	0.91	82.7	6.2	3.93

Another important contribution in accounting for the variance in maize yields for Region 2, is that of November and December rainfall which account for 26.3 and 11.7 per cent of the variation (Fig. 4.4). The final model which accounts for 82.7 per cent of the total variation in yield (Eq.4.2) does not include any temporal effects on yields.

$$Y = 11.71 + 0.03 \text{ Nov (R)} + 0.02 \text{ Dec (R)} + 0.01 \text{ Mar (R)} - 0.87 \text{ Mar (T)} \quad (4.2)$$

The overall fit of the regression equation is significant at the 99 per cent level.

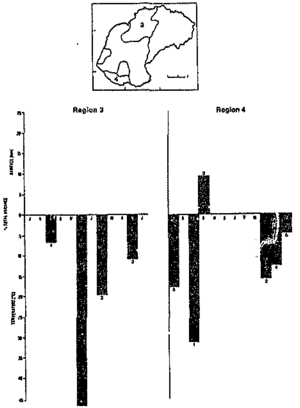


Fig.4.5- Relative contribution of rainfall and temperature months to the total variance in maize yields ; numerals denoting order of importance in the regression, that is, order in which they were taken into the regression.

Maize yields in the northern part of the Orange Free State, that is Region 3, are more sensitive to temperature than rainfall. The effect of December and February temperatures in accounting for variance in maize yields, is by no means insignificant; they account for a total of 66.5per cent ofyield variations (Table 4.8).

TABLE 4.8- THE CONTRIBUTION OF COMBINED PREDICTOR VARIABLES TO THE TOTAL VARIANCE IN MAIZE YIELDS OF REGION 3; R_n relates to rainfall ; T is temperature(R^2 is cumulative).

Variable Entered	Multiple Correlation		Increase in R^2(%)	F value to enter
	Coefficient(R)	R^2(%),variance		
Dec.(T)	0.68	46.8	46.8	12.29
Feb.(T)	0.82	66.5	19.7	7.65
May (T)	0.88	77.7	11.2	6.02
Sep.(T)	0.92	84.5	6.9	4.88

Fig. 4.5 illustrates the relative contributions of the different temperature months to the total yield variation. Although the final regression function contains a mere 4 independent variables (Eq.4.3) it accounts for a sizeable 84.5 per cent of the variance in maize yields.

$$Y = 56.09 + 0.49 \text{ Sep (T)} - 1.20 \text{ Dec (T)} - 1.67 \text{ Feb (T)} - 0.86 \text{ May (T)} \quad (4.3)$$

Interesting compensation between early and late season climatic effects is exhibited by Region 4. The early season effects include July and September temperatures and October rainfall; the late season effects are April, May and June temperatures (Table 4.9). These results may seem rather odd, but bearing in mind that harvesting is up to and including June (Garlipp, 1976), these results are quite acceptable.

TABLE 4.9 - CONTRIBUTION OF RAINFALL AND TEMPERATURE MONTHS *IN ACCOUNTING FOR YIELD VARIANCE IN REGION 4:* R_n denotes rainfall and T refers to temperature (R^2 is cumulative).

Variable Entered	Multiple Correlation Coefficient(R)	R^2(%), variance	Increase in R^2(%)	F value to enter
Sep.(T)	0.51	25.9	25.9	4.90
Apr.(T)	0.64	41.6	15.6	3.48
Oct.(R_n)	0.71	50.7	9.2	2.23
May(T)	0.80	63.4	12.6	3.79
Jul.(T)	0.90	81.0	17.7	9.32

These interrelationships between climatic variables are illustrated in Fig. 4.5. The overall fit of the regression function (Eq. 4.4) is significant at the 99 per cent level and it accounts for 81 0 per cent of the total variance in maize yields.

$$Y = 10.46 + 0.05 \text{ Oct } (R) + 0.69 \text{ Jul } (T) - 0.77 \text{ Sep } (T) -0.89 \text{ Apr } (T) + 0.76 \text{ May } (T) \qquad (4.4)$$

Region 5 shows several similarities to Region 1: both indicate that climatic factors close to planting are of most importance in determining final maize yields. In Region 5 , October , January and December rainfall account for 27.1, 17.5 and 11.2 per cent of the variance respectively while 16.1 per cent of the variation is due to June temperatures (Table 4.10).

Fig. 4.6 illustrates the relationships of climatic variables to each other and their contributions to the total variance in maize yields.

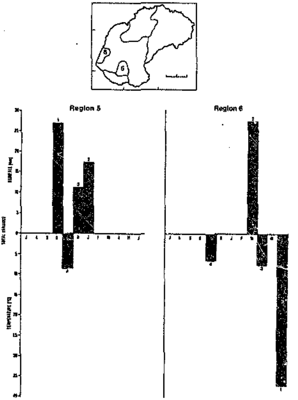


Fig.4.6 Relative contribution of rainfall and temperature months to the total variance in maize yields; numerals denote order of importance,that is, order in which months were pulled into the regression.

TABLE 4.10 - CONTRIBUTION OF ALL PREDICTOR VARIABLES TO THE TOTAL VARIANCE IN MAIZE YIELDS FOR REGION 5; R_n is Rainfall T is temperature (R^2 is cumulative).

Variable Entered	Multiple Correlation		Increase in R^2(%)	F value to enter
	Coefficient(R)	R^2(%),variance		
Oct.(R_n)	0.52	27.1	27.1	5.21
Jan.(R_n)	0.67	44.6	17.5	4.12
Time	0.77	59.0	14.3	4.19
Dec.(R)	0.84	70.2	11.2	4.14
Jun.(T)	0.93	86.3	16.1	11.78
Nov.(T)	0.97	94.9	8.6	15.21

Another similarity between Region 5 and Region 1 is that in both regions, time accounts for a sizeable fraction of the total variation in maize yields (refer to Table 4.4). This clearly serves to increase the coefficient of determination and the overall fit of the model in both regions. The final regression function for Region 5, given in Eq.4.5 explains 94.9 per cent of the total variation in maize yields.

$$Y = 4.09 + 0.08 \text{ Time} - 0.01 \text{ Oct (R)} + 0.02 \text{ Dec (R)} + 0.01 \text{ Jan (R)} - 0.39 \text{ Nov (T)} + 0.42 \text{ Jun (T)} \qquad (4.5)$$

Region 6 is quite different from other regions in that maize yields are quite indifferent to climatic conditions between July and December. Late season rainfall and temperatures contribute 74.4 per cent of the total variation in yields (Table 4.11). This surprising result is illustrated in Fig. 4.6 which enhances the differential contributions of rainfall and temperature to the total variance in yields.

TABLE 4.11 - CONTRIBUTION OF ALL PREDICTOR VARIABLES TO THE TOTAL VARIANCE IN MAIZE YIELDS OF REGION 6; R_n is rainfall; T is temperature (R^2 is cumulative).

Variable Entered	Multiple Correlation		Increase in R^2(%)	F value to enter
	Coefficient(R)	R^2(%),variance		
Jun.(T)	0.61	37.0	37.0	8.22
May (R_n)	0.82	67.1	30.1	11.87
Apr.(T)	0.86	74.4	7.3	3.41
Nov.(T)	0.90	81.0	6.6	3.84

The final model for Region 6 is shown in Eq. 4.6

$$Y = -15.57 - 0.07 \text{ May (R)} - 0.48 \text{ Nov (T)} + 0.44 \text{ Apr (T)} + 1.34 \text{ Jun (T)} \qquad (4.6)$$

This model accounts for 80.9 per cent of the total variation in yields which is the lowest of all the regions but still a sizeable amount of the variance.

* * *

The regression models developed for each region are efficient in accounting for large fractions of the total variance in maize yields, varying from 80.9 per cent in Region 6 to 95.0 per cent in Region 5. It is well known that rainfall and temperature are important to maize yields (Azzi, 1956). Contrary to expectations, the findings of this analysis have shown that the contribution of climatic conditions to yields is not limited to certain particular stages in the growth cycle of the maize plant: for instance, the importance of moisture during the flowering stage of the maize plant has been stressed (Garlipp, 1976; Kiriakov, 1938). Instead this study has revealed that different months of rainfall and temperature are limiting factors to maize yields in different regions of the Orange Free State. *In addition temporal changes*, due to improved farming practice and technology are frequently overemphasized. Of the 6 regions studied only Regions 1 and 5 exhibited temporal variation to be of significance in determining yields; other regions appear to be unaffected by such changes.

CHAPTER 5

MODEL EXTENSION FOR FORECASTING PURPOSES

5.1 INTRODUCTION

Using stepwise multiple regression, linear models for predicting maize yields have been developed for each region. It was shown that climatic factors and time dependent changes did not account for all of the variance in maize yields. By examining the residuals from the regressions it is possible to investigate that fraction of the variance in yields not accounted for by the regression models. Ideally the residuals should be randomly distributed; in the present study the presence or absence of non - random behaviour, and the nature thereof, is examined. The maximum entropy method of spectral analysis will be used for this analysis. The presence of non-random behaviour in the residuals, could have serious implications on yield estimates obtained using these regression models. This reason alone is sufficient justification for the analysis.

It has been shown that only certain of the original predictor months were included in the final regression models for each of the 6 maize regions. In other words, stepwise regression has proved efficient in forming economical subsets of the original independent variables. It would be of advantage if these models could be used to obtain a forecast for the current crop as early in the season as possible. In this chapter the possibility of deriving an equation that provides a forecast in January will be investigated.

5.2 EXAMINATION OF RESIDUALS FROM THE REGRESSION

The resi[illegible] the regression have been subjected to spectral ana[illegible]. The maximum entropy method has been used rather *than conventional* spectral analysis because of its suitability for analysing short time series. Instead of the assumed random distribution of the residuals, significant peaks have been found in the spectra (Table 5.1). The relative importance of peaks is assessed on a basis of a signal to noise ratio greater than 3 (Currie, 1974c). Using this criterion all peaks except three are significant.

TABLE 5.1- WAVELENGTHS AT WHICH PEAKS ARE CENTERED IN THE SPECTRA OF THE TEMPORAL SERIES FOR REGIONS 1 to 6. INSIGNIFICANT PEAKS ARE MARKED THUS *.

Region	Peaks (wavelengths in years)			
1	10.00		3.33	2.31
2	12.00*	5.00	3.33	2.22*
3	8.57		4.00	2.22*
4	12.00	6.67	3.16	2.14
5		6.67	3.53	2.40
6	10.00		4.00	2.22

The clearly defined peaks revealed in the spectra show a regional distribution in their occurrence (Fig. 5.1). Regions 2 and 4 exhibit a weak oscillation with a mean period of 12 years; peaks with wavelengths of 8 - 10 years are found in Regions 1,3 and 6; Regions 2,4 and 5 exhibiting signals with a mean period of 6 years. *Bearing in mind that the length of yield* series is 16 years; oscillations of longer wavelengths have been taken care of in the regressions.

At the higher frequency end of the spectra, an oscillation of 3 to 4 years is ubiquitous. It is especially strong in the

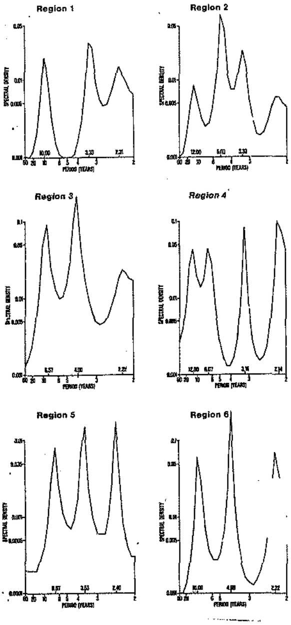


Fig. 5.1- Spectra of residuals from regressions with Regions 1 to 6 showing ubiquitous distribution of the 3-4 year and the quasi-biennial oscillation and spatially dependent oscillations in the longer wavebands.

northern and southwestern parts of the Orange Free State (Regions 3 and 6) where it peaks on 4.00 years. This correlates well with the 3-4 year oscillation in rainfall, the ubiquity of which has been reported by Tyson (1972) and Tyson, *et al* (1975).

An oscillation in the wavelength of 2 to 3 years is also apparent. The peaks in this waverange are strongest in the south and southwest (Regions 4,5 and 6) with signals weakening in a northerly direction. Although the spectral peaks in this waverange for Regions 2 and 3, are not statistically significant, the fact remains that they do exist. It has been argued that peaks in the high frequency region of the spectrum are due to aliasing. In addition to the widespread documentation of the quasi-biennial oscillation (Godin, 1972; Trenberth, 1976; Wright, 1968), the problem of aliasing diminishes when sampling is not from instantaneous points on a continuous time scale. The isolation of the 2 - 3 year oscillation in maize yields accords with the fluctuation of similar wavelength found in South African rainfall series (Tyson,1972; Tyson *et al*, 1975; Dyer, 1975b). It is likely, therefore, that the observed 2 to 3 year oscillation is in some way linked to the biennial oscillation found in various geophysical parameters.

The analysis has shown that the residuals are not random but contain oscillatory components. If the regression models are to be used for practical purposes, such as forecasting current maize yields in the Orange Free State, then it will be necessary to incorporate a trigonometric term into the regression model. The inclusion of such a non-linear term however, goes beyond the bounds of the present dissertation. Instead, use of stochastic models in forecasting yields is limited to the derivation of multiple regression functions based on data from July to December inclusive. This enables a forecast of the potential yield of the current crop to be given in January.

5.3 A METHOD OF MIDSEASON FORECASTING OF MAIZE YIELDS

In an attempt to use statistical models for forecasting purposes, a model for giving estimates in January, has been developed. This has been achieved by including only monthly climatic data between July and December inclusive, as the set of predictor variables in the regressions. In addition to the inclusion of temporal variation as a potential predictor, a new predictor variable has been introduced: the yield of the previous year has been added and may indicate the potential fertility of the land.

It was found that the feasibility of estimating maize yields in January varied from one region to another. In Regions 1 and 3 it is highly likely that the regression given in Table 5.2, could be used for forecasting yields in January. The amount of total variance in maize yields accounted for by the models is 95.4 and 84.6 per cent respectively. In these two models the yield of the previous year contributes substantially to variation in current yields. Thus the introduction of this new predictor variable nullifies the need to include, in Regions 1 and 3, a trigonometric term in the models. December rainfall and August temperature in both cases contribute to the high coefficients of determination (R^2). These results are most worthwhile because considering the Orange Free State as a whole, Regions 1 and 3 are well within the southern boundary of the maize quadrangle (Fig. 5.2). Midseason forecasting in these areas would therefore be most useful.

In comparison with Regions 1 and 3, the results found in Regions 2 and 6 were mediocre. Of the total variation in yields for each region, 50.6 and 61.1 per cent respectively were accounted for by the models (Table 5.2). Although only 2 to 3 predictor variables are included in these models, a fair amount of the variance in yields has been accounted for.

Regions 4 and 5 exhibited the least affinity for midseason forecasting of maize yields. In both regions single predictor months were significant at the 10 per cent level and accounted for 31.6 per cent (Region 4) and 21.6 per cent (Region 5) of the variation in maize yields. Clearly it would not be prudent

TABLE 5.2- LINEAR REGRESSION MODELS FOR FORECASTING CURRENT MAIZE YIELDS IN JANUARY. THE FEASIBILITY OF USING THESE MODELS IN PREDICTIONS IS INDICATED BY THE PERCENTAGE OF THE VARIANCE ACCOUNTED FOR (R^2).

Region	Midseason Linear Forecasting Models	Multiple Correlation	
		Coeff. (R)	R^2 (%)
1	Y = 3.03 - 0.46 Time -0.29 $Yield_{-1}$ - 0.02 Sep (R) + 0.01 Nov (R) - 0.02 Dec (R) + 0.87 Aug (T)	0.98	95.4
2	Y = - 2.93 + 0.03 Nov (R) + 0.02 Dec (R)	0.71	50.7
3	Y = 37.30 + 0.39 $Yield_{-1}$ - 0.03 Dec (R) + 0.89 Aug(T) - 2.01 Dec (T)	0.91	84.6
4	Y = 12.66 - 0.58 Sep (T)	0.56	31.6
5	Y = 0.86 - 0.02 Oct (R)	0.47	21.6
6	Y = - 6.37 + 0.61 Jul (T) - 1.10 Aug (T) + 0.58 Dec (T)	0.78	61.1

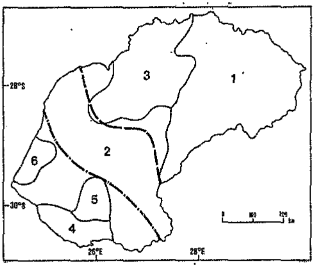


Fig. 5.2 - Generalized boundaries of homogeneous maize regions (1 - 6) in relation to the approximate limits of the maize quadrangle.

— — — is the limit of the major maize growing area.

-·-·-· *border of secondary maize growing area.*

to attempt to forecast yields in January in these regions with the models in their present form. The addition of a trigonometric term however, may enhance these results.

5.4 TESTING THE MODELS

When sufficient data is available it is sometimes divided into two parts; these are not necessarily equal. The first section is usually longer and is used to estimate the parameters in the model; the second, often shorter section, is used for what is commonly known as testing the model. In fact one is not testing the model as this has already been done by considering the coefficient of multiple correlation or the coefficient of determination. Instead it is the stability of the data-generating process that is being tested. This policy which simulates future observations is particularly useful when forecasts for several time units ahead are required (Montgomery and Johnson, 1976). To apply this procedure rigidly when data is sparse however, is neither advisable nor is it necessarily worthwhile.

In this dissertation regression models have been developed for providing a forecast of maize yields in January. In other words, the parameters estimated in the model originated from the period July to December inclusive. Hence using these models a forecast for only the current year can be entertained, because the predictor variables are generated from the previous year. It is reasonable therefore to assume that, provided the coefficient of determination attains an acceptable level, the generating process will remain sufficiently stable for the model to give a reliable forecast. It will also be recalled that a confidence interval can be placed on the expected value of a forecast (Draper and Smith, 1966). In the present situation testing the generating process is considered to be unjustified. The risk being taken is that the underlying generating process may involve a step function leading to unexpected results. In particular, a step function would bring about a sudden alteration in the observed values and not a gradual change. There is no evidence to suggest that this is likely to occur in the data under analysis here. Bearing in mind the small sample size available for testing stability, the risk is therefore worth taking.

*   *   *

It has been shown that it is possible to use multiple regression models with sets of predictor months truncated after December, to provide forecasts of maize yields in January. The most reliable forecasts were found in northern areas situated within the limits of the maize quadrangle (Regions 1 and 3). The accuracy of yield estimates decreased as one progressed southwards; it was suggested however, that the addition of a trigonometric term might well enhance their reliability. It was decided that there was no evidence to suggest that a step function in the generating process may occur, hence testing the accuracy of the estimate on independent data was considered unjustified, considering the small size of the sample.

CHAPTER 6

CONCLUSIONS

The effects of climatic factors and variations thereof on agricultural production have become a popular topic for discussion. This interest stems from a [illegible]er of climatic disturbances in recent years and their effects on man's activities, in conjunction with our ever-increasing world population. Current agro-climatological research is concentrated on the large scale monitoring of crop expectations and the anticipation of food shortages. In South Africa relatively few workers have been concerned with the quantitative prediction of crop yields using climatic data. Those that have, have dealt with single production systems in mathematical models rather than large scale analyses using statistical modelling techniques.

In this dissertation attempts have been made to show the relevance of crop yield analyses in relation to climatic factors over a geographical region in South Africa. The initial objective of the dissertation was to examine structural relationships existing between maize yields and climate in the Orange Free State; secondly to explore the feasibility of using stochastic models to forecast maize yields. The specific results of the investigation are best summarized as follows:

1. Principal component analysis has been used to replace the 38 magisterial districts with a set of orthogonal variables, which account for most of the variance in the original data set. Vector rotation achieved further clarification of the new variables and permitted the delimitation

of 6 homogeneous and uncorrelated maize growing regions within the Orange Free State.

2. These homogeneous maize regions provided the input data matrix for using a two-way fixed effects model to isolate the independent components of annual maize yields. These components are: the regional *temporal effect*, *regional spatial effect* and the local or micro-effect. Only the temporal effect for each region was subjected to further analysis. Using this approach the number of variables to be analysed was effectively reduced from time series for 38 magisterial districts, to single time series for each of the 6 maize regions.

3. For the purpose of studying maize-climate relationships, the temporal series of each region was subjected to stepwise multiple regression. Annual yield was the dependent variable; the predictor variables included monthly values of rainfall and temperature, together with a term for temporal variation.

4. The regression models developed for each region included at most, 6 predictor variables, at a 95 per cent level of significance. These regression functions accounted for a sizeable amount of the total variance in maize yields, the lowest being 80.9 per cent and the highest 95.0 per cent of the total variation in yields.

5. The multiple regression showed that maize yields are strongly related to rainfall and temperature on a monthly basis. Results showed that the contribution of climatic conditions to yields is not limited to certain stages of growth in the maize plant. Instead, different months of rainfall and temperature are important throughout the year in different regions of the Orange Free State.

6. Temporal variations were of significance in determining maize yields in Regions 1 and 5. These time

dependent changes include improved technology such as the introduction of high yielding varieties of maize, and other cultivation practices. On a basis of these results, the effects of temporal changes in determining yield variations can be said to be frequently overrated.

7. The residuals from the regression were analysed using the maximum entropy method of spectral analysis. This analysis provided a means of investigating that fraction of the variance in maize yields unaccounted for by the regression models. The residuals were not randomly distributed but exhibited oscillatory behaviour patterns which were regionally dependent in their occurrence, particularly those of longer wavelengths. A 3 to 4 year signal and the quasi-biennial oscillation were ubiquitous in their spatial distribution.

8. For practical purposes of forecasting yields in the Orange Free State, it is recommended that in some regions a trigonometric term be added to the regression models; this would eliminate the oscillatory components in the yield series and enhance the forecasting accuracy of the regression equation.

9. The feasibility of forecasting in January the current crop yield has been tested, and was found to vary between one region and another. It was evident that this method would be most suited to the north and northeastern parts of the Orange Free State. As one moves further south however, the plausibility of using this approach in predicting yields decreases, and in the far south this method is unacceptable. However, a spectral analysis of the regression residuals has shown that yield estimates in these areas may be improved by the addition of a trigonometric term in the regressions.

* * *

The theme of this dissertation was an agroclimatological analysis of maize yields in the Orange Free State. The results have shown that distinct relationships between climatic factors, namely, rainfall and temperature, temporal variations and maize yields are present on a regional basis. By developing regression models for giving yield estimates in January it has been possible to show that the monitoring of crop expectations is not an over-optimistic goal. A major advantage of using regional time series rather than the numerous individual district records is that the labour involved in the analysis is effectively reduced. Thus, given broadscale crop and climatic data, and employing objective quantitative techniques, there are exciting possibilities for the future development of agroclimatology in South Africa.

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.

Australian Acad. of Science, 1976: *Report of a Committee on Climatic Change.* Rep. No.21, March, Canberra.

Azzi, G., 1956: *Agricultural Ecology.* Constable and Co., London.

Bagrov, N.A., 1959: The analytical representation of the sequence of meteorological fields by means of natural orthogonal components. *Moscow, Tsent. Inst. Prog. Trudy. Vyp.,* 7,3-24.

Baier, W., 1973 a: Crop- weather analysis model; review and model development. *Jour. Appl. Met.,* 12, 937-947.

Baier, W., 1973 b: Crop-weather analysis model; summary. *Intern. Jour. Biometeor.,* 17, 313-320.

Baier,W., 1974: The challenge to agricultural meteorology. *W.M.O. Bull.,* 23, 221-224.

Baier, W., and Robertson, G.W., 1967: Estimating yield components of wheat from calculated soil moisture.

Can. Jour. Plant Sc., 47, 617-630.

Baier, W., and Williams, G.D.V., 1974: Regional wheat yield predictions from weather data in Canada. *Proc. Symp. on Agrometeorology of the wheat crop,* Braunshweig, Fed. Rep. Germany.

Baier, W.,* 1962: Make better use of rainfall. *Farming in S. Afri.,* 38, 47-49.

Baier, W.,* 1963: *Studies on macroclimates and microclimates and their influence on crops with special reference to Potchefstroom and the Highveld region.* Unpub. M.Sc. thesis, Univ. of Pretoria, Pretoria.

Barber. F.H., 1910: Is South Africa drying up? *Agric. Jour.,* Cape of Good Hope, 86, 167-170.

Bhargawa, P.N., Pradhar, A.,and Das, M.N., 1974: Influence of rainfall on crop productic). *JNKVV Res. Jour.,* 8 , 22-29.

Bourke, A., 1975: The resurgence of agricultural meteorology-(editorial). *Agric. Meteor.,* 15, 321-323.

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

Bryson, R.A., 1974: Hoyuppskera: an heuristic model for hay yield in Iceland.*Int. Env. Sci., Report*, No. 22 Univ. of Wisconsin, Madison.

Bryson, R.A., 1973: *World food prospects and climatic change.* Unpub. memorandum. Centre for Clim. Res., Univ. of Wisconsin. Madison.

Bunting, A.H.,*et al,* 1975: Seasonal rainfall forecasting in West Africa. *Nature,* 253, 622-623.

Bunting, A.H.,*et al,* 1976: Rainfall trends in the West African

* different authors

Sahel. *Quart. Jour. Roy. Met. Soc.,* 102, 59-64.

Burg, J.P., 1968: *A new analysis technique for time series data.* Paper presented at NATO Advanced Study Inst. on Signal Processing, Enschede, Netherlands.

Burg, J.P., 1970: *New concepts in power spectra estimation.* Paper presented at the 40th meeting, Soc.Explor. Geophys., New Orleans.

Burg, J.P., 1972: The relationship between maximum entropy spectra and maximum likelihood spectra. *Geophysics,* 37, 375-376.

Buys, M.E.L., 1973: Agroclimatic evaluation in the Rep. of South Africa. Registered projects No (A) S-Pr 50/2 to (A) S-Pr 50/8 of Dep. of Agriculture, S.Afr.

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

Child, D., 1970: *The essentials of factor analysis.* Holt, Rinehart and Winston, London.

Clarke, P.C., 1970: Arctic sea ice in summer. *Weather,* 25, 215-218.

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. S. Africa,* 12, 295-299.

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

Craddock, J.M., 1973: Problems and prospects for eigenvector analysis in meteorology. *The Statistician,* 22, 133-145.

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.

Crafford, D.J., and Nott, R.W., 1970: Soil, rainfall and maize yields in the highveld region. *Farming in S. Afri.*, 46, 33-43.

Crafford, D.J., and Nott, R.W., 1971: Soil, rainfall and maize silage yields. *Farming in S. Afri.*, 47, 11-13.

Currie, R.G., 1973 a: Geomagnetic line spectra - 2 to 70 years. *Astrophys. Space Sci.*, 21, 425-438.

Currie, R.G., 1973 b: Pacific region anomaly in the geomagnetic spectrum at ∿ 60 years. *S. Afr. Jour. Sci.*, 69, 379-383.

Currie, R.G., 1974 a: Period and Q_W of the Chandler wobble. *Geophys. Jour. Roy. Astr. Soc.*, 38, 179-186.

Currie, R.G., 1974 b: Solarcycle signal in surface air temperature. *Jour. Geophys. Res.*, 36, 5657-5660.

Currie, R.G., 1974 c: Harmonics of the geomagnetic annual variation. *Jour. Geomag. Geoele. tr.*, 26, 319-328.

Dale, R.F., and Hodges, H.F., 1972: *Weather and corn yield study.* Progress Rep., Dep. of Agronomy, Purdue Univ., Layfayette.

De Bruyn, L.P., 1970: Maize and moisture. *Farmer's weekly*, 119, 20-22.

Degtyareva, G.V., 1973: *Forecasting the spring wheat harvest* *before planting by inertial factors.* U.S. Joint Public. Res. Service, Arlington, Va.

De Jager, J.M., 1971: How to beat the weather. *Farming in S. Afri.*, 47, 40-46.

De Jager, J.M., 1974: *Putu* a dynamic seasonal maize crop growth model. *Can. Int. Biol. Program Res. Rep.*, 306-308.

De Jager, J.M., and Mallett, J.B., 1972: Effect of moisture stress upon maize production and its economic significance. *S. Afr. Jour. Sci.*, 68, 182-186.

Draper, N.R., and Smith, H., 1966: *Applied Regression Analysis.* John Wiley, London.

Dyer, T.G.J., 1975 a: Solar activity and rainfall variation over Southern Africa. *S. Afr. Jour. Sci.*, 71, 369-372.

Dyer, T.G.J., 1975 b: *Secular variation in South African rainfall.* Unpub. Ph.D. thesis, Univ. of Witwatersrand, Johannesburg.

Dyer, T.G.J., 1975 c: The assignment of rainfall stations into homogeneous groups: an application of principal component analysis. *Quart. Jour. Roy. Met. Soc.*, 101, 1005-1012.

Dyer, T.G.J., 1976 a: On the components of time series; the removal of spatial dependence. *Quart. Jour. Roy. Met. Soc.*, 102, 157-165.

Dyer, T.G.J., 1976 b: Expected future rainfall over selected parts of South Africa. *S.Afr. Jour. Sci.*, 72, 237-239.

Dyer, T.G.J., 1976 c: On the application of some stochastic models to precipitation forecasting. *Quart. Jour. Roy. Met. Soc.*, 102, 157-166.

Dyer, T.G.J., 1976 d: Meridional interactions between rainfall and surface pressure. *Nature*, 264, 48-49.

Fabricius, A.F., 1971 a: Something on the rainfall pattern of the South Eastern Transvaal. *Lect. Sers., Dept. Agric. Tech. Serv.*, 2, 4-12.

Fabricius, A.F., 1971 b: A meteorological picture of the Northern Transvaal. *Lect. Sers., Dept. Agric. Tech. Serv.*, 3, 4-14.

Federov, E.K., 1974: Effect of weather conditions on the yield of agricultural crops and the meteorological service to agriculture. *Problemy bor'by s sasukhoi i rost proizvodstva sel' skokhozyaisivennoiproduktsi (Kolos, Moscow)*, 60-79 (in Russian).

Fisher, R.A., 1924: The influence of rainfall on the yield of wheat at Rothamsted. *Phil. Trans Roy. Soc., London*, B 213, 89- 142.

Fullard, H., (ed.) 1976: *The Geographical Digest.* Philip and Son, London.

Garlipp, G.W.E.A., 1976: The influence of rainfall on maize grain yield. *Agroplantae*, 8, 21-22.

Ghildyal, B.P., and Jana, R.K., 1967: Agrometeorological environment affecting rice yield. *Agron. Jour.*, 59, 286-287.

Gilman, D.H., 1957: *Empirical orthogonal functions applied to 30-day forecasting.* Sci. Rept. No.1, Contract AF 19 (604) - 1283, Dept. of Meteor., M.I.T., Cambridge, Mass.

Glover, J., Robinson, R., and 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 Univ.Press, Liverpool.

Gribbin, J., 1974: Weather warning: you are now experiencing a climatic change. *Nature,* 252, 182-183.

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

Gross, E.R., and Rust, R.H., 1972: Estimation of corn and soybean yields utilizing multiple curvilinear regression methods. *Soil Sci. Soc., Amer., Proc.,* 36, 316-320.

Hammond, A.L., 1975: Crop forecasting from space: toward a global food watch. *Science,* 188, 434-436.

Hanks, R.J., 1974: Model for predicting plant yield as influenced by water use. *Agron. Jour.,* 66, 660-665.

Hannes, G., 1976: Factor analysis of coastal air pressure and water temperature. *Jour. Appl. Met.* 15, 120-126.

Haylett, D.G., 1968: Population growth and food resources in South Africa. *S. Afr. Jour. Sci.,* 64,369-374.

Hofmeyr, W.L., and Schulze, B.R., 1963: Temperature and rainfall trends in South Africa during the period of meteorological records. *Proc. Rome Conf. on Changes of Climate with special reference to arid zones,* UNESCO, Paris.

Huda, A.K.S., *et al,* 1975: Contribution of climatic variables ·. predicting rice yield. *Agric. Meteor.,*15, 71-86.

Huda, A.K.S., Ghildyal, B.P., and Tomar, V.S., 1976: Contribution of climatic variables in predicting maize yields under monsoon conditions. *Agric. Meteor.,* 17, 33-47.

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

Jagannathan, P., and Parthasarathy, B., 1973: Trends and periodicities of rainfall over India. *Mon. Wea. Rev.*, 101, 371-375.

Jakovleva, N.I., Cuvasina, I.E., and Kudaskin, G.D., 1968: The refinement of natural orthogonal functions for the pressure field over the Northern hemisphere. *Glav. Geof. Obs. Trudy*, 201, 60-71 (in Russian).

Jarrett, P.A., Altona, R.E., and Hyam, G.F.S., 1966: Effect of planting date on the growth and yield of maize. *S. Afr. Jour. Sci.*, 62, 80-84.

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

Jenkinson, A.F., 1975: Some quasi-periodic changes in rainfall in Africa and Europe. *Proc. of WMO/IAMAP Symp. on Long-term climatic fluctuations*, Norwich

Jolliffe, I.T., 1972: Discarding variables in a principal component analysis, I : artificial data. *Appl. Statis.*, 21, 160-173.

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. Stat. Psychology*, 14, 1-12.

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

Keen, C.S., 1971: *Rainfall spectra and the delimitation of rainfall regimes in South Africa.* Unpub. M.Sc. thesis, Univ.

of Witwatersrand, Johannesburg.

Kendall, M.G., Sir, 1975: *Multivariate Analysis.* Charles Griffin, London.

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

Kidson, J.W., 1975: Eigenvector analysis of monthly mean surface data. *Mon. Wea. Rev.,* 103, 177-186.

Kiriakov, K., 1938: Effect of the variations of the water regime on the growth, development and yield of maize. *La Meteorologia Pratica,* 19 No. 6.

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

Kristjannson, L., 1969: The ice drifts back to Iceland. *New Scientist,* 41, 508-509.

Kukla, G.J., and Kukla, H.J., 1974: Increased surface albedo in the Northern hemisphere. *Science,* 183, 709-714.

Kupper, L.L., 1972: Fourier analysis and spherical harmonic equations. *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.

Kutzbach, J.E., 1970: Large scale features of monthly mean Northern hemisphere anomaly maps of sea-level pressure. *Mon. Wea. Rev.,* 98, 708-716.

Lacoss, R.T., 1971: Data adaptive spectral analysis methods. *Geophysics,* 36, 661-675.

Lamb, H.H., 1975 a: Remarks on the current climatic trend and its perspective. *Proc. of the WMO/ IAMAP Symp. on Long-term Climatic Fluctuations,* Norwich.

Lamb, H.H., 1975 b: Long-term climatic fluctuations. *W.M.O. Bull.,* 25, 3-9.

Landsberg, H.E., 1974: Hearings on the world food supply. *Bull. Amer. Met. Soc.,* 55, 1135-1136.

Landsberg, H.E., 1975: Sahel drought : change of climate or part of climate. *Archiv. für Met. Geoph. Biokl., Ser.B.,* 23, 193-200.

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.

Lewin, J., and Lomas, J., 1974: A comparison of statistical and soil moisture modeling techniques in a long-term study of wheat yield performance under semi-arid conditions. *Jour. Appl. Ecol.,* 11, 1081-1090.

Livingstone, D., 1857: *Missionary travels and research in South Africa.* London.

Lomas, J., and Shashoua, Y., 1973: The effect of rainfall on wheat yields in an arid region. *Proc. UNESCO Symp. on Plant Response to Climatic Factors,* Uppsala, 1970.

Lomas, J., Schlesinger, E., and Lewin, J., 1974: Effects of environmental and crop factors on the evapo-transpiration rate and water-use efficiency of maize. *Agric. Meteor.,* 13, 239-251.

Lorenz, E.N., 1956: *Empirical orthogonal functions and statistical weather prediction.* Scien. Rept. No. 1, Statis. Forecasting Project, Dep. of Meteor., M.I.T., Cambridge, Mass.

Mallett, J.B., 1973: *The use of climatic data for maize grain yield predictions.* Unpub. Ph.D. thesis (Agric.), Univ. of Natal, Pietermaritzburg.

Mane, V.D., and Dastane, N.G., 1971: Effect of water regimes on growth and yield of NP 130 and IR.8 rice. *Indian Jour. Agron.*, 16, 1-5.

Mason, B.J., 1976: Towards the understanding and prediction of climatic changes - Symons Memorial Lecture. *Quart. Jour. Roy. Met. Soc.*, 102, 473-498.

Mitchell, J.M., *et al*, 1966: Climatic change. *W.M.O. Tech. Note*, 79.

Moffat, R., 1842: *Missionary labours and scenes in Southern Africa.* London.

Moll, J., 1969: Plant maize according to rainfall rigures. *Farmer's Weekly*, 117, 113.

Montgomery, D.C., and Johnson, L.A., 1976: *Forecasting and Time Series Analysis.* Mc Graw Hill, London.

Murata, Y., 1975: Estimation and simulation of rice yield from climatic factors. *Agric. Meteor.*, 15, 117-132.

Murata, Y., and Togari, Y., 1972: Analysis of the effect of climatic factors upon the productivity of rice at different localities in Japan. *Proc. Crop. Sci. Soc. Japan*, 41, 373-387.

Myburgh, J., 1971: *Temperature study of the Orange River area with regard to crop production.* Unpub. M.Sc. thesis (Agric.) Univ. of the Orange Free State, Bloemfontein.

National Oceanic and Atmospheric Administration, 1973: *The influence of weather and climate on U.S. grain fields:*

mper crops or droughts. A report to the Administrator, NOAA.

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

Newman, J.E., 1974: Applying meteorology to agriculture - (editorial). *Agric. Meteor.,* 13, 1-3.

Newman, J.E., and Pickett, R.C., 1974: World climates and food supply variations. *Science,* 186, 877-881.

Nix, H.A., 1974: Climate and crop productivity in Australia. *Proc. Conf. on Climate and Rice,* Intern. Rice Res. Inst., Los Banos, Philippines.

Nix, H.A., and Fitzpatrick, E.A., 1969: An index of crop water stress related to wheat and grain sorghum yields. *Agric. Meteor.,* 6, 321-337.

Nuttonson, M.Y., 1961: *The physical environment and agriculture of the Union of South Africa with special reference to its winter rainfall regions containing areas climatically and latitudinally analogous to Israel.* Amer. Inst. of Crop Ecology, Washington.

Peres, M.A., 1930: Preliminary investigations of the rainfall of Lourenco Marques. *S. Afr. Jour. Sci.,* 27, 132-135.

Pochop, L.O., Cornis, R.L., and Becker, C.F., 1975: Prediction of winter wheat from short term weather forecasts. *Agron. Jour.,* 67, 4-7.

Polevoj, A.N., and Myzina, T.I., 1973: The basis of a method for agrometeorological forecasting of the average areal oats yield capacity in the non-chernozem zone of the European part of the country. *Meteor. Gidrol.,* Leningrad, 12, 85-92 (in Russian).

Ratkowsky, D.A., and Martin, D., 1974: The use of multivariate analysis in identifying relationships between disorder and mineral element content in apples. *Aust. Jour. Agric. Res.,* 25, 783-790.

Revelle, R., 1974: Food and population. *Scient. Amer.,* 231, 161-170.

Rinne, J., 1971: Investigation of the forecasting error of a simple barotropic model with the aid of empirical orthogonal functions. *Geophysica,* Helsinki, 2, 185-239.

Robertson, G.W., 1974 a: World Weather Watch and wheat. *W.M.O. Bull.,* 23, 149-154.

Robertson, G.W., 1974 b: Wheat yields for 50 years at Swift Current, Saskatchewan in relation to weather. *Can. Jour. Plant Sci.,* 54, 625-650.

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

Runge, E.C.A., 1968: Effects of rainfall and temperature interactions during growing season on crop yield. *Agron. Jour.,* 60, 503-507.

Russell, E.W., 1976: Agricultural meteorology and tropical agriculture (editorial). *Agric. Meteor.,* 16, 1-3.

Russell, J.S., 1973: Yield trends of different crops in different areas and reflections on the sources of crop yield improvement in the Australian environment. *Jour. Aust. Inst. Agric. Sci.,* 39, 156.

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 remedy. *S. Afr. Jour. Sci.,* 15, 139-190.

Smillie, K.W., 1966: *An Introduction to Regression and Correlation.* Academic Press, London.

Smith, L.P., 1976: *Methods in Agricultural Meteorology* Elsevier, Amsterdam.

Smylie, D.E., Clarke, G.K.C., and Ulrych, T.J., 1973: Analysis of the irregularities in the Earth's rotation, in Bolt B.A. *et al* (eds.): *Methods in computational physics.* Vol. 13 Academic Press, New York.

South African Weather Bureau: *Annual reports of Meteorological Data,* 1949-1966, Pretoria.

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

Sreenivasan, P.S., 1974: Influence of rainfall on wheat varieties at Jalgam and Niphad. *Agric. Meteor.,* 13, 267-278.

Sreenivasan, P.S., and Banerjee, J.R., 1973: The influence of rainfall on the yield of rainfed rice at Karjat, India. *Agric. Meteor.,* 11, 285-292.

Stidd, C.K., 1967: The use of eigenvectors for climatic estimates. *Jour. Appl. Met.,* 6, 255-264.

Thompson, L.M., 1975 a : Weather variability, climatic change and grain production. *Science,* 188, 535-541.

Thompson, L.M., 1975 b: World weather patterns and food supply. *Jour. of Soil and Water Conserv.,* 30, 144-147.

Tidmarsh, C.E.M., 1968: Bioclimatology with special reference to agriculture. *S. Afr. Jour. Sci.,* 64, 456-459.

Tredoux, J., 1962: How weather studies benefit agriculture. *Farming in S. Afri.*, 38, 29-31.

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.

Trenberth, K.E., 1976: Fluctuations and trends in indices of the Southern hemispheric circulation. *Quart. Jour. Roy. Met. Soc.*, 102. 65-75.

Tripp, W.B., 1888: Rainfall of South Africa, 1842-1886. *Quart. Jour. Roy. Met. Soc.*,14, 108-123.

Tyson, P.D., 1971: Spatial variation of rainfall spectra in South Africa. *Annals Ass. Amer. Geog.*, 61, 711-720.

Tyson, P.D., 1972: Rainfall spectra and recent climatic variation in Southern Africa. *International Geography* 1972, Toronto Univ. Press, 202-204.

Tyson, P.D., Dyer, T.G.J., and Mametse, M.N., 1975: Secular changes in South African rainfall : 1880 to 1972. *Quart. Jour. Roy. Met. Soc.*, 101, 817-833.

Ulrych, T.J., *et al*, 1973: Predictive filtering and smoothing of short records using maximum entropy. *Jour. Geophys. Res.*, 78, 4959-4964.

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., 1925: Note on the apparent regularity of the occurrence of wet and dry years in South West Africa. *S. Afr. Jour. Sci.*, 22, 94-95.

Van Jaarsveld, A.S., 1973: The influence of rainfall on the yield of rain-fed winter wheat at Bloemfontein . *Fert. Soc. S. Afr. Jour.*, 1, 7-11.

Veldman, D.J., 1967: *Fortran programming for the behavioural sciences*. Holt, Rinehart and Winston, New York.

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

Wallace, J.M., and Dickenson, R.E., 1972: Empirical orthogonal representation of time series in the frequency domain, Part 1. *Jour. Appl. Met.*, 11, 887-892.

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

Wellington, J.H., 1960: *Southern Africa, a geographical study*, Vol. II. Cambridge Univ. Press, Cambridge.

Whitmore, J.S., 1948: Agroclimatology. The scope of its application to South African agriculture. *Farming in S. Afri.*, 23, 293-303.

Whitmore, J.S., 1950: The marginal areas of the Orange Free State: Pt. I A climatic study; Pt. II Climate and crop production; Pt.III Climate and the use of the land. *Farming in S. Afri.*, 25, 23-28; 47-52, 58; 85-89.

Williams. G.D.V., 1969: Weather and prairie wheat production. *Can. Jour. Agric. Econ.*, 17, 99-109.

Williams, G.D.V., Joynt, M.I., and Mc Cormick, P.A., 1975: Regression analyses of Canadian prairie crop- district cereal yields 1961-1972, in relation to weather, soil and trend. *Can. Jour. of Soil Sci.*, 55, 43-54.

Wilson, J.F., 1865: Water supply in the basin of the River