DEVELOPMENT OF AN OBJECTIVE STATISTICAL SYSTEM TO FORECAST SUMMER RAINFALL OVER SOUTHERN AFRICA

by

Mark R Jury[1]
Henry M Mulenga[2], **Simon J Mason**[3], **Annabela Brandao**[4]
Climate and Weather Research Lab
Oceanography Dept
Rondebosch 7700 South Africa

[1]currently: Climate Impact Predictions, Geography Dept,
Univ Zululand, KwaDlangezwa, 3886
[2]on leave from: Zambia Meteorological Service, Lusaka
[3]Climatology Research Group
University of Witwatersrand, Johannesburg
4Statistical Sciences Dept
University of Cape Town

WRC Report No: 672/1/97
ISBN No: 1 86845 362 6

DEVELOPMENT OF AN OBJECTIVE STATISTICAL SYSTEM TO FORECAST SUMMER RAINFALL OVER SOUTHERN AFRICA

ACKNOWLEDGEMENTS

A report to the Water Research Commission produced by Prof M R Jury, WRC Project Leader, formerly Climate and Weather Research Lab, Oceanography Dept, Univ Cape Town. The analysis and products contained herein were contributed by Mr H M Mulenga, WRC project PhD bursar from the Zambian Meteorological Service; Dr S J Mason, consultant to the WRC project, CRG, Univ Witwatersrand; and Dr A Brandao, consultant to the WRC project, Statistical Science Dept, Univ Cape Town. This team was supported by inputs from the Climate and Weather Research Lab: E A Post, data and systems analyst; K M Levey, C McQueen amd E Mpeta, researchers; and C Lander, part-time research assistant.

The advice and insights contributed by the steering committee members ensured success of the project. The members were Dr G Green, Mr H Maaren, Mr S Biljon and Mr F Marais from the WRC; Mr G Coetzee and Mr W Landman of the SAWB; Prof J vanHeerden, Univ Pretoria; Prof J deJager, Univ OFS; and Prof G Brundrit, Univ Cape Town.

CONTENTS:

	Pages
Acknowledgements	2
Executive Summary	4
a. BACKGROUND	
b. USER NEEDS AND TARGETS	
c. MODEL DEVELOPMENT	
d. SKILL TESTS	
e. CONCEPTUAL UNDERSTANDING	
f. IMPLEMENTATION	
g. CONCLUSION	
h. FUTURE DIRECTIONS	
i. DATA AND ASSISTANCE	
j. REFERENCES	
k. PAPERS PUBLISHED DURING PROJECT	
Chapter 1 - Introduction	11
1.1 BACKGROUND AND NEEDS	
1.2 PREDICTORS	
1.3 FORMULATION OF TARGETS	
1.4 TIME SCALES	
1.5 MODEL FORMULATION	
1.6 INFLUENCE OF ENSO PHASE	
1.7 PROJECT OBJECTIVES	
Chapter 2 - Targets and Predictors	18
2.1 DEVELOPMENT OF PROJECT METHODOLOGY	
2.2 FORMULATION OF PREDICTORS AND TARGETS	
2.3 DYNAMICAL STUDIES	
2.4 OPERATIONAL INPUTS AND NETWORKING	
Chapter 3 - Models and Skill Tests	25
3.1 OPTIMUM MODELS (22 yrs, 4 pred)	
3.2 OPTIMUM MODELS (37+ yrs, 5 pred)	
Chapter 4 - Conclusions	30
4.1 PRODUCTS AND RECOMMENDATIONS	
4.2 SUMMARY AND FUTURE DIRECTIONS	
5. References	33
6. Papers published during project	
FIGURES	35

Executive Summary

The project's objectives as set out in the WRC contract were to:

• identify atmospheric and oceanic precursor patterns which anticipate summer rainfall over the plateau of southern Africa,

• formulate reliable predictors of summer rainfall and associated climatic impacts for use in training statistical models,

• develop multi-variate algorithms to skilfully predict area-rainfall and other climatic impacts, and

• analyse dynamical mechanisms underlying seasonal rainfall over the southern African plateau.

Products envisaged from the WRC project included:

• reliable predictive models of area-rainfall and other water-related climate targets for widespread external use, and

• diagnostic analyses of drought / flood scenarios for internal use in pattern-recognition and conceptualisation.

a. BACKGROUND

A statistical system to objectively forecast southern African summer rainfall one season in advance is useful in the managing of water resources and agricultural enterprize, and would benefit anyone in southern Africa whose livelihood is impacted by year-to-year fluctuations in climate. There exists a high potential for seasonal predictions because of links between the global El Nino-Southern Oscillation (ENSO) and southern African weather systems. These links can be explored using global data sets. ENSO signals over Africa have been documented (Lindesay 1988; Jury et al 1994) and hemispheric-scale precursor patterns have been determined using pair-wise correlation maps (Rocha 1992; Pathack 1993; Jury 1996).

Global data sets employed in the WRC project include sea surface temperature (SST) at 5 deg resolution for the period 1950-1995 from the UK Meteorological Office, upper winds from the National Centres for Environmental Prediction (NCEP) and European Community Medium-range Weather Forecast gridded products, surface air pressure and winds from the Comprehensive Ocean-Atmosphere Data Set, ENSO and quasi-biennial oscillation indices from the CAC, and satellite cloud estimates via outgoing longwave radiation (OLR) and highly reflective cloud albedo from CAC extending to 1971. Key areas identified by previous researchers such as Mason (1992), Rocha (1992) and Pathack (1993) offering high correlation at one season lead time were considered. The principal modes of regional SST were analysed using normalised departures. Because of the importance of satellite cloud indices in preliminary analyses, the development of statistical models was confined to the period 1971-1993, represented by 264 consecutive months. As annual statistics are used for model training, the sample size is 22. For some historical targets and a sub-set of the predictors, longer 37-41 year models could be considered.

b. USER NEEDS AND TARGETS

Forecasts are required in the austral spring for management purposes, so predictors for the July to November period were considered. To reduce in-season 'noise', three month means were used. Exploratory analyses enabled over 100 candidate predictors to be identified. Target rainfall data were obtained from national weather services in the region. Emphasis was placed on areas with significant water resources and rain-fed agricultural production. Rainfall station data were grouped according to annual total, seasonal cycle, topographic height, environmental field patterns and cross-correlations. Area-averaged rainfall targets were about 300 km diameter (10^5 km^2) and slightly larger than those of Landman (1995). It was considered that a wider area-average improves the potential for predictability, up to a spatial limit defined by local users. Normalised departures of monthly station data were averaged into early and late summer seasons: November to January and January to March, respectively. January was weighted by half to reduce its influence as a transition month. The averaging period represents a compromise between 'noisy' single month targets and user needs for guidance on seasonal rainfall distribution.

Whilst over 100 candidate predictors were available, ultimately 23 predictors were selected by 9 South African rainfall models. Of these predictors, about one-third were significantly correlated with ENSO indices, another one-third (some overlapping) relate to the transition of the Indian NE monsoon, and about one-quarter reflect conditions over the South Atlantic which may control temperate latitude atmospheric westerly waves. Of the SST inputs, the Indian and Atlantic Oceans contributed equally to all models, whilst Pacific SSTs provided only limited inputs at lead times of one season. The OLR, winds and pressure predictors made sizeable contributions to the models, as did global indices such as the southern oscillation index and the quasi-biennial oscillation. To improve forecast skill, lead times of 2 to 3 months were considered.

c. MODEL DEVELOPMENT

Multi-variate linear regression statistical models were developed using standard statistical software.. Each historical area-rainfall index was fitted by an optimum mix of statistically significant predictors for the preceding spring, using a forward step-wise technique. This was done independently by Dr Jury and A Brandao using different statistical packages, Statgraphics and Genstat, respectively. To prevent over-fitting and artificial skill, predictors were restricted to four or five. This provided up to 17 degrees of freedom. Models were formulated to achieve a maximum hindcast adjusted r square.

A number of potential candidate models were developed for each target and submitted for skill validation tests. Models with co-linear predictors of opposing sign were screened out. The final selection of model for each target was based on an optimum skill-test correlation giving a maximum hindcast fit.

The predictors used in the South African summer rainfall models (9) consisted of global indices (SOI and QBO, each selected 3 times), SSTs in the Atlantic (selected 5 times), Indian Ocean convection (OLR) and air pressure, selected 5 times each, and southern ocean sea surface temperature (2). Significant contributions to the South African models were also made by Indian Ocean SST (4), and surface Indian and upper Atlantic winds (6). In the climate impact models (12) for river run-off etc., nearly equal contributions from Atlantic (8) and Indian (10) SST were noted. Secondary contributions came from Indian pressure (8), Indian surface wind (6), Indian OLR (4) and pressure/wind off Angola (4), with minor inputs from other predictors. The climate impact models had a r^2 fit of 72.7% indicative of high levels of predictability consistent with late summer rainfall.

d. SKILL TESTS

The multi-variate models were validated using the jack-knife technique. These statistical validation tests were done independently by A Brandao and Dr Mason using different software. The technique involves removing each year individually in the 22 year record and predicting it based on parameters defined from the remaining years. The process is repeated by removing the next year and so on until 22 forecasts have been made. Differences between observed and predicted values for each 'removed' year were evaluated. Mean correlations between observed and 2-3-month-lead-time predicted values for 4-predictor models were: 60.7% for early summer and 72.2% for late summer models; predictability for late summer is higher, probably because of the stronger influence of ENSO phase and tropical climate after December.

The statistical 'miss' rate of the models was assessed by considering the number of years the observed minus jacknife-predicted difference was > 0.5 times the standard deviation, and of opposite sign. Based on this method, the early summer models mean miss rate is 25%, for the late summers it is 18%, or about one in every five years.

e. CONCEPTUAL UNDERSTANDING

Conceptual models can be used to facilitate an understanding of the climate dynamics and ocean - atmosphere coupling. According to the late summer models for the NE South African plateau, including: the NW Province, Free State, Gauteng and adjacent portions of Mpumalanga and North Provinces, the September to November predictors have the following anomalies in respect of increased late summer rainfall:

- global SOI positive, Pacific Nino3 SST negative, and tropical Atlantic 200 hPa winds easterly (negative);

- QBO in west phase (east phase previous year);

- tropical Indian Ocean SST negative in the central area, surface winds easterly (JAS season) and SST anomaly positive in east, and convection reduced (+OLR) in the west;

- pressure positive in the south Indian Ocean and SST positive to the east; and

- south-west Atlantic Ocean SST below normal in mid-latitudes.

A subset of the model predictors are in agreement with Hastenrath et al (1995).

f. IMPLEMENTATION

In operational application, climatic data mainly from the National Centre for Environmental Prediction (NCEP) are received via the Internet and seasonal predictor anomalies are extracted for model input. Most predictors require a three-month mean to be calculated and the anomaly normalised by division of the standard deviation at the same location. For SST patterns the patterns are assessed by calculation of principal components which are normalised and smoothed by Dr S Mason. Results of the multi-variate algorithms are posted on the website of the Cape Town Weather Watch during spring (August to December) at the URL address: < http://os2.iafrica.com/weather > under the forecast menu, with the current URL address:

< http://www.sea.uct.ac.za/weather/forecast/ctww_seasonal_outlook.html >

The seasonal long-lead forecast service is maintained in support of water resource managers, farmers, government departments and regional businesses. Models for other countries in the SADC have been developed in collaboration with the respective national weather services. The outlook is updated regularly in spring, and provides expected departures of area-rainfall over southern Africa; and forecasts of other climate impacts such as maize yield, malaria incidence, major river flows, etc. Post-season analysis and verification is done and forecast errors are analysed with a view to refinement of models.

g. CONCLUSION

Seasonal outlooks based on the WRC project results have been successfully implemented to help mitigate the impacts of southern Africa's fluctuating climate. The Internet website seasonal outlook is accessed up to 1000 times per month, in addition to a fax dissemination service to over 200 local users. The skill of the various models has been assessed over the past three decades. The results meet international practice for reliability, with tercile 'hit' rates in excess of 66% for most models. It is believed that an advance warning of drought risk and seasonal rainfall prospects will improve the economic growth potential of southern Africa and provide additional security for food and water supplies.

h. FUTURE DIRECTIONS

Further work on seasonal predictability could involve use of re-analysed global model products such as NCEP monthly fields from 1957. The model training period could thus be extended, giving increased degrees of freedom and statistical confidence. Studies based on global model re-analysed products and their principal components would enable a wiser choice of independent predictors submitted for model optimisation. Because skill correlations average ~ 66 % for most models, the predictability 'gap' could be closed further through additional work. Some degree of uncertainty will remain (estimated at 25 %) owing to the random nature of climate particularly in neutral or non-ENSO years.

The statistical models described here are based on past history and the potential for climate 'drift' suggests that the models could become gradually obsolete. To minimise this, regular updating of predictors and model algorithms is necessary. Better observations and modelling of Atlantic and Indian Ocean El Ninos would greatly benefit prediction efforts and could enable accurate forecasts up to two seasons in advance.

i. DATA AND ASSISTANCE

Data were supplied by the NCEP, UK Meteorological Office, ECMWF, SADCO South Africa, CCWR South Africa, and the national weather services of South Africa, Zimbabwe, Namibia, Zambia, Tanzania, Madagascar and Mauritius. COADS and OLR data were quality checked and re-analysed by Prof B Wang, Univ Hawaii. Many of the climatic targets and predictors are available in the form of spreadsheet files from Prof Mark Jury at the Geography Dept, Univ Zululand, KwaDlangezwa, 3886, South Africa.

j. REFERENCES

Hastenrath, S, Greischar, L, and VanHeerden, J, 1995, Prediction of the summer rainfall over South Africa, *J Climate*, 8, 1511-1518

Landman, W A, 1995, A cannonical correlation analysis model to predict South African summer rainfall, *Exp Long-lead Frcst Bull*, 4, 4, 23-24

Lindesay, J, 1988, The Southern Oscillation and atmospheric circulation changes over southern Africa, *PhD Thesis*, Univ Witwatersrand, Johannesburg, 283 pp.

Jury, M R, McQueen, C A, and Levey, K M, 1994, SOI and QBO signals in the African regoin, *Theor Appl Climatol*, 50, 103-115

Jury, M R, 1996, Regional teleconnection patterns associated with summer rainfall over South Africa, Namibia and Zimbabwe, *Intl J Climatol*, 16, 135-153
Mason, S J, 1992, Sea surface temperatures and South African rainfall variability, PhD Thesis, Univ Witwatersrand, Johannesburg, 235 pp
Pathack, B M R, 1993, Modulation of South African summer rainfall by global climatic processes, *PhD Thesis*, Univ Cape Town, 214 pp.
Rocha, A M C, 1992, The influence of global sea surface temperatures on southern African summer climate, *PhD Thesis*, Univ Melbourne, 248 pp.

k. PAPERS PUBLISHED DURING PROJECT

Jury, M R, Pathack, B M R, Rautenbach, C deWet, and VanHeerden, J, 1996, Drought over southern Africa and Indian Ocean SST: statistical correlations and GCM results, Global Ocean Atmos Sys,4, 47-63
Jury, M R, 1996, Regional teleconnection patterns associated with summer rainfall: South Africa, Namibia and Zimbabwe, Intl J Climatol, 16, 135-153
Jury, M R, and Levey, K M, 1997, Vertical structure of wet spells over southern Africa, Water S A, 23, 51-55
Jury, M R, Weeks, S, and Gondwe, M P, 1997, Satellite vegetation as an indicator of climate variability over southern Africa, S. Afr. J Science, 93, 34-38
Jury, M R, and Majodina, M, 1997, A climatology of southern Africa extreme weather: 1973-1992, Theor Appl Meteorol, 56, 103-112
Levey, K M, and Jury, M R, 1996, Composite intra-seasonal oscillations of convection over southern Africa, J Climate,9, 8, 1910-1920
Gondwe, M P, and Jury, M R, 1997, Sensitivity of vegetation (NDVI) to climate over southern Africa: relationships with rainfall and OLR, S A Geogr J, (in press)
Jury, M R, 1997, Southeast Atlantic warm events: composite evolution and consequences for southern African climate, S Afr J Marine Sci, (in press)
Jury, M R, 1997, Inter-annual climate modes over southern Africa from satellite cloud OLR 1975-1994, Theor Appl Meteorol, (in press)
Jury, M R, and Levey, K M, 1997, Vertical structure of mid-latitude rossby waves during summer wet spells over southern Africa, S Afr J Science, (in press)
Majodina, M, and Jury, M R, 1997, Composite winter cyclones south of Africa: evolution during eastward transit over the Agulhas warm pool, S Afr J Marine Sci, (in press)
Makarau, A, and Jury, M R, Predictability of Zimbabwe summer rainfall, Intl J Climatol, (in press)
Mason, S J, and Jury, M R, 1997, Climatic fluctuations in southern Africa: a reflection on underlying processes, precursors and predictability, Prog Phys Geogr, (in press)

Chapter 1 - Introduction

1.1 BACKGROUND AND NEEDS

Rainfall over southern Africa is limited and highly variable from year to year and agricultural production constitutes up to 75% of the economic activity in rural areas. Maize production is particularly prone to drought, varying by a factor of 10 from good to bad years, despite technological advancements. Water storage systems have periodically failed during global El Nino events, such as in 1983 and 1992. Forecasts of summer rainfall more than one season in advance, if sufficiently reliable and of the appropriate space - time scale, may assist the management of environmental resources and reduce national debt. Thus a need arises to define the limits of predictability in southern Africa and to establish user requirements for an early warning system.

Predictability in southern Africa is relatively high because links between the global El Nino and regional weather systems are robust and well understood. When sea surface temperatures (SST) in the central and eastern tropical Pacific Ocean increase above normal, westerly upper level winds are enhanced downstream over the tropical Atlantic Ocean, bringing dry air and strengthening the high pressure cell over Botswana. The central tropical Indian Ocean SST changes in sympathy with the eastern Pacific, with a lag of 1-2 months. During El Nino warm phase years, increased monsoon flow recurvature into tropical cyclones of the SW Indian Ocean produces a 'dipole'. Increased convective rainfall east of Madagascar is offset by drought over southern Africa. Zonal circulation cells are often invoked as explanations for these phenomena. The regular alternation of stratospheric zonal winds known as the quasi-biennial oscillation, also plays a role, albeit more obscure, in regulating rainfall over southern Africa. In addition to the known El Nino connections, research has indicated that the austral spring is a more predictable time of the year than the boreal spring, so favouring long-range forecasts for the southern hemisphere summer over those of the more economically active northern hemisphere. In the following sections, optimum methods of formulating long-range statistical prediction models will be outlined in the context of user requirements.

Numerous studies have probed the limits of predictability for southern Africa. Relationships between antecedent climate fields and seasonal rainfall over southern Africa have been identified using monthly data sets gridded over the globe (Rocha 1992, Pathack 1993). A number of researchers have looked at sea surface temperature (SST) patterns, surface and upper level winds, Southern Oscillation and Quasi-Biennial Oscillation indices and satellite out-going longwave radiation (a proxy for cloud depth). Much of the work has focussed on pair-wise correlation maps between an area-rainfall 'target' and the field variable over the period 1960 to present. Correlation values peak at -0.6 for the central Indian Ocean SST vs SE African late summer rainfall. For satellite OLR the correlations are slightly higher and positive. In general, pair-wise correlation functions attain maximum value 2 to 3 months before the season. Exploratory analyses have enabled climatologists to assemble a wide selection of predictors to develop statistical multi-variate models. Because forecasts are required in the austral spring for management purposes, predictors are often considered for the July to November period.

1.2 PREDICTORS

Because of the influence of the ENSO phase on southern Africa climate, it is generally held that a majority of predictors will come from the tropical latitudes. In these latitudes near the equator, eastward moving convective waves cause 40-60 day changes to the circulation and ocean - atmosphere coupling processes. Although some research has suggested an interaction with seasonal transitions of the Indian monsoon, most Madden-Julian Oscillations, as they are known, tend to be random. To alleviate these sources of weather-induced 'noise', most researchers use running monthly averages of predictors such as the Southern Oscillation Index. As 5 month averages tend to be too long for operational purposes, 3 months seems to be the common choice. Hence predictor values for early summer targets are drawn from July to September averages, whilst September to November averages would provide inputs to late summer models. Use of predictors for months prior to July appears to reduce forecast skill for most lag-correlation statistical model formulations.

Spatial domains of the predictors are usually determined through principal component analysis or correlation maps. Once analysed, the principal component time score or a time series for a key box area of high lag-correlation is extracted for use as a predictor. In most cases the areas are larger than 10° latitude x 20° longitude, hence covering a domain large enough to have a lasting impact on regional climate - weather coupling. Commonly used predictors include those which obtain statistically significant pair-wise correlation at lags 2 - 6 months, or the larger mode principal component time scores. Thus variables which explain minor field variance or obtain lower lead-time significance are rejected. On the other hand, it may be useful to submit vast numbers of predictors which have little *a priori* significance, in the hopes of obtaining a skillful model which does not fit a pre-determined conceptual framework. It may be that pair-wise assessments do not proclaim the appropriate predictors for use in multi-variate models. Another approach is to identify commonly used predictors, such as tropical SST in the central Indian Ocean, and formulate statistical or dynamical models to predict their trend. Success in this realm could offer longer lead times, for example Zimbabwe's maize yield using Pacific Nino3 SST forecasts (Cane et al 1994).

1.3 FORMULATION OF TARGETS

The homogeneity and size of the target area are considered at the outset. In general, smaller areas yield uneven time series which reflect the patchy nature of the summer rainfall distribution over southern Africa. The random, small-scale component is less likely to be linked to global events and ocean - atmosphere coupling. There may be an optimum target size which is a compromise between reducing 'noise' and creating a homogeneous mix of station data. The grouping will be guided by seasonality of rainfall, station elevation and response to inter-annual variations of climate. Optimum sizes may vary from 3 x 4 degree ovals as used by Jury (1996), to large river catchments, to large districts as defined by Nicholson et al (1988), to smaller districts (South African Weather Bureau, SAWB), or to principal components as deduced from gridded rainfall or OLR data sets. For statistical purposes most rainfall station data are converted to normalised departures by subtracting the mean and dividing by the standard deviation. The individual time series can be combined for a selected area to produce a spatially smoothed index.

In addition to rainfall, other parameters which integrate the effects of climate variability may offer useful targets. Natural streamflows, crop yields such as maize, climate-related health statistics such as malaria incidence are but a few of the climate-impact predictands which users require forecasts for. In some cases it may be necessary to de-trend these time series, if global warming or advancing technology result in impacts which are not driven by year-to-year climate variability.

1.4 TIME SCALES

In most long-range forecast applications, the monthly time scale is considered. For agricultural users, an ideal rainfall distribution is spread across the season, with a peak in late January. For run-off and replenishment of water resources, more concentrated flooding is needed. Therefore consideration of area-rainfall time scales is necessary. Recent studies of area-averaged rainfall have identified a common 20-35 day cycle (Levey 1993, Makarau 1995). Wet spells of about one week duration occur at approximately monthly intervals from November to March. However in any given year, it is likely that considerable fluctuations will occur. Wet spells may spread across two months, or not occur at all. The tropical Madden-Julian Oscillation has been implicated in some events, whilst other wet spells are seen to be driven by easterly and/or westerly waves impinging onto the region from the SW Indian Ocean sub-tropics and the South Atlantic mid-latitudes respectively. Further work has demonstrated that the early summer wet spells are embedded in a circulation regime distinctly different to their late summer counterparts.

In grouping the predictand rainfall targets temporally, it is advisable to take these factors into account. Just as predictors of single month duration are noiser than their seasonal counterparts, predictands may be optimised by averaging into 2 or 3 month periods. In reducing the noise in the time series, it is expected that more skillful and reliable models will result. However there is an upper limit which is defined differently by various user groups, wherein forecasts for the whole summer season and for areas too diverse will be of reduced value. Compromises need to be made and it is thought that at least the early and late summer be divided and considered independently. This is relevant for many parts of southern Africa which experience a mid-summer dry spell.

1.5 MODEL FORMULATION

Long-range forecast lead times of more than two months have enhanced economic value because management decisions may be implemented. An indication of the state of the ENSO phase is available from NCEP and other international organisations, and long-range forecasters can publicise these results at longer lead times. Statistical model lead times which are so long as to result in unskillful products due to reduced predictive potential are avoided. Again a compromise solution emerges: forecasts for the coming summer (November to March) are to be made public by September and refined and finalised in November.

All forecast models which receive national and regional attention are skill-tested by objective means. For statistical models based on 22 years the jack-knife skill test is used, whereas for 37+ year models independent training and validation periods may be considered.The models are expected to achieve a hindcast 'fit' r^2 value in excess of 50% and skill test correlation values of a similar magnitude. Miss rates can be evaluated by a number of tests, for example a tercile category scheme. Here an observed minus jacknife-predicted residual value > 0.5 standard deviation of opposite sign (> 50% departure of forecast from the actual value) is taken as a clear indication of a poor forecast.

To prevent artificial skill by over-fitting, the number of predictors is limited such that the degrees of freedom is optimised. For 22 year models this means that predictors are limited to 4, so the degrees of freedom exceeds 16. Another important point is that statistical models need to be trained on the most up-to-date information. Hence forecasts for each summer are to be based on algorithms updated to at least two years previous. This places a burden on long-range forecasters to continually update their data sets and refine their statistical models accordingly. However, the benefits in terms of more highly skilled and relevant models will be evidenced when users view the forecasts as reliable and are willing to make management decisions based thereon. Indeed forecast skill is the highest priority of users. Miss rates or false alarms of more than one-in-four years are likely to be considered unacceptable by users. Forecaster and institutional credibility is at stake in this regard.

1.6 INFLUENCE OF ENSO PHASE

Many of the predictors currently in use by long-range forecasters in southern Africa are significantly correlated with the Southern Oscillation Index and eastern Pacific (Nino 3) sea surface temperatures. Hence it is the ENSO phase which is being exploited to forecast climate impacts in the region. It is interesting to note the response that GCM experiments obtain for various predictors during summer. Increases in SST of either or both the tropical central Indian Ocean and eastern Pacific contribute to increased convection in the SW Indian Ocean at the expense of southern Africa. The known Pacific ENSO signals can be monitored during austral spring, together with regional ENSO influences within the Indian monsoon circulation and convection, and equatorial Atlantic upper zonal winds (Jury et al 1994). It may be advisable to develop independent long-range forecast models for strong ENSO phase years, and for non-ENSO years as conceptualised by Walker (1989). The CRG, Univ of Witwatersrand have discriminated models based on the phase of the QBO (Mason 1992). However, fragmentation of time series would result in a loss of statistical confidence, at this stage.

1.7 PROJECT OBJECTIVES

It is useful to reiterate the objectives of the WRC project here. The project ambitiously sought to:

- identify atmospheric and oceanic precursor patterns which anticipate summer rainfall over the plateau of southern Africa,
- formulate reliable predictors of summer rainfall and associated climatic impacts for use in training statistical models,
- develop multi-variate algorithms to skilfully predict area-rainfall and other climatic impacts, and
- analyse dynamical mechanisms underlying seasonal rainfall over the southern African plateau.

The WRC Project was expected to deliver the following products:

• reliable predictive models of area-rainfall and other water-related climate targets for widespread external use, and

• diagnostic analyses of drought / flood scenarios for internal use in pattern-recognition and conceptualisation.

Chapter 2 - Targets, Predictors and Methods

To formulate long-range forecast models, selected target time series are needed. These are often represented by area-averaged summer rainfall departures for a historical period of say 25 to 40 years. Predictors taken from the surrounding region, selected on the basis of an extensive history of research, need to be available for a similar length of time. Many predictors such as upper winds and cloud depth anomalies are only available in a globally gridded format from the early 1970s onward. The selection of predictors from these global sets can take three approaches:

• correlation maps of the gridded products with the target time series at lead times of 2 to 6 months can guide the researcher to key areas (as shown in Figure 1a,b);

• composite maps of gridded products prior to extreme wet and dry summers can be subtracted, offering similar key areas; or

• the predictor gridded products can be assessed for unique patterns of field variability using principal components analysis.

In this way the major inter-annual signals in a particular field such as SST can be extracted. Typically the top five or so 'modes' are extracted for use as predictors, depending on the variance they explain in relation to the rest. Seasonal means of the principal component time scores for the July to November period are used in many statistical model forecasts (Mason 1992).

2.1 DEVELOPMENT OF PROJECT METHODOLOGY

During the WRC project a number of data gathering and analysis exercises were carried out to bring together and reduce vast quantities of environmental statistics which describe historical fluctuations in the ocean and atmosphere relevant to southern Africa.

- ***predictands formulated***: 40 area-rainfall and climate impact indices were carefully computed by averaging monthly statistics such as rainfall from the South African Weather Bureau and nearby national weather services over specified areas and seasons. Rainfall data were obtained from the SA Weather Bureau and other sources, averaged over areas similar to that suggested by Pathack (1993) and Landman (1995\. Figure 2a shows a subset of the area-averaged rainfall targets and table 1 in chapter 3 describes these in more detail.

- ***predictor data bases enhanced***: UKMO sea surface temperature (SST) principal component fields 1950-1992 were provided by Dr S J Mason - Univ Wits, South African Data Centre for Oceanography (SADCO) ship data from the tropical Atlantic and the southern oceans were obtained and processed, Indian Ocean predictors were obtained from Prof B Wang, Univ Hawaii (during 1994 on a sabbatical visit by Prof Jury) and were processed; the total number of candidate predictors for submission was 166 by the end of 1996. Figure 2b provides an example of how environmental fields in the Indian Ocean were divided into key area predictors. Table 2 in chapter 3 describes a sub-set of predictors selected by South African rainfall models in further detail.

- ***correlation functions and inter-relationships investigated:*** for most predictors pair-wise tests with rainfall 'targets' were performed and relationships with known ENSO variables tested. Co-linearity between predictors was assessed. The use of principal components for SST reduces this problem as each component is unique. For other variables certain *a priori* selection criteria were applied, for example key areas were identified based on pair-wise correlation maps. The aim was to ensure that each predictor represented a widespread and unique signal. Many predictors composing the multi-variate models were associated with the ENSO phase, as shown in Figure 2c.

- ***meteorological dynamics studied***: inter-annual and intra-seasonal analyses of outgoing longwave radiation (OLR - cloud depth) and satellite normalised difference vegetation index (NDVI) data were completed and disseminated in the form of an atlas. Composite analysis of intra-seasonal wet spells anchored to a convective region over SE Angola was undertaken using pentad European Community Medium-range Weather Forecast (ECMWF) data. A detailed study of hemispheric scale principal components for SST, OLR and wind was done over the 1980-1994 period using gridded data for 168 consecutive months. After removal of the seasonal cycle, the first five modes of the rotated departures were analysed and inter-relationships were noted. Multi-variate models were formulated to explain associations between patterns of oceanic anomalies, atmospheric circulation and convection. Examples of these are illustrated in Figure 3a,b.

- ***objective forecast models developed***: for each target predictand two or three models involving up to 5 predictors were formulated and tested. These were implemented during the 1995 forecast cycle and refined during 1996. Comprehensive jack-knife skill tests were conducted on second and third generation models to validate their ability to predict rainfall etc 2-3 months in advance. Typical skill score correlations were of order 70%.

- ***resources enhanced***: all computers in the lab were upgraded with additonal memory and faster motherboards, a colour printer was added; all at low cost. Some of the computer and data resources were re-located to the Univ of Zululand where the project leader has been situated from 1997 onwards.

2.2 FORMULATION OF PREDICTORS AND TARGETS

Predictor variables for input to statistical models were averaged into July-September (JAS) and September-November (SON) periods and include:

- globally representative values for the southern oscillation index (SOI) in the form of Tahiti-Darwin air pressure differences from 1950-1995;

- the stratospheric quasi-biennial oscillation (QBO) based on 30 hPa Singapore zonal wind anomalies in the preceding year from 1955-1995;

- tropical Atlantic upper zonal wind anomalies at 200 hPa in the area 40W-0, 5N-10S from the NCEP (1968-87) and ECMWF (1980-94);

- sea surface temperature (SST) prinicipal component time scores for individual ocean basins: South Atlantic, Indian and Pacific; extending to 35°S from the UKMO SST dataset from 1950-1995;

- sea surface temperature (SST) prinicipal component time scores for a combined Atlantic + Indian Ocean area extending to 40S from UKMO SST dataset from 1950-1991;

- pressure, SST and surface winds in the tropical SE Atlantic to the west of Angola: 0-15E, 0-10S from 1968-1994;

- pressure, SST, outgoing longwave radiation (OLR) and surface winds in the Indian Ocean region from 12N-30S, 40-100E, gridded from the Comprehensive Ocean-Atmosphere Dataset (COADS) to 5 lat. x 15 long. areas from 1950-1992. Satellite OLR data originally from NCEP were re-calculated by Prof B Wang, Univ Hawaii and were available from 1971-1992. This dataset proved essential but limited the historical training period for most models;

- southern ocean pressure, SST and surface winds averaged over the area 40-55S, 40W-60E from relatively sparse ships' data over the period 1960-1995;

- ECMWF weather data at 7 vertical levels at 2.5 deg resolution from 70W-100E, 20N-60S from 1980-1994. The dataset was used for dynamical studies; being too short for statistical model development.

Most predictors were available from 1950 except for upper winds and OLR which commenced in 1968 and 1971 respectively. Because of their significant influence on many models, statistical training periods for second and third generation models were limited to 22 years (1971-1993). In the 4th round of model development a training period in excess of 37 years was considered, albeit with a limited set of predictors. These proved to have significantly lower skill (~20%) and have not been implemented in operational long-range forecasts disseminated by the project leader. The longer models are of value in assessing decadal cycles and predictability for the 50s and 60s.

The target predictands used to describe climate variability include:

- A rainfall index for the Eastern Cape Province coastal belt made up of SAWB districts 13,22,23,24,27,28,29,41,42,43;

- NW - FS Provinces = SAWB districts 60,70,71,72,73,74,81,82,83,84,85,89,90,91,92,93;

- former Transvaal highveld = SAWB districts 46,47,48,49,50,61,62,63,64,65,74,75,76,77,86,87; and

- Swaziland - KZN = SAWB districts 25,31,32,33,34 + Maputo, Mozambique rainfall data.

These rainfall indices were formulated by averaging monthly district totals and compared favourably with those of Pathack (1993). Areas of southwestern South Africa with annual totals < 300 mm or a non-summer seasonality were excluded from consideration.

- Monthly rainfall standardised departures were averaged for groups of stations in central Zimbabwe, N Namibia, SW Zambia, N Tanzania, and the Madagascar highlands. From the monthly gridded rainfall product of Hulme (Univ E Anglia) data were extracted for central Botswana. Most rainfall predictands extended from 1950-1995.

- Additional indices were developed for the number of tropical cyclone days in the SW Indian Ocean near Mauritius. Maize yield statistics from the NW and FS provinces, the former Transvaal highveld and the national average, and for central Zimbabwe were obtained. Malaria incidence in the South African lowveld, divided by the local population was added.

- Streamflow anomalies were obtained for the Vaal River, the Zambezi at Victoria Falls, the Okavango River near Rundu, and a flood index was developed for the NE Drakensberg mountains. All these targets were added and models were subsequently formulated.

In the case of rainfall for South Africa, Namibia and Zimbabwe the area-rainfall predictands were divided into early and late summer seasons: where 'early' was defined as (Nov + Dec + (Jan x 0.5)) / 2.5; and 'late' = ((Jan x 0.5) + Feb + Mar)) / 2.5. The time series therefore includes at least two major wet spells per annum (Levey 1993), affording a better linkage between local weather and regional climate. Another advantage is that January's influence is divided in accordance with Pathack (1993), who suggested its limited predictability was a function of varying mid-latitude and tropical regimes. The early summer rainfall is driven by baroclinic westerly wind shear, whilst late summer is dominated by barotropic convective instability (Makarau 1995).

2.3 DYNAMICAL STUDIES

To better understand the dynamics underlying regional climate variability, principal components for SST, geopotential height, OLR, 200 hPa wind fields, etc were analysed over the period 1980-1994, based on UKMO, ECMWF and CAC gridded products (168 consecutive months). The seasonal cycle was removed by subtracting the historical mean from each month of data. Spatial loading and temporal scores for rotated departures were analysed for the 1st 5 modes of each variable and inter-relationships were studied.

The first mode of upper level winds is represented by an alternation of zonal tropical flow from easterly in La Nina years to westerly in El Nino years, in response to meridional temperature gradients. Upper wind fluctuations over the tropical Atlantic are closely associated with tropical Pacific SSTs (as shown in Figure 3c) and may offer high levels of predictability. The second and third modes of upper wind variability takes the form of standing sub-tropical waves in the southern westerly belt (Figure 3a,b). The 2nd mode has a NE-SW orientation emanting from the tropical Atlantic and is associated with dry conditions and El Nino. The 3rd mode has a NW-SE orientation and is associated with wet conditions and La Nina. Ridging of the South Atlantic High south of the country and development of a continental vortex is facilitated in this mode. Convective modes from OLR principal components distinguish the tropical Atlantic and Indian monsoon zones as important features.

The dynamical studies suggest that heating anomalies of tropical SST affect the disposition and intensity of standing wave trains in the circulation and convection. These wave trains alternately cause wet or dry summers and shift the distribution of rainfall to early and late summer periods.

2.4 OPERATIONAL INPUTS AND NETWORKING

The rapid input of predictors to the models during the forecast cycle from August to November depends on Internet data flow. Of particular use are NCEP global anomalies of SST, pressure, upper and lower winds and OLR. Some useful predictor anomaly products come from Australia. European inputs are limited as a result of privatisation of their meteorological services. It seems adviseable to shift attention to NCEP products in subsequent follow-up studies. Verification of model forecasts is done in collaboration with the national weather services in the region and via NCEP African Desk products.

Consideration of climate impacts gave rise to analysis of proxy variables such as satellite NDVI (vegetation) and OLR (cloud depth), surface temperature, and evaporation, in addition to the stream flow and regional maize yield statistics already mentioned. In many cases data sets were either too short to train models or too patchy to be considered reliable. In the case of streamflow, man-made dams would alter long-term records. Cause-effect relationships between targets and principal components were assessed and models were developed to predict predictors commonly selected by the models, such as central Indian Ocean SST. This has improved our understanding of regional ENSO signals.

As part of a wider effort to mitigate climate impacts in southern Africa, the project leader initiated contacts across southern Africa to develop a network of expertise. A number of workshops were attended which brought together users and forecasters. User opinions on the value and type of seasonal outlooks were considered. Many users prefer small target areas, short time intervals and long lead times, whereas climatologists maximise predictive skill using large target areas, long time intervals, and short lead times. Compromises were identified. The problem of conflicting forecasts and the potential for confusion amongst the user community was considered. The networking enabled additional contacts and criticisms to be evaluated and incorporated into the project results.

Chapter 3 - Models and Skill Tests

In this section the models formulated to make seasonal outlooks are presented. The models are based on an optimum mix of predictors and two time periods are considered. The main set of models are developed for the period 1971-1992 (22 years), whilst a second longer set are developed over the period 1955-1992. The multi-variate regression equations are formulated using the forward step-wise method considering all candidate predictors within Statgraphics and Genstat software packages. The criteria for predictor retention was a significance level < 0.10 (90% confidence limit). This narrowed down the 100+ predictors to about 10 for each predictand. Finally, skill was evaluated and models were reduced to a maximum of four predictors for 22 year models to preserve the degrees of freedom in jackknife tests. For models based on the longer period up to five predictors were permitted. In most cases the optimum 22 year model with four predictors gave a correlation in excess of +0.6 between model forecast and observed. It was necessary to screen out co-linear predictors which corrupted a number of models. Although this could eliminate conspiring El Nino signals, no artificial skill would be introduced. Predictors closely associated with the ENSO phase are graphically illustrated in Figure 4a,b.

3.1 OPTIMUM MODELS (22 yrs, 4 predictors)

JFM RAIN	MULTI-VARIATE ALGORITHMS		r sq hindcast fit
E CAPE	+.91(oSOI)-.54(oQBO_)-.41(aCiv)+.25(aArBolr)		**= 81**
NW-FS	-.75(oQBO_)+.57(oWolr)+.40(oSOI)-.44(aElu)		**= 73**
TRVL	-.67(oCist)+.29(oATpc3)-.43(oATpc4)-.36(oQBO_)		**= 64**
Swazi-KZN	-.68(aATpc4)-.37(oSocNs)+.56(aMaurV)+.61(oNIv)		**= 71**
NAMIB	-.75(oCIst)+.48(aEolr)-.45(oPC10in)+.41(oSTang)		**= 72**
.ZIMB	+.65(oSTang)-.59(oCIst)+.39(aSWv)-.33(oSocNs)		**= 59**
ZAMB	-.75(aAtlW)+.62(aATpc2)-.44(aCIst)+.31(aSWp)		**= 80**
jBOTS	-.37(oAtlW)+.64(aWolr)+.38(aMauRv)-.56(oArBp)		**= 69**
jMAD	+.53(oAtlw)-.55(aSWp)-.77(oNEolr)-.40(aSTang)		**= 79**
jMalawi	-.43(aAtlw)-.58(aEist)+.46(aSEolr)+.38(oIpc3)		**= 66**
		S. Afr. Avg	*= 72.2*

NDJ RAIN	using JAS PREDICTORS	r sq hindcast fit
E CAPE	-.82(AtlW)+.51(Ipc3)+.38(ATpc4)-.42(ATpc3)	= 67
NW-FS	-.89(Pac1)+.33(SWp)+.32(WCi-ABp)-.22(oSocNs)	= 62
TRVL	+.94(SWp)+.57(SOI)+.94(ArBolr)-.79(ArBp)	= 56
Swazi-KZN	+.70(Wiv)+.89(aArBolr)+.48(SWp)-1.1(SIp)	= 58
NAMIB	-.85(ArBp)+.49(BeNst)+.56(ATpc2)-.30(ATpc5)	= 82
ZIMB	+.57(Niv)+.53(ATpc2)-.84(ATpc5)-.64(Pang)-.53(Ipc2)	= 71
TANZond	-.62(Ciu)-.61(Vang)-.44(SOI)-.29(Niv)	= 56
	S. Afr. Avg	*= 60.7*

CLIMATE IMPACTS		r sq hindcast fit
DBGrf	+.49(oATpc3)-.89(aATpc5)-.53(aArBp)+.39(aPac2)	= 70
SA mz dt	-.29(oQBO_)-.31(aIpc2)+.56(oColr)+.35(oSIst)	= 59
OFSmz	+.45(oSTang)-.60(oSWv)+.24(aWCI-ABp)-.28(aAtlW)	= 74
TRVLmz	+.35(aSWp)+.51(oWolr)-.41(oPC9ai)+.39(oSIst)	= 71
ZIM mz dt	-.94(oPang)+.40(mQBO)+.51(aCiv)+.26(aWIu)	= 77
MALpc	+.34(aSTang)+.83(aUang)+.34(aVang)-.29(aWIst)	= 86
TC day	+.54(oQBO_)+.47(oSocNp)-.37(aWCi-ABp)+.28(oCiv)	= 66
SUG tn	-.59(oAtlW)+.85(oWIp)+.43(aCIst)+.51(oNIv)	= 71
SUG yd	+.85(oNIv)+.31(oSEolr)+.31(aATpc1)+.44(oSOI)	= 84
ANCH	-.50(oIpc3)+.83(aVang)+.29(aUang)-.31(aCIv)	= 66
PILCH	-.39(oArBp)-.74(aCIst)+.58(oSWv)-.39(oScNu)	= 70
F-OIL	+.09-.86(aPC7)-.55(aEolr)-.48(oScnS)+.39(aArBp)	= 63
VICF	-.32(aMauRp)-.36(aArBst)-.37(oAtlW)+.73(oATpc2)	= 81
VAALr	-.48(oSIp)+.40(oPC4ag)+.39(aPC5ai)+.71(oPC9ai)	= 78
OKArun	+.07-.87(oBeNst)+.61(oScNu)-.75(aWi-Sip)+.36(oPC10in)	= 68
Gauteng Temperature using RF predictors		
mjT	+.54(oQBO)-.41(aPang)+.62(aArBp)-.51(oMolr)	= 62
jjT	-.40(aATpc2)+.52(oATpc5)+.62(oVang)-.31(aWIst)	= 54
jaT	+.28(oMauRv)-.27(oArBolr)-.47(oPang)+.41(aIpc3)	= 65
jjT%	+.42(oEolr)+.32(aMauRv)+.56(oCIv)-.44(oSocNp)	= 60
djfTEMP	+.92(aArBp)-.87(aColr)-.24(aMauRv)-.21(aPang)	= 80

Table 1 defines all target regions and table 2 the predictors used by South African rainfall and maize yield models. The climate impact models have a mean hindcast fit of 72.7 %. Jack-knife skill tests are, in most cases, very consistent with hindcast fit and confirm the validity of the model formulations. The model skill tests, which compare the 2 month lead forecast after jacknifing with actual value, are shown for a selection of models in Figures 5 to 10. Climate impact targets are more predictable than early summer rainfall. Figure 11a,b illustrates the predictors selected by climate impact models and the jacknife skill test model miss rate. Indian and Atlantic sea surface temperatures are the most commonly selected predictors in these models. Early summer rainfall models will yield poor forecasts about once every four years, whilst late summer and climate impact models achieve more reliable forecasts. Models based on a longer historical basis are outlined in the section below.

3.2 OPTIMUM MODELS (37+ yrs, 5 predictors)

JFM RAIN	MULTI-VARIATE ALGORITHMS	r sq hindcast fit
E CAPE38	+.56(oSOI)-.33(oQBO_)+.43(oATpc1)-.39(aMaurV)-.32(aSeiV)	**= 63**
NW-FS38	-.06+.42(oSOI)-.45(oQBO_)+.47(aWist)-.27(oCist)-.25(aNiV)	**= 45**
TRVL38	-.38(oQBO_)-.38(oATpc4)-.38(oCIst)+.38(oIpc4)+.27(oCiv)	**= 45**
ZIM37	-.08+.53(oPC4ag)-.26(oQBO_)+.55(aWiU)-.34(aCiU)-.46(oCist)	**= 40**
jBots37	-.26(oQBO_)+.40(oPC4ag)-.37(oSiP)-.37(aMaurV)-.44(aEiV)	**= 53**
NDJ RAIN	using JAS PREDICTORS	
E CAPE45	+.32(aATpc1)+.36(aATpc4)+.28(aIpc3)-.37(aPac1)-.29(aArBp)	**= 38**
NW-FS44	-.01+.48(aSOI)-.34(aSIp)+.22(aATpc4)+.22(aIPC1)+.34(aWCi-ABp)	**= 46**
TRVL44	+.02+.43(aSOI)+.25(aWCi-AbP)+.46(aWI-SiP)+.27(aWist)+.28(aSwV)	= 32
ZIM37	-.02-.45(aPac1)+.58(aWCi-ABp)+.49(aWi-Sip)+.40(aSist)-.41(aNiU)	**= 68**
CLIMATE IMPACTS		
SA mz42	-.05+.50(oSIst)+.58(aWiv)-.14(aWCi-AbP)-.27(aPC9ai)-.27(oCist)	**= 51**
SA mz dt39	+.02-.53(aPac2)+.24(aWi-Sip)+.21(oATpc1)+.30(oATpc3)+.23(oATpc5)	**= 37**

Mean hindcast fit of these models is 47%. This is appropriate for a longer training period where climatic teleconnections with local rainfall may vary. Problem models are those for early summer rainfall over the TRVL (Gauteng) area, late summer rainfall over Zimbabwe, and detrended (dt) South African maize yield.

Table 1: Target area definition; refer also to figure 2 a

Target	Parameter / size	Area / borders	Lat. long.	Time (months)
E Cape	rainfall avg. 10^5 km^2	province	32 S, 28 E	NDJ, JFM
NW-FS	rainfall "	provinces x 2	27 S, 26 E	NDJ, JFM
TRVL	rainfall "	multi-province	25 S, 29 E	NDJ, JFM
Swazi-KZN	rainfall "	multi-provinces	28 S, 32 E	NDJ, JFM
NAMIB	rainfall "	northern region	20 S, 18 E	NDJ, JFM
ZIMB	rainfall "	central highlands	18 S, 30 E	NDJ, JFM
ZAMB	rainfall "	Zambezi valley	15 S, 25 E	NDJFM
jBOTS	rainfall "	central region	22 S, 23 E	DJF
jMAD	rainfall "	central highlands	20 S, 46 E	DJF
jMalawi	rainfall "	central region	13 S, 32 E	DJF
TANZond	rainfall "	northern region	5 S, 35 E	OND
DBGrf	rainfall exceedence	Drakensberg > 1500 m	28 S, 30 E	DJF
SA mz dt	maize yield (detrended)	South Africa	27 S, 27 E	April
FS mz	maize yield / ha	province	29 S, 28 E	"
TRVL mz	maize yield / ha	province	25 S, 29 E	"
ZIM mz dt	maize yield (detrended)	Zimbabwe	18 S, 30 E	"
MAL pc	malaria incidence	eastern South Africa	28 S, 31 E	"
TC day	tropical cyclone	SW Indian Ocean	18 S, 60 E	DJFM
SUG tn	sugar cane prod	SE'ern South Africa	28 S, 32 E	following JJA
SUG yd'	sugar cane yield / ha	"	"	"
ANCH	anchovy catch / effort	S. Benguela	32 S, 16 E	following JJA
PILCH	pilchard catch / effort	S. Benguela	"	"
FOIL	fish oil yield	S. Benguela	"	"
VICF	Victoria Falls flow	Zambezi Valley	18 S, 26 E	JFMA
OKArun	Okavango streamflow	Andara Namibia	18 S, 21 E	JFMA
TEMP	temperature	Gauteng highveld	26 S, 28 E	DJF etc

Table 2: Predictor area definition (South African rainfall and national maize models)

NOTE: all times are JAS or SON months; most have areas > 10^6 km^2; see figure 2 b

Predictor / selected		Parameter	Area / borders	Lat. long.
SOI	3	southern oscil. index	Tahiti-Darwin pressure	west-central Pacific
QBO	3	quasi-biennial oscil.	zonal wind 30 hPa	Singapore
ArBolr	3	outgoing longwave rad.	Arabian Sea	10 N, 55 E
ATpc4	3	sea surface temp.	SW. Atlantic	40 S, 45 W
SocnS	2	sea surface temp.	Southern Ocean	45 S, 10-40 E
ATpc3	2	sea surface temp.	N. Atlantic dipole	10 N, 35 W
SWp	2	air pressure	SW Indian Ocean	27 S, 47 E
Pac1	1	sea surface temp.	central Pacific	0, 160 W
Sip	1*	air pressure	south Indian Ocean	12 S, 77 E
Atlw	1*	upper zonal wind	central Atlantic Ocean	2 S, 20 W
ArBp	1*	air pressure	Arabian Sea	10 N, 55 E
EiU	1*	surface zonal wind	eastern Indian Ocean	2 S, 77 E
Colr	1	outgoing longwave rad.	central Indian Ocean	2 S, 62 E
Cist	1	sea surface temp.	central Indian Ocean	2 S, 62 E
Wolr	1	outgoing longwave rad.	western Indian Ocean	2 S, 47 E
CiV	1	sfc. meridional wind	central Indian Ocean	2 S, 62 E
MaurV	1	sfc. meridional wind	Mauritius region	20 S, 62 E
NiV	1	sfc. meridional wind	northern Indian Ocean	10 N, 70 E
WCi-ABp	1	air pressure	south-north dipole	2 N, 55 E
Ipc2	1	sea surface temp.	south Indian dipole	35 S, 50 E
Ipc3	1	sea surface temp.	SE'ern Indian Ocean	25 S, 90 E
WiV	1	sfc. meridional wind	western Indian Ocean	2 S, 47 E
Sist	1	sea surface temp.	south Indian Ocean	12 S, 77 E

* highly correlated with global ENSO signals: SOI and Pac1 (nino3)

4. Conclusions and Recommendations

Long-range forecasts based on multi-variate models from this WRC project have been formulated and tested. The models meet accepted levels of reliability as demonstrated by skill validations. Models have been developed for numerous rainfall targets and for other climate-impacted resources such as maize yield and streamflow. A unique approach has been taken in that atmospheric predictors, in addition to SST, have been used.

Given the robust jacknife model validations (averaging 66.5 % for South African rainfall targets), it can be said that atmosphere-ocean interactions which teleconnect with weather systems over southern Africa may be anticipated through the predictors assembled here. Southern Africa is fortunate to be able to exploit the global ENSO phase so successfully in the prediction of climate. In utilising the project results, the impacts of fluctuating rainfall can be mitigated through strategic planning. An advance warning of drought and flood risk and seasonal rainfall prospects will provide additional security for food and water supplies, so they are less likely to inhibit future economic growth in southern Africa.

4.1 PRODUCTS AND RECOMMENDATIONS

The implementation of reliable long-range forecasts requires that understandable and consistent messages reach the users and public. Forecast messages can be disseminated through individual channels, networks, institutions, etc, in a variety of forms and formats. Long-range forecasters can check for consensus during the operational forecast cycle (August to November) via the Internet, etc. The South African Long-lead Forecast Forum has been set up to provide a common ground and peer-review of the methodologies employed. Partnerships between researchers and operational scientists can be pursued so that new directions explored in the academic environment are transferred to public institutions.

Long-range forecasters can encourage the public to use output products. Limits of predictability need to be stressed, particularly for the more ambiguous (non-ENSO) modes of climate variability. Forecast products should include the skill, confidence and error bars associated with the models. A range of independent model forecasts or categorical rainfall probabilities can be provided for each target.

Many long-range forecasters in southern Africa rely on climatic data from global sources such as the NCEP via the Internet. Monthly anomalies and patterns are extracted for model input. Model results and a forecast discussion are sent via fax to users aligned to each of the long-range forecast groups affiliated to the Southern African Long-Lead Forecast Forum. Some outlooks are posted on public access websites, as done for the products developed here. The Internet website 'Cape Town Weather Watch' is frequently consulted, as indicated by data requests averaging over 10 000 per day in 1997.

Prior to 1995 the WRC project leader based long-range forecasts on subjective assessments of a host of environmental variables. Since then, objective models have been developed and their outputs have been included in publicly disseminated seasonal outlooks. These outlook bulletins list the percentage departure from the training period (1971-1992) distinguished by target area and sub-seasonal time period. Since 1991, long-range forecasts have been issued and verified. During the 1995/96 and 1996/97 seasons the new objective forecasts performed well, with few exceptions.

Because of the potential economic impact the results can have, it is recommended that long-lead forecasts be provided free of charge to users, so that no liability is created by the service provider. Forecast groups can interchange user lists so that multiple opinions are supplied to the market. It is more likely that users will make a management decision based on a consensus of multiple forecasts. Forecaster need to understand how managers can hedge their strategic plans to mitigate climate impacts. Long-range forecasters need to consider how their message can reach the mass of population who have limited access to electronic media and are engaged in backyard subsistence farming. Simple radio messages in home languages could make an impact if backed by community outreach efforts at government level.

4.2 SUMMARY AND FUTURE DIRECTIONS

Resource managers will begin to plan for inter-annual climate variability once it is recognised that seasonal outlooks are reliable. Compromises on predictand time and space scale are needed: sub-seasonal forecasts over districts of 300 x 300 km are preferred and scientifically viable. Users will be more comfortable with forecasts that are clear and unambiguous. Simple terminology and consensus forecasts are the goal.

Forecasters need to liaise regularly with users to re-define targets and to appreciate the management decisions likely to arise from their products. Targets which directly impact the economy, such as maize yield and natural streamflow, can be considered equally to rainfall, and appear to offer higher levels of predictability because of their accumulative nature. Public awareness of climate impacts and predictability can be explored through development of community outreach projects and educational training resources. It is believed that agriculture and other climate-impacted economic activities can benefit from long-range forecasts based on ENSO teleconnections, so improving the economic growth potential of southern Africa.

Further work on seasonal predictability could involve use of re-analysed global model products such as NCEP monthly fields from 1957. The model training period could thus be extended, giving increased degrees of freedom and statistical confidence. Studies based on global model re-analysed products and their principal components would enable a wiser choice of independent predictors submitted for model optimisation. Because skill correlations average ~66% for most models, the predictability 'gap' could be closed further through additional work.

The statistical models described here are based on past history and the potential for climate 'drift' suggests that the models could become gradually obsolete. Therefore regular updating of predictors and model algorithms is necessary. Effort directed at predicting the predictors and understanding Atlantic and Indian Ocean El Nino events, would provide useful inputs and increased forecast lead times.

5. REFERENCES

Cane, M A, Eshel, G, and Buckland, R W, 1994, Forecasting Zimbabwean maize yield using eastern equatorial Pacific sea surface temperatures, *Nature*, 370, 204-205

Jury, M R, McQueen, C A, and Levey, K M, 1994, SOI and QBO signals in the African regoin, *Theor Appl Climatol*, 50, 103-115

Jury, M R, 1996, Regional teleconnection patterns associated with summer rainfall over South Africa, Namibia and Zimbabwe, *Intl J Climatol*, 16, 135-153

Landman, W A, 1995, A cannonical correlation analysis model to predict South African summer rainfall, *Exp Long-lead Frcst Bull*, 4, 4, 23-24

Levey, K M, 1993, Intra-seasonal oscillations of convection over southern Africa, *MSc Thesis*, Oceanogr Dept, Univ Cape Town, 226 pp

Makarau, A, 1995, Intra-seasonal oscillatory modes of the southern Africa summer circulation, *PhD Thesis*, Oceanogr Dept, Univ Cape Town, 324 pp

Mason, S J, 1992, Sea surface temperatures and South African rainfall variability, *PhD Thesis*, Univ Witwatersrand, Johannesburg, 235 pp

Nicholson, S E, Kim, J, and Hoopingarner, J, 1988, *Atlas of African rainfall and its interannual variability*, Florida State Univ, Tallahassee

Pathack, B M R, 1993, Modulation of South African summer rainfall by global climatic processes, *PhD Thesis*, Oceanogr Dept, Univ Cape Town, 214 pp.

Rocha, A M C, 1992, The influence of global sea surface temperatures on southern African summer climate, *PhD Thesis*, Univ Melbourne, 248 pp.

Walker, N D, 1989, Sea surface temperature-rainfall relationships and associated ocean-atmosphere coupling mechanisms in the southern African region, *PhD Thesis*, Oceanogr Dept, Univ Cape Town, 173 pp

6. PAPERS PUBLISHED DURING PROJECT

Jury, M R, Pathack, B M R, Rautenbach, C deWet, and VanHeerden, J, 1996, Drought over southern Africa and Indian Ocean SST: statistical correlations and GCM results, Global Ocean Atmos Sys,4, 47-63

Jury, M R, 1996, Regional teleconnection patterns associated with summer rainfall: South Africa, Namibia and Zimbabwe, Intl J Climatol, 16, 135-153

Jury, M R, and Levey, K M, 1997, Vertical structure of wet spells over southern Africa, Water S A, 23, 51-55

Jury, M R, Weeks, S, and Gondwe, M P, 1997, Satellite vegetation as an indicator of climate variability over southern Africa, S. Afr. J Science, 93, 34-38

Jury, M R, and Majodina, M, 1997, A climatology of southern Africa extreme weather: 1973-1992, Theor Appl Meteorol, 56, 103-112

Levey, K M, and Jury, M R, 1996, Composite intra-seasonal oscillations of convection over southern Africa, J Climate, 9, 8, 1910-1920

Gondwe, M P, and Jury, M R, 1997, Sensitivity of vegetation (NDVI) to climate over southern Africa: relationships with rainfall and OLR, S A Geogr J, (in press)

Jury, M R, 1997, Southeast Atlantic warm events: composite evolution and consequences for southern African climate, S Afr J Marine Sci, (in press)

Jury, M R, 1997, Inter-annual climate modes over southern Africa from satellite cloud OLR 1975-1994, Theor Appl Meteorol, (in press)

Jury, M R, and Levey, K M, 1997, Vertical structure of mid-latitude rossby waves during summer wet spells over southern Africa, S Afr J Science, (in press)

Majodina, M, and Jury, M R, 1997, Composite winter cyclones south of Africa: evolution during eastward transit over the Agulhas warm pool, S Afr J Marine Sci, (in press)

Makarau, A, and Jury, M R, Predictability of Zimbabwe summer rainfall, Intl J Climatol, (in press)

Mason, S J, and Jury, M R, 1997, Climatic fluctuations in southern Africa: a reflection on underlying processes, precursors and predictability, Prog Phys Geogr, (in press)

FIGURES (following pages)

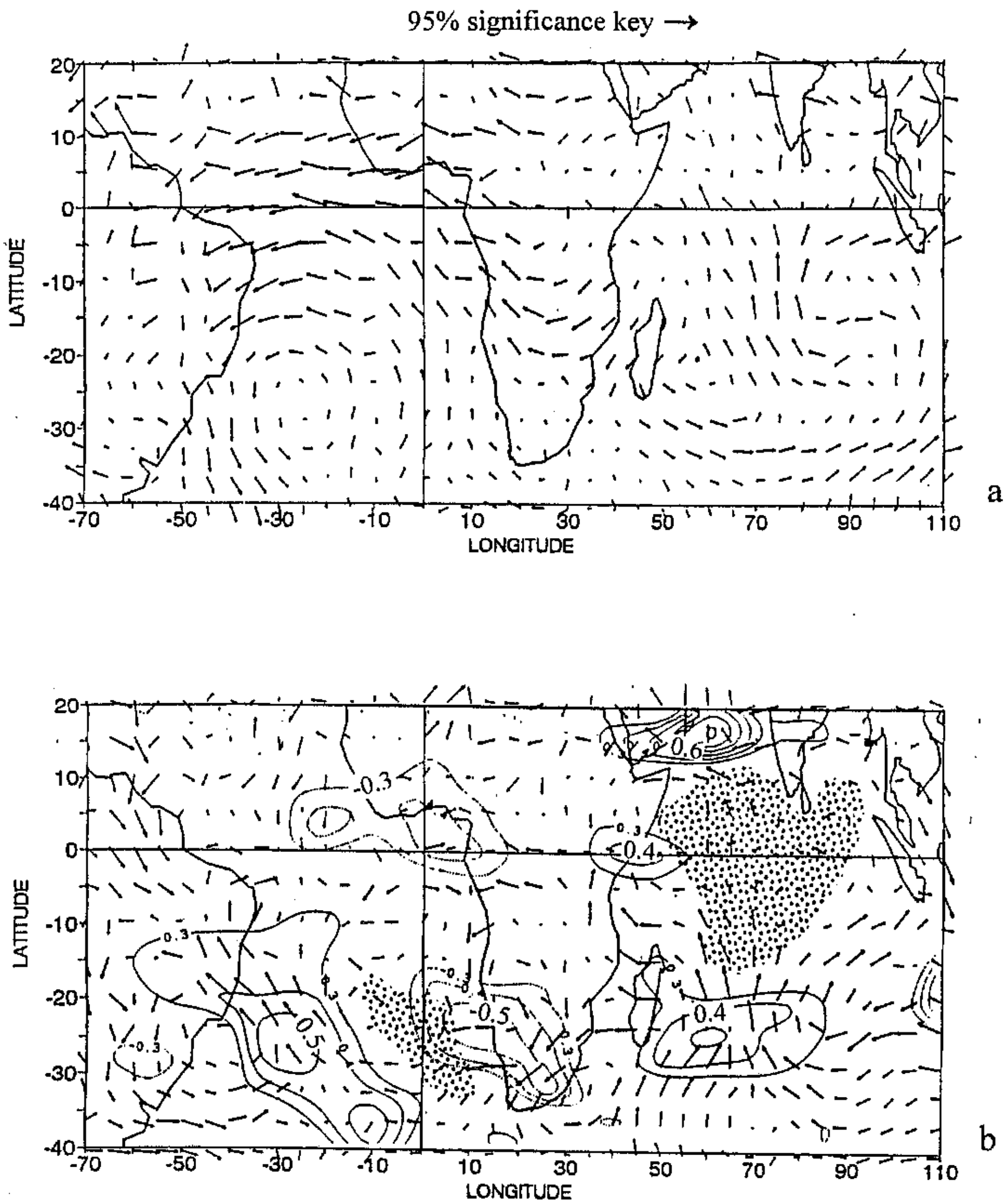


Figure 1 a (top) Pair-wise correlation map of South African highveld summer rainfall versus four month lead time (September) 200 hPa (12 km) winds. Correlation vectors are constructed at each grid point from individual U and V wind correlations; b (lower) is the same correlation map in respect of 700 hPa (3 km) winds at zero lag (January) with OLR (isolines) and SST (shaded) significant negative correlations superimposed.

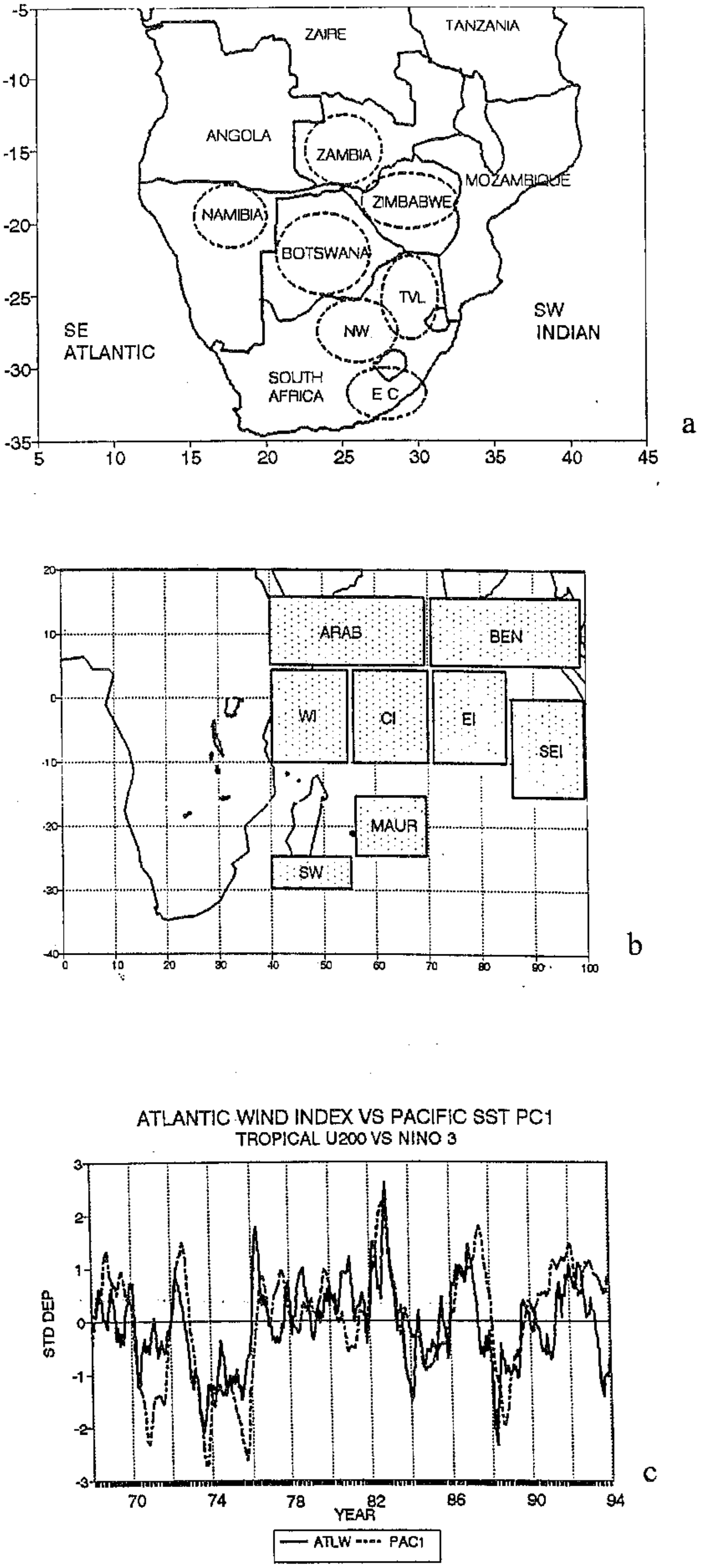


Figure 2 a (top) rainfall target areas used in the study to develop forecast models. Each area is formed from an average of station data within the specified domain, b (middle) an example of predictor target areas used to develop forecast models. Environmental variables are averaged in specified areas to form predictor indices, c (bottom) comparison of predictor indices for Atlantic winds and Pacific SST, revealing ENSO events.

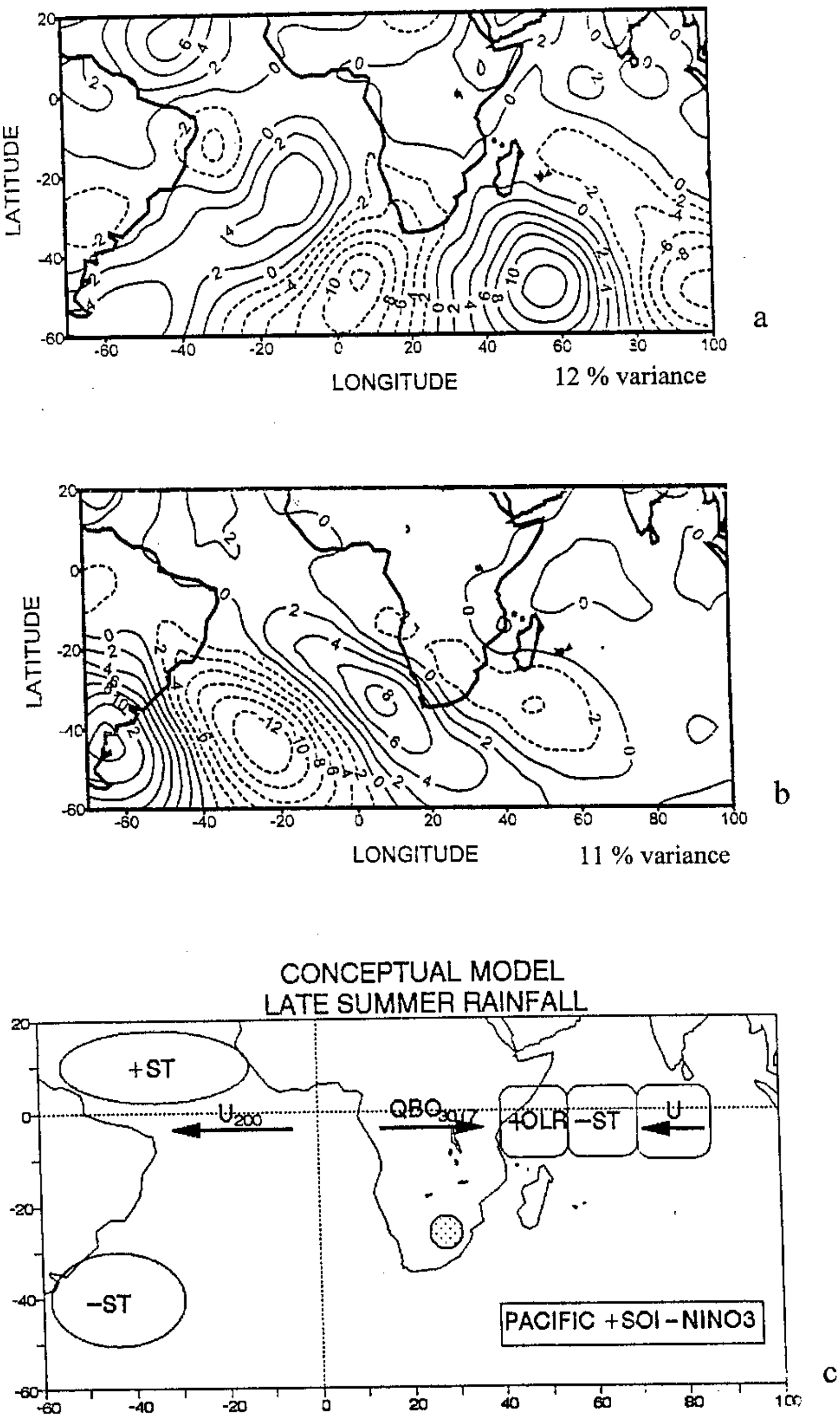


Figure 3 a, b (top, middle) principal component spatial loading patterns for the 2nd and 3rd upper wind modes, expressed as scalar meridional components, showing the forward (top) and backward leaning patterns involving standing waves in the sub-tropical jet stream - constructed from correlation matrices, varimax rotated; c (lower) schematic representation of predictor indices favouring increased late summer rainfall over the South African highveld. ST= sea temperature, U_{200} is the Atlantic wind index, OLR = outgoing longwave radiation, a number of ENSO variables should indicate cool phase for increased rainfall.

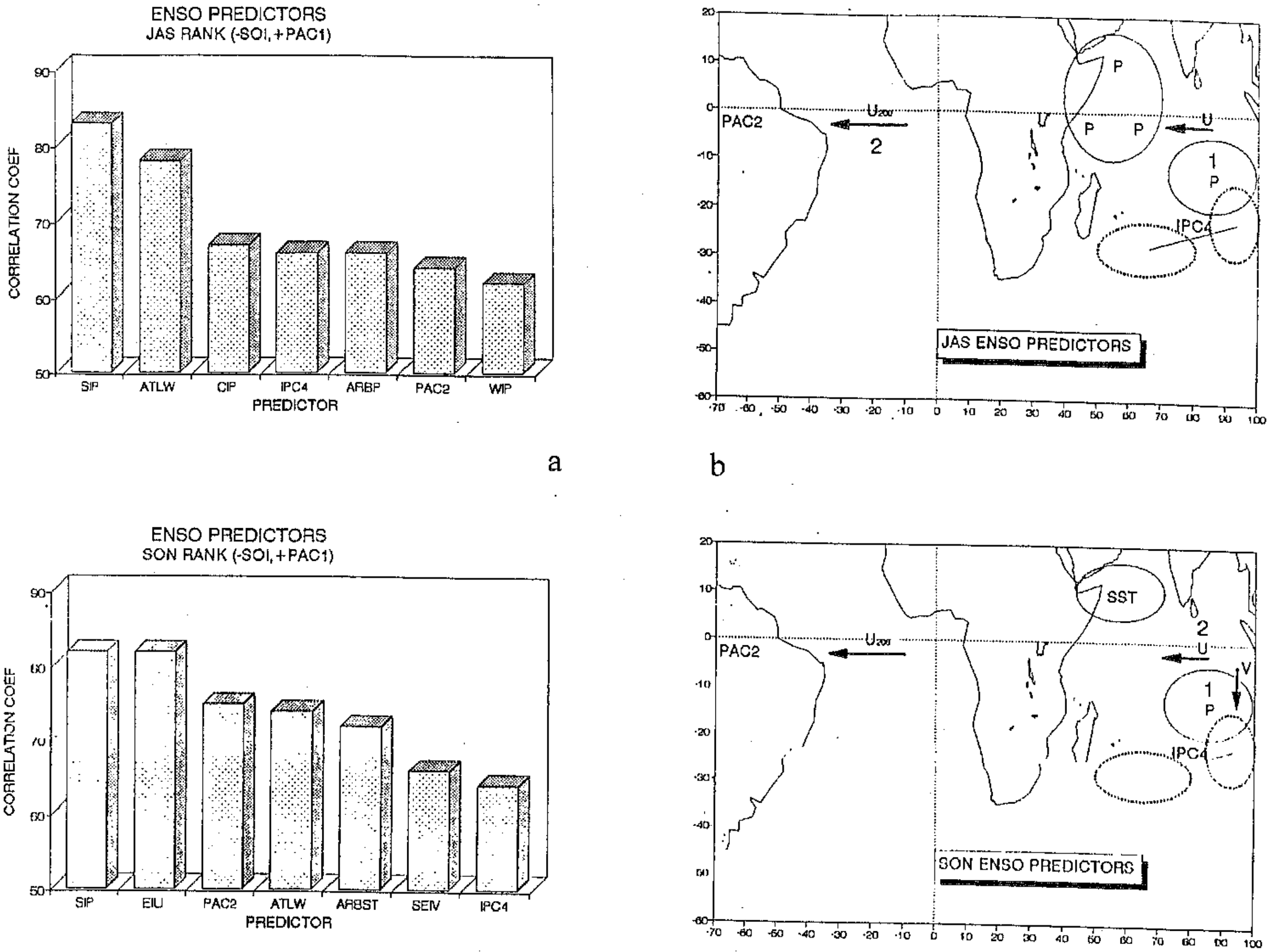
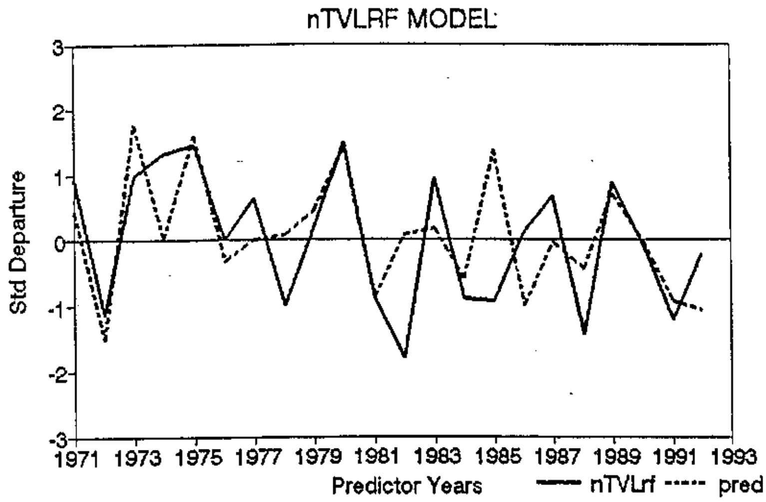


Figure 4 (left) environmental variables surrounding southern Africa which are closely associated with the SOI and central Pacific SST. Variable locations and meaning are indicated at right, where P = pressure, IPC4 = SST pattern and vectors show sense of wind direction in cool phase ENSO. Large number refer to correlation rank.



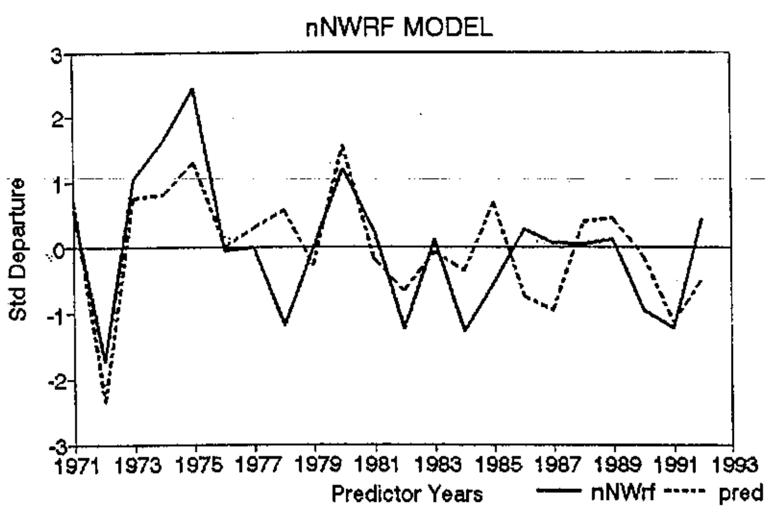
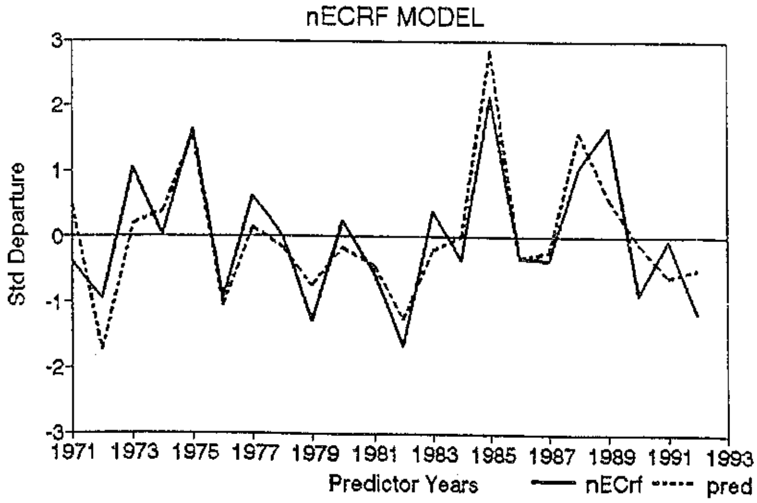


Figure 5 Jack-knife skill tests for 4 predictor 22-year models, showing jacknife-predicted (dashed) and observed (solid) values for former Transvaal highveld (top) and NW+FS provinces of South Africa (lower), both for early summer (November to mid-January).



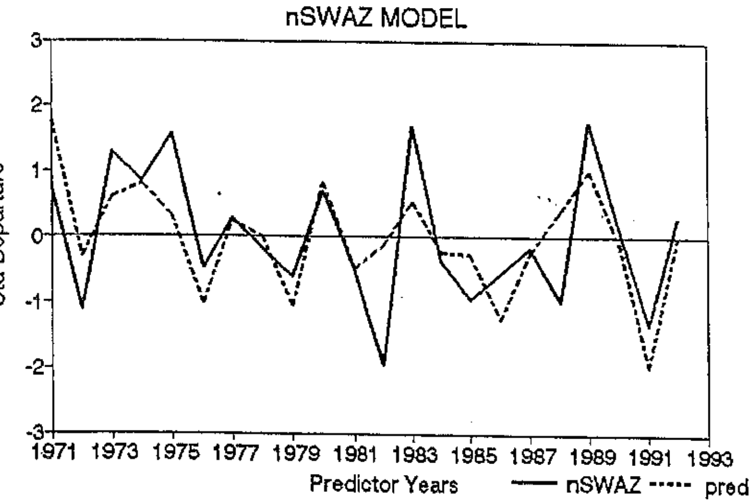
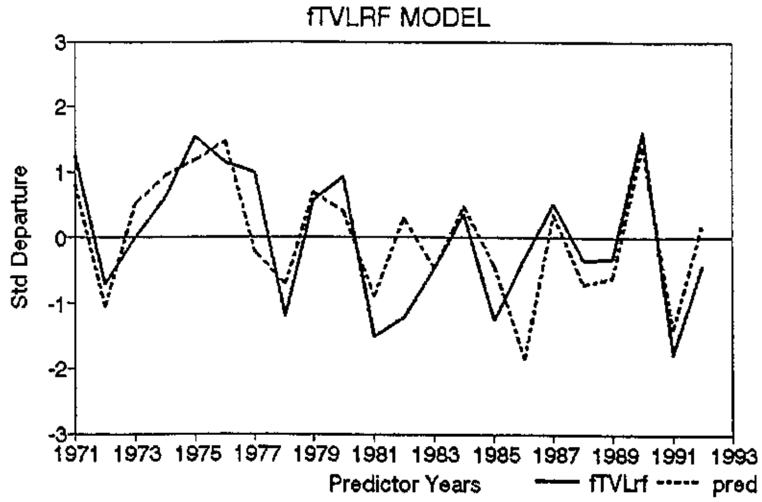


Figure 6 Jack-knife skill tests for 4 predictor 22-year models, showing jacknife-predicted (dashed) and observed (solid) values for eastern Cape coastal plains (top) and lowveld Swaziland area (lower), both for early summer (November to mid-January).



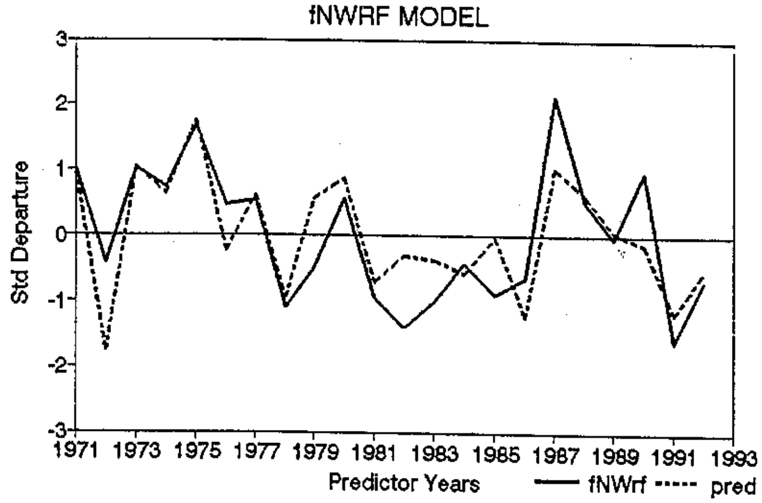
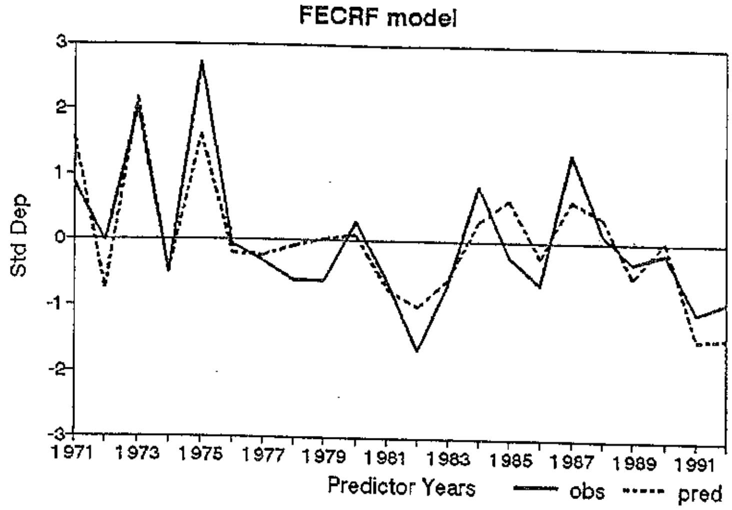


Figure 7 Jack-knife skill tests for 4 predictor 22-year models, showing jacknife-predicted (dashed) and observed (solid) values for former Transvaal highveld (top) and NW+FS provinces of South Africa (lower), both for late summer (mid-January to March).



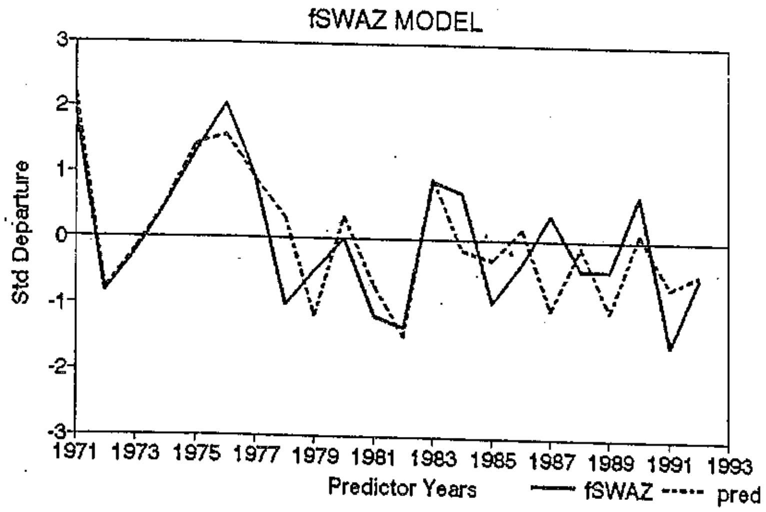
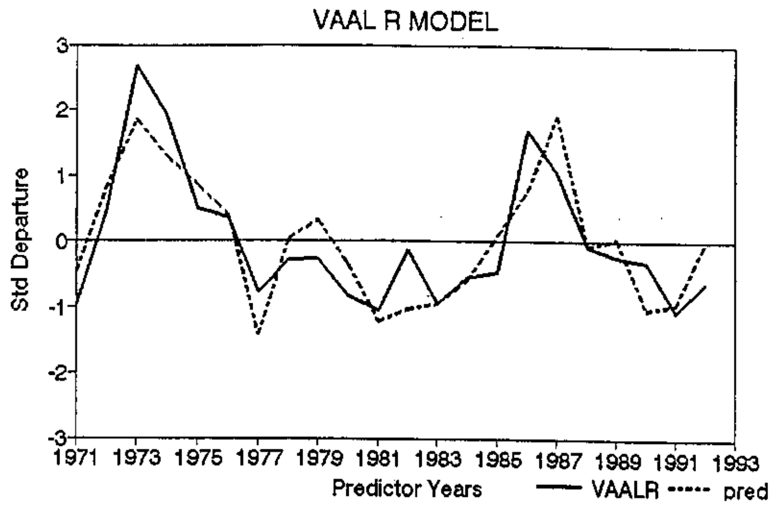


Figure 8 Jack-knife skill tests for 4 predictor 22-year models, showing jacknife-predicted (dashed) and observed (solid) values for eastern Cape coastal plains (top) and lowveld Swaziland area (lower), both for late summer (mid-January to March).



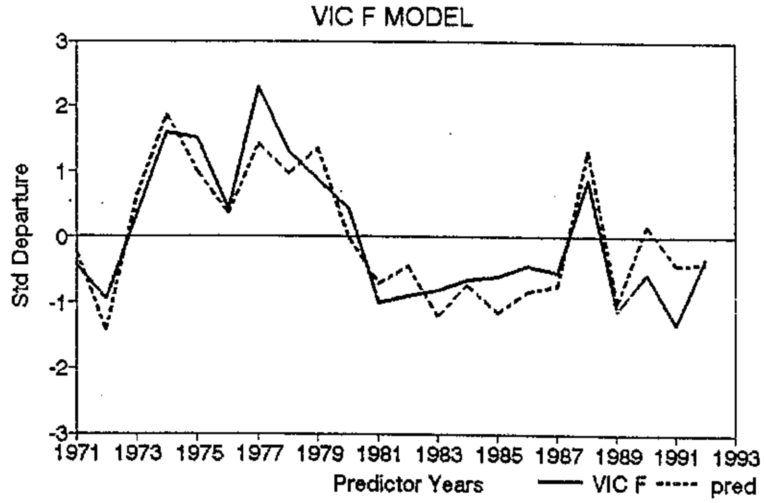
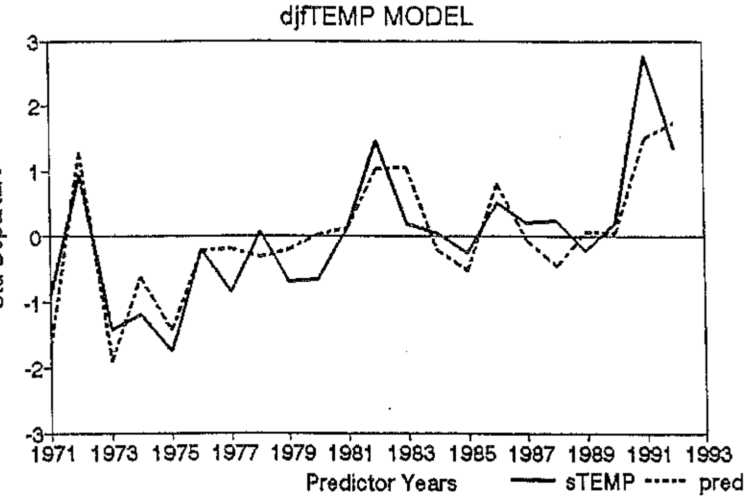


Figure 9 Jack-knife skill tests for 4 predictor 22-year models, showing jacknife-predicted (dashed) and observed (solid) values for Vaal River runoff (top) and Victoria Falls streamflow (lower), both for late summer (January to March).



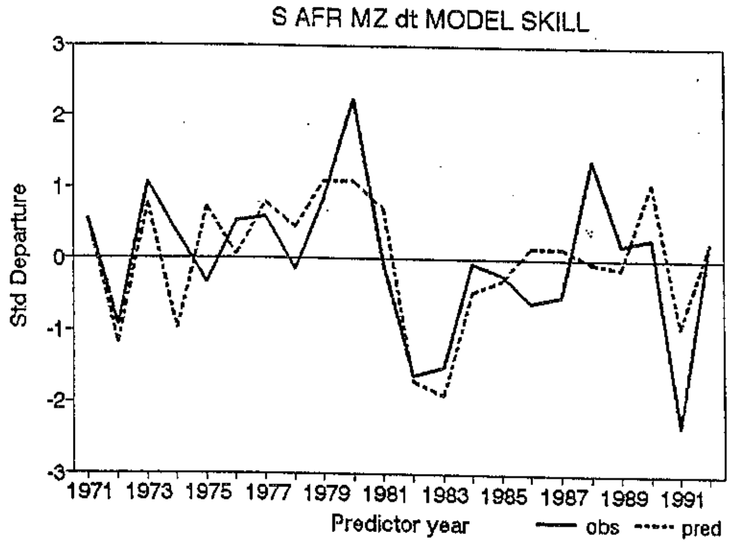


Figure 10 Jack-knife skill tests for 4 predictor 22-year models, showing jacknife-predicted (dashed) and observed (solid) values for Gauteng maximum summer tempertures (top) and South African national maize yield (lower).

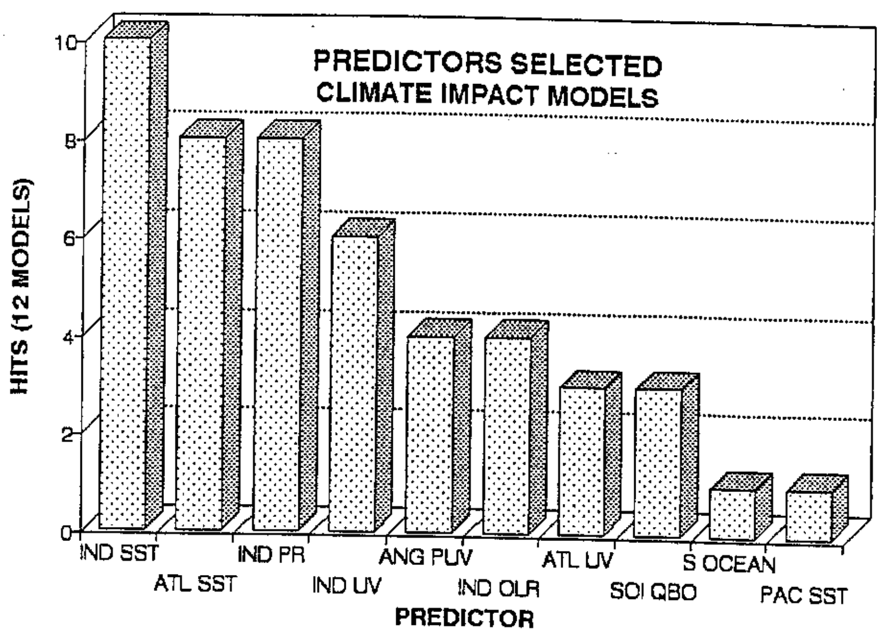


Figure 11 Summary figures showing predictors selected by the climate impact models (top) and the model miss rate in skill tests for the climate impact and rainfall (RF) models. This indicates that streamflow and crop yield models out-perform rainfall models. Late summer is more predictable than early summer (~25 % miss rate).