Verification of South African Weather Service operational seasonal forecasts

by

Peggy Seanokeng Moatshe

DISSERTATION

submitted in partial fulfillment of the requirement for the degree

MASTER OF SCIENCE

in the

FACULTY OF NATURAL AND AGRICULTURAL SCIENCES

at the

UNIVERSITY OF PRETORIA

SUPERVISOR: Prof W.A Landman

July 2008

DECLARATION

I declare that the dissertation that I hereby submit for the MSc degree in Meteorology at the University of Pretoria has not previously been submitted by me for degree purposes at other university.

Signature:………………………… Date:………………………

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

ACKNOWLEDGEMENTS

The author wishes to express her appreciation to the following persons and organizations for their assistance and contribution to make this dissertation possible.

- Dr Simon Mason for implementing the verification program codes at South African Weather Service
- Prof Willem Landman for his advice and for creating a suitable working environment at the University of Pretoria and at South African Weather Service
- Mrs. Anna Bartman for her support with homogenization, data manipulation and verification program codes
- Mrs. Mary-Jane Kgatuke for her support and encouragement
- South African Weather Service Librarians for providing the scientific papers and books cited in the dissertation
- Last but not least, my gratitude goes to my family especially my husband and friends for their persistent encouragement and Almighty God for giving me the strength to challenge this complicated and unexpected but interesting profession

Verification of South African Weather Service operational forecasts

Peggy Seanokeng Moatshe

Supervisor:	Prof W.A Landman
Department:	Department of Geography, Geoinformstics and Meteorology
Degree:	Master of Science
Faculty:	Faculty of Natural and Agricultural Sciences
University:	University of Pretoria

Summary

The South African Weather Service rainfall seasonal forecasts are verified for the period of January-February-March to October-November-December 1998-2004. These forecasts are compiled using different models from different institutions. Probability seasonal forecasts can be evaluated using different skill measures, but in this study the Ranked Probability Skill Score (RPSS), Reliability Diagram (RD) and Relative Operating Characteristics (ROC) are used. The RPSS is presented in the form of maps whereas the RD and ROC are analyses are presented in the form of graphs. The aim of the study is to present skill estimates of operational seasonal forecasts issued at South African Weather Service

A limited number of forecasts show positive RPSS value throughout the validation period. From RD and ROC analysis, there is no skill in predicting the normal category as compared to below-normal and above-normal categories. Notwithstanding, the frequency diagrams show that the normal category was often given a large weight in the operational forecasts.

The value of verifying seasonal forecast accuracy from the user's perspective is important. The understanding of seasonal forecast performance helps decision makers to determine when and how to respond to expected climate anomalies. Therefore the frequent update of the seasonal forecast verification is important in order to help Users make better decisions.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

j	Category
J	Number of categories
O_j	Components of the vector
O_m	Cumulative observation vector
P_j	Probability assigned to the $_j{}^{th}$ category
P_m	Cumulative forecast vector
$P(o_j \vert y_j)$	Conditional distribution of observation given the forecasts

LIST OF ABBREVIATIONS

ABI	Amalgamated Beverages Industry
AGCMs	Atmospheric General Circulation Models
AMJ	April-May-June
CCM32	Model developed at the National Centers for Atmospheric Research from different atmospheric general circulation model
CRG	Climate Research Group
CSAG	Climate System Analysis Group
DKRG	Deutsches Klimrechenzentrum-German Climate Computing Center
ECHAM3.6	The European Community-HAMbug model developed at the Mark Plank Institute for Meteorology in Germany from different atmospheric circulation Model
ECMWF	European Center for Medium-Range Weather Forecast
ENSO	El Niño Southern Oscillation
GCMs	General Circulation Models
GFCSA	Global Forecasting Center for Southern Africa
JAS	July-August-September
JFM	January-February-March
IRI	International Research Institute
L	Transvaal Lowveld
NCEP-MRF9	Atmospheric general climate model developed at the National Centers for Environmental Prediction in the United States
NT	Northern Transvaal
OND	October-November-December

RD	Relative Diagram
RGSCS	Research Group for Seasonal Climate Studies
ROC	Relative Operating Characteristics
RPS	Ranked Probability Score
RPS_{clim}	Ranked Probability Score for Climatology
RPSS	Ranked Probability Skill Score
SAWS	South African Weather Service
SST	Sea-Surface Temperature
UKMO	United Kingdom Meteorological Office

Chapter 1

Introduction

1.1 Winter and summer rainfall region of South Africa

Several studies have been done on the demarcations of the regions of South Africa. Schulze (1965) classified the climatic region of South Africa based on geographical consideration such as mountain ranges, rivers and political boundaries. By applying the above mentioned method, Southern Africa was demarcated into 15 regions. These regions were named according to their political boundaries, e.g. L and NT for Transvaal Lowveld and Northern Transvaal, respectively. The method was declared as outdated because some of the political boundaries do not exist anymore. Therefore, in 1900 Koppen (Sanderson, 1999) followed with a different classification method.

The climate data of South Africa used for Koppen classifications is for the period 1961 to 1990. By applying the above mentioned classifications, South Africa was demarcated into dry and relatively wet regions. The Western Cape receives the highest amount of rainfall in winter while rainfall occurs throughout the year in the Eastern Cape Province. Winter rainfall over the most southern and western region, as well as the eastern coastal regions, is due to frontal-type weather. The above change abruptly to drier areas is due to the Cape mountain ranges which form a natural barrier between the humid area and the dry area in the north. KwaZulu-Natal Province which is located in the eastern part of the country is humid. Limpopo Drakensberg mountain range which causes very high orographic rainfall on the one side of the Mpumalanga province while drier conditions are experienced on the other side of the mountain. A greater part of Limpopo and Northwest Provinces is characterized by semi-dry conditions. The eastern part of the Northwest Province (e.g. Rustenburg) is more of humid while the far west (e.g. Kuruman) is dry. In Limpopo Province, the far northern part (e.g. Musina) is drier than the southern part.

1.2 Sea-Surface Temperature associations with South African rainfall

The strongest links between Sea-Surface Temperature (SST) patterns and mean seasonal rainfall and temperature are found in tropical regions and seasonal forecasts in tropical regions are most skillful (Mason et al., 1994). The best known ocean-atmosphere link is the El Niño Southern Oscillation phenomenon. El Niño occurs when SSTs in the tropical pacific are above average with an anomaly greater than 0.5 for three consecutive months. El Niño can disrupt the normal pattern of seasonal rainfall and temperature around the globe, bringing for example large changes in seasonal rainfall causing droughts in some regions and floods in others. Although the strongest links between SST and seasonal climate are found in the tropics, there is good evidence that similar, if not weaker, links are present in other parts of the globe including South Africa.

Comprehensive studies were made on the link between the El Niño and La Niña events and summer rainfall over South Africa (e.g. Walker and Bliss, 1930 and Van Heerden et al.,1988). It was only after the 1982/1983 El Niño event which caused large scale drought conditions over South Africa that the importance of studying the oscillation in detail was realised. During an El Niño event, South Africa generally experiences below-normal summer rainfall conditions, but this does not occur during all the events.. The amount of rainfall received over the country differs considerably from one event to the other. The 1996/1997 El Niño season is a good example of a strong event where normal rainfall conditions were experienced over a large part of South Africa. Below normal rainfall conditions in South Africa are not always tied to an El Niño event and the same is the case with La Niña events, i.e. above-normal rainfall conditions can be experienced in the absence of a La Niňa events (Kruger, 1998).

SST anomalies of the oceans adjacent to South Africa are also related to southern Africa rainfall variability (Walker, 1990) and may help to improve predictions of rainfall variability over southern Africa (Landman and Mason, 1999). Mason (1990) found that the annual rainfall total of the summer rainfall region varies closely in phase with the SST gradient intensity in the southwestern and southeastern Atlantic Ocean. Mason (1995) discovered that when Subtropical Atlantic Ocean SSTs are anomalously cold (warm), southern Africa is characterised by dry (wet) conditions with the strongest rainfall SST association over the central southern Atlantic ocean. For equatorial Indian Ocean, if the SSTs are warm (cool) to the southeast of Africa, wet (dry) condition are experienced over southeast Africa (Jury, 1992).

Above average SSTs over the tropical Indian Ocean together with other conditions can lead to the development of tropical cyclones and an increase in the frequency of cyclones. The association between SSTs in the Equatorial Indian Ocean and seasonal rainfall over South Africa is nonlinear (Mason and Jury, 1997). The rainfall-SST associations vary during the summer rainfall season defined as October to March. The central Equatorial Indian Ocean SSTs show very significant associations with February and March South African rainfall (Pathack et al., 1993). The western tropical Indian Ocean seems to be an important source of atmospheric moisture, especially during the second half of summer where most of the summer regions receive tropical moisture. Therefore if the SSTs over the western tropical Indian Ocean anomalously high then there will be an enhancement of rainfall over South Africa (Landman and Mason, 1999). Apart from the Equatorial Indian Ocean and Atlantic Ocean SSTs, there is also the Agulhas current that contributes to the rainfall over South Africa. The variability of Agulhas current together with the atmospheric circulation account for some of the variability of the summer rainfall over South Africa (Mason, 1990).

1.3 Seasonal forecast models and methods

Weather forecasts provide information of weather (e.g. prediction of rainfall, frontal passages) expected over the next few days, but beyond about a week ahead it is not possible to predict these day-to-day changes in detail (Bartman, 2002). However, mean conditions over a number of months can be predicted and these are called seasonal forecast. Seasonal forecasts are made possible by the Earth's surface conditions, e.g. SSTs that fluctuate slowly fluctuations leaving some memory in the atmosphere and therefore making the atmosphere partly predictable. The slow fluctuations of SST can be predicted, to some extent, up to about 6 months ahead. The links between SST and atmospheric conditions can be represented in computer models of the atmosphere and ocean. Computer models developed at different institutions, similar to those used in making daily forecasts and climate change prediction.

Major understanding of the predictability of the atmosphere at seasonal-to-interannual time scales has been achieved during the most recent years or so (Carson, 1998). In determining the future behaviour of the climate system, for example the state of the phase of ENSO or the mean global circulation pattern, from knowledge of its present state and past behaviour, two approaches are used, namely an empirical- and physically-based models (Trenberth, 1992; Mason et al., 1996). Empirical models rely on past statistical associations between atmospheric and Oceanic parameters. Physical models attempt to forecast the time-average of future atmospheric conditions by simulating the dynamic and thermodynamic processes, which determine the state of the non-linear atmosphere.

Statistical methods use techniques such as regression equations to relate the atmospheric conditions to a set of independent variables such as SSTs or earlier states of the atmosphere (Ward and Folland, 1991; Barnston, 1994; Drosdowsky, 1994; Hastenrath et al., 1995; Huang et al., 1996). The first principles approach utilises equations believed to represent the physical, chemical and biological processes governing the climate system. These models are called General Circulation Models (GCMs) and are being used extensively in seasonal forecasting (Palmer and Anderson, 1994; Ji et al., 1994; Hunt, 1997; Mason et al., 1999; Gates et al., 1999). The Statistical forecast techniques have predictive skill that can be ascribed to low frequency changes in SSTs of the oceans, particularly the tropical Pacific Ocean. Graham and Barnett (1995) discovered that statistical models perform better than physical models in many regions of the globe but the climate variables such as rainfall are very difficult to model statistically (Wilks 1995). On the other hand, physical models at long-range climate forecasting perform badly if they are used without the appropriate SST forcing.

A combination of Physical Ocean models, Atmospheric General Circulation Models (AGCMs) and statistical models is helpful in predicting seasonal rainfall and temperature because different models address different problems. Coupled dynamic ocean-atmosphere models are successfully used in the forecasting of climate parameters such as SSTs (Kirtman et al., 1997).

Meteorological institutes around Africa and other continent use several AGCMs that produce seasonal climate predictions as input to the final seasonal forecast. During the period 1997-2000, CCM3.2 developed at the National Centers for Atmospheric Research (Hack et al. 1998) and run at the IRI; ECHAM3.6, developed by the Max-Planck Institute for Meteorology (DKRZ 1992) and the NCEP-MRF9 developed by the National Centers for Environmental Prediction (Livezey et al 1996) and run by collaborators at the Queensland Center for climate Applications in Australia where some of the AGCMs used by the South African Weather Service (SAWS). Multi-decadal simulations of approximately 50 years, using observed SSTs, have been produced for each of these models, each of which has produced at least 10 ensemble members. These long historical runs provide estimates of model potential predictability and characteristics of model climatology that are essential to interpreting the seasonal predictions from each model (Mason et al., 1999).

1.4 Seasonal forecasting at South African Weather Service (SAWS)

Since the early 1990s different research groups within South Africa were involved in research to investigate suitable seasonal climate prediction methods (Mason et al., 1996). This was motivated by the loss of life and damages that were experienced due to flooding and drought in the country (Klopper, 1997). University of Pretoria (Rautenbach and Smith, 2001), the University of Zululand (Jury, 1995), the Climate Research Group (CRG) at the University of Witwatersrand (Mason, 1998), and the Climate System Analysis Group (CSAG) at the University of Cape Town (Tennant and Hewiston, 2001) were part of this research.

At SAWS, which was then called South African Weather Bureau, the Research Group for Seasonal Climate Studies (RGSCS) was founded in 1994 and its main goals were to investigate statistical relationship of rainfall over southern Africa and other climate system variables (Klopper, 2002). Its task is also to issue a regular seasonal rainfall forecasts for South Africa, Namibia and Botswana. The RGSCS continued with research and development of statistical model and later expanded its work to include dynamical models to improve the long-term climate prediction in South Africa. The Long-Range Forecasting Group was subsequently established at the SAWS. This group makes use of the empirical, dynamical and empirical-dynamical forecasting techniques to compile forecasts.

The Long-Range Forecasting Group at the SAWS and Climate Systems Analysis Group based at the University of Cape Town are currently the two main institutions in South Africa issuing seasonal forecast. A number of institutions in South Africa are involved in the production and dissemination of long-range seasonal forecasts and the Global Forecasting Center for Southern Africa

(GFCSA, www.gfcsa.net) serves as a focal point for such activities. South African Weather Service also uses product from three institutions from overseas namely United Kingdom Meteorological Office (UKMO), International Research Institute for Climate and Society (IRI) in the United State of America and the European Center for Medium-Range Weather Forecast (ECMWF) in Europe.

The Long-Range Forecasting Group meets every month to make seasonal forecasts for the next seasons using information from the above-mentioned institutions (Klopper, 2002). The consensus takes place and at the end, there will be a final seasonal forecast in the form of a map (figure1.1) given as the probability (in percentage) for each of three equi-probable categories (above-normal, normal and below-normal) expected to occur over a certain region (Klopper, 2002). The Long-Range Forecasting Group issues probabilistic rather than deterministic forecasts in order to account for the uncertainties in the climate system. The category with the highest probability is the most likely to occur but there are also lesser probabilities assigned to each of the remaining two categories and they are never small enough to be disregard.

Expected Total Rainfall for the period August-September-October 2006

AREA 1:

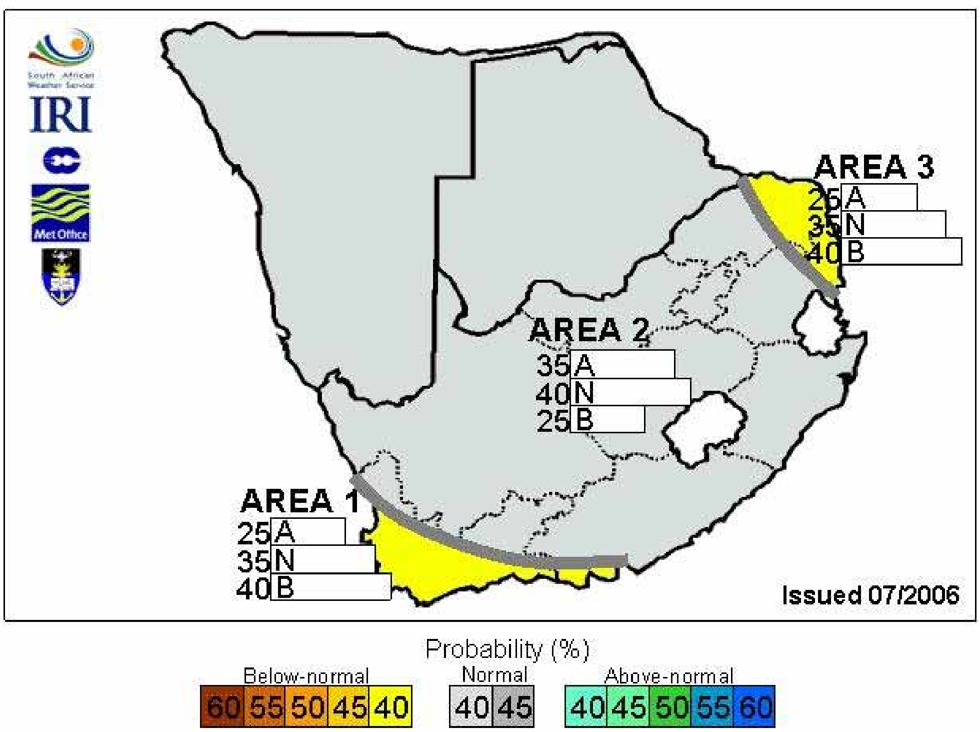
25% chance that the total rainfall for this period will be ABOVE-normal.
35% chance that the total rainfall for this period will be normal.
40% chance that the total rainfall for this period will be BELOW-normal.

AREA 2:

35% chance that the total rainfall for this period will be ABOVE-normal.
40% chance that the total rainfall for this period will be normal.
25% chance that the total rainfall for this period will be BELOW-normal.

AREA 3:

25% chance that the total rainfall for this period will be ABOVE-normal.
35% chance that the total rainfall for this period will be normal.
40% chance that the total rainfall for this period will be BELOW-normal.



Please send comments to longrange@weathersa.co.za

Figure 1.1 Probability seasonal forecasts issued by South African Weather Service.

1.5 Importance of seasonal forecasts

Seasonal forecasts possess a certain level of economic value for some sectors even though they exhibit considerable uncertainty (Klopper, 2002). The economical benefits of climate forecasts were shown during the past years (Sonka et al., 1988) especially during the past ENSO events. On longer time scales social, environmental and economic activities will partly govern the sensitivity towards climate events and determine the type of information to respond to (Klopper, 2002). Social benefits directly associated with the use of climatological information include stability or improvement of the environment, living, traveling and working conditions. For people to travel from one place to another they must have the climate information of the area. For example if people want to travel to Zimbabwe, they must know the seasonal forecast because if the season will be wet, then it will be characterized by malaria. Again if the area is likely to be cold, they must be prepared and carry warm clothes. Climate information is also used by contractors. If a contractor has a project to build a shopping mall, he/she must know what the season will be like. During a wet season, the project will be delayed because concrete may take a longer time to dry. Climate information can also be of use by Amalgamated Beverages Industry (ABI). If a summer season is expected to be characterised by many cold days, then that implies that ABI should not produce more drinks than usual.

There are many users of climate information in South Africa but the agricultural sector is affected the most by climate variability (O'Loughlin, 1988). Farmers use climate information as one of the tools in the planning process for example when to plant, fertilizer applications, stock management etc (Klopper, 2002). The farmer needs to know the right time to plough. During a season when below-normal rainfall conditions are experienced, seasonal forecasts can help farmers save money by planting fewer crops.

1.6 What makes a good seasonal forecast?

In order to determine whether or not the forecast is valuable, there is a need to monitor the quality of the seasonal forecasts and its use (Thorne and Stephenson, 2001). There must be a clear link between quality and value. Therefore combination of quality and value statistics could be used by users to choose the best seasonal forecast provider and to set limits for performance related contracts.

Stanski et al. (1989) reviewed six attributes of a weather forecast that make up the total quality which are reliability, accuracy, skill, resolution, sharpness and uncertainty. They also make an important point that no single verification measure provides complete information about the quality of a product. Value is the degree to which the forecast helps the decision-maker to realise some incremental economic and/or other benefit (Jolliffe et al., 2003). Unlike quality, the value of seasonal forecast depends on user requirements (Thorne and Stephenson, 2001). It shows that there is a difference between seasonal forecast quality and seasonal forecast value. A seasonal forecast has a high quality if it predicts the observed conditions well according to some objective or subjective criteria and it has value if it helps users to make better decisions.

The value of seasonal forecasts to a particular activity is measured by the expected increase in economic benefits arising from the use of these seasonal forecasts in the decision making process (Klopper, 1999). A seasonal forecast structure has economic value if user's decisions are influenced by various seasonal forecasts. If the quality of the seasonal forecast is such that the user makes the same decision with or without the seasonal forecast, then the seasonal forecast is of no value. In general, seasonal forecasts of a variety of weather variables over a wide range of time scales possess positive economic value for a spectrum of decision makers (Klopper, 1999).

Brier and Allen (1951) classified the value and quality of seasonal forecast based on administrative, scientific and economic aspects. They call these a three way classification. According to them no classification is perfect and they have common characteristics. A common important characteristic is that any verification measure should be informative and it should be chosen to answer the questions of interest and not simply for reasons of convenience (Jolliffe et al., 2003).

For administration purposes, there is a need to have some statistics of performance of seasonal forecasts (Jolliffe et al., 2003). This shows effectiveness of seasonal forecast using a small number of forecasting models and various numbers of forecasting models. It also measures the performance and the involvement of forecasters. For this purpose, a small number of overall verification measures of seasonal forecast performance are usually preferred. A verification measure can then be used to justify whether the seasonal forecast is better with involvement of forecasters or not. It also determines the value of seasonal forecasts using a small number or a bigger number of forecast models (Jolliffe et al., 2003). Therefore, they can guide strategy for future investment of resources in seasonal forecasting.

The scientists will be more concerned with understanding and improving the seasonal forecast system (Jolliffe et al., 2003). Many meteorological scientists are busy implementing new forecasting models and verification can deduct whether a forecast has quality or not. There is a need for using different skill scores in verifying the seasonal forecast and the result can show whether these models have problems or not. Sometimes the models can lack skill in predicting the season and this can lead to lack of information. Therefore, seasonal forecasts verification can lead to an improvement in the scientific understanding of the underlying processes to improve models, and eventually to improve seasonal forecasts (Klopper, 2002).

Economic use is usually taken to mean something closer to the users of the seasonal forecasts (Jolliffe et al., 2003). The verification structure should be kept as simple as possible in terms of communicating results to users but complexity arises because users have different interests. For example, seasonal forecasts of summer rainfall that can be of interest to both a farmer, and to an insurance company covering risks of the event cancellations due to wet weather (Jolliffe et al., 2003). However different aspects of the seasonal forecast are relevant to each. The farmer will be interested in the total rainfall, and its distributions across the season, whereas the insurance company's concern is mainly restricted to information on the likely number of wet days (Klopper, 2002).

The economic view of seasonal forecast verification needs to take into account the economic factors underlying the user's needs for seasonal forecasts when devising a verification measure. It is sometimes known as 'customer based verification', as it provides information to be understood by the 'customer'. Another aspect of forecasting for specific users is the extent to which users prefer a simple, less informative seasonal forecast (Klopper, 2002). The seasonal forecast is more informative (e.g. a probability seasonal forecast) but is difficult to interpret. It confuses users because all three categories are assigned a certain probability.

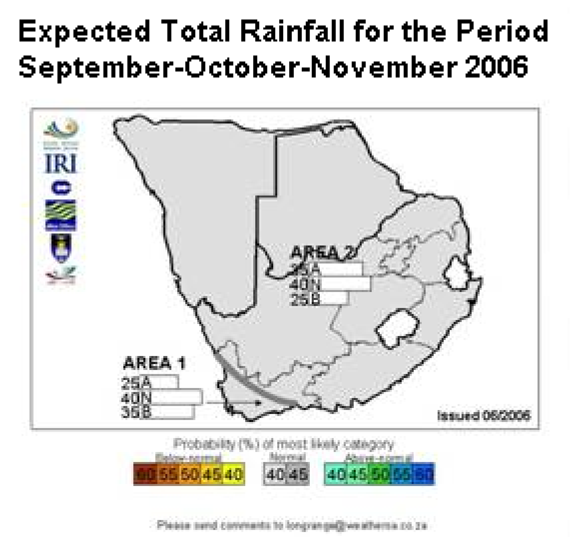
1.7 How to verify the seasonal forecast

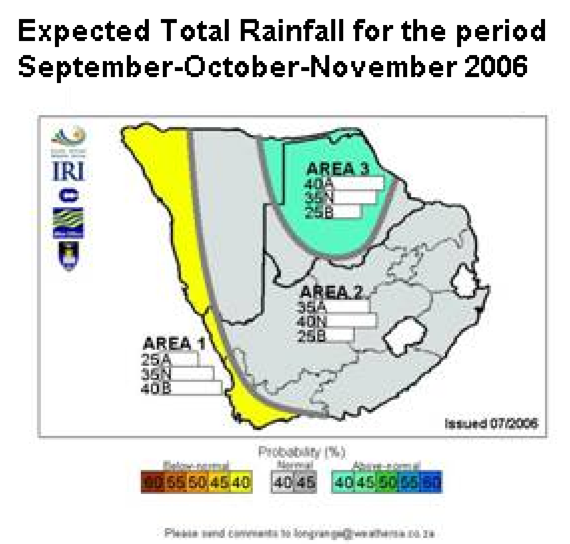
It is more appropriate to judge probabilistic seasonal forecast in the aggregate rather than individually but a time series of all seasonal forecasts should help users visualize aggregate performance (Klopper, 2002). The demonstrated seasonal forecast skill, or lack thereof, provides a basis of experience for exploring the implications of seasonal forecast performance. Seasonal forecast quality can have implications for prioritising scientific efforts, realizing competitive advantages, adjusting management processes, and changing seasonal forecasting efforts (Klopper, 2002). With the verification structure in place, users will realise the potential of seasonal forecasts.

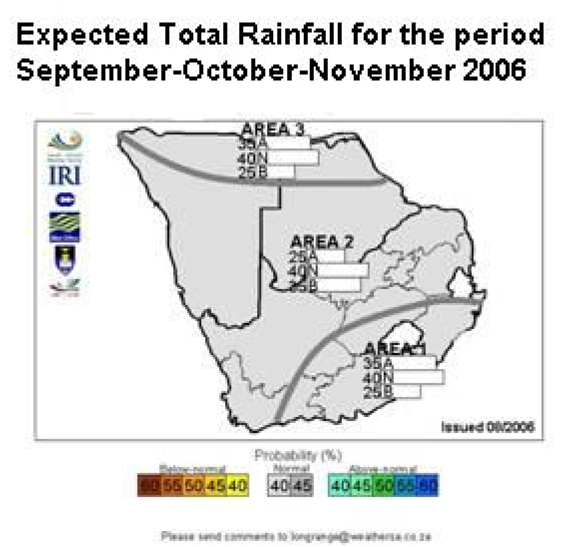
1.8 The truth when verifying the seasonal forecast

When verifying seasonal forecast, there must be an observational dataset that comes from an observational data source (Jolliffe et al., 2003). These could be rain-gauge measurements, satellite derived cloud cover, geopotential height analyses and so on. In many cases it is difficult to know the exact truth because there are errors in the observations. Sources of uncertainty include random and bias errors in the measurements themselves when the observational data are analyzed or otherwise altered to match the scale of the forecast (Jolliffe et al., 2003). Rightly or wrongly, most of the time, errors in the observational data sets are ignored. This happens only if the errors in the observations are much smaller than the expected error in the forecast (high signal to noise ratio). Even skewed or under-sampled verification data can give us a good idea of which seasonal forecast products are better than others when inter-comparing different seasonal forecast methods.

1.9 Eye-ball verification







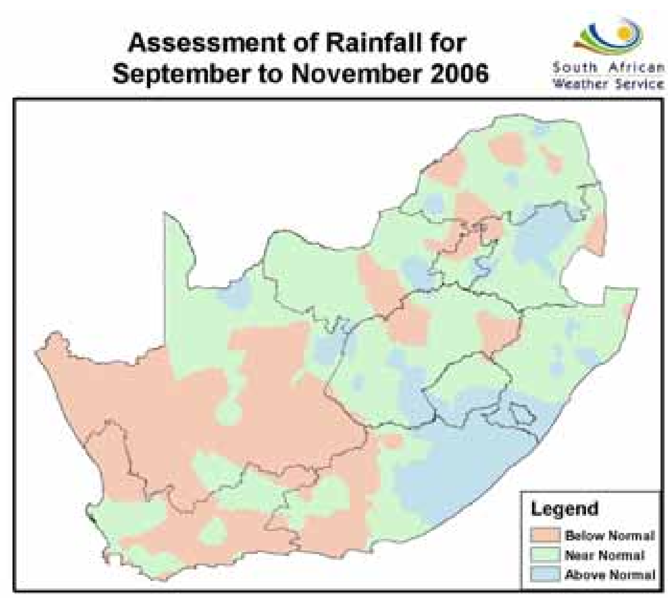


Table 1: Eye-ball verification with different lead-times (SAWS)

Eye-ball verification is the method that looks at the forecast and observations side by side and uses human judgment to distinguish the forecast errors, Joliffe et al., 2003 Eye-ball verification is a verification method that is currently used by South African Weather Service and it can be found on www.weathersa.co.za/FcastProducts/LongRange/ViewSeasonEyeBall.jsp. The method is subjective because it does not give quantitative verification statistics. The judgments come from different individuals with different views. For example one user can see a seasonal forecast useful and others not, depending on their judgment. Eye-ball verification can sometimes mislead people because of lack of statistics that can show quality and value of the forecast. Therefore verifying forecasts in terms of skill measures showing statistics is important.

1.10 Objective of the research

The objective of the study is to present skill estimates of SAWS operational seasonal forecast for the period 1998 to 2004.

The verification structure provides several methods that accommodate variations in user's interpretive abilities. It offers trade-offs between different levels of informativeness and understandability, and enables users to increase the cleverness of their understanding about seasonal forecast, their reliability, and implications of using them for decision making. SAWS currently uses eye-ball verification to judge the seasonal forecast. Therefore, there is a need of performing a more comprehensive verification of seasonal forecast products, which must be done in a rational manner such that the administrative, scientific and economic needs are met. This has become a requirement, since SAWS could begin to generate income from its seasonal forecast services. SAWS issues their operational seasonal forecasts probabilistically. The seasonal forecast must be verified and should be available to potential users of seasonal forecast outlooks (Goddard et al, 2003). Goddard et al (2003) found that the skill level associated with the normal category is generally low but this also needs to be tested using the SAWS operational seasonal forecasts.

Chapter 2

Data and Method

2.1 Introduction

Verification is the process of determining the quality of forecasts (Jolliffe et al., 2003). A wide variety of seasonal forecast verification structures exist, but all involve measures of the relationship between a seasonal forecast or a set of seasonal forecast and the corresponding observation(s) of the predictand. Any seasonal forecast verification method thus necessarily involves comparisons between matched pairs of seasonal forecasts and the observations to which they pertain. There is a need to verify the forecast in order to present skill estimates of operational probabilistic seasonal forecast.

2.2 Verification data

In this study the South African Weather Service seasonal forecast outlook issued for October-November-December 1998 to October-November-December 2004 are analysed. The seasonal forecast is issued at different lead-times but the study only concentrates on 0-lead-time, which is the seasonal forecast issued a month prior to the target season. This study only concentrates on 0-lead time because it is the only lead-time that has a complete set of forecasts for the seven year period being studied. The observation dataset is the averaged monthly total rainfall data from January 1998 to December 2004 for 963 stations.

2.3 Verification measure

The seasonal rainfall forecast is verified using three skill measures which are the Ranked Probability Skill Score (RPSS), Reliability Diagram (RD) and Relative Operating Characteristics (ROC). These verification measures take account of the probability assigned for each category using the forecast climatology as a reference. The skill measures are more appropriate to measure the quality of a probabilistic forecast, since the magnitude of the error depends on the probabilities assigned to the different categories.

2.3.1 Ranked Probability Skill Score (RPSS)

The RPSSs are calculated from Ranked Probability Scores (RPSs) values. A RPS is a verification measure that is sensitive to the difference between the probabilities assigned to each category and the category observed (Wilks, 2006). If a high probability was assigned to a certain category and that category occurred a high score will be assigned, however if the category is not observed, a penalty will be given. Therefore RPS can be calculated by the square errors with respect to cumulative probabilities in the forecast and observation distribution of categories (Wilks, 1995).

Let p_j be the probability assigned to the j^{th} category; the cumulative forecast vector P_m for the first m categories can be defined as

$$P_m = \sum_{j}^{m} pj \qquad \text{i}$$

Similarly, 0_m defines the cumulative observation vector for the first m categories

$$O_m = \sum_{j=1}^{m} oj \qquad \text{ii}$$

The components of the vector o_j are all zero except for the category in which the observation occurs. The RPS is defined as the sum of the squared differences between the components of the cumulative forecast and observation vectors of Equations (i) and (ii)

RPS = $\sum_{m=1}^{J}(Pm - Om)$ iii

where J represents the number of categories. A perfect forecast would assign a probability equal to one to the category that is observed, and a probability equal to zero to all the other categories. Therefore, the vectors P_m and O_m will be the same and RPS will be zero. If the forecast departs from perfection, then the square differences will depart from zero and RPS will be greater than zero.

Equation (iii) was used to calculate the RPS of every seasonal forecast and averaged the Ranked Probability Scores in order to obtain the mean RPS value. Therefore, the Ranked Probability Skill Score is expressed relative to the reference probability $RPS_{reference}$ as follows

RPSS= $\frac{\overline{RPS} - \overline{RPSreference}}{0 - \overline{RPSreference}} = 1 - \frac{\overline{RPS}}{\overline{RPSreference}}$ iv

Where reference can be any point, but in this study climatology is the reference.

2.3.2 Reliability diagram (RD)

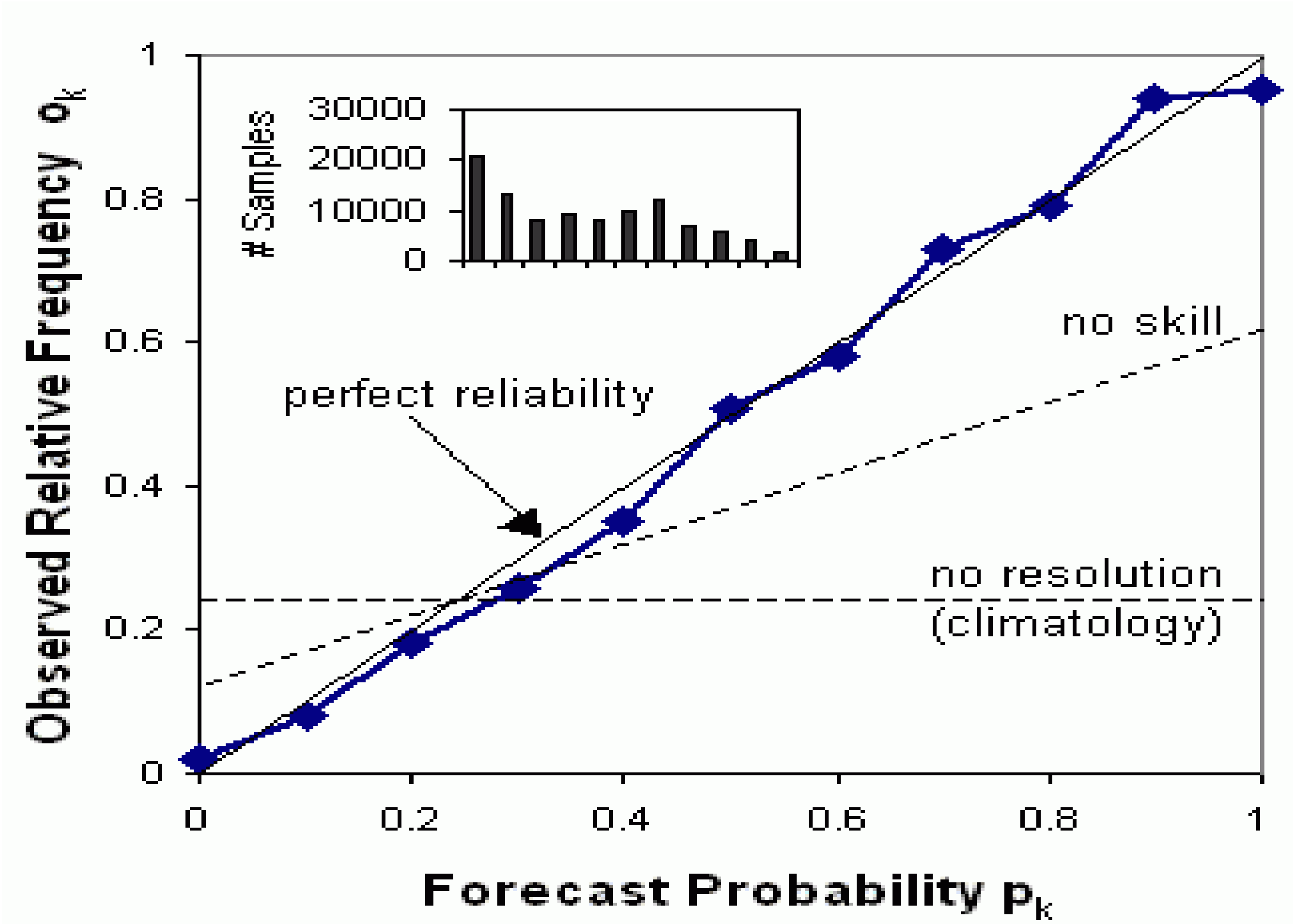


Figure 2.1 Reliability diagram (www.bom.gov.au/bmrc/wefor/staff/eee/verif/verif_web_page.html)

The Reliability diagram is used to graphically represent the performance of probability forecasts of dichotomous events (Wilks, 1995). A reliability diagram consist of only the plot of observed relative frequency as a function of forecast probability, the 1:1 diagonal perfect reliability line, and a summary of the frequency of use of each value (figure 2.1). The reliability diagram allows a more prominent display of the frequency of use of the forecasts (Wilks, 2006). This is an important consideration, since the plotted points on the reliability diagrams represent the conditional distribution of observations given the forecasts, $p(o_j \mid y_j)$, and the frequency of use of the forecasts is just the unconditional distribution of the forecasts, $p(y_j)$. Thus the reliability diagram is a compact display of the full distribution of forecasts and observations and is a more informative representation of forecast performance than single scalar scores, which in the study is RPSS. The graph in the top left corner of figure 2.1 is known as sharpness diagram which shows the frequencies of the forecast probabilities divided by the total number of forecasts (Wilks, 2006). The sharpness diagram gives estimate of the marginal probability distribution of the forecast probabilities.

2.3 Relative Operating Characteristics (ROC)

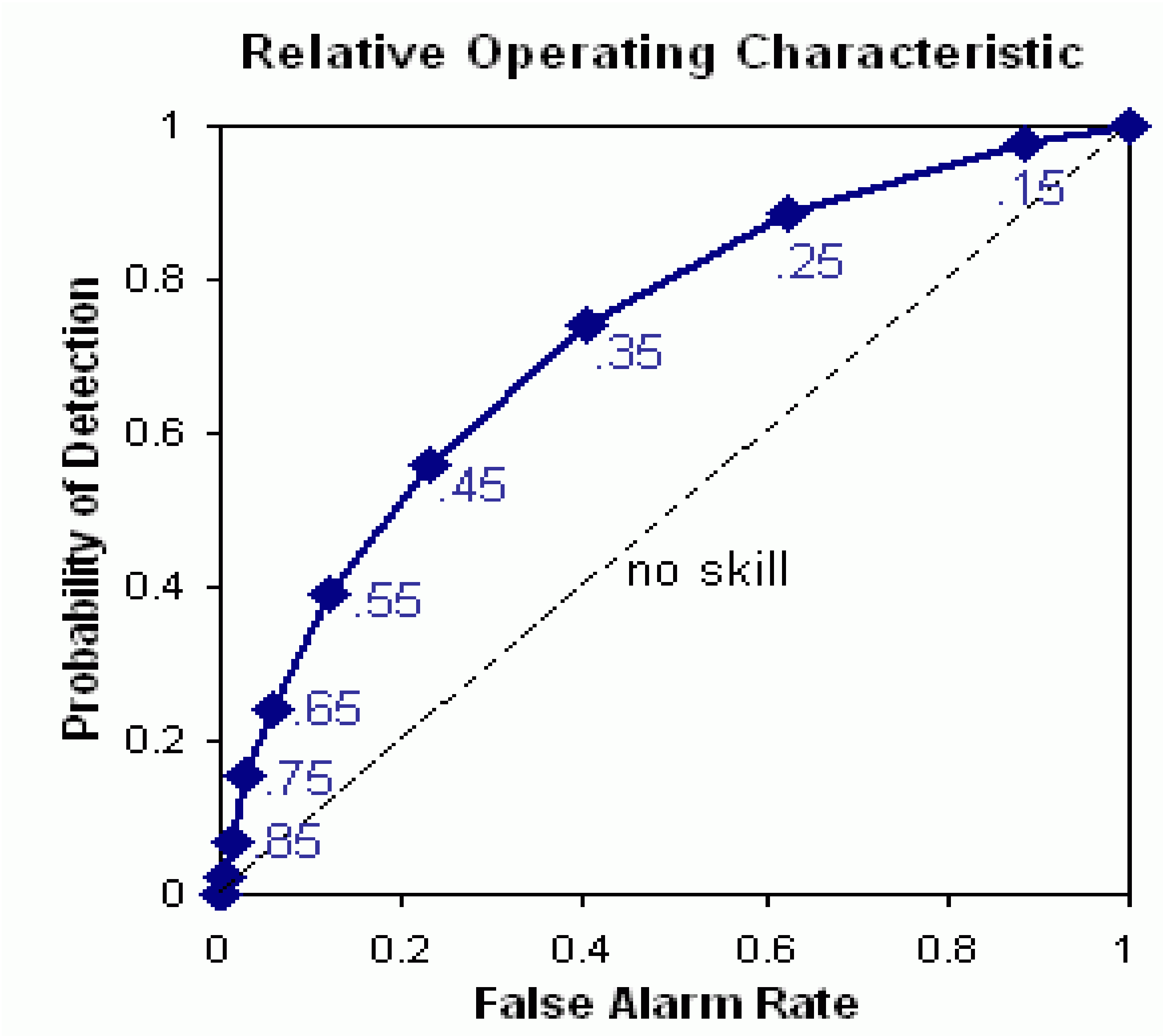


Figure 2.2 *Relative Operating Characteristics*

(www.bom.gov.au/bmrc/wefor/staff/eee/verif/verif_web_page.html*)*

Relative Operating Characteristics (ROC) comes from quality control and signal detection theory where the quality of performance is assessed by the relation between probability of detection (POD) and false alarm rates (Swets 1973; Egan 1975; Mason 1982). The graph (figure 2.2) of POD against false alarm rates within the range of probability thresholds is called the relative operating characteristics. The POD and false alarm rates are closely related to the threshold used in transforming from probabilistic forecasts to yes/no forecasts. The POD can be increased by reducing the probability threshold, but at the same time the false alarm rate is increased (Zhang and Casey, 2000). Similarly, reducing the false alarm rate is at the expense of reducing the POD. In the ROC curve, the seasonal forecast is considered to be perfect if the seasonal forecasts are located at the point (0,1). In this case, the score will be either 0% or 100% for probabilistic seasonal forecasts. The worst seasonal forecast locates at the point (1,0) in which the seasonal forecast gives either 0% or 100% probabilistic forecasts where the constant value forecasts and random forecasts will locate on the straight line between (0,0) and (1,1).

The shape of the ROC curve gives a total description of the skill of the seasonal forecasts at all probability thresholds. Seasonal forecast with a good skill will have its ROC curve lying above and to the left of the (0,0) to (1,1) diagonal. However, a seasonal forecast with low skill will have its ROC curve lying below and to the right of the (0,0) to (1,1). ROC can be quantified in two different methods (Zhang and Casey, 2000). The first method is the area beneath the ROC curve whereas the second method is POD versus the false alarm rate (Green and Swets, 1966). The study will only use the area beneath the ROC curve because it gives one value for a score and is easily understood. The larger the area, the better the seasonal forecast skill. If the area of the seasonal forecast is less than 0.5 of the whole (unit area), then the seasonal forecast is less skillful than a random or constant forecast.

Chapter 3

Ranked Probability Skill Score (RPSS) Results

3.1 Introduction

In this study the Ranked Probability Skill Score (RPSS) is the Ranked Probability Skill (RPS) of the forecast compared with the RPS of the forecast of climatology (RPS_{clm}) that assigns 0.33 for each of three categories. The value of RPS_{clm} depends on which category value was observed, being lower for the middle category than the two outer categories. RPSS gives credit for forecasting the observed category with high probabilities, and penalties for forecasting the wrong category with high probabilities is substantial (Goddard et al., 2003). The maximum RPSS is 1, but a score of 1 could only be obtained by forecasting the observed category with a 100% probability. The RPSS is used to verify South African Weather Services operational seasonal forecast. Only seasonal forecasts issued a month before the target season (0-lead-time) are verified over a 7-year period (1998 to 2004). If the RPSS for the forecasts is 0 then there is no skill in the seasonal forecasts, because that is the same score one would get by consistently issuing a seasonal forecast of climatology (0.33 0.33 0.33). On the other hand, a negative score suggests that forecasts of climatology are better than seasonal forecasts. The results are shown in figure 3.2 to 3.5 for the nine provinces of South Africa (figure 3.1).

Figure 3.1 Nine provinces of South Africa (Map from SAWS website)

3.2 January-February-March (JFM)

The Ranked Probability Skill Scores (RPSSs) for seasonal rainfall forecasts of the January-February-March (JFM) late summer seasons for the period 1998-2004 are shown in figure 3.2. For the JFM seasons, most of Limpopo province and Northern Cape, western and southwestern parts of North West receive between fifteen and twenty percent of the annual total rainfall. The rest of the provinces receive between ten and fifteen percent of the annual total rainfall except the south and south-western coast which receive between zero and ten percent of their annual total rainfall. The RPSS for JFM seasonal forecasts is positive over the eastern parts of the Western province and most parts of KwaZulu-Natal. There is a positive score over the borders of Lesotho and the central areas of Free State and North-west. For Limpopo, Gauteng, Mpumalanga, Northern Cape, Eastern Cape and most parts of the North-West and Free State, the RPSS for the JFM seasonal forecasts is negative. Overall, the SAWS forecast was worse than climatology, which suggests that users would have been better of using climatology than these forecasts.

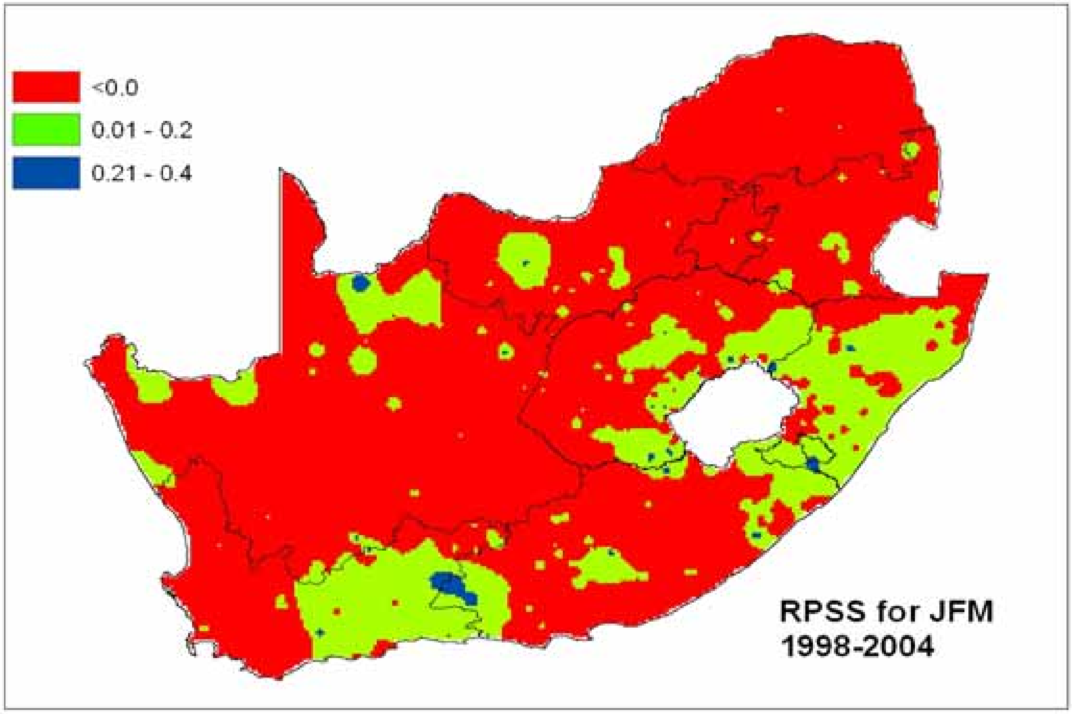


Figure 3.2 Ranked Probability Skill Score for January-February-March

3.3 April-May-June (AMJ)

The Ranked Probability Skill Scores (RPSS) for April-May-June (AMJ) seasonal forecasts for the period 1998-2004 are shown in figure 3.3. This season is the beginning of winter season where the western parts of Northern and Western Cape Provinces receive between ten and fifteen percent of annual total rainfall. The eastern part of Western Cape, the central part of Northern Cape and western part of Eastern Cape receive between five and ten percent of the annual total rainfall. The rest of the country receives between zero and five percent of the annual total rainfall. The RPSS for AMJ seasonal forecast is positive over the entire Gauteng Province, the most part of Mpumalanga Province, western of Eastern Cape Province and along the border of Western and Northern Cape. The rest of the country shows negative RPSS values. The seasonal forecast is in most of the areas worse than the forecast of climatology especially in the area where the rainfall is expected to be high i.e. Western Cape Province.

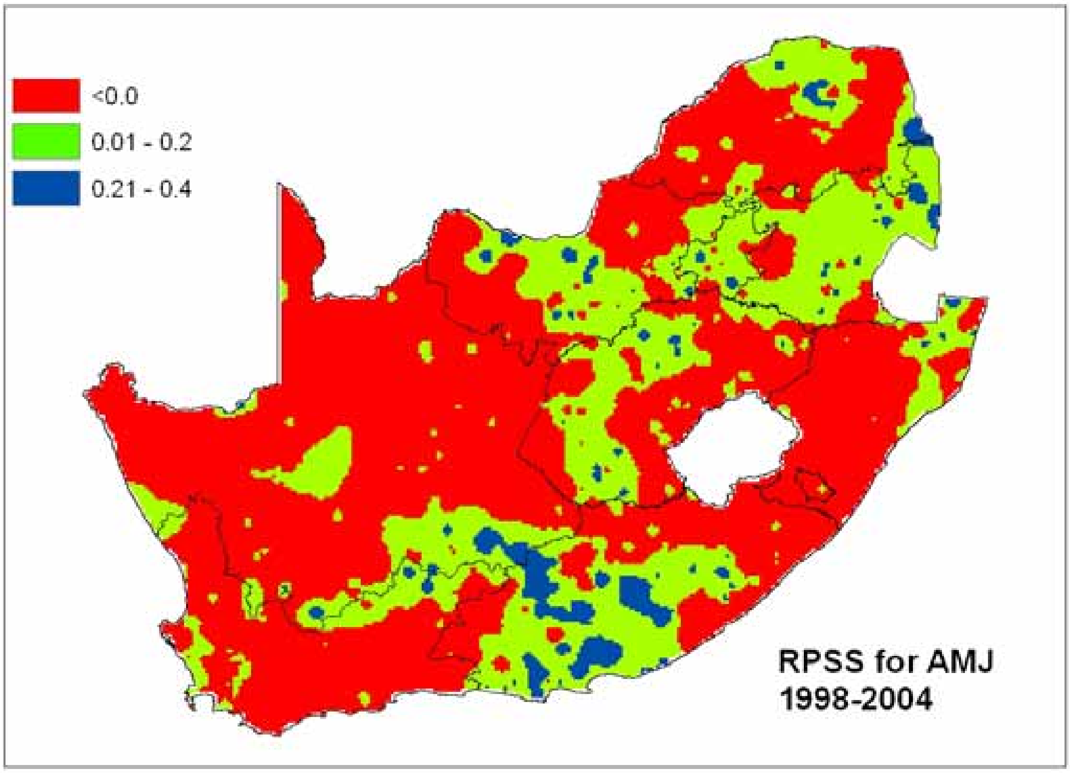


Figure 3.3 Ranked Probability Skill Score for April-May-June

3.4 July-August-September (JAS)

The Ranked-Probability-Skill-Scores (RPSS) for rainfall seasonal forecast of the July-August-September (JAS) season for the period 1998-2004 are shown in figure 3.4. For the JAS season, the areas of highest rainfall are the western parts of the Northern and Western Cape Provinces with ten to fifteen percent of annual total rainfall. Five to ten percentage of the annual total rainfall is received in the Western interior of the Northern Cape Province, the central interior of the Western Cape Province and the south coast of the Eastern Cape Province. The larger part of the country receives between zero and five percent of their annual total rainfall. The skill of the JAS seasonal forecasts is positive over the western part of the Western Cape and the larger part of the Northern Cape Province with a few patches of positive values over the Limpopo Province. Most of the area in the country shows negative RPSS. The entire Eastern Cape, KwaZulu-Natal, Mpumalanga, Gauteng, most parts of the Limpopo and Northwest province show negative RPSS values. Even though most of the country behaves worse than a forecast of climatology, there are a few areas that show positive RPSS values.

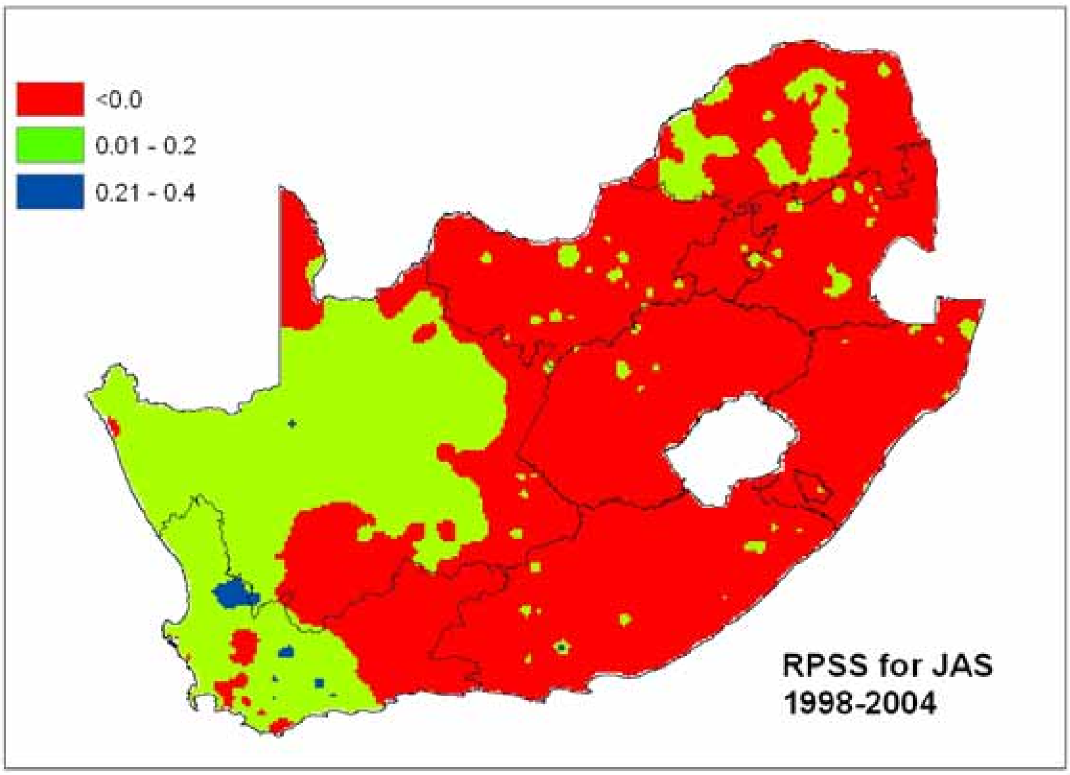


Figure 3.4 Ranked Probability Skill Score for July-August-September

3.5 October-November-December (OND)

For the October-November-December (OND) season, south of Mpumalanga along the Free State Province north of Gauteng Province and the Northwest Province receive fifteen to twenty percent of annual total rainfall. The Western part of Northern and Western Cape Provinces receive respectively zero and five percent of the annual total rainfall whereas the most part of Northern Cape Province, the eastern parts of Western Cape Province, western part of the Eastern Cape Province and along the eastern coast of KwaZulu-Natal receive five to ten percent of the annual total rainfall. The rest of the country receives between ten to fifteen percent of the total annual rainfall. The RPSS of the OND seasonal forecasts (figure 3.5) is positive over the most part of the Northern Cape, Northwest, Limpopo, Eastern Cape, KwaZulu-Natal and Free State Provinces and eastern part of the Western Cape. There are only a few areas that show negative Ranked Probability Skill Score (RPSS) which include the whole of Gauteng, western part of the Western Cape, and south of the Eastern Cape. For the OND season, the forecast was better for the most part than the forecast of climatology.

The SAWS seasonal forecasts are more skillful in this season compared to other seasons. The reason for an improvement in the skill in this season is the high association of rainfall in this season with the ENSO. The summer season of 1997/1998 was an El Nino season, and it was followed by La Nina events from 1998 to 2000. Dynamical models are more skillful in forecasting rainfall during ENSO years and therefore that skill will automatically transfer to subjective forecasts where there would be a lesser confusion when all models agree.

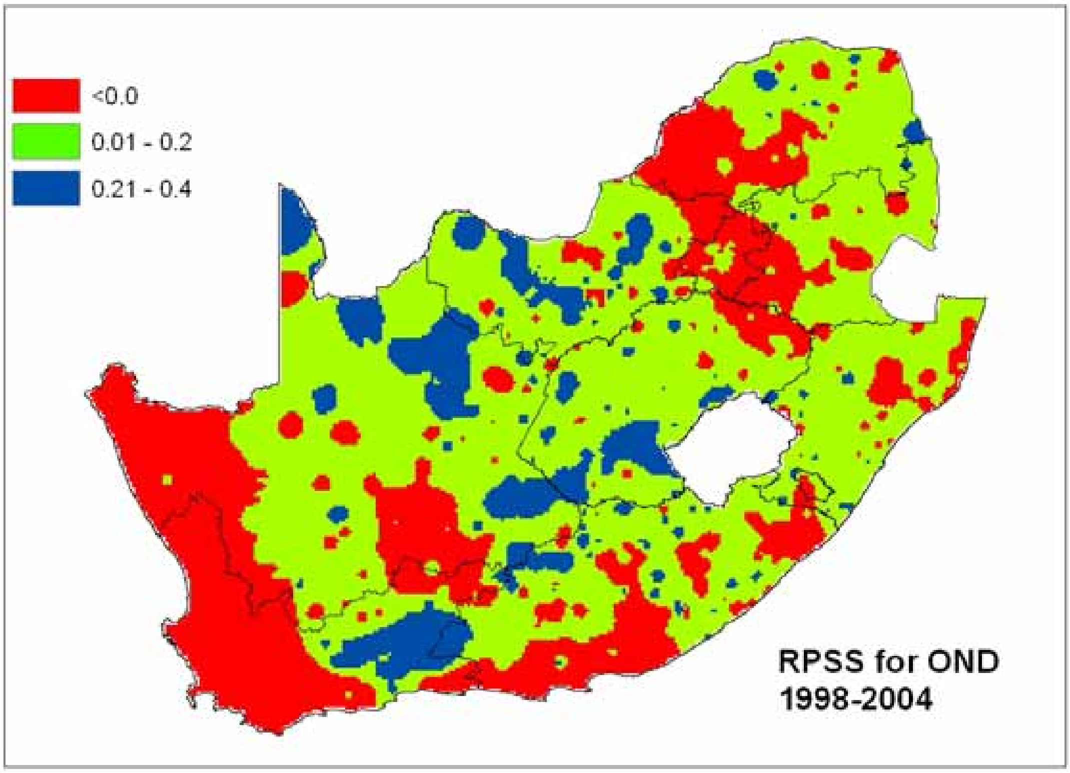


Figure 3.5 Ranked Probability Skill Score for October-November-December

Summary

The South African Weather Services operational seasonal forecasts verified in this study is for the period 1998 to 2004. These probabilistic three-category forecasts for seasonal rainfall result from the use of forecast models from different institutions both run in South Africa and abroad. The seasonal forecast issued by South African Weather Service for the period 1998 through 2004 were assessed here using the Ranked Probability Skill Score (RPSS), a measure that considers the probabilistic content of the forecast. The seasonal forecasts show more coherent coverage of the skill over most of the region of October-November-December (OND) seasons than the other three seasons (January-February-March (JFM), April-May-June (AMJ) and July-August-September (JAS)) because of the strong link between OND and ENSO. A verification study done at the IRI for multi models also found the average RPSS model for South Africa to be low as compared to other countries for JFM, AMJ and JAS (Goddard et al, 2003).For JAS, the skill is higher over the winter rainfall area than for summer rainfall areas. This may be attributable to the fact that for each season forecasters pays more attention to a region which receives the most rainfall. Seasonal forecasts over regions for which positive skill exists should be considered by users because the seasonal forecasts are potentially useable for input to decision making (Goddard et al, 2003). However, the forecasts were subjected to a very strict measure of forecast performance and regions that exhibit good skill in this analysis are likely to appear skillful when subjected to other verification measures i.e. RD and ROC. Forecasters definition of forecast quality may even vary among users.

Negative RPSS implies that there is a need to add more models and improve the existing models. For meantime, users can be cautioned when using the forecast that has low skill because there might be cost loss in terms of economy. When taking a decision they can at least rely on climatology for their decision making.

Chapter 4

Reliability Diagram (RD) Results

4.1 Introduction

The Reliability diagram examines the joint frequency distribution of forecasts and observations in order to diagnose particular strengths and weakness of a set of forecasts (Wilks, 2006). It also scores each category individually i.e. below-normal, near-normal and above-normal. An unconditional bias is when the forecast is consistently too high or too low and conditional bias is when the systematic forecaster has over- or under-confidence (Wilks and Godfrey, 2002). In particular, unbiased forecast exhibiting an appropriate level of confidence produce reliability diagrams whose points fall close to the 1:1 line. If the SAWS operational seasonal forecasts exhibit points predominantly left of the 1:1 line, then the seasonal forecast is under-forecast or if the seasonal forecast exhibits points predominantly right of the 1:1 line, then the seasonal forecast is over-forecast. The Reliability diagrams for the SAWS operational seasonal forecasts are shown in figure 4.2(a-c) to figure 4.5(a-c), separately for the below-normal, near-normal and above-normal categories. The thin lines through each calibration function show weighted least-squares regressions (Murphy and Wilks, 1998) that help guide the eye through the irregularities that are in the seasonal rainfall forecasts. The bars in each refinement distribution identify the forecast probabilities. The averaged forecasts are shown by triangles on the horizontal axes and the average observations are indicated by the pentagons on the vertical axes. The arrows in horizontal and vertical axes indicate the forecasts of climatology.

4.2 January-February-March (JFM)

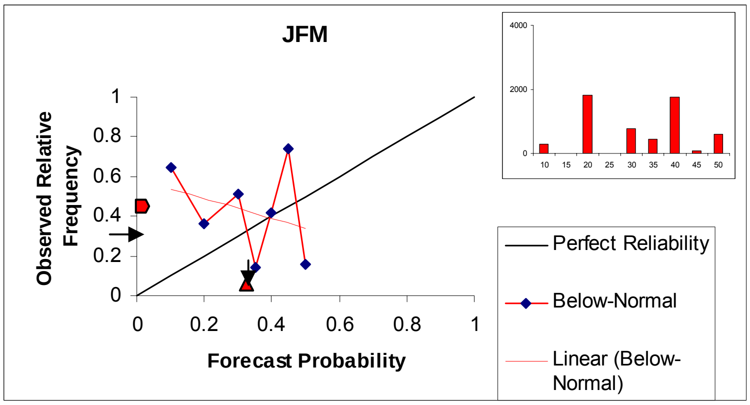


Figure 4.2a Reliability Diagram for the below-normal category

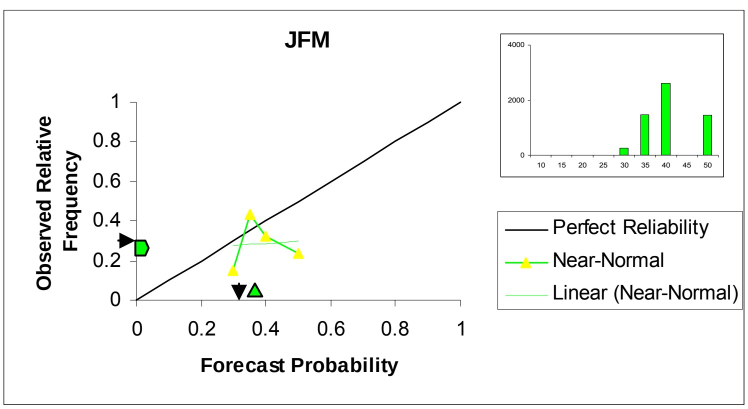


Figure 4.2b Reliability Diagram for the near-normal category

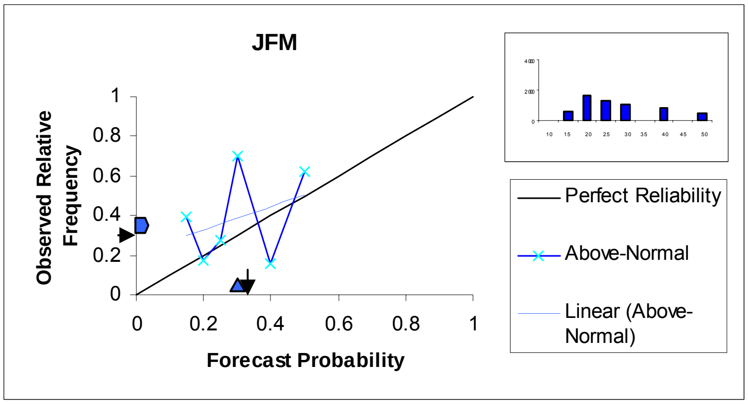


Figure 4.2c Reliability Diagram for the above-normal category

The Reliability diagrams for the January-February-March (JFM) seasonal rainfall forecasts are shown in figure 4.2 (a-c) separately for the below-normal, near-normal and above-normal outcomes for 0-month lead-time. A forecast of climatology (33.3%, 33.3%, 33.3%) was never issued for this season. The forecast of climatology is issued when there is a big uncertainty amongst all models used for the seasonal operational forecast. The above-normal JFM rainfall seasonal forecast (fig 4.2c) exhibit a good calibration function because the regression line lies close to the 1:1 line. It is even shown (Fig 4.2c) that there is not much difference between the average observation (horizontal axes) and the average forecast (vertical axes). For the below-normal JFM rainfall seasonal forecast, there are forecasts that are below and above the forecast of climatology. The forecasts that are below forecast of climatology are reasonably well calibrated, while forecast above forecast of climatology do not resolve differences in the event outcomes. The near-normal rainfall JFM seasonal forecast exhibits essentially no skill, approximately flat calibration function (Wilks and Godfrey, 2002). The JFM seasonal forecast was in favour of the above-normal category and the forecast of above-normal category should have been given more weight.

The reliability diagrams scored the forecast better than the RPSS as discussed in the precious chapter. This is because the RPSS scores all categories together while the reliability diagrams score the different categories separately and from the results it is evident that the problem is with the above-normal category. This category should be forecast with bigger probabilities than what the forecasters assigned to it during the study period. The problem here is not with the model forecasts as such but rather forecasters seem to be hesitant of forecasting this category with more weight. Forecasters should therefore forecast this probability as directed by the models and not be hesitant.

4.3 April-May-June (AMJ)

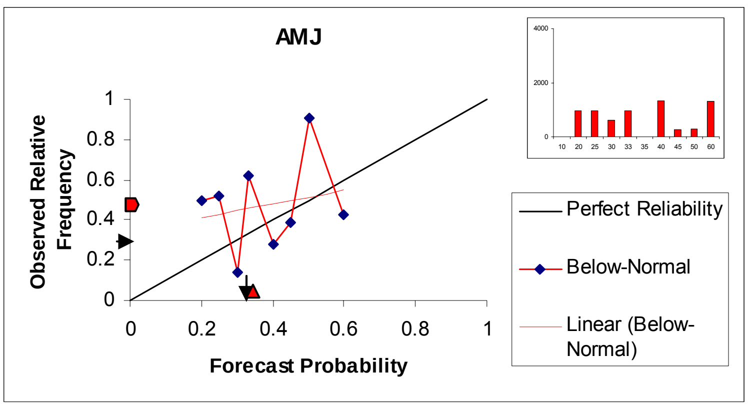


Figure 4.3a Reliability Diagram for the below-normal category

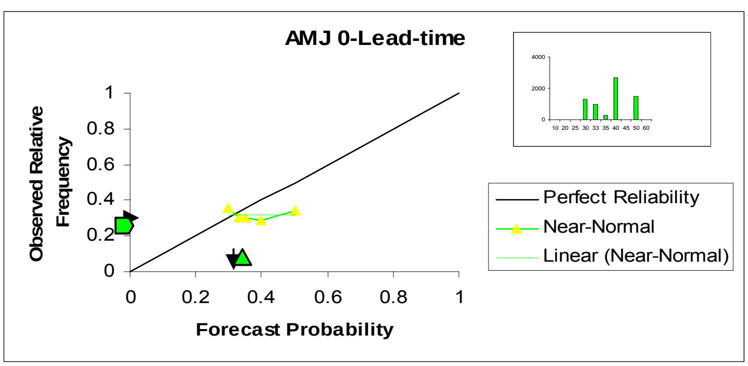


Figure 4.3b Reliability Diagram for the near-normal category

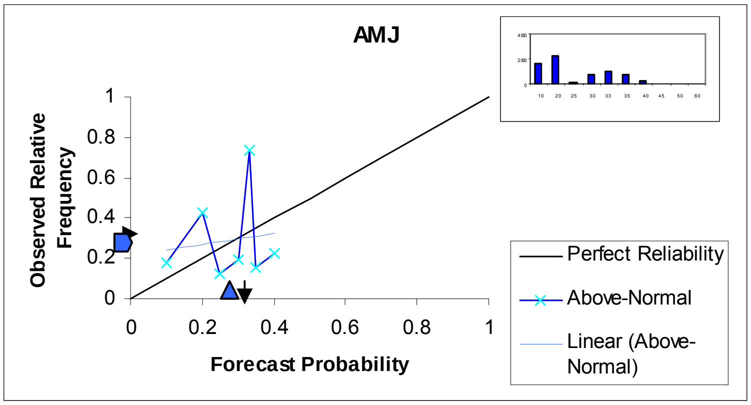


Figure 4.3c Reliability Diagram for the above-normal category

The Reliability diagrams for the April-May-June (AMJ) seasonal rainfall forecasts are shown in figure 4.3 (a-c) separately for the below-normal, near-normal and above-normal outcomes for 0-month lead-time. The seasonal forecast of climatology was issued several times during 1998 to 2004. The below-normal AMJ rainfall seasonal forecast (fig 4.3a) exhibit less resolution where there are seasonal forecasts below the forecast of climatology and there are also forecasts above forecast of climatology. The near-normal AMJ rainfall seasonal forecast exhibits essentially no resolution (Wilks and Godfrey, 2002). The near-normal category was over-forecast because most of the time it was forecast (horizontal axes) but not observed (vertical axes). The above-normal AMJ rainfall seasonal forecasts are reasonably well calibrated because the forecast are smaller than the climatological forecast (Wilks, 2002). The above-normal category was observed most of the times when it was forecast. (figure 4.2c).

The results suggest that the near-normal category should be forecast as less often as possible because it exhibits no skill. The outer-categories must be given more weight especially the below-normal category for this season. The RPSS showed that the forecast does not have skill for this AMJ season. The forecast models need to be improved to help users in terms of decision making especially over the winter-rainfall region knowing that this is the beginning of winter-rainfall season. This will also help forecasters not to issue forecast of climatology.

4.4 July-August-September (JAS)

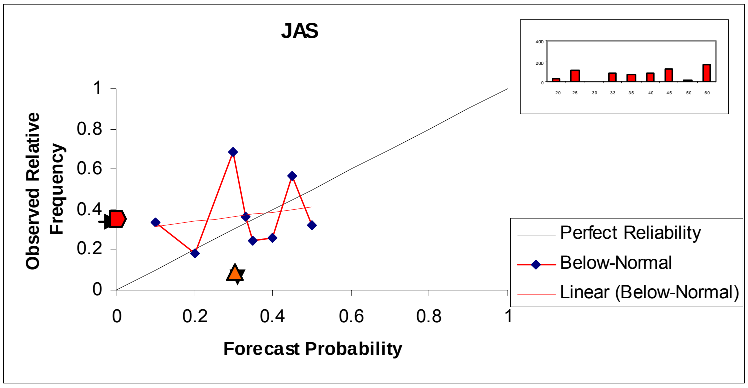


Figure 4.4a Reliability Diagram for the below-normal category

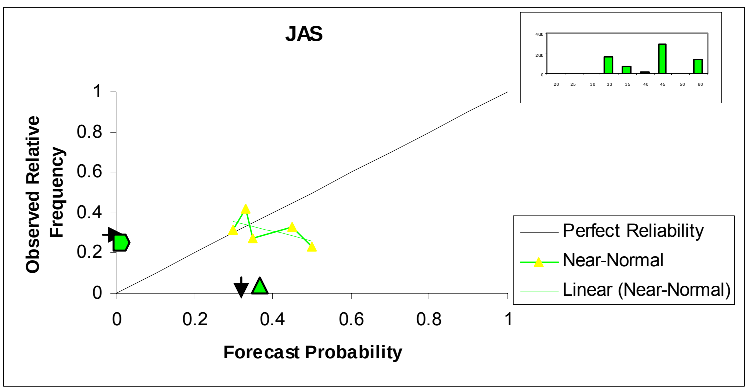


Figure 4.4b Reliability Diagram for the near-normal category

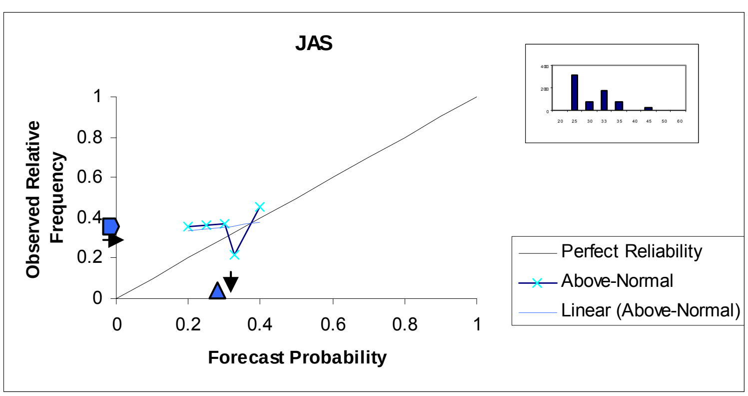


Figure 4.4c Reliability Diagram for the above-normal category

The Reliability diagrams for the July-August-September (JAS) seasonal rainfall forecasts are shown in figure 4.4 (a-c) separately for the below-normal, near-normal and above-normal outcomes for 0-month lead-times. A forecast of climatology was issued in this season for the period 1998 to 2004. The below-normal JAS rainfall seasonal forecast (fig 4.4a) exhibits less resolution. There was a small number of times where the below-normal category was observed (vertical axes) but not forecast (horizontal axes). The near-normal JAS rainfall seasonal forecasts are reasonably well calibrated because the forecasts are smaller than the forecasts of climatology (Wilks, 2002). There is a slight difference between the averaged forecast (horizontal average) and the average observation (vertical axes). The above-normal rainfall category for the JAS seasonal forecast is essentially where the line of regression is under the perfect line (1:1).

There is usually no skill in forecasting the near-normal category, but for these JAS seasons there was skill. This category should however still be avoided because it is a narrow category. The outer categories should be forecast instead. Forecast models need to be improved because a forecast of climatology was issued in one of the instances. Skillful forecast will help users decide whether to plant or not in the winter-rainfall region since this is the mid-winter season. It will also help the summer-rainfall region in terms of coldness especially the electricity consumption since there is a serious electricity shortage in the country. If the winter will be too cold there will too much of electricity consumption and therefore the skillful forecast is needed for preparation in advance.

4.5 October-November-December (OND)

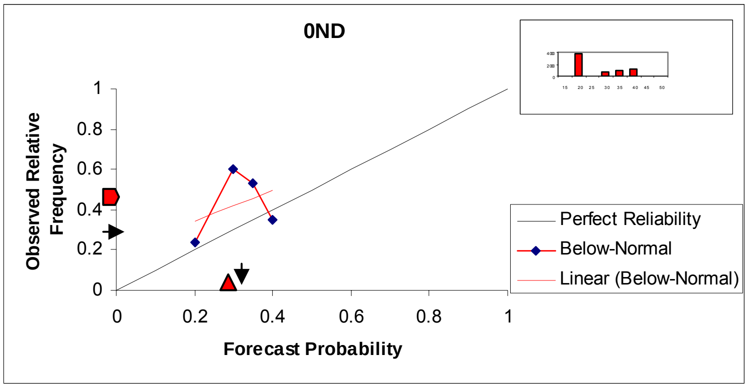


Figure 4.5a Reliability Diagram for the below-normal category

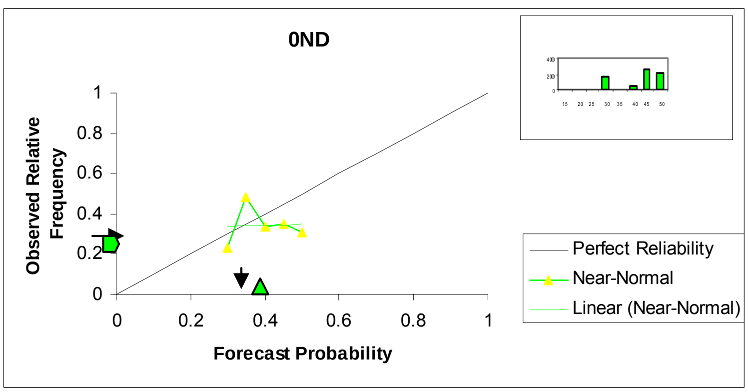


Figure 4.5b Reliability Diagram for the near-normal category

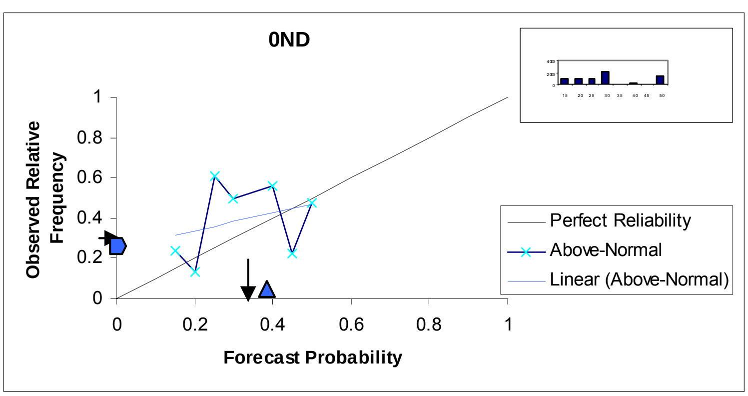


Figure 4.5c Reliability Diagram for the above-normal category

The Reliability diagrams for the October-November-December (OND) seasonal rainfall forecasts are shown in figure 4.5 (a-c) separately for the below-normal, near-normal and above-normal outcomes for 0-month lead-times. In this season the forecast of climatology was never issued for the three categories (figure 4.5 (a-c)). The below-normal category for the OND rainfall seasonal forecast (fig 4.5a) is under-forecast. The above-normal category for the OND rainfall seasonal forecast (fig 4.5b) exhibit forecasts above and below climatology forecasts. The category is under-forecast because the category was observed (horizontal axes) but not forecast (vertical axes). The near-normal rainfall OND seasonal forecast exhibits essentially no resolution. The near-normal category is over-forecast because it was forecast a number of times but not observed.

Out of this four season, OND is the only forecast that shows skill from the result of RPSS but the outer categories still need to be given more weight. The forecasters must not use the middle category when issuing the forecast since it does not have skill. The forecast models are not that bad since the forecast of climatology was never issued for this season. The summer-rainfall region users can use this forecast for decision making shown by the result of RPSS but for winter-rainfall region more has to be done to improve the forecast. The forecast shows skill over the summer rainfall region since the OND season has correlation with ENSO.

4.6 Summary

January-February-March (JFM) is the late summer season. A dry spell occurred mainly in the years 1998 to 2000 during the JFM season.(Wilks and Godfrey, 2002). The below-normal rainfall was observed even though the above-normal category was forecast most of the time. The averaged seasonal forecast for below-normal outcomes was equal to forecast of climatology. April-May-June (AMJ) is the beginning of the winter season and rainfall is not expected over a bigger part of the eastern part of the country. October-November-December

(OND) is the first half of summer rainfall season but the below-normal category was observed most of the times. The near-normal category was forecast a lot of times but not observed as frequently as forecast.

Forecasters need to improve their forecasts to advice users correctly. The outer-categories need to given more weight in most of the season. The winter-rainfall region needs to be improved because in most of the season there was no skill. Users can be advised to base their decision on skillful OND season but other seasons need to be improved. It was found that the skill in OND is caused by the correlation with ENSO and indirectly associated with changes in Indian Ocean Sea Surface Temperatures (SSTs)

Chapter 5

Relative Operating Characteristics (ROC) Results

5.1 Introduction

The Relative Operating Characteristics (ROC) assesses the performance of the fisting system that distinguishes between the intrinsic discrimination capacity and the decision threshold of the system. It is also a pure index of accuracy in the sense of the inherent capacity of the system to discriminate one state from another. It is a quantitative estimate of the probabilities of forecast outcome for any decision threshold that the system might use and the tradeoff between probabilities as the decision threshold. An index of the decision threshold which makes it possible to incorporate climatology probabilities, the value and costs of the various forecast outcome to determine the threshold that is optimal for the forecast in a given situation. A potential advantage of the skill measure is that it is indirectly related to a decision theoretic approach and so can be easily related to the economic value of probability forecast for forecast users. Therefore it is technically a skill measure, as it doesn’t compare with the reference.

For this study, a forecast skill is evaluated using the area under the ROC curve. The area under the ROC curve is favoured because it assesses the seasonal forecast using a scalar value. The seasonal forecasts for each of the eight homogenous regions (figure 5) are verified. The homogenous regions are constructed from 970 rainfall stations whereby the stations are grouped according to their seasonal climate variability to make up the eight regions. If a ROC score is one then the score area is consider to be perfect since the ROC curve for a perfect seasonal forecast pass through the upper-left corner. However, if the ROC curve for a random forecast lies along 45° diagonal of the unit square then it will give the area of 0.5 therefore indicating no skill. There is also no skill if the area is less than 0.5. The ROC of the SAWS operational seasonal forecasts are shown in figure 5.1(a-c) to figure 5.4(a-c), separately for the below-normal, near-normal and above-normal for the eight homogenous regions.

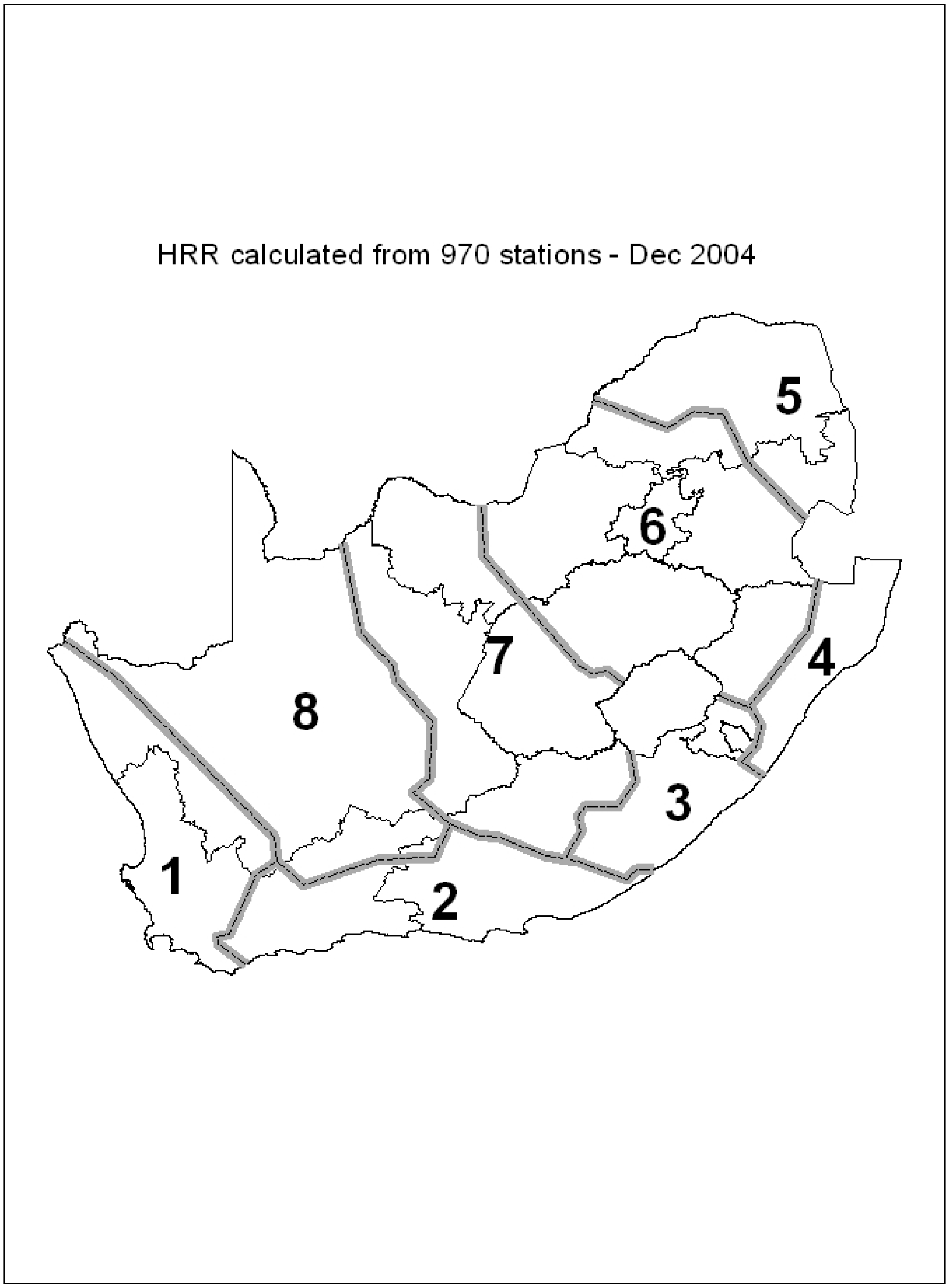


Figure 5.1 Eight homogenous regions (SAWS)

5.2 January- February-March (JFM)

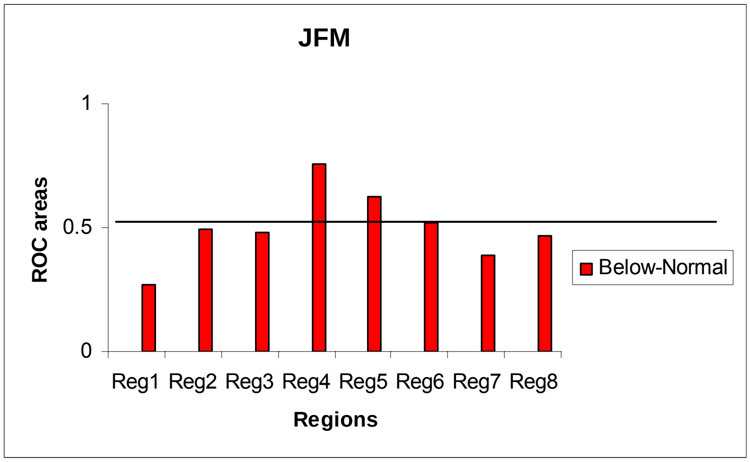


Figure 5.2a Relative Operating Characteristics area score for below-normal category

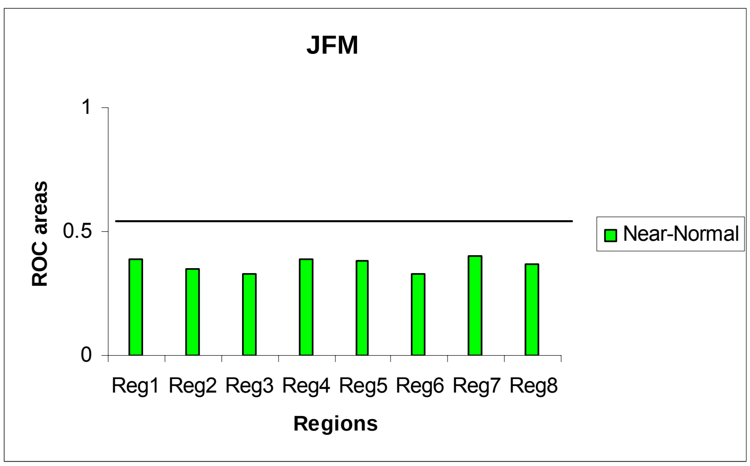


Figure 5.2b Relative Operating Characteristics area score for near-normal category

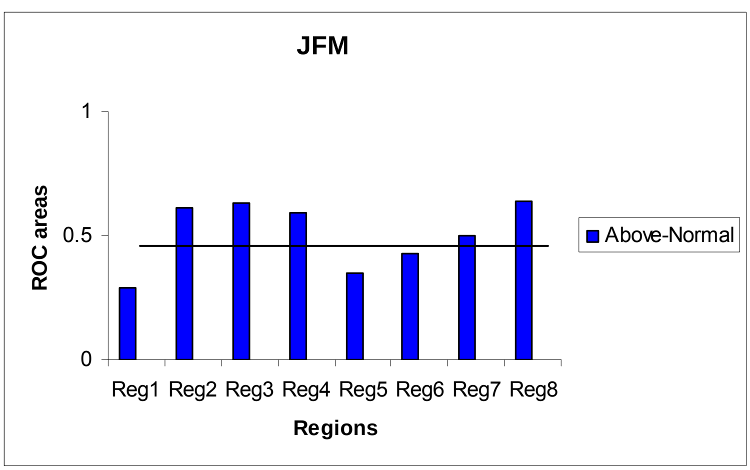


Figure 5.2c Relative Operating Characteristics area score for above-normal category

The Relative Operating Characteristics (ROC) area score for the January-February-March (JFM) seasonal rainfall forecasts are shown in figure 5.2 (a-c) separately for the below-normal, near-normal and above-normal categories for 0-month lead-time. For JFM, there is skill in predicting below-normal rainfall for stations in regions 4, 5 and 6 (figure 5.2a). The area under the ROC curve for stations in those regions is greater than 0.5. The area under the ROC curve for stations in other regions is under 0.5 which implies that there is no skill in predicting the below-normal category for those stations in the remaining 5 regions. There is no skill in predicting the near-normal category for stations in all eight homogenous regions (figure 5.2b). The area under the ROC curve is less than 0.5 for all the regions (figure 5.2b). There is skill in predicting above-normal category for stations in regions 2, 3, 4, 7 and 8 (figure 5.2c). The area under the ROC curve is greater than 0.5 as shown in figure 5.2c. The area under the ROC curve for stations in other regions is under 0.5 which implies that there is no skill in predicting the above-normal category for stations in regions 1, 5, and 6.

The result shows that, there is skill in predicting summer rainfall for regions that receive their rainfall in summer with the above-normal category and a few regions with the below-normal category. For winter rainfall regions, there is no skill in predicting rainfall for the JFM season. This finding may be attributed to the fact that forecasters do not pay much attention to the winter rainfall regions during summer because this season is not an important rainfall season for the region. Forecasters should either mark out the area that is dry and not produce forecasts for this region, or they should pay attention to these regions in as much as they do for summer rainfall regions.

5.3 April-May-June (AMJ)

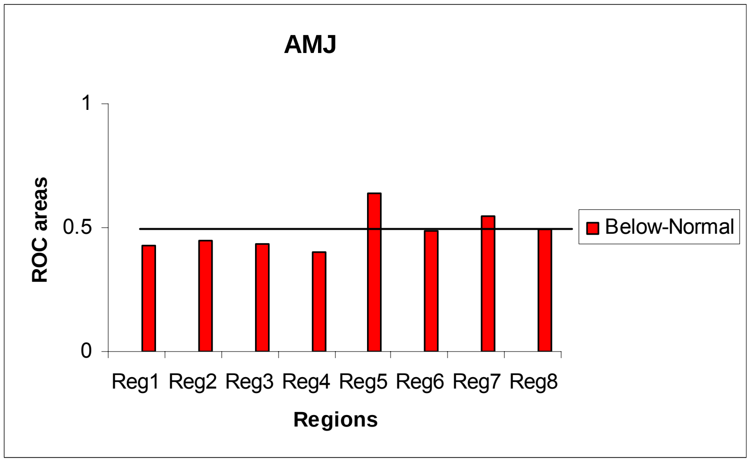


Figure 5.3a Relative Operating Characteristics area score for below-normal category

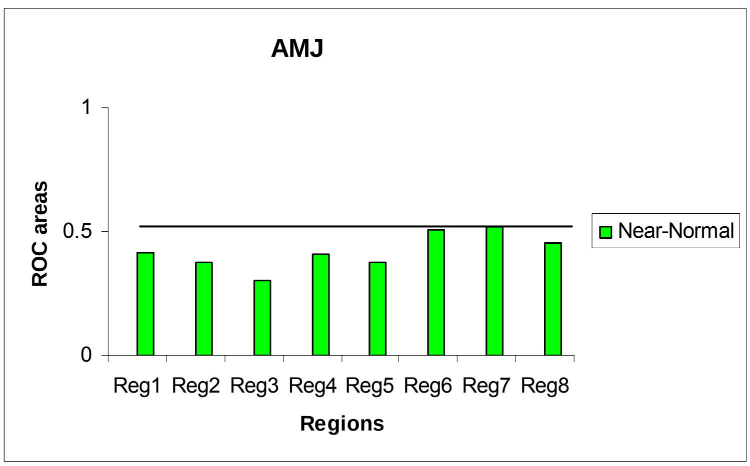


Figure 5.3b Relative Operating Characteristics area score for near-normal category

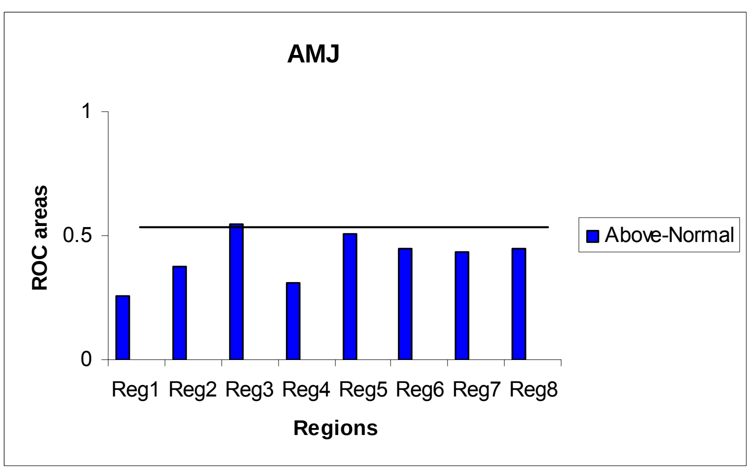


Figure 5.3c Relative Operating Characteristics area score for above-normal category

The Relative Operating Characteristics (ROC) area score for the April-May-June (AMJ) seasonal rainfall forecasts are shown in figure 5.3 (a-c) separately for the below-normal, near-normal and above-normal outcomes for 0-month lead-time. For AMJ, there is skill in predicting below-normal category for stations in regions 5, 6, 7 and 8 (figure 5.3a). The area under the ROC curve for stations in those regions is greater than 0.5 as shown in figure 5.3a. The area under the ROC curve for stations in other regions is under 0.5 which implies that there is no skill in predicting the below-normal category for stations in regions 1, 2, 3 and 4. There is no skill in predicting the near-normal category for stations in most of the eight homogenous regions (figure 5.3b) except in regions 6 and 7. The area under the ROC curve is less than 0.5 in most of the eight homogenous regions (figure 5.3b). There is skill in predicting above-normal category for stations only in regions 3 and 5 (figure 5.3c). The area under the ROC curve for stations in those regions is greater than 0.5 as shown in figure 5.3c. The area under the ROC curve for stations in other regions is under 0.5 which implies that there is no skill in predicting the above-normal category for those 6 stations.

As AMJ is the beginning of winter, it shows that the area that receives rainfall throughout the year does not have a skill with below-normal category and also the winter rainfall region. Models need to be improved and forecasters must not use middle-category. Above-normal category is favoured in the summer rainfall region even in the beginning of winter.

5.4 July-August-September (JAS)

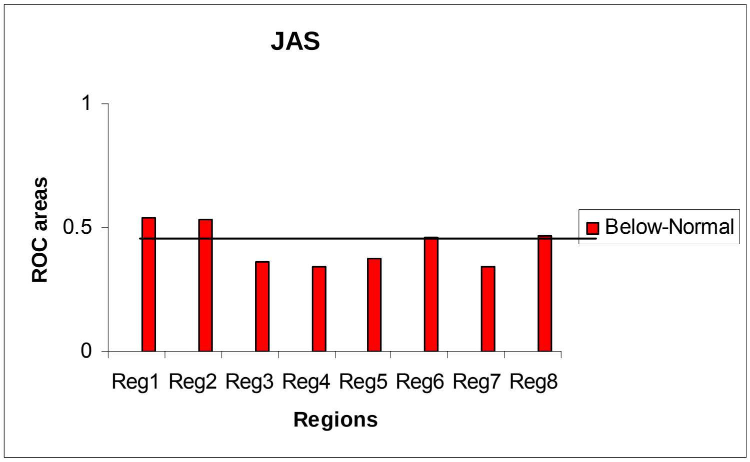


Figure 5.4a Relative Operating Characteristics area score for below-normal category

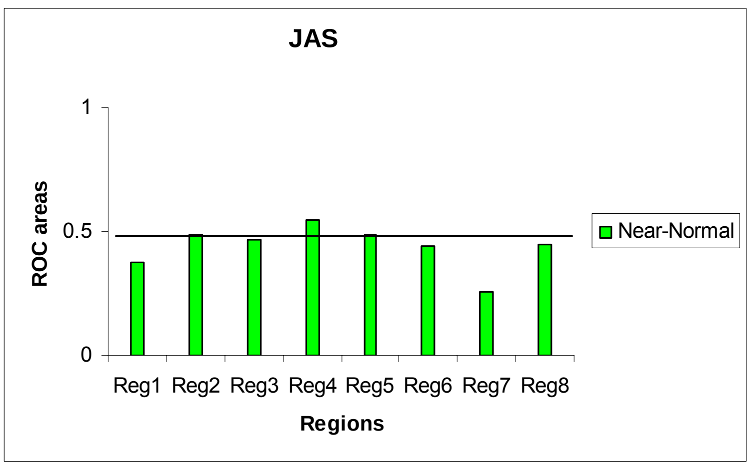


Figure 5.4b Relative Operating Characteristics area score for near-normal category

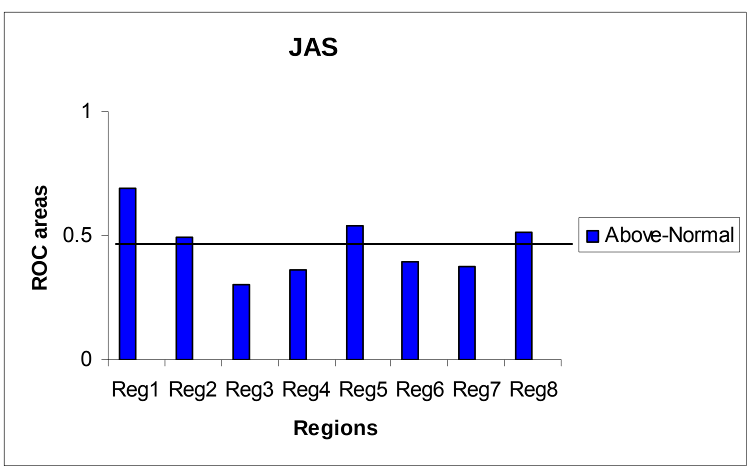


Figure 5.4c Relative Operating Characteristics area score for above-normal category

The Relative Operating Characteristics (ROC) area score for the July-August-September (JAS) seasonal rainfall forecasts are shown in figure 5.4 (a-c) separately for the below-normal, near-normal and above-normal outcomes for 0-month lead-time. For JAS, there is skill in predicting the below-normal category for stations in regions 1 and 2 (figure 5.4a). The area under the ROC curve for stations in those regions is greater than 0.5 as shown in figure 5.4a. The area under the ROC curve for the rest of the stations in other regions are under 0.5 which implies that there is no skill in predicting the below-normal category for those stations in regions 3, 4, 5, 6, 7 and 8. There is no skill in predicting the near-normal category for stations in most of the eight homogenous regions (figure 5.4b) except in region 4. The area under the ROC curve is less than 0.5 in most of the eight homogenous regions (figure 5.3b). There is skill in predicting the above-normal category for stations in regions 1, 5 and 8 (figure 5.4c). The area under the ROC curve for stations in those regions is greater than 0.5 as shown in figure 5.4c. The area under the ROC curve for stations in other regions is under 0.5 which implies that there is no skill in predicting the above-normal category for those stations in the other remaining regions.

Winter rainfall regions show skill with the below-normal category but for summer-rainfall regions, models need to be improved because in most of the regions there is no skill. Forecasters are advised not use the middle category because it lacks skill in most of the time.

5.5 October-November-December (OND)

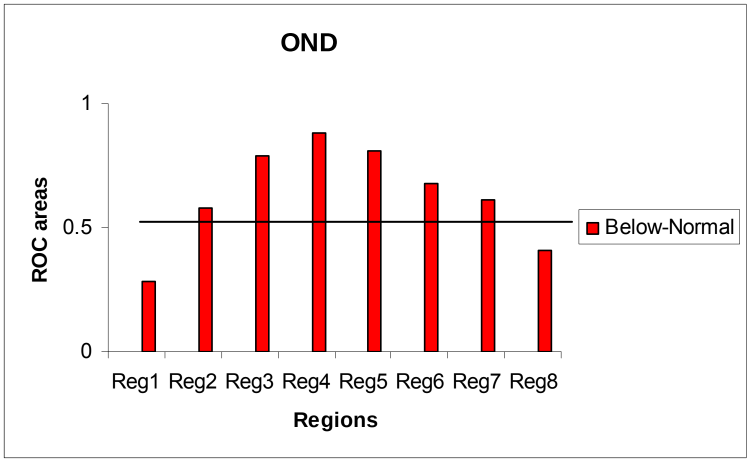


Figure 5.5a Relative Operating Characteristics area score for below-normal category

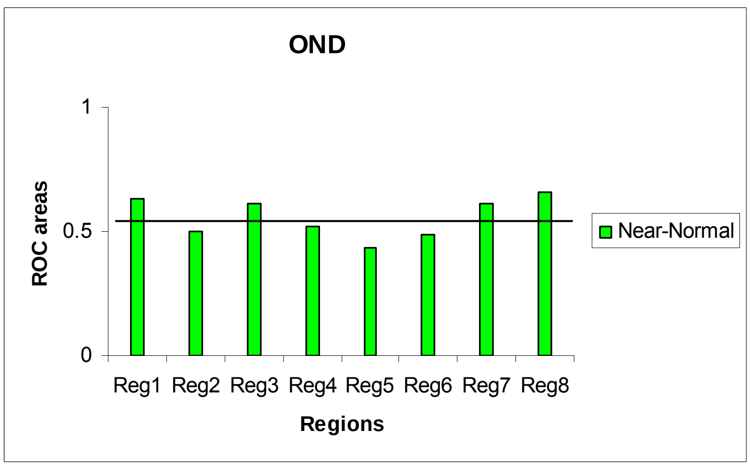


Figure 5.5b Relative Operating Characteristics area score for near-normal category

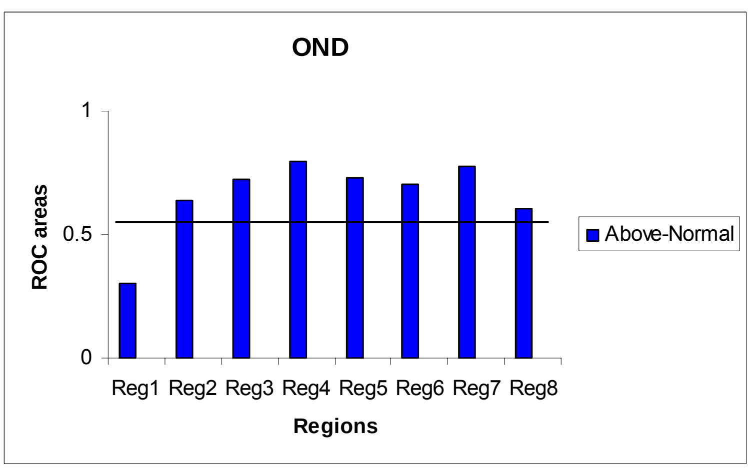


Figure 5.5c Relative Operating Characteristics area score for above-normal category

The Relative Operating Characteristics (ROC) area score for the October-November-December (OND) rainfall seasonal forecasts are shown in figure 5.5 (a-c) separately for the below-normal, near-normal and above-normal outcomes. For OND, there is skill in predicting the below-normal category for stations in regions 2, 3, 4, 5, 6 and 7 (figure 5.5a). The area under the ROC curve for stations in those regions is greater than 0.5 as shown in figure 5.5a. The area under the ROC curve for stations in other regions is under 0.5 which implies that there is no skill in predicting the below-normal category for those stations in regions 1 and 8. OND is the only season that has skill when predicting the near-normal category for stations in most of eight homogenous regions (figure 5.5b) which are regions 1, 3, 4, 7 and 8. The area under the ROC curve is greater than 0.5 in the most of the eight homogenous regions (figure 5.5b). There is no skill in predicting the near-normal category for stations in regions 2, 5 and 6. There is skill in predicting above-normal category for stations in most of the regions except stations in region 1 (figure 5.5c). The area under the ROC curve is greater than 0.5 as shown in figure 5.5c.

OND shows skill in most of regions for the three categories, but the use of the middle-category is not a good idea as it was supported by other skill measure, RD and RPSS.

5.6 Summary

Relative Operating Characteristics (Mason, 1982) is considered by the World Meteorological Organization (WMO) as a recommended method of indicating the skill of probabilistic seasonal forecasts. The area under the ROC curve is a simple index summarizing the skill of a forecast system but is sensitive to the number of points that are plotted. For January-February-March (JFM), the forecast of below-normal rainfall category can be issued with confidence as compared to the forecast of above-normal category in this season. April-May-June (AMJ) rainfall is associated with low skill because the below-normal rainfall category was observed but not forecast. The area under the ROC curve has shown that the South African Weather Service operational seasonal forecast is more successful for the above-normal category as compared to the below-normal category for October-November-December (OND). Such information is valuable because a forecaster can provide higher level of confidence in seasonal forecasts for above-normal category for this time of the year.

In general, the use of the near-normal category is not a good idea, since there was no skill in most of the seasons except OND. Models need to be improved especially in terms of seasons, when is mid-summer or winter. Forecasts show skill in summer-rainfall region during summer and also happen with winter-rainfall region where it shows skill in winter. This need to be looked at because it disadvantages Users especially when they are in the summer-rainfall region and during winter and other way round. They also need forecast for some other reason, i.e. for coldness, if the winter will be too cold, then they need to prepare for that especially since South Africa experiences shortage of electricity lately.

Chapter 6

Conclusion

The South African Weather Service (SAWS) started issuing probabilistic seasonal rainfall and temperature forecasts in 1994. Seasonal forecasts are issued probabilistically to reflect the amount of forecast uncertainty involved in this time range. When the forecasts were started, the main contribution to the forecasts was from statistical models. In 1997, output from a number of GCMs run at SAWS as well as one run at IRI was used as part of input to the final SAWS seasonal forecast. With time forecast from the ECMWF and UKMO models were used and now model output from the Universities of Pretoria and Cape Town are also considered. The forecasts that are available for the whole year that had input from both GCMs and statistical models started in 1998. Until recently the SAWS seasonal forecasts were not verified. In 2004, eye-ball verification was introduced; however there is a need for forecasts to be verified using statistical skill measures because eyeball verification is very subjective.

The objective of the research is to use statistical measures to verify the SAWS operational seasonal forecasts. Verification is done for the years 1998 to 2004. Verification is made for four seasons that are defined by three months following each other from January until December. Verification can be done using a number of skill measures, however, in this study only three measures are used namely Ranked Probability Skill score (RPSS), Reliability Diagram (RD) and Relative Operating Characteristic (ROC). Each of these skill measures provides valuable information about the skill of the forecasts. The RD includes a full depiction of forecasts and observations whereas RPSS gives credit for forecasting the observed category with high probabilities, and penalties for forecasting the wrong category with high probabilities is substantial. Therefore to evaluate the seasonal forecasts with different skill measures is advantageous.

An RPSS of 1 reflects a perfect forecast, 0 shows that the forecast is not performing better than climatology, while a score less than 0 indicates that forecasts are worse than climatology. The forecasts for the season October-November-December (OND) are more skillful as compared to the other three seasons. During the JAS season the forecasts are skillful over the western parts of the country where most of the rainfall is expected during this season. During the other two seasons skill scores of less than 0 dominate the country which means that the forecasts performed worse than forecast of climatology.

RD shows skill when the seasonal forecast lies on a perfect line (1:1). If the seasonal forecast lies above of the perfect line then it is over-forecast, but if it is under the perfect line then it is under-forecast. The forecast of climatology was never issued for JFM and OND which are summer seasons but for JAS and AMJ the forecast of climatology was issued during the period of 1998 to 2004. There is skill in predicting JFM and OND seasons with the above-normal category. The line of regression is close to perfect reliability; however, the average observation is more than the averaged forecast. On the other hand, there is skill in predicting the below-normal category for AMJ and JAS.

For the ROC diagram, a score of one represents perfect forecasts. The area of less than 0.5 of the forecasts shows no skill because the area given by a random forecast is 0.5. The ROC curve shows that the above-normal category is forecast with more skill as compared to the below-normal category for the JFM and OND seasons. In three of the four seasons there is no skill in predicting the season with the near-normal category.

Among the three skill measures, RPSS shows negative values in the in most of the seasons except in OND. Most of the positive scores are shown in the summer rainfall season of OND. The rest of the season shows small area of positive values. For the RD and ROC, below-normal rainfall was observed even though above-normal was forecast most of the time. The RD and ROC show almost the same pattern. Both of them have larger magnitudes of positive skill than RPSS. This is because of the fact that the ROC score and RD do not distinguish the severity of errors across categories while the RPSS give different penalties in terms of categorical distance errors. Again in both RD and ROC, there is no skill in prediction for the near-normal category.

From RD and ROC, it shows that there is skill in predicting the two seasons with outer-categories. This also reflected by RPSS where OND has a skill because of outer-categories but JFM didn't have skill because of not using the above-normal category in most of the times. For AMJ and JAS, below-normal category is favoured and it shows skill but in RPSS result, the skill is reflected especially in the winter-rainfall regions. Forecasters need to be advised to use the outer-categories but a lot need to be done in terms of models. The skill is mainly in summer rainfall region when is summer and same applies to winter-rainfall region where there skill happen in winter but for beginning of summer (OND) and mid-winter (JAS). For mid-summer (JFM) and beginning of winter (AMJ) lot need to be done because there is luck of skill. This was reflected by all the skill measured and it also shows that the use of outer-categories could improve the forecasts. Forecasters need to be advised not to use the middle category because of luck of skill as shown by the skill measures. They must also be advised not to issue the forecast of climatology since does not help Users in terms of decision making.

The largest areas for which the SAWS seasonal rainfall forecasts show skill for all the rainfall categories is found during the summer season of OND. In winter, the forecasts are more skillful over the western coast and the adjacent interior. Therefore the higher skills are found during the rainfall seasons associated with the highest climatology means. When the analysis for the three categories is made separately, it becomes obvious that there is more skill in predicting the above-normal and the below-normal categories than the normal category. The above finding is in agreement with the results that were found at the IRI (Goddard et al., 2003). The above results suggest that the SAWS seasonal forecasts provide useful information on above-normal category (summer rainfall regions) and below-normal category (winter rainfall regions). However, the forecasts are still very far from perfect and so there is still a lot of research that has to be directed towards improving operational forecasts at the SAWS.

The seasonal forecasts are made using models as well as human intervention. It is not clear whether or not human interventions improves the model forecasts, or worsen them. Verification based on the seasonal forecasts made using the different models and an objective method to combine the forecasts, and those that are combined subjectively, should be made to determine which one performs the better. Research done at the IRI for models show that some models have no skill but this cannot be the only contributing factor to errors in the forecasts, because SAWS does not only use the IRI models for their decision making. Regional Climate Models could also be used because they represent small scale events better than Global models. Forecasters should be careful with the probabilities assigned to the near-normal category, because as shown by the ROC and RD skill measures, there is no skill in predicting this category for most of the regions.

Updated forecast skill scores using different skill measures should be available to users of seasonal forecast. It enables users to be informed and understand seasonal forecasts and its limitations, and this should increase the usefulness of seasonal forecasts for decision making. Decision makers can begin to determine essential forecast attribute requisite performance thresholds and relationship among the quality of forecast and their usefulness in decision making and ultimately their economic value. Ultimately, a meteorologist's determination of skillful forecast is only valuable to the extent that the forecast can provide benefit to those incorporating the information into their decision process. However, forecast provider's definition of forecast quality may vary greatly from user's definition of quality and such definition may even vary among different users

(Hartmann et al, 2002). Some countries have started to verify the forecast according to users needs (Hartmann et al, 2002) and South African Weather Service can follow the lead and verify the forecast according to these needs.

References

Barnett, T.P., 1984: Prediction of El Niño of 1982-83.*Mon. Wea. Rev*., **112**, 1403-1407.

Barnston, A.G., 1994: Linear statistical short-term climate predictive skill in the Northern Hemisphere. *J.Climate*, ***7**, 1513-1564.*

Barnston, A.G., Mason, S.J., Goddard, L., DeWitt, D.G., and Zebiak, S.E., 2003: Multi-model ensembling in seasonal forecasting at IRI. Bull. *American Meteor. Soc*., **84**, 1783-1796.

Brier, G.W and Allen, R.A., 1951: Verification of weather forecasts. In: MaloneTF(ed). Compendium of Meteorology. *American Meteor. Soc*. 841-848.

Carson, D.J., 1998: Seasonal forecasting. *Quarterly.J.Royal. Meteor. Soc*, **124**, 1-26.

Drosdowsky, W., 1994: Analog (non-linear) forecasts of the Southern Oscillation Index time series. *Wea.Forecasting*, ***9**, 78-84*.

Egan, J.P., 1975: Signal detection theory and Relative Operating Characteristics analysis. *Academic Press*, 277pp.

Goddard, L., Barnston, A.G. and Mason, S.J., 2003: Evaluations of the IRI"s "net assessment" seasonal climate forecasts. *Amer. Meteor. Soc*., 1761-1781.

Green, D.M and Swets, J.A., 1996: Signal detection theory and psychophysics. Wiley, 455pp (Reprinted by Robert E Krieger Publishing Co., 1974)

Hastenrath, S., Greischar, L. and van Heerden, J., 1995: Prediction of summer rainfall over South Africa. *J.Climate*, **8**, 1511-1518.

Huang, J., Van den Dool, H.M and Barnston, A.G., 1996: Long-Lead seasonal temperature prediction using optimal climate normals. *J.Climate*, **8**, 809-817.

Hunt, B.G., 1997: Prospects and problems of multi-seasonal predictions. *Int.J.Climatol*, **17**, 137-154.

Ji, M., Leetmaa, A., and Kousky, V.E., 1996: Coupled model prediction of ENSO during the 1980s and 1990s at the National Centers for Environmental Prediction. *J. Climate*, **9**, 3105-3120.

Jolliffe, I.T. and Stephenson, D.B., 2003: Forecast verification: A practitioner's guide in atmospheric science. 240pp.

Jury, M.R., 1995: A review of research on ocean-atmosphere interactions and South African climate variability. *SA.J.Sci*., **91**, 289-294.

Jury, M.R. and Mason, S.J., 1999: Exploratory log-range model to estimate summer climate variability over southern Africa. *J.Climate*, **12**, 1892-1899

Klopper, E., 1999: The use of seasonal forecasts in South Africa during the 1997/98 rainfall season. *Water S.A.*, **25**, 311-316.

Klopper, E., 2002: The use and benefits of seasonal climate forecasts in South Africa.

Kruger, A.C., 1998: The relationship between ENSO, seasonal rainfall and circulation patterns in South Africa.

Landman, W.A. and Mason, S.J. 1999: Operational long-lead prediction of South African rainfall using canonical correction analysis. *Int. J. Climatol*., **19**, 1073-1090.

Landman, W.A. and Tennant, W.J., 2000: Statistical downscaling of monthly forecasts. *Int. J. Climatol*., **20**, 1521-1532.

Lindesay, J.A, 1988: South African rainfall, the southern hemisphere semi-annual cycle. *J. Climatol*., **8**, 17-30.

Livezey, R.E., Masutani, M., and Ji, M., 1996: Sea-Surface Temperature-forced seasonal simulation and prediction skill for versions of the NCEP/MRF model. Bull. *Amer. Meteorol. Soc*., **77**, 507-517.

Mason, I.B., 1982: A model for assessment of weather forecasts. *Australian Meteorol.* Magazine, **30**, 291-303.

Mason, S.J., 1990: Temporal variability of sea surface temperature around southern Africa. *SA. J.Sci*, **86**, 243-252.

Mason, S.J., 1995: Sea-Surface Temperature-South African rainfall associations, 1910-1989. *Int. J. Climatol*., **15**, 119-135.

Mason, S.J., 1996: The role of meteorology in the national economy. *Mon.Wea.Rev*, November, 382-393.

Mason, S.J and Jury, M.J. 1997: Climate variability and change over southern Africa: a reflection on underlying processes. *Progress in Physical Geography*, **21**, 23-50.

Mason, S.J., 1998: Seasonal forecasting of South African rainfall using a non-linear discriminant analysis model, *J. Climatol.*, 147-164.

Mason. S.J., Goddard, L., Graham, N.E., Yelaeva, E., Sun, L. and Arkin, P.A., 1999: The IRI seasonal climate prediction system and the 1997/98 El Nino event. *Bull. Amer. Meteor. Soc*, **80**, 1853-1873.

Murphy, A.H., 1996: The Finley affair: a signal event in the history of forecast verification. *Wea. Forecasting*, **11**, 3-20.

Murphy, A.H., and Wilks, D.S., 1998: A case study in the use of statistical models in forecast verification: precipitation probability forecast. *Wea. and forecasting*, **13**, 795-810.

Palmer, T.N. and Anderson, D.L.T., 1994: The prospects of seasonal forecasting. *Quaterly.J.Royal.Meteor. Soc*, ***120**, 755-793.*

Pathack, B.M.R, Jury, M.R., Shillington, F.A and Courtney, S. 1993: South African summer rainfall variability and its association with the marine environment. *Water Research Commission Report No 278.1/94*.

Rautenbach, C.J. deW, and Smith, I.N., 2001: Tele-connections between global sea-surface temperatures and the inter-annual variability of observed and model simulated rainfall over southern Africa. *J. Hydrology*, **254**, 1-15.

Sanderson, M., 1999. The classification of climates from Pythagoras to Koppen. *Bull. Amer. Meteor. Soc*, **80**, 669-673

Schulze, B.R., 1965. Climate of South Africa. Part 8. General Survey. *WB28*. South African Weather Bureau, Pretoria.

Sonka, S.T, Changnon, S.A and Hofing, S., 1988: Assessing climate information use in agribusiness, Pt. 2, Decision experiments to estimate economic value. *J. Climate, Boston*, **1**, 766-774. 1988.

Stanski, H.R., Wilson, L.J and Burrows, W.R., 1989: Survey of common verification methods in meteorology. *World Weather Watch Technical Report,* **8**. WMO, Geneva.

Swets, J.A., 1973. The Relative operating characteristics in psychology. *Science*, **182**, 990-100.

Trenberth, K.E., 1992: Climate system modeling. *Cambrige.Uni.Press*, 788pp.

Walker, N.D., 1990: Links between South African summer rainfall and temperature variability of the Agulhas and Benguela currents systems. *J.Geophysical.Research*, **95**, 3297-3319.

Van Den Dool, H.M. and Toth, Z., 1991: Why do forecasts for"Near Normal "often fail? *Wea. Forecasting*, **6**, 76-85.

Van Heerden, J., Terblanche, D.E and Schulze, G.C., 1988: The Southern Oscillation and South African summer rainfall, *J. Climatol*., **8**, 577-597.

Ward, M.N. and Folland, C.K., 1991: Prediction of seasonal rainfall in the north Nordeste of Brazil using eigenvectors of sea-surface temperatures. *Int. J. Climatol*., **11,** 711-743.

Wilks, D.S., 1995: Statistical method in the atmospheric sciences*: An into. Press*. 464pp.

Wilks, D.S., 2002: Realizations of daily weather forecast seasonal climate. *J. Hydrometeor.*, **3**, 195-207.

Wilks, D.S., 2006: Statistical method in the atmospheric sciences: Second Edition. *Int. Geo. Series*, **91**, 627pp.

Wilks, D.S., and Godfrey, C.M., 2002: Diagnostic verification of the IRI new assessment forecasts, 1997-2000.*J. Climate*, **15**, 1369-1377.

Zhang, H. and Casey, T., 2000: Verification of categorical probability forecasts. *American Meteorol. Soc.*, **15**, 80-89.