ASSESSING THE PREDICTABILITY OF RAINFALL IN SOUTHERN AFRICA WITH FOCUS ON ESWATINI USING GLOBAL OCEAN AND ATMOSPHERE INDICATORS

By

NCONGWANE Musa Sibusiso

I56/9856/2018

Department of Meteorology

School of Physical Sciences

University of Nairobi

A Research Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Meteorology, University of Nairobi, Kenya.

June 2019

DECLARATION

I declare that this dissertation is my original work and has not been submitted elsewhere for examination, the award of a degree or publication. Works of other people used in this dissertation have been duly sited as references in accordance with the requirements of the University of Nairobi.

Signature **Date**

NCONGWANE Musa S. …….. ……

I56/9856/2018

Department of Meteorology

University of Nairobi

This dissertation was developed and written with our approval as University Supervisors:

Signature **Date**

Dr. Joseph N. Mutemi …….. ... …….

Department of Meteorology

University of Nairobi

P.O. Box 30197-00100

Nairobi, Kenya

Dedicated to my late Supervisor and Mentor

The late Prof. Joseph M. Ininda

Department of Meteorology

University of Nairobi

P.O. Box 30197-00100

Nairobi, Kenya

DEDICATION

I dedicate this dissertation to all my family, friends, my wonderful mother, my precious daughter Lwakhile and my fiancée.

ACKNOWLEDGEMENTS

First and foremost, I am grateful to God the creator of the universe for blessing me with my current job which led me to getting this wonderful opportunity to study at the University of Nairobi. I am also humbled to the Lord for keeping me healthy and safe during my studies and furthermore for keeping my family back home healthy while I was away.

My sincere gratitude goes to my supervisors Dr. J. N. Mutemi and the late Prof. J. M. Ininda for their valuable support and guidance throughout my research period. Their mentorship and thought provoking discussions made it possible for me to complete my dissertation.

Special appreciation goes to the entire staff at Eswatini Meteorological Service especially the forecasting office, for encouraging me to go and study at the University of Nairobi. Special thanks are forwarded to my supervisor at work Mr. E. Seyama for passionately encouraging me to go for studies and for his belief in my capabilities to succeed. I am also grateful to Mr. S. Nzalo for providing the rainfall data used in this study.

My appreciation also goes to the Government of Eswatini for granting me study leave to pursue this degree. I cannot forget to pass my special thanks to the World Meteorological Organisation for awarding me this fellowship to study for my Master of Science in Meteorology at the University of Nairobi.

Special appreciation goes to the Department of Meteorology under the chairmanship of Prof. A. Opere. I also extend my sincere appreciation to my classmates and friends for their valuable support and encouragement in compiling this dissertation.

Last but not least, I am very grateful to Pamela and Mark Majodina the representative of WMO for Eastern and Southern Africa, for their support during my studies.

MAY THE PRECIOUS GRACE OF GOD THE ALMIGHTY BE WITH YOU ALL!

ABSTRACT

Southern Africa and Eswatini in particular rely heavily on rainfall for agricultural production which supports socio-economic activities. Thus, skilful and timely rainfall predictions are very imperative, since it is only the prediction information that can be used to trigger actions that lead to reduction of climate related hazards. The predictive potential of rainfall in southern Africa and more especially in Eswatini using dominant ocean and atmosphere indicators was established while seeking to investigate the potential for predicting rainfall in Eswatini with increased time leads. The study was motivated by the significant role of rainfall variability in influencing social and economic activities in the sub-region and the need to avail skilful climate forecasts in advance of the target season.

Rainfall data from stations in Eswatini and gridded rainfall data from the Global Precipitation and Climatology Centre were used in the analysis, together with global wind, mean sea level pressure and sea surface temperature data sets drawn from the NCEP/NCAR reanalysis. The study used principal component analysis (PCA) or empirical orthogonal functions (EOFs) as the main analysis method while principal component regression was used for developing statistical rainfall prediction models. The ENSO phenomenon was found to cause the second greatest variability in the global SSTs after the general warming of the global oceans which had the largest variance.

A notable outcome from the study is that circulation at upper levels of the atmosphere have an important contribution to the predictability of rainfall in Eswatini. Using the station rainfall data, models for predicting December-January-February (DJF) rainfall anomalies were developed and a lead time of 3 months using June-July-August (JJA) predictors was found to be viable with good and useful results.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS AND ABBREVATIONS

CPT	Climate Predictability Tool
DJF	December-January-February
ENSO	El Niño-Southern Oscillation
EOF	Empirical Orthogonal Function
FAO	Food and Agriculture Organisation
GPCC	Global Precipitation Climatology Centre
IOD	Indian Ocean Dipole
ITCZ	Inter-tropical Convergence Zone
JFM	January-February-March
JJA	June-July-August
MAM	March-April-May
mb	millibar
MJO	Madden-Julian Oscillation
MLR	Multiple Linear Regression
MSLP	Mean Sea Level Pressure
NCEP/NCAR	National Centers for Environmental Prediction/ National Center for Atmospheric Research
NOAA	National Oceanic and Atmospheric Administration
PC	Principal Component
PCA	Principal Component Analysis

PCR	Principal Component Regression
ROC	Relative Operating Characteristic
SARCOF	Southern African Regional Climate Outlook Forum
SIOD	Subtropical Indian Ocean Dipole
SON	September-October-November
SSTs	Sea Surface Temperatures
SWIO	South-West Indian Ocean
TCs	Tropical Cyclones
TTTs	Tropical-Temperate Troughs
U	Zonal wind
WMO	World Meteorological Organisation

CHAPTER ONE

1.0 INTRODUCTION

The southern Africa region is characterized by varied topography associated with a variety of climate patterns. Agriculture in this region, which supports social and economic development, is significantly reliant on seasonal rainfall. Generally, over seventy-five percent (75 %) of the annual rainfall is received during the austral summer period (October to March), a season when most of the growing spell takes place. Variations in the amount of rainfall received especially during the agricultural growing season may have serious implications to the economies of this region which comprises largely of developing countries. A majority of the population lives in rural areas and relies heavily on rainfall-fed subsistence farming. The rainfall in the region is greatly uneven in space and time. Moreover, the rainfall pattern is associated with floods as well as droughts that have augmented frequencies (Mwafulirwa and Jury, 2002). For example, according to a report by Eswatini Economic Policy Analysis and Research Centre, the recent El Niño influenced drought of 2015/16 cost the economy of Eswatini, the country formerly called Swaziland E3.843 billion (about $289.6 million).

Apart from the unpleasant effects of rainfall variation to agriculture, seasonal rainfall anomalies also impact negatively on numerous social and economic activities that include nature conservation, power production, transportation, tourism, forestry and many others. Hence, timely and reliable seasonal rainfall forecasts would significantly contribute to effective scheduling, management and administration of all social and economic activities in the region which depend on rainfall. Several studies on rainfall variability have been conducted for the southern Africa region with the aim of improving seasonal rainfall forecasts (Jury, 1996; Nakamura, 2011; Fauchereau *et al.,* 2003; Todd *et al.,* 2011).

Even though dynamical climate prediction models have been developed in an attempt to improve forecasting, empirical statistical prediction models remain valuable, sometimes performing better than dynamical models in operational prediction (Landman *et al.*, 2009), especially in seasonal rainfall prediction where longer lead times are desired and where small spatial scales are considered (Shongwe *et al*, 2006). In this study statistical models were developed with the aim of assessing their predictive potential with increased time leads. This is

imperative for timely planning and resource mobilisation especially for disaster management agencies.

1.1 Statement of the problem

Seasonal rainfall forecasts in the region are currently produced (by the Southern African Regional Outlook Forum-SARCOF) around September thus providing limited time for the forecast information to be directly usable in taking decisions which may lead to reducing climate related impacts. This study hence aimed to assess the predictability of rainfall over the area of study, while aiming to improve the lead time for seasonal rainfall prediction. Long lead forecasts are desirable in enhancing mitigation measures especially against extreme climate events, like the recent extreme 2015 drought in the region. The economy and livelihood of Eswatini and southern Africa as a whole depend largely on rain-fed agriculture hence a good understanding of the rainfall variability cannot be over emphasized.

1.2 Objective of the study

This section presents the main objective and specific objectives of the study.

1.2.1 Main objective

The main objective was to assess the predictability potential of seasonal rainfall over southern Africa and Eswatini in particular using the dominant ocean and atmosphere patterns which are the major drivers of year to year rainfall variability over the southern Africa region.

1.2.2 Specific objectives

The study was based on the following specific objectives:

(a) To determine the major space and time modes of the potential predictors influencing rainfall variability.
(b) Identification of the nature of atmospheric circulation modes and sea surface temperature associated with various rainfall anomaly patterns in southern Africa and Eswatini in particular.
(c) To develop regression models for prediction of seasonal rainfall using the modes obtained in (b).

1.3 Justification of the study

The study is motivated by the important role of seasonal rainfall variability in influencing social and economic activities in the region. In particular, climate extremes like flooding and drought have significant impacts on the livelihood activities and economies of southern Africa. In Eswatini, failure of seasonal rainfall and droughts cause huge socio-economic problems like failure of maize crop which is the stable food crop, water shortages for domestic and hydropower generation among other adverse impacts. When rainfall is surplus, the opposite types of impacts maybe realised. The region is dominated by developing countries with limited resources for preparing and responding to natural disasters in form of floods and water shortages induced by droughts. Hence, skilful and timely rainfall predictions are very important, in that it is only the prediction information that can be used to trigger early actions that lead to reduction of the climate related hazards. The study investigated the potential for predicting rainfall in the region with increased time leads and possibly enhances early warning capacity for the region and Eswatini in particular. This will go a long way in the disaster risk reduction strategies and effective management of water resources among others.

1.4 Area of study

This section introduces the area where the study was conducted. Figure 1 shows southern Africa (left) and a detailed map of Eswatini (right) highlighting the physical features. Southern Africa as considered in the study lies between 16^{o}S to 38^{o}S and 12^{o}E to 38^{o}E, however, the focus was on Eswatini which had stations rainfall records.

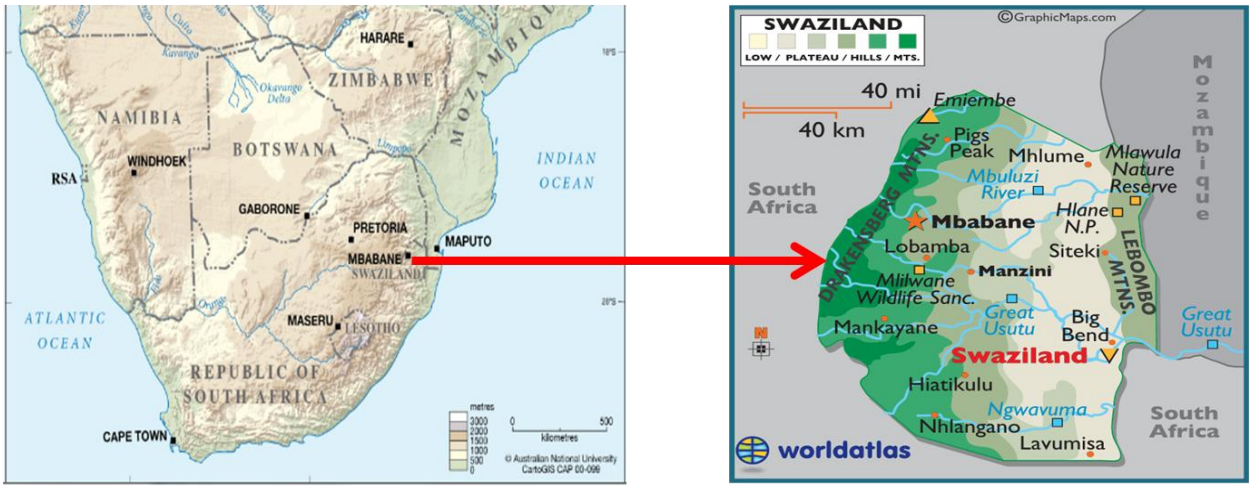


Figure 1: The map of southern Africa and Eswatini (formerly Swaziland) as defined in the study.

Sources:http://asiapacific.anu.edu.au/mapsonline/base-maps/southern-africa and https://www.worldatlas.com/webimage/countrys/africa/lgcolor/szcolor.htm (right panel).

1.4.1 Climate of southern Africa

Southern Africa's climate is arid to tropical. The arid and semi-arid regions are most affected by droughts, while intense rainfall after drought episodes has the potential to result in devastating flooding. Southern Africa is affected by some dominant wind systems like the southeasterly wind system and the inter-tropical convergence zone. The south easterlies bring moisture from the Indian Ocean resulting in rainfall especially along the eastern highlands, while the inter-tropical convergence zone has a contribution to both the dry and wet seasons experienced by the tropical regions. Apart from seasonal variation, the region's rainfall fluctuation is also known to be influenced by ENSO (Nicholson and Kim, 1997).

Rainfall in the region generally occurs in the austral summer months (October to March) except for the coastal south and southwest South Africa where most of the rainfall is received during austral winter (Reason, 2017).

1.4.2 Geography and climate of Eswatini

The Kingdom of Eswatini (formerly Swaziland) is a landlocked country found in southeastern Africa. It shares borders with Mozambique on the east and South Africa on the north, west and south. The total area covered by the country is approximately 17 360 km^2. The country is mainly hilly and mountainous on the west and east. There are some moderately sloping areas in the Lowveld. The country has four agro-ecological zones namely; the Highveld, Middleveld, Lowveld and Lubombo Plateau. These zones are based on geology, soils, vegetation, rainfall amount received and elevation.

Rains occur mainly in summer and the climate is humid subtropical. The climate varies mainly due to altitude with more rainfall occurring on the western side where the altitude is higher. The summer season spans from around October to March with about 75% of the precipitation falling during this period. The rainfall amounts received by each of the ecological zones are shown in table 1 (FAO, 2005).

Table 1: Ecological zones in Eswatini with annual rainfall ranges received by each zone.

Ecological Zone	**Rainfall (mm)**
Highveld	700-1550
Middleveld	550-850
Lowveld	400-850
Lubombo Plateau	550-850

Source: FAO, 2005

Essentially from table 1, in terms of rainfall, the ecological zones can be reduced to 3 since the Middleveld and the Lubombo Plateau receive similar amounts.

1.4.2.1 Climatology of Eswatini

The annual cycle of rainfall in Eswatini (example of Mbabane station rainfall) is shown in figure 3. As shown, figure 3 captures well the unimodal rainfall characteristic with a peak in January for Eswatini. All other stations in the country possess a rainfall pattern similar to that of Mbabane shown in figure 3.

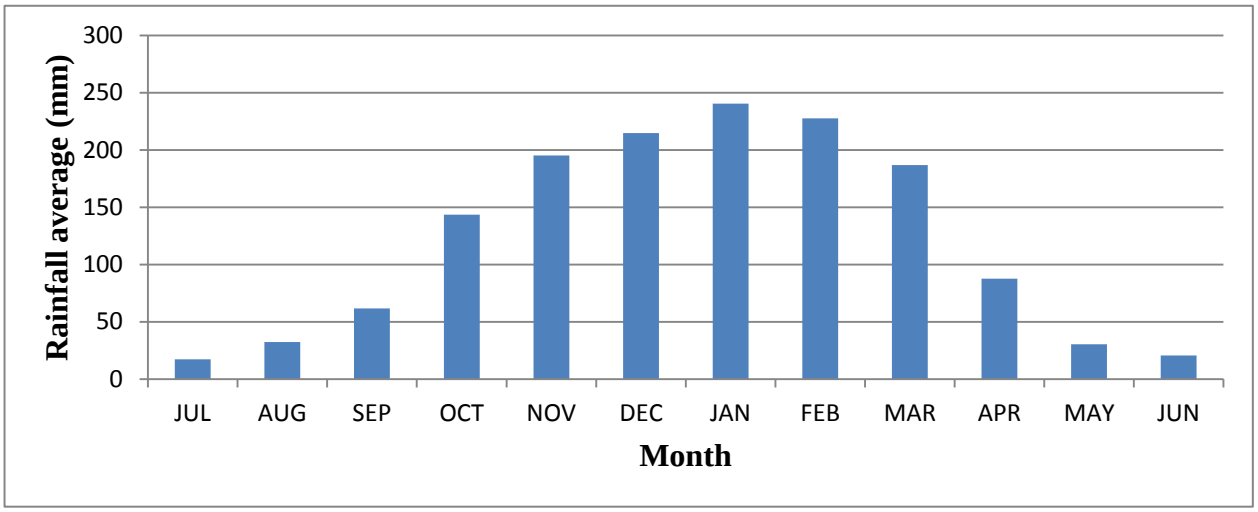


Figure 2: The monthly mean rainfall of Eswatini; an example of Mbabane station rainfall.

CHAPTER TWO

2.0 LITERATURE REVIEW

One of the most challenging issues in dynamical climate models is to forecast precipitation. This is due to the fact that apart from it being influenced by large-scale dynamics of the atmosphere, it is also influenced by weather systems operating at small-scale and by a number of atmospheric variables and local topography (Gitau *et al*., 2014). However, significant progress has been made in forecasting precipitation (especially rainfall) at seasonal time scales in many parts of the world (e.g. Nicholson, 2014; Barnston and Korecha, 2007; Hasternrath *et al.*, 1995 and Maldonado *et al.* among others). Many studies have linked the predictive potential of rainfall over southern Africa to its association with sea surface temperatures (SSTs) and ENSO. These include Kenyon and Hegerl (2010), Tyson *et al.* (2002), Mason and Graham (1999) and Sharon (1987) among others.

Nicholson and Kim (1997) confirmed the correlation between southern African rainfall and ENSO. On recalibrating General Circulation Model (GCM) forecasts over southern Africa, Landman and Goddard (2002) found that the 850 mb geopotential height has the greatest predictive strength over the region. While investigating the regional circulation features influencing southern African rainfall during the positive phase of ENSO (El Niño), Blamey *et al*. 2018, found two low level circulation features and a mid-level circulation feature that play imperative roles in controlling the rainfall through either their intensification or weakening. They discovered that a strong El Niño is associated with a weak Angola low and a weak Mascarene High in the south Indian Ocean which results in less moisture being advected into the subcontinent hence less rainfall. On the other hand, they found that during the same period the mid-level Botswana high is intensified leading to suppressed rainfall conditions. Using Generalized Linear models, a study by Ambrosino *et al*. (2011) found that major factors influencing rainfall in southern Africa include relative humidity and ENSO. Jury (1996), found significant correlation between southern African rainfall and SSTs located in central equatorial Indian Ocean. Positive SST anomalies in the neighbouring South Atlantic were found to increase summer rainfall in some parts of southern Africa (Barnston *et al*., 1996).

Moreover, some researchers have made attempts to investigate seasonal rainfall predictive potential over different parts of southern African region. Using an atmospheric general circulation model to conduct an ensemble of experiments, warming SSTs over the south western Indian Ocean were found to be linked with enhanced rainfall over central and large parts of eastern South Africa (Mulenga and Reason, 1999). On applying model output statistics to a general circulation model (GCM), Landman *el at*. (2005), found positive prediction skill for the seasons DJF and MAM (March-April-May) when predicting extreme rainfall over southern Africa.

Furthermore, some researchers have attempted to explore the potential for long lead forecasts for different areas within southern Africa. For example, Jury and Makarau (1997) investigated the predictability of summer rainfall in Zimbabwe and found viable summer rainfall forecasts with a lead time of one season. Also, only particular places in southern Africa were assessed to have usable skill (>0.3) during the January-February-March (JFM) summer period with 1 to 4 months lead forecasts (Barnston *et al.*, 1996).

Most studies based on developing statistical models for predicting rainfall over southern Africa have only used sea surface characteristics and the current study considered atmospheric variables. Moreover, studies similar to the current one have not been conducted in Eswatini before hence the study focused on Eswatini's rainfall.

2.1 Major influencers of rainfall variability in southern Africa

2.1.1 Tropical Temperate Troughs (TTTs)

These result from the interaction between tropical and extra-tropical systems that produces cloud bands with a northwest-southeast orientation over southern Africa. TTTs are regarded as the main rain-producing systems at synoptic scale over southern Africa (Cook *et al*., 2004; James *et al.*, 2018). TTTs also contribute to the observed contrasting rainfall conditions over southern Africa and equatorial East Africa during ENSO events whereby the former becomes anomalously dry while the latter becomes anomalously wet (Nicholson and Kim, 1997). While studying the contribution of TTTs to variability in rainfall over southern Africa, Hart *et al*. (2013) found that on average TTTs contribute around 30-60% with a well observed spatial variability. However, during La Niña occurrences a substantial number of cloud bands form

while fewer develop during occurrences of ENSO (Hart *et al*., 2018). The TTTs tend to migrate from west to east into the Mozambique Channel from southern Africa (Crétat *et al*., 2012; Ratna *et al*. 2013).

2.1.2 Tropical Cyclones

While studying the contribution of tropical cyclones (TCs) to annual rainfall in southern Africa's eastern interior (Limpopo basin), Malherbe *et al*. (2012), found a less than ten percent contribution, however, about fifty percent of devastating rainfall occurring over the north-eastern part of South Africa is caused by them. Using TC climatology data for the South-West Indian Ocean (SWIO) basin, Mavume *et al*. (2009) found that out of 64 cyclones making landfall 48 of them affected Madagascar while 16 made landfall over Mozambique. This makes tropical cyclones to be the most devastating natural hazard to the people and economy of Madagascar (Nash *et al.*, 2015). Ash and Matys (2012), while investigating the trajectories of TCs in the SWIO discovered that both the subtropical Indian Ocean dipole (SIOD) and ENSO have a significant influence on the paths taken by the TCs. Island countries in SWIO and mainly those in southern Africa's eastern interior are repeatedly affected by cyclones forming in the basin (Malherbe *et al*., 2012) while about 50% of TCs that formed in the Mozambique channel between 1948 and 2010 made landfall (Matyas, 2015).

2.1.3 Sea Surface Temperatures (SST)

Apart from well established teleconnections between SSTs in the Pacific Ocean and rainfall variability in southern Africa, some studies have found links between southern African rainfall and neighbouring Atlantic and Indian Oceans. For example, Williams *et al*. (2008), found that warming SSTs over southwestern Africa and cooling SSTs over the central part of south Atlantic lead to increased extremes in rainfall and daily rainfall over southern Africa. Reason and Mulenga (1999) found that when the SSTs located over south-west Indian Ocean are warmer, South Africa experiences wetter conditions on the eastern and central parts while colder SSTs are accompanied by opposite effects.

2.1.4 Inter-tropical Convergence Zone (ITCZ)

The Inter-tropical Convergence Zone (ITCZ) is a low pressure region encircling around the earth near the equator where north easterly and south easterly winds converge. It is characterised by a series of thunderstorms. Rainfall in southern Africa around January and February is associated with the ITCZ (Todd *et al*., 2004). The inter-tropical convergence zone's association with wet summers in southeastern Africa is such that it (ITCZ) shifts to south and strengthens while during dry summers it shifts to the north and weakens (Cook *et al.,* 2004). The ITCZ reaches its southern most position around central Mozambique late in the austral summer (Reason, 2017).

2.1.5 Madden-Julian Oscillation (MJO)

The MJO comprises of a west-east propagation of convective clusters on a large scale from the Indian Ocean to western Pacific Ocean around the equator. Over the tropical atmosphere MJO is known to be the major mode influencing variability (Pohl *et al*., 2007). However, the influence of MJO over southern Africa is spatially heterogeneous with sharp periodicities in the 30-60 days range in only tropical regions (Pohl *et al*., 2007).

2.1.6 El Niño Southern Oscillation (ENSO)

Under normal conditions the tropical Pacific Ocean easterly trade winds blow persistently, from the eastern Pacific where a region of high pressure exists to a region around Indonesia dominated by lower surface pressure. As the winds blow westward they drag surface water to the western Pacific where it piles up as warm water. However, every 2 to 7 years a breakdown of the surface pressure occurs whereby surface pressure decreases over eastern Pacific and increases over western Pacific. This pressure change causes the trades to weaken and during strong pressure reversal events, westerly winds replace the easterlies leading to warm water piling up on the eastern Pacific.

As the warming process nears an end the atmospheric pressure on the western Pacific starts to fall and begins to rise over the eastern Pacific. This pattern of alternating surface air pressure decrease and increase over the two ends of the Pacific basin is referred to as Southern Oscillation (SO). Since the reversals in surface air pressure and warming of the ocean occur nearly at the same time, the phenomenon has been termed El Niño-Southern Oscillation (ENSO). ENSO has

been shown to influence precipitation around the globe and especially over the tropical areas. For example, Nicholson and Kim (1997) found that ENSO causes variability in rainfall over a large proportion of the African continent. Ropelewski and Halpert (1987) found ENSO influences rainfall in several regions including North America, Australia, South America, Africa, Indian subcontinent and Central America. In southern Africa ENSO is known to play a dominant role in modulating variations in rainfall (Camberlin *et al*., 2001).

2.1.7 Subtropical anticyclones

These are high pressure regions forming around the 30 degrees latitudes north and south of the equator as a result of the Hadley circulation. Two of the most influential subtropical anticyclones over southern Africa are the south Atlantic (St Helena) and south Indian (Mascarene) anticyclones in the southern hemisphere. The south Atlantic and south Indian Ocean anticyclones are stronger during the winter season in the southern hemisphere especially over the oceans basins (Lee *et al*., 2013). During strong El Niño events the South Indian Ocean anticyclone may be weakened as in the intense El Niño of 2015/2016 leading to less moisture advection into southern Africa hence resulting into reduced rainfall (Blamey *et al*., 2018). On the other hand typically, during strong El Niño the south Atlantic High tends to be intense and shifts eastward, hindering the development of a strong Angola low thus reducing the amount of moisture transported into southern Africa (Colberg and Reason, 2004). In austral winter, as the subtropical anticyclones migrate northward they bring substantial rainfall to the southwest while a large part of the subcontinent remains relatively under subsidence (Reason, 2017).

2.2 Conceptual framework

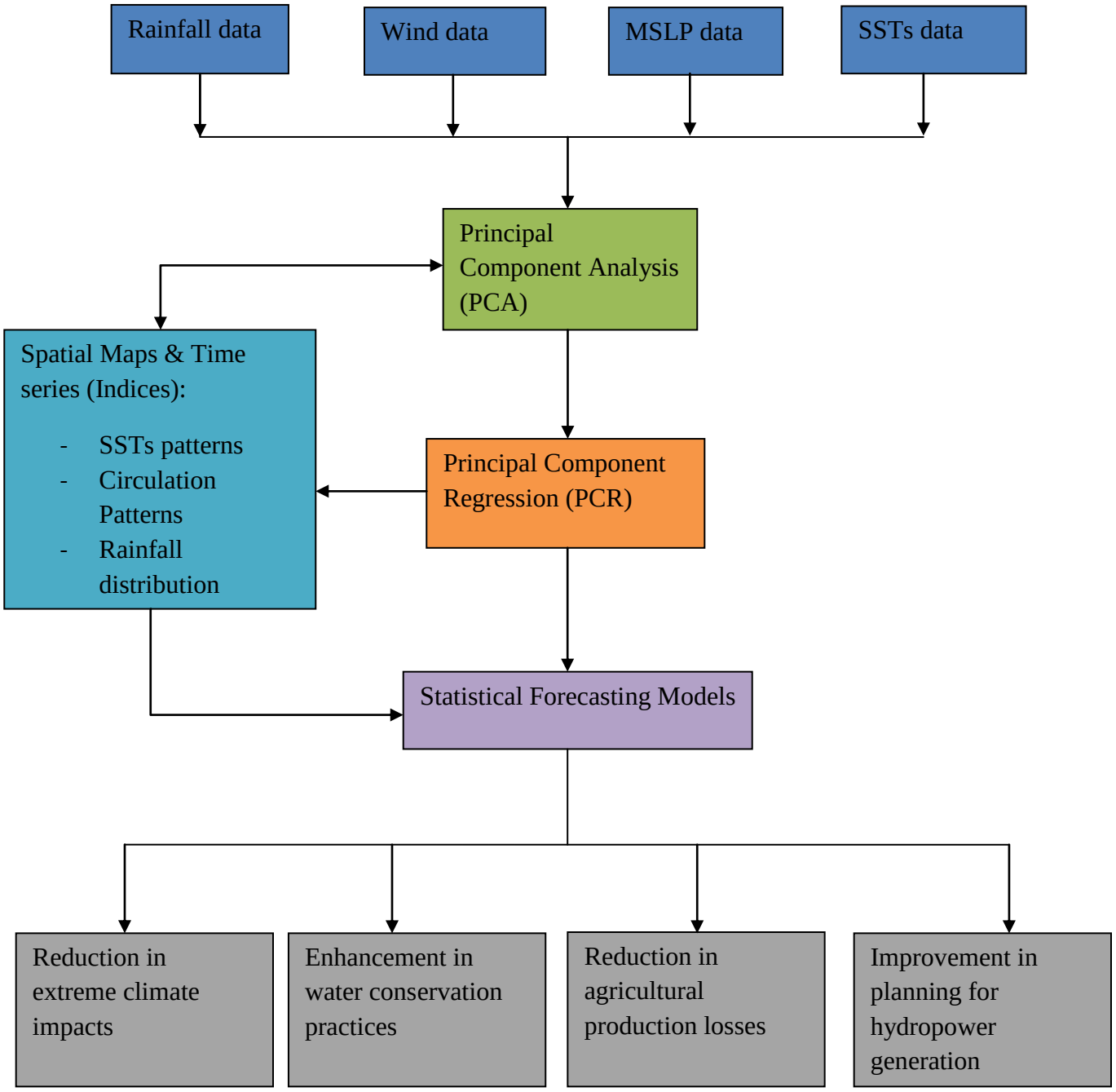


Figure 3: Conceptual framework.

CHAPTER THREE

3.0 DATA AND METHODOLOGY

In this chapter the data and major methods of analysis used are discussed. The data included station rainfall and gridded rainfall from Global Precipitation Climatology Centre (GPCC), global Sea Surface Temperatures (SSTs), Mean Sea Level Pressure (MSLP) and reanalysis winds (at 1000, 850,700 and 200 mb level) from National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis. These levels adequately represent the lower, middle and upper levels of the atmosphere. The methods employed on the other hand consisted of Pearson Correlation, Standardized Anomaly Indices, Linear regression used in the sense of Empirical Orthogonal Functions (EOFs) as spatial modes and Principal Components, which were used as the rainfall predictors. The SST, MSLP and zonal wind (U) data were averaged as 3 month seasons.

3.1 Data used in the study

3.1.1 Rainfall data

Monthly rainfall data from the stations in the area of study was utilized. The analysis base period from 1969 to 2018 was used. Since most of the time rainfall stations do not uniformly represent the area of interest, gridded monthly rainfall from GPCC data set was also used in the study. This data set spans from 1961 to 2013 on a $2.5^{o}\times2.5^{o}$ global grid. This data set cautiously combines data from satellites, sounding and rainfall stations observations to produce a more complete rainfall analysis over the oceans and adds the most desired spatial detail to analyses of rainfall over land masses. GPCC data sets are sufficiently discussed by Becker *et al.* (2013) and Schneider *et al*., 2017 among others. The GPCC data set has been successfully used for example by Rubel and Kottek (2010), in developing world climate shifts projected up to 2100 based on the climate classification by Köppen-Geiger.

The stations used in the study covered Eswatini and had varying data record lengths and are shown in table 2. Figure 4 shows the distribution of the rainfall stations within the country. There were twelve (12) available stations and three (3) of them had more than ten percent (10%) missing data hence, could not be used for further analysis. The correlation method was used to fill any data gaps among the remaining nine stations.

Table 2: Rainfall stations used in the study

Station No.	Station	Latitude	Longitude	Period
1	**Mbabane**	-26.34	31.14	1961-2018
2	**Malkerns**	-26.55	31.16	1961-2018
3	**Matsapha**	-26.52	31.32	1968-2018
4	**Pigg's Peak**	-25.98	31.26	1961-2018
5	**Big Bend**	-26.86	31.93	1961-2018
6	**Nhlangano**	-26.09	31.19	1965-2018
7	**Mananga**	-26.00	31.76	1961-2018
8	**Khubuta**	-26.81	31.43	1961-2018
9	**Mpisi**	-26.38	31.53	1961-2018

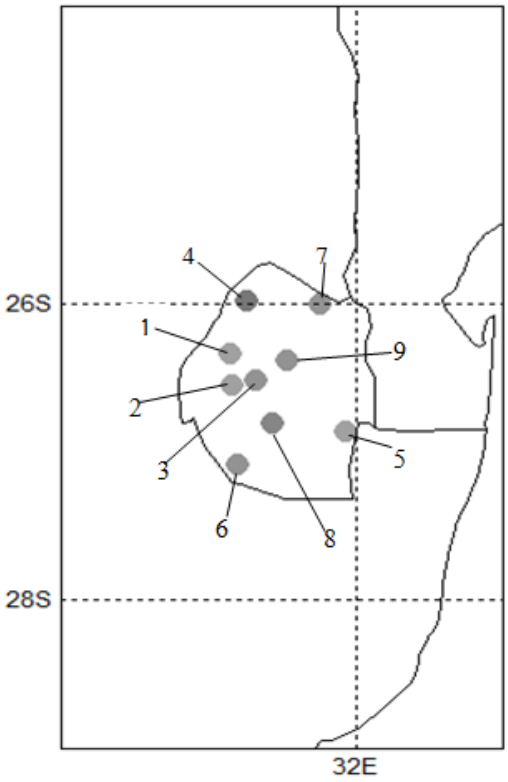


Figure 4: Distribution of the rainfall stations shown in table 2 used in Eswatini (formerly called Swaziland).

However, there are numerous challenges with records for rainfall. Often at times the data records suffer from gaps or missing data which require to be estimated. Primary errors in rainfall records include station rainfall time series heterogeneity, misprint errors and errors inherited during data transmission. The probable causes of heterogeneity in records for rainfall stations may be as a result of a change in station location as well as/or instrument change. Hence, it is imperative to do a quality control on the station records before any meaningful analyses can be conducted. Therefore, data quality control measures taken are discussed under the section of methodology.

3.1.2 Re-analysis wind

Wind data from a global data set was used in the study. This data set is from the National Centers for Environmental Prediction/ National Center for Atmospheric Research (NCEP/NCAR) reanalysis, and was downloaded from their website *https://www.esrl.noaa.gov/psd/*. The reanalyzed zonal winds used are for 1000, 850, 700 and 200 mb levels, spanning a period from 1961 to 2018. Detailed description of the NCEP/NCAR reanalysis datasets is found in Kalnay *et al.* (1996).

3.1.3 Global sea surface temperature

The sea surface temperature (SST) is another data set that was used in the present study. This data set was obtained similarly, to wind data from the National Centers for Environmental Prediction/ National Center for Atmospheric Research (NCEP/NCAR) reanalysis, and was downloaded from their website *https://www.esrl.noaa.gov/psd/*, spanning from 1961 to 2018. Detailed dataset description can be found in Kalnay *et al.* (1996).

3.1.4 Global mean sea level pressure

Global Mean Sea Level Pressure (MSLP) data set was also downloaded via the National Centers for Environmental Prediction / National Center for Atmospheric Research (NCEP/NCAR) reanalysis, and was downloaded from their website *https://www.esrl.noaa.gov/psd/*. This dataset spans from 1961 to 2018. Detailed dataset description can also be found in Kalnay *et al.* (1996).

3.2 Methodology

Methods implemented in this study are discussed in this section. The methods encompass Pearson Correlation, Standardized Anomaly Indices, Principal Component Regression (PCR) Modelling.

3.2.1 Data quality control and estimation of data

Too often station records suffer from missing data. A number of data estimation methods have been implemented by diverse researchers. Some of these methods include arithmetic mean, isopleths, inverse distance-weighting, thiessen polygon, correlation and regression etc. For the current study the correlation method was used with a restriction of up to 10% of the total data for a given area to be estimated. Under this method a neighbouring station that is significantly correlated with the station (A) with missing records is first identified. This is achieved through correlating the station A with neighbouring stations. The correlation equation is given in equation 4. The station (B) with the greatest positive correlation coefficient and consistent records is then used to estimate the missing value (X) of station A through the following equation:

$$X_{Ai} = \frac{\overline{X_A}}{\overline{X_B}} X_{Bi} \quad \ldots\ldots 1$$

Where;

X_{Ai} is missing record for station A for the i^{th} year

X_{Bi} is rainfall record for station B with consistent records for the i^{th} year

$\overline{X_A}$ is the long term average for station A based on available records

$\overline{X_B}$ is the long term average for station B based on the period of available records in A.

After filling in missing values to form a complete data set the next step was to test whether the data was consistent or not. The single mass curve was used for this purpose. The single mass curve method involves plotting accumulated parameters such as rainfall, against time. A straight line indicates consistent data while a change in slope lasting for more than five consecutive points indicates inconsistency of the data set.

3.2.2 Standardisation of data

The region of southern Africa is characterised by varied topography which results in average rainfall amounts that differ significantly from one place to another. This implies that variations in means and standard deviations in space are great. This therefore warrants the use of standardised values. In essence, the standardisation of the data ensures that the influences of spread and location are removed. The standardised anomalies (z) are simply calculated by first subtracting the sample mean from each record value and then divide the result by the standard deviation of the sample. The standardised anomalies hence have a mean of 0 and standard deviation of 1. The equation for calculating the standardised anomalies is given by:

$$z = \frac{X - \bar{X}}{S_X} = \frac{X'}{S_X} \quad \ldots\ldots 2$$

Where;

z is the standardised anomaly

X is the observed data

$\bar{X}$ is the mean

X' is the anomaly and

S_X is the standard deviation.

3.2.3 Principal Component Analysis (PCA)

Principal component analysis (PCA) is a statistical technique used to reduce a data set containing a large number of variables into one that contains a smaller number of variables, with the smaller number being the most important in the representation of the information which the variables contain. The smaller numbers of variables resulting from the PCA are linear combinations of variables in the original data set and are uncorrelated to each other. In addition to reducing large data sets, the PCA also finds patterns in the data and retains the maximum variability contained in the original data. A number of researchers have also used the PCA method to study intricate relationships among numerous variables (Ininda, 1994; Nyakwada, 2009). A good and easy to

understand mathematical description of PCA/ EOF (Empirical Orthogonal Function) analysis is given by Wilks (2006). The mathematical basis of the PCA is given in equation 3.

$$Z_j = \sum_{k=1}^{m} a_{jk} F_k \qquad ; (j = 1,2,3, \ldots, n) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 3$$

Where:

Z_j = standardised variable j.

$a_{jk}F_k$ = standardised multiregression coefficient of variable *j* on factor *k* (factor loading)

F_k = hypothetical factor *k* (principal component)

m = number of common factors

n = number of variables

There are advantages in using the EOF/PCA technique in redundancy, multi-collinearity and noise removal (Wilks, 2006) in a given large data set which include:

- The ability to expose temporal as well as the spatial patterns which might be related to known physical processes involved in the generation of the data.
- The capability to reduce the large amount of data by substituting the measured variables together with their inter-connected variables by smaller number of uncorrelated variables. Uncorrelated implies that each is orthogonal to the other. Since the PCs come from data, and often the data is observations, it implies different observations give different PCs, and hence the term empirical orthogonal functions (EOFs).
- No obligation for the observation points to be equidistant from each other as compared to other orthogonal functions.
- The orthogonality in time of the PCs.

The PCA can be done based on the correlation matrix or variance-covariance matrix. The PCA based on covariance matrix is more suitable for analysis whose main purpose is to isolate strongest variations in the data set. On the other hand if the variables are measured in different units the PCA based on correlation matrix is of preference (Wilks, 2006). In the present study the PCA used was based on the variance-covariance matrix.

Due to the fact that PCA is used with the major purpose to obtain new variables that represent the greatest fraction of the variance in the original data, the redundancy of some variability in individual variables is not a surprise. Hence, it is crucial to select principal components which efficiently portray the greatest part of the variance of the original data as well as the fundamental physical processes.

Methods for selecting significant components:

(a) Kaiser criterion – according to this selection criterion all components with eigenvalues greater than one are retained (Kaiser, 1960). One shortfall of this criterion is that it can lead to over factoring.
(b) Scree plot – for this method the component numbers are plotted against the eigenvalues. The cut-off is made just before the graph starts to level off. The disadvantage of this criterion is its subjectivity.
(c) Proportionality of variance accounted for – for this criterion a percentage of variance in the data set is fixed as a criterion for retaining significant components (Wilks, 2006).

For the current study the climate Predictability Tool (CPT) was used for the purpose of retaining a few major components since the other methods retained a large number of components. The CPT is developed by the International Research Institute for Climate and Society and is tailored to produce seasonal climate forecasts using data sets like SST, global winds and dynamic model output among others. Cross-validated forecasts produced using the first principal component, are correlated to the rainfall anomalies and the correlation coefficient is used to calculate a goodness index. Next, the second principal component is added to make the cross-validated forecasts which are again correlated with the rainfall anomalies and a new goodness index is calculated. Subsequent principal components are added one at a time following the same procedure with the new goodness index produced at each stage compared to the previous one. The best goodness

index is retained under the optimum column together with the corresponding number of principal modes.

3.2.4 Correlation analysis

Correlation describes the linear relationship between a pair of variables. It is bounded by -1 and +1. A value of +1 indicates a perfect linear positive relationship while a value of -1 indicates a perfect linear negative relationship. Zero indicates no correlation at all. The formula for calculating the correlation coefficient is given by the equation:

$$r_{XY} = \frac{\sum_i^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i^n (X_i - \bar{X})^2 (Y_i - \bar{Y})^2}} \quad \ldots\ldots 4$$

Where,

n= total number of observations

$\bar{X}$= Mean for variable X

$\bar{Y}$= Mean for variable Y

X_i along with Y_i are independent variables.

Testing the significance of correlation

The correlation coefficient was tested for significance using the student t-test with the degree of freedom *n-2*. The following equation for that purpose is given by;

$$t_{n-2} = r\sqrt{\frac{n-2}{1-r^2}} \quad \ldots\ldots 5$$

Where;

t-is the value of the student-t-test

n- number of observations (data points)

r- correlation coefficient being tested.

Facts of the student-t test are found in many publications (like WMO, 1983). To draw conclusions about the significance of the correlation the null hypothesis (H_0: r = 0) was formed which simply states that the sample correlation is not significantly different from zero. The alternate hypothesis (H_a: $r \neq 0$) which means that the correlation coefficient is different from zero was also formulated. The p-value method was used to assess the significance of the correlation coefficient. The p-value based on a significance level ($\alpha = 0.05$) was used to decide whether to reject the null hypothesis or fail to reject the null hypothesis. A p-value less than 0.05, leads to the decision that the correlation coefficient is significantly different from zero hence, statistically significant. A p-value greater than 0.05 means that the correlation coefficient is not significantly different from zero, thus the correlation is not statistically significant.

3.2.5 Regression analysis

Regression is used to find and fit a linear equation that relate variables and can be simple linear or multiple as described in the following sub-sections.

3.2.5.1 Simple linear regression

Simple linear regression is used for finding a straight line that best-fit through pairs of related variables. This line of best-fit is referred to as a regression line. It allows for computation of values for one variable from known values of another variable. Simple linear regression can only be used when one has two continuous variables – an independent variable and a dependent variable. The independent variable is called a predictor and the dependent variable is called a predictand. The equation for a simple linear regression is given by equation 6:

$$\hat{Y} = a + bX \quad \ldots\ldots 6$$

Where;

$\hat{Y}$ is the dependent variable

a is the regression constant,

b is the slope of the equation and

X is the independent variable.

3.2.5.2 Multiple linear regression

Multiple linear regression (MLR) is used to determine a mathematical relationship between a predictand and more than one predictor. For a relationship between a dependent variable Y and some k independent variables ($X_1, X_2, ..., X_k$), a multiple regression equation for the prediction of Y may be written as follows:

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + ... + b_kX_k \quad \text{..... 7}$$

Where;

b_0 is the constant of regression

b_k is the coefficient of regression and

X_k is the principal component (predictor).

Not all potential predictors, however, are required in the regression equation (model) as too many predictors can lead to model over-fitting. A statistical selection procedure is thus required to trim the number of predictors. The procedure for this purpose is stepwise regression. This procedure automatically decides which predictors are the best in order of their importance. There are two types of stepwise multiple regressions that can be used namely, forward selection and backward elimination. The forward selection technique was used in the study. This technique involves beginning without any variable in the model, then testing and adding any variable that improves the model the most based on a chosen model comparison criterion. Forward selection is continued provided no more predictors with $p \leq 0.05$ could be included. Predictors that are significantly correlated ($p \leq 0.05$) to any of the preceding chosen predictors are excluded from the model, to curtail multicollinearity.

3.2.6 Validation of statistical models

Long-lead forecasts have an inherent uncertainty thus; the forecasts are commonly represented probabilistically (Goddard *et al.*, 2001). For the study the Relative Operating Characteristic scores (ROC) were used to assess the performance of the statistical models. The Relative Operating Characteristics have been used in a number of studies to verify probabilistic seasonal forecasts (e.g. Landman and Beraki, 2012; Hall *et al*., 2017; Phakula, 2016). ROC scores reveal

the forecast system's ability to distinguish between individual categories and they provide the forecast skill for each individual category. ROC scores are sufficiently discussed in Mason and Graham (1999). However, the main concerns to the society are the extreme climate occurrence. It follows that only results for extreme categories (below-normal and above-normal) are reported in the study. ROC scores above 0.5 (50%) indicate a forecast system that is skilful. Less or no helpful information is implied by ROC scores equivalent to 0.5 while those less than 0.5 imply a negative skill of the forecast system.

3.2.7 Wind Plots

Plots for zonal mean wind circulation for the September-October-November (SON) season were plotted for each of the levels (1000, 850, 700 and 200 mb) using plotting tools available in the National Oceanic and Atmospheric Administration (NOAA) website. The plots covered the period used for analysis in the study.

CHAPTER FOUR

4.0 RESULTS AND DISCUSSION

Results and discussion of the results obtained from the study are presented in this chapter.

4.1 Records estimation and data consistency test

The test for data consistency presented in chapter three was used to check if the data (after filling missing records) were consistent throughout the period of analysis. Figure 5 shows the results of the single mass curve for the station of Mbabane. The trend line is straight indicating that the data records are consistent in time. Similar results were obtained for all the stations that were used in the analysis.

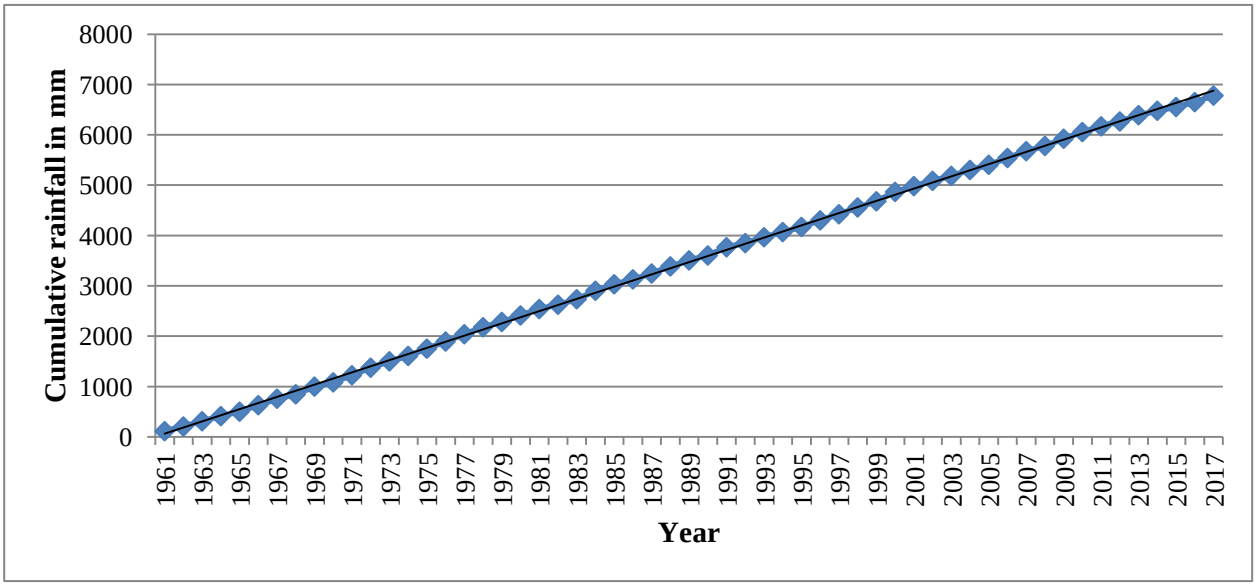


Figure 5: Single mass curve for Mbabane.

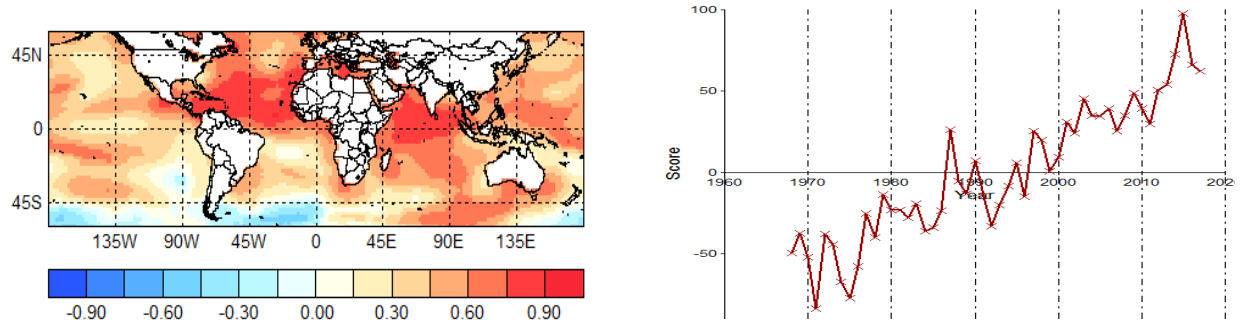
4.2 Major spatial and temporal modes of variability in the potential predictor variables

This section introduces the major space and time modes representing variability in the potential predictor variables hence, influence rainfall in Eswatini.

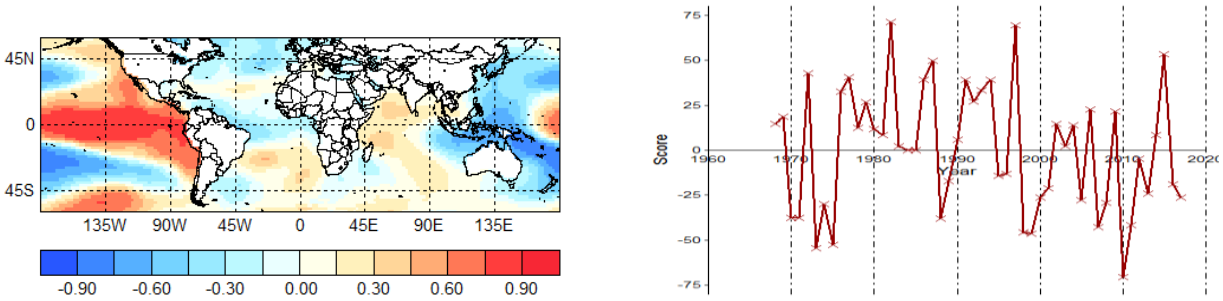
4.2.1 Major modes of variability in SON averaged SSTs

The first major spatial mode of variability in SSTs having an influence on DJF rainfall in Eswatini as shown in figure 6 (a), represents a general warming in the global tropical oceans. The second spatial mode as shown in part (b) of figure 6 shows strong positive loadings (associated with warming SSTs) over the eastern Pacific and strong negative loadings (associated with cooling) over the western Pacific Ocean. This clearly represents an ENSO pattern and to some extent a positive Indian Ocean dipole (IOD). This second spatial mode or ENSO pattern has also been reported by Chavez and Messie (2011) when studying global modes of sea surface temperature variability in relation to regional climate indices. Lastly, the third spatial mode as shown in figure 6 (c) represents, a concentration of positive loadings associated with warming in south Atlantic and north Indian oceans.

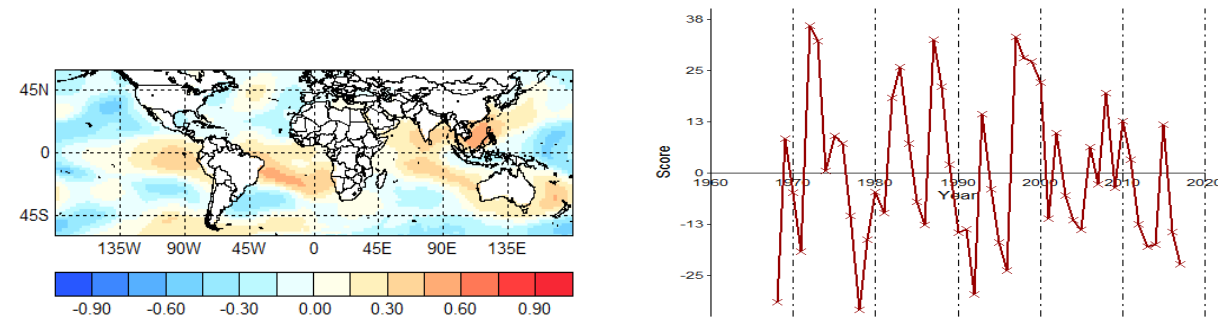
On the other hand, the first major temporal mode of variability in figure 6 (a) right panel, shows a positive trend. This means there is a general upward trend in the warming of the global oceans associated with the first spatial mode of variability. The second principal mode in the right panel of figure 6 (b) shows some of the strong ENSO (El Niño) episodes that occurred in the years 1982, 1998 and 2016. El Niños are usually associated with droughts in southern Africa and Eswatini in particular. The third principal mode, shown on the right panel of figure 6 (c) shows a decrease in its amplitudes between positive phase and negative phase over the years. This mode explains 5% of the variability in SSTs over the global oceans. The first and second major SST modes explain respectively, 26% and 17% of the variability in the sea surface temperatures.



(a)



(b)



(c)

Figure 6: First 3 spatial modes of variability in SST together with their corresponding principal components; (a) first mode, (b) second mode and (c) third mode**.**

4.2.2 Major modes of variability in SON averaged MSLP

Figures 7 (a) to (c), show the major modes of variability in Mean Sea Level Pressure (MSLP). The first major spatial mode of variability in figure 7 (a) shows strong positive loadings over the tropical Indian, Atlantic Oceans and land areas between 45 °N/S of the equator while over the Pacific Ocean it is dominated by negative loadings. Its corresponding temporal mode on the right panel reveals two major negative peaks in 1970 and 2010 while in between these years it has been dominated by positive peaks. The second major spatial mode of variability (shown in figure 7 (b)) in MSLP is dominated by weaker (as compared to the first mode) positive loadings over the areas between 45 °N and 45 °S of the equator. Its corresponding principal component reveals a sharp upward trend before 1980 and a gentle downward trend thereafter. The third major spatial mode in figure 7 (c), shows negative loadings over the Atlantic Ocean, Africa north of the equator, North and South America. Positive loadings are concentrated over the western Pacific and over Australia. The corresponding third major temporal mode shows an increasing trend with one large positive peak in 1997.

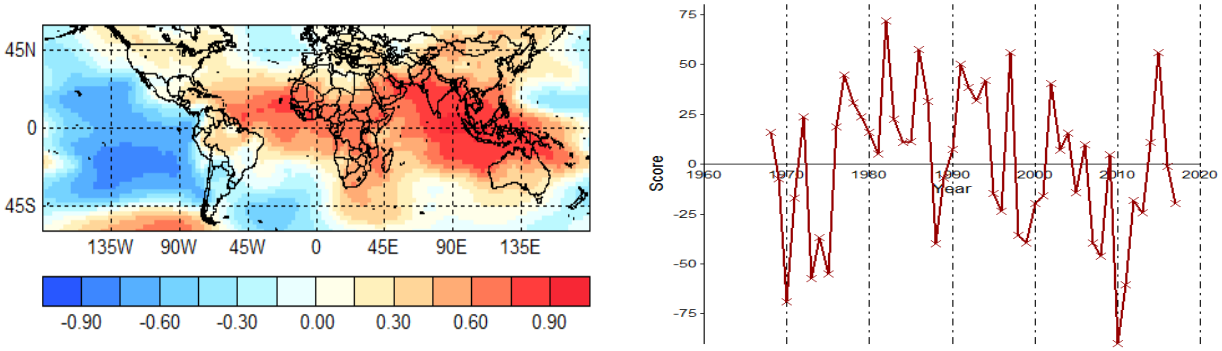


Figure 7 a: First spatial mode of variability in MSLP during SON together with its corresponding principal component.

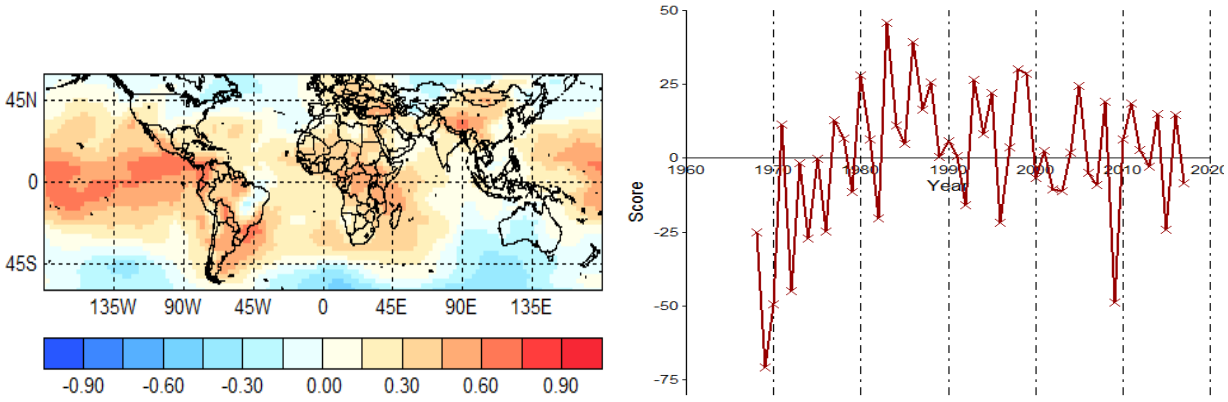


Figure 7 b: Second spatial mode of variability in MSLP during SON together with its corresponding principal component.

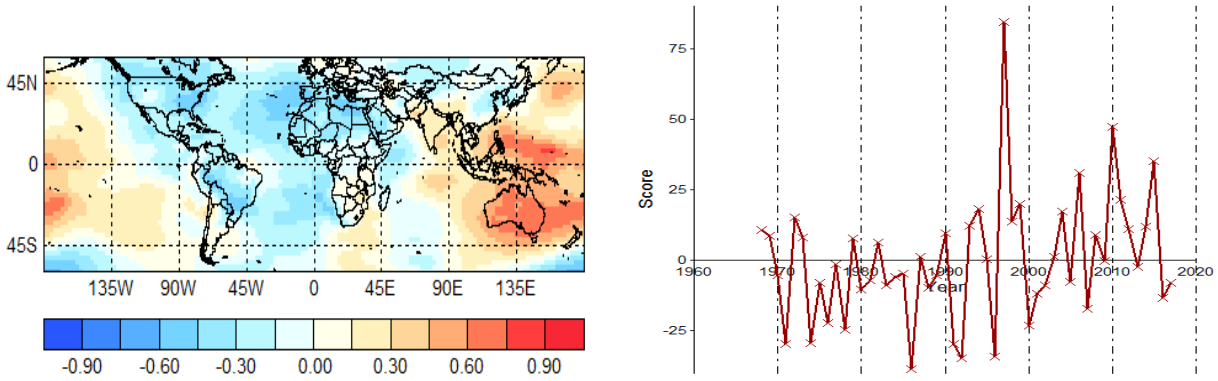


Figure 7 c: Third spatial mode of variability in MSLP during SON together with its corresponding principal component.

4.2.3 Major modes of variability in SON averaged zonal wind at 1000 mb level

The major modes of variability in zonal wind at the 1000 mb level are shown in figure 8 (a) to (e). The first spatial mode in figure 8 (a) shows strong positive loadings mainly over the Pacific Ocean. The tropical Indian Ocean is dominated by negative loadings. The corresponding first principal component reveals a downward trend. The second major spatial mode in figure 8 (b) shows a concentration of positive loadings over the equatorial Pacific Ocean, equatorial Atlantic and Indian Ocean below the equator. Negative loadings can be seen more clearly over Australia. The corresponding second principal component shows that the second spatial mode has a positive trend over the years.

In the third major spatial mode displayed in figure 8 (c) the Atlantic Ocean is dominated by negative loadings. Over the Indian Ocean positive loadings can be seen along and over the equator while below the equator negative loadings dominate that region. In the Pacific Ocean positive loadings concentrate below the equator next to the South American continent while part of the Ocean over New Zealand negative loadings dominate. The corresponding principal component shows that the third spatial loading was in negative phase before 1975 and changing phases thereafter. This principal component displays on average an upward trend. The fourth major spatial component as displayed in figure 8 (d) shows positive loadings over much of the Indian Ocean, negative loadings over a large part of the Atlantic Ocean and negative loadings along the equatorial Pacific Ocean. According to the corresponding fourth principal component the fourth spatial mode had very strong positive phases in the years 1968, 1976 and 2016 while strong negative phases occurred in 1972 and 1985. Figure 8 (e) displays the fifth major spatial mode which shows positive loadings dominating along the 45 °S latitude. The corresponding fifth principal component has a positive trend.

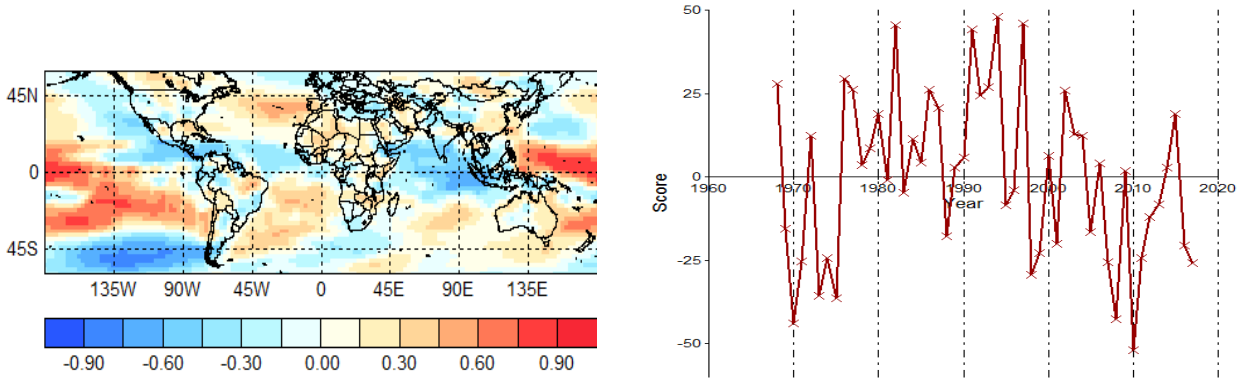


Figure 8 a: First spatial mode of variability in zonal wind in the 1000 mb pressure level during SON together with its corresponding principal component.

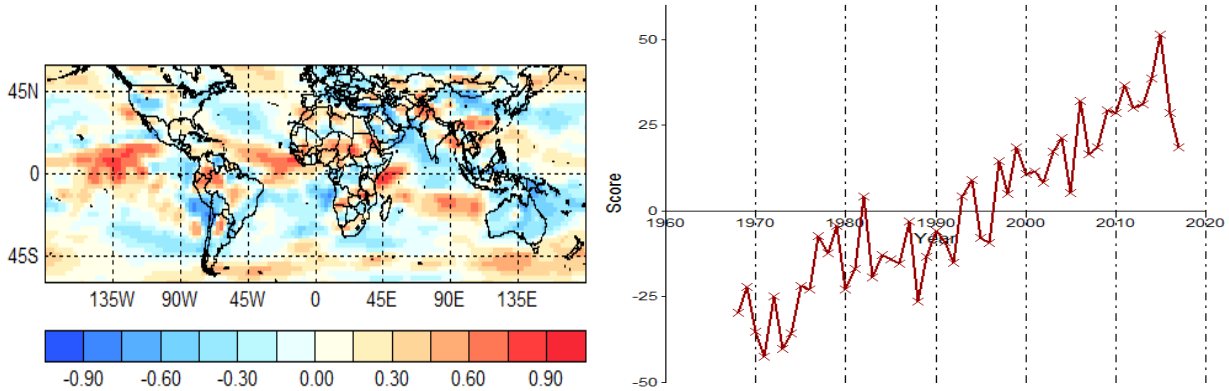


Figure 8 b: Second spatial mode of variability in zonal wind in the 1000 mb pressure level during SON together with its corresponding principal component.

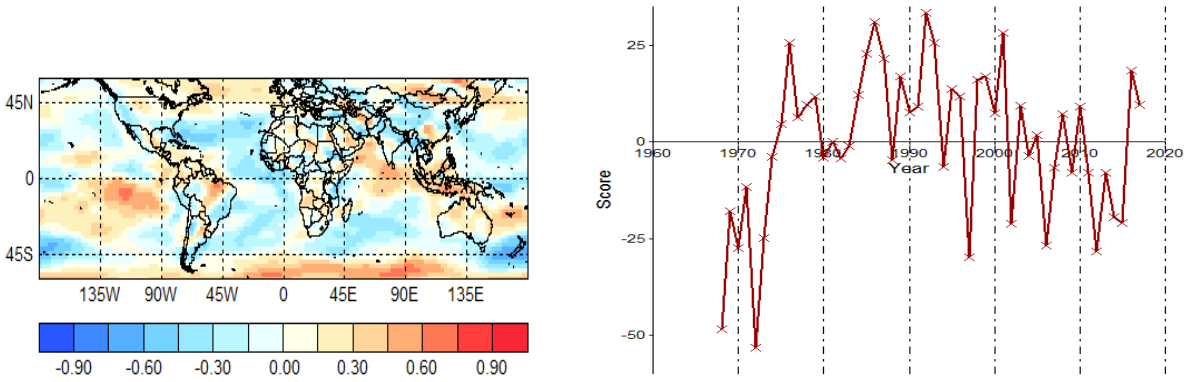


Figure 8 c: Third spatial mode of variability in zonal wind in the 1000 mb pressure level during SON together with its corresponding principal component.

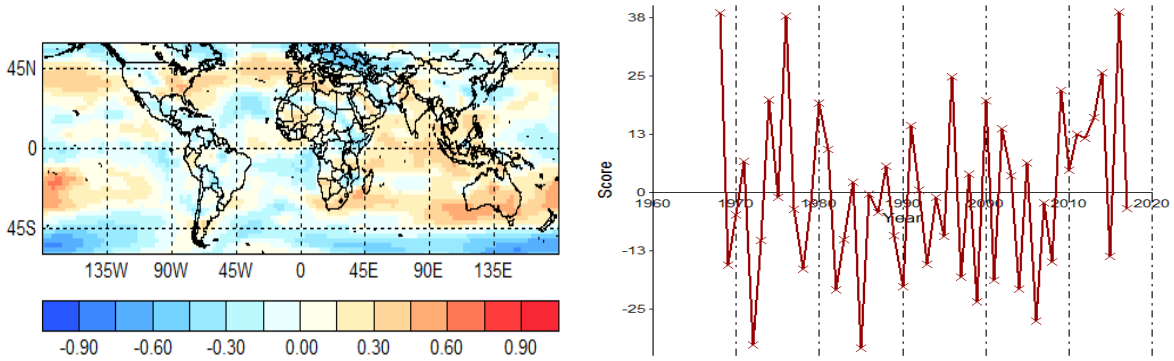


Figure 8 d: Fourth spatial mode of variability in zonal wind in the 1000 mb pressure level during SON together with its corresponding principal component.

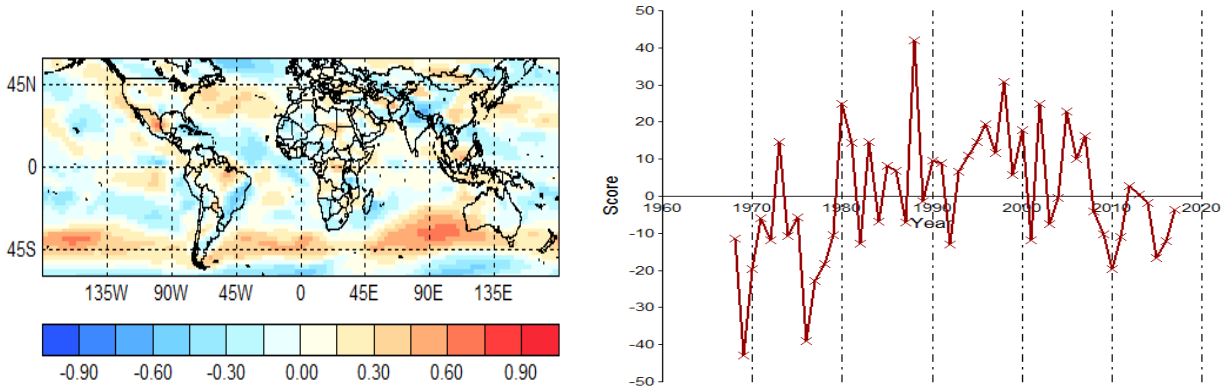


Figure 8 e: Fifth spatial mode of variability in zonal wind in the 1000 mb pressure level during SON together with its corresponding principal component.

4.2.4 Major modes of variability in SON averaged zonal wind at 850 mb level

The SON averaged major modes of variability in zonal wind at the 850 mb pressure level are shown in figures 9 (a) to (c). The first major spatial mode displayed in figure 9 (a) shows negative loadings along equatorial Indian Ocean and Atlantic Ocean. The positive spatial loadings are concentrated over the Pacific Ocean. The corresponding principal component has a negative trend. Figure 9 (b) shows the second major spatial mode which has strong positive loadings along the equatorial Atlantic and Pacific Oceans while the equatorial Indian Ocean is dominated by strong negative loadings. The corresponding second principal component shows a strong positive trend of the second major spatial mode meaning that this spatial mode is turning more into its positive phase. The third major spatial mode of variability is displayed in figure 9 (c) and it reveals positive loadings over much of the Atlantic Ocean and Indian Ocean below the equator. The Pacific Ocean is dominated by negative loadings over the eastern part of the basin. The corresponding third principal component shows a negative trend and the highest positive phase of the spatial mode which occurred in 1968.

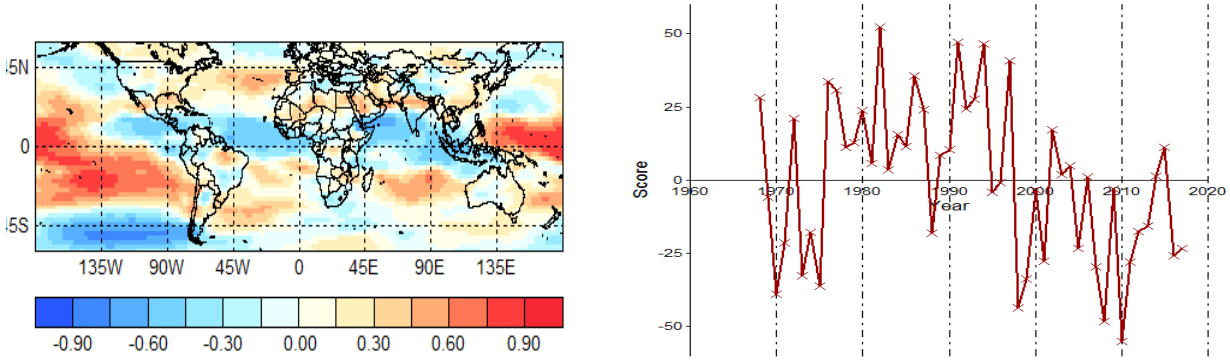


Figure 9 a: First spatial mode of variability in zonal wind in the 850 mb pressure level during SON together with its corresponding principal component.

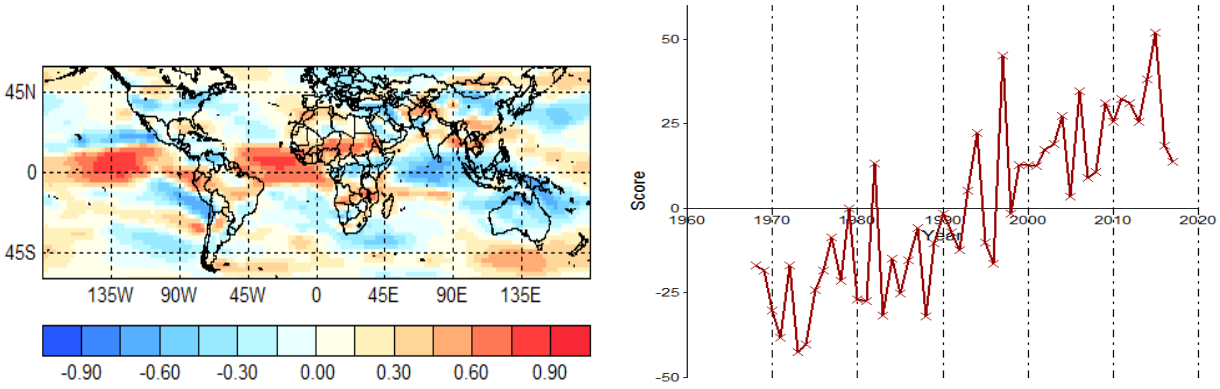


Figure 9 b: The second spatial mode of variability in zonal wind in the 850 mb pressure level during SON together with the corresponding principal component.

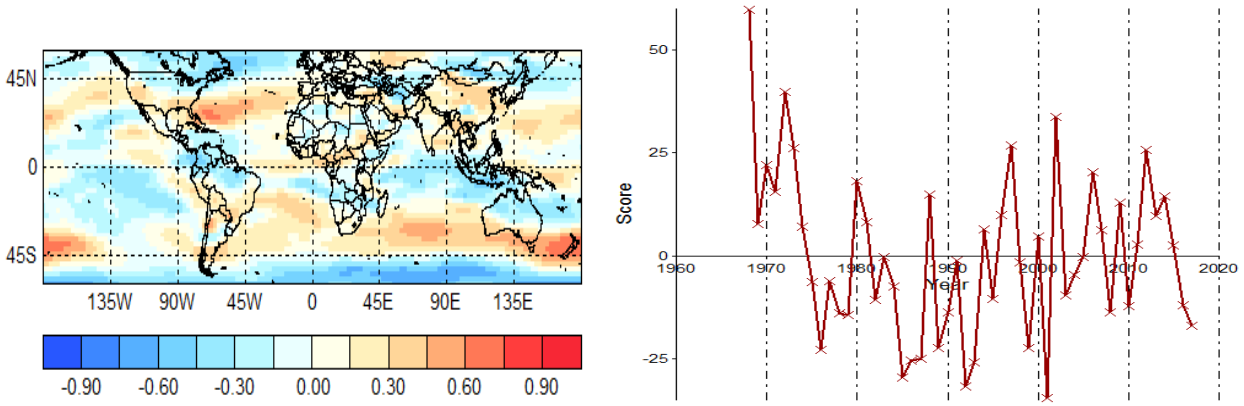


Figure 9 c: The third spatial mode of variability in zonal wind in the 850 mb pressure level during SON together with the corresponding principal component.

4.2.5 Major modes of variability in SON averaged zonal wind at 700 mb level

Figures 10 (a) to (e) display the SON averaged major modes of variability in zonal wind at 700 mb pressure level. The first major spatial mode as shown in figure 10 (a) is similar in structure to the first major mode for zonal wind at 850 mb shown in figure 9 (a). Likewise the second major mode shown is figure 10 (b) is similar in spatial and temporal structure to the second major mode for zonal wind at 850 mb shown in figure 9 (b). The third major spatial mode (as shown in figure 10 (c)) displays weak but positive loadings along the equatorial oceans. Stronger positive loadings are observed over the northern part of Australia while the southern part of Australia into the Indian Ocean shows negative loadings. The corresponding third principal component shows a positive linear trend.

Figure 10 (d) shows the fourth major spatial mode which is almost similar to the previous third spatial mode. The corresponding fourth principal component has a positive linear trend. The fifth spatial mode as shown in figure 10 (e) reveals stronger positive loadings over southern and central Africa. Negative loadings can be seen over eastern equatorial Pacific Ocean and along the 45 °S latitude. The corresponding fifth principal component shows a generally decreasing linear trend.

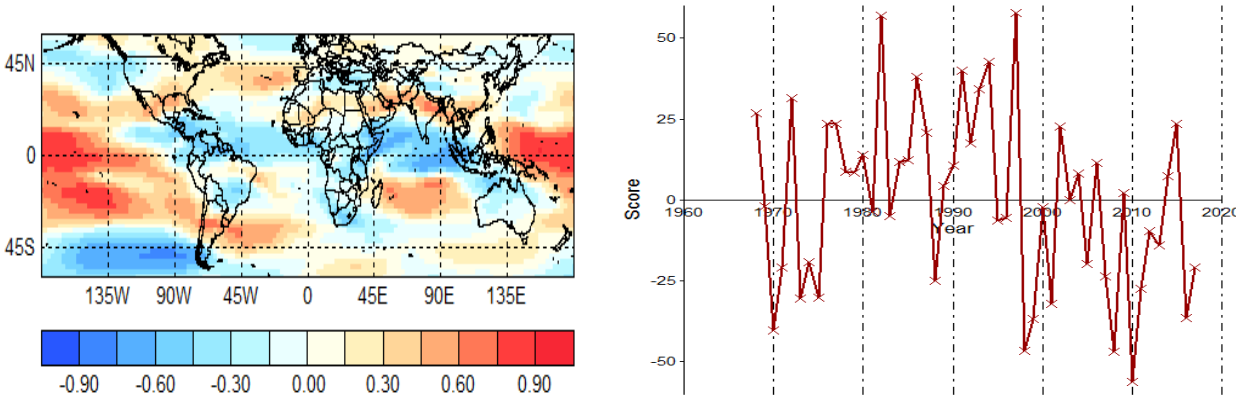


Figure 10 a: The first spatial mode of variability in zonal wind in the 700 mb pressure level during SON together with the corresponding principal component.

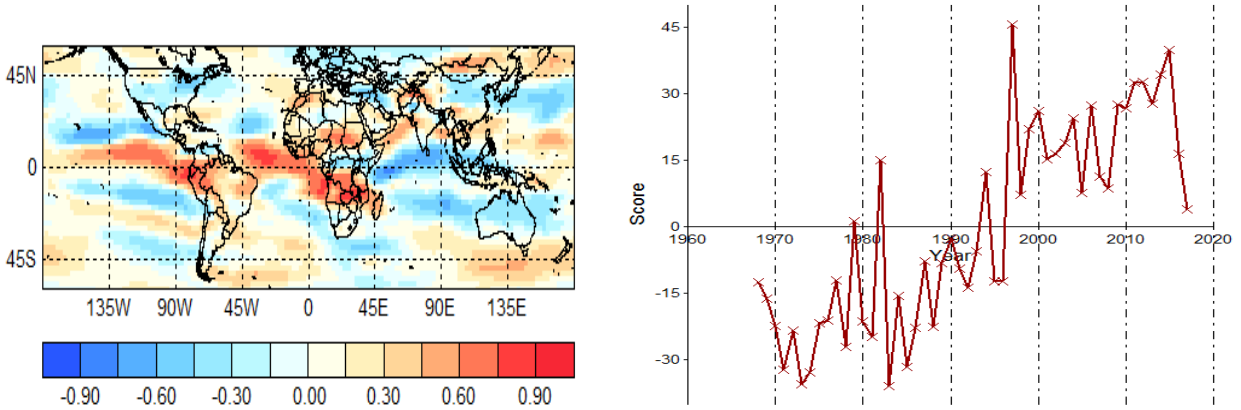


Figure 10 b: The second spatial mode of variability in zonal wind in the 700 mb pressure level during SON together with the corresponding principal component.

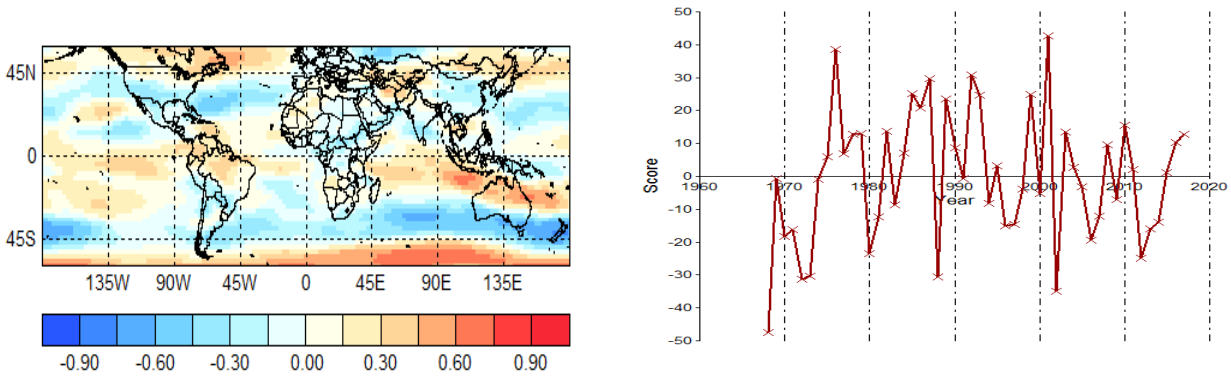


Figure 10 c: The third spatial mode of variability in zonal wind in the 700 mb pressure level during SON together with the corresponding principal component.

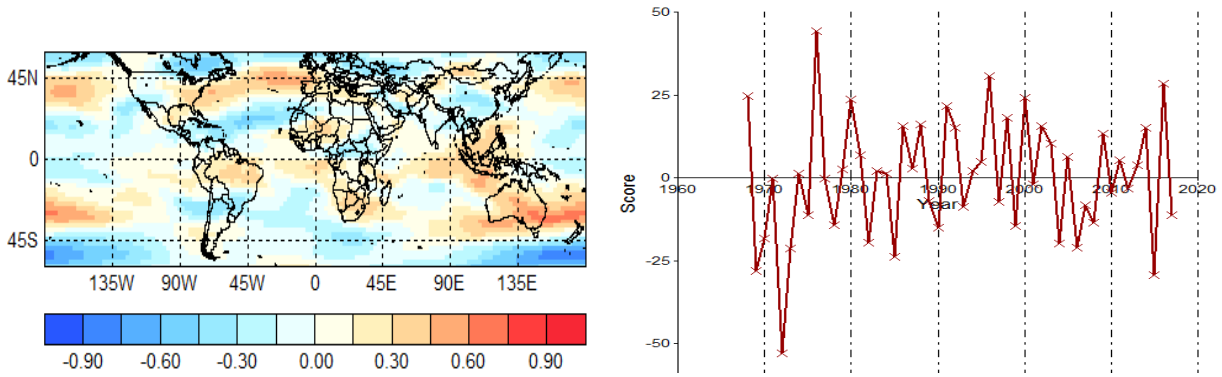


Figure 10 d: The fourth spatial mode of variability in zonal wind in the 700 mb pressure level during SON together with the corresponding principal component.

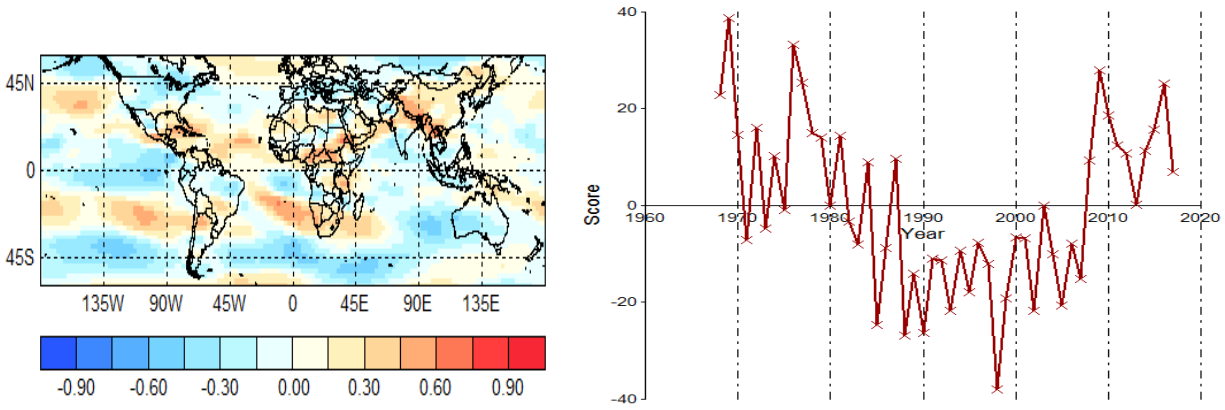


Figure 10 e: The fifth spatial mode of variability in zonal wind in the 700 mb pressure level during SON together with the corresponding principal component.

4.2.6 Major modes of variability in SON averaged zonal wind at 200 mb level

Figure 11 displays the major mode of variability in zonal wind (averaged over September-November) at the 200 mb pressure level. It can be noticed that over most of tropical Indian Ocean and equatorial Atlantic, the loadings are negative while the equatorial Pacific Ocean possesses strong positive loadings. This is similar to mode 1 loading in the 850 mb level, but larger in spatial extent. The corresponding principal component shows no trend.

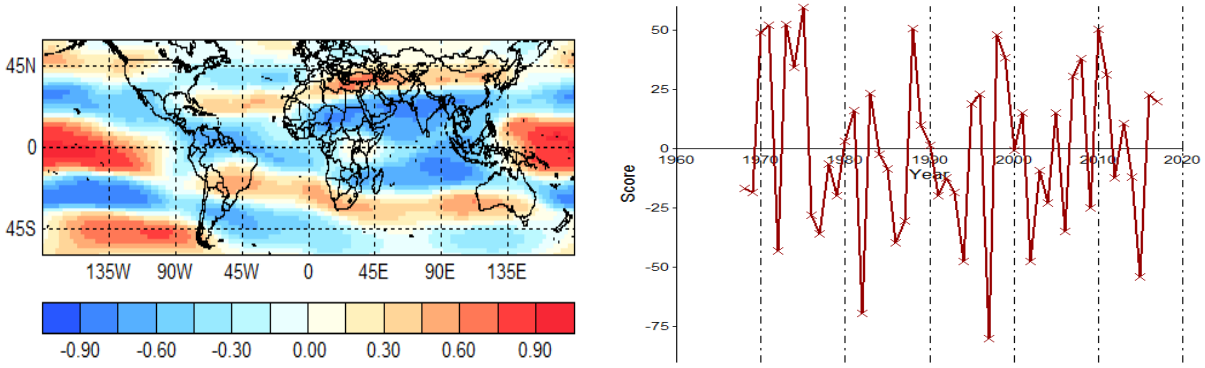


Figure 11: The first spatial mode of variability in zonal wind in the 200 mb pressure level during SON together with the corresponding principal component.

4.3 Wind circulation

The circulations of winds during SON season at the different levels were plotted. These are given in figures 12 (a) to (d). In the 1000 mb level westerly winds blow from the south Atlantic into the southern parts of South Africa. On the eastern side in places like Eswatini and Mozambique easterly winds bring in moisture from the southwest Indian Ocean. In the 850 mb level easterly winds can be seen blowing into a large part of southern Africa, north of South Africa from the Indian Ocean. In the 700 mb level (figure 12 (c)) westerly winds dominate from about 20 °S latitude towards the southern coastal areas of the sub-region. The southern Africa region is generally dominated by westerly winds in the 200 mb level and seen in figure 12 (d).

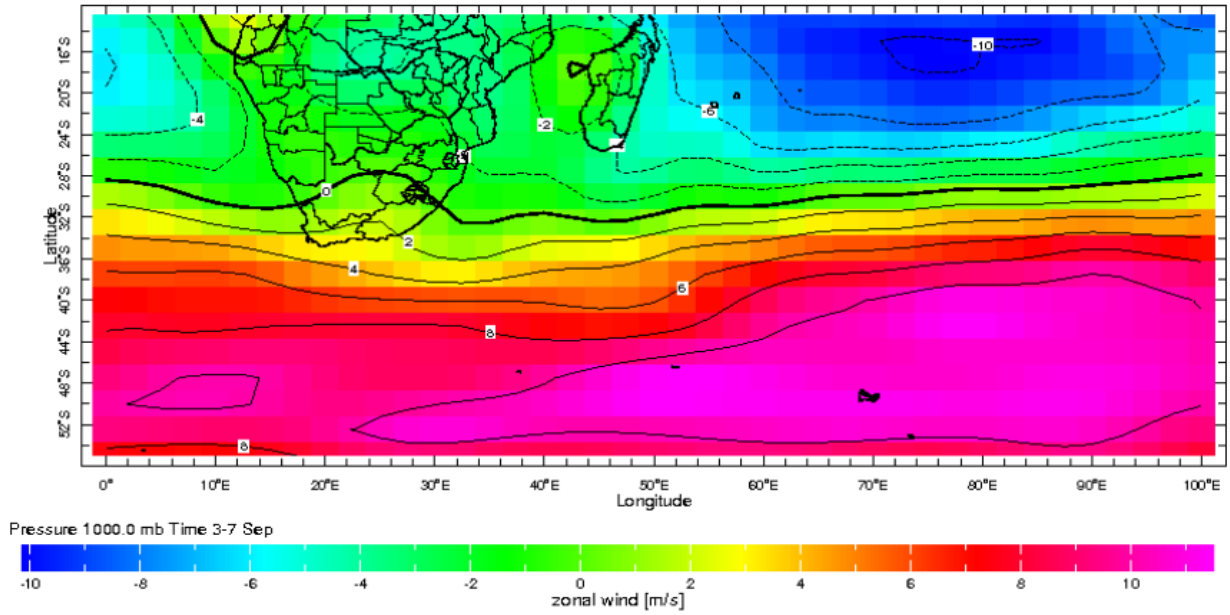


Figure 12 a: Climatological mean wind circulation during SON for 1000 mb pressure level.

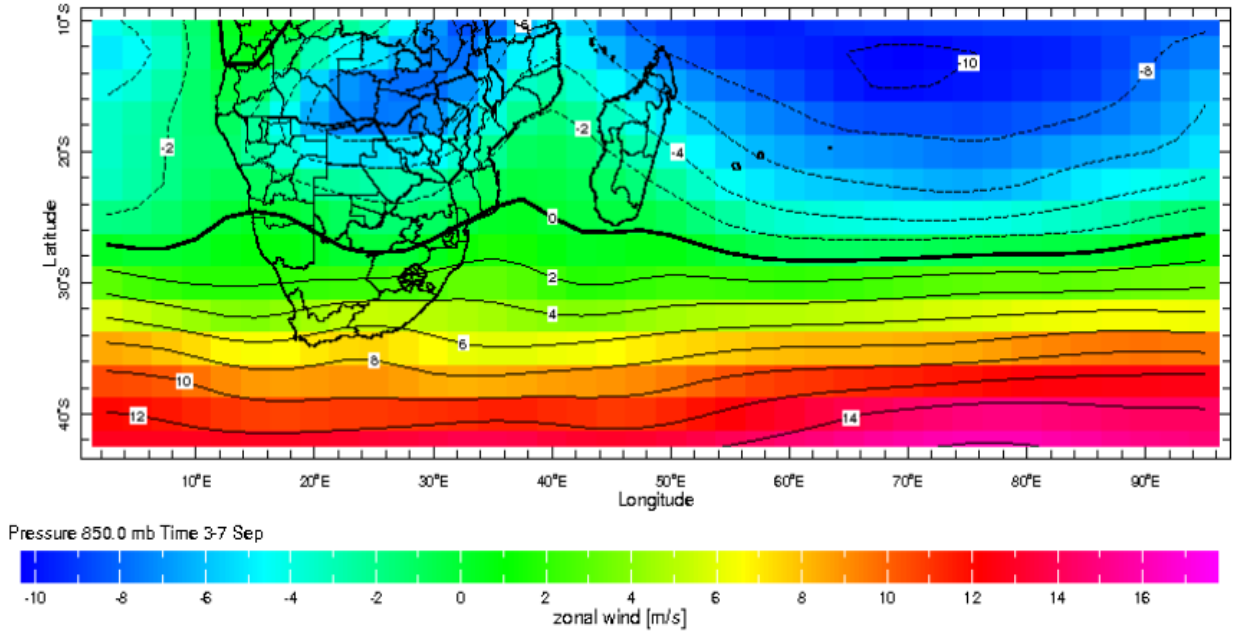


Figure 12 b: Climatological mean wind circulation during SON for 850 mb pressure level.

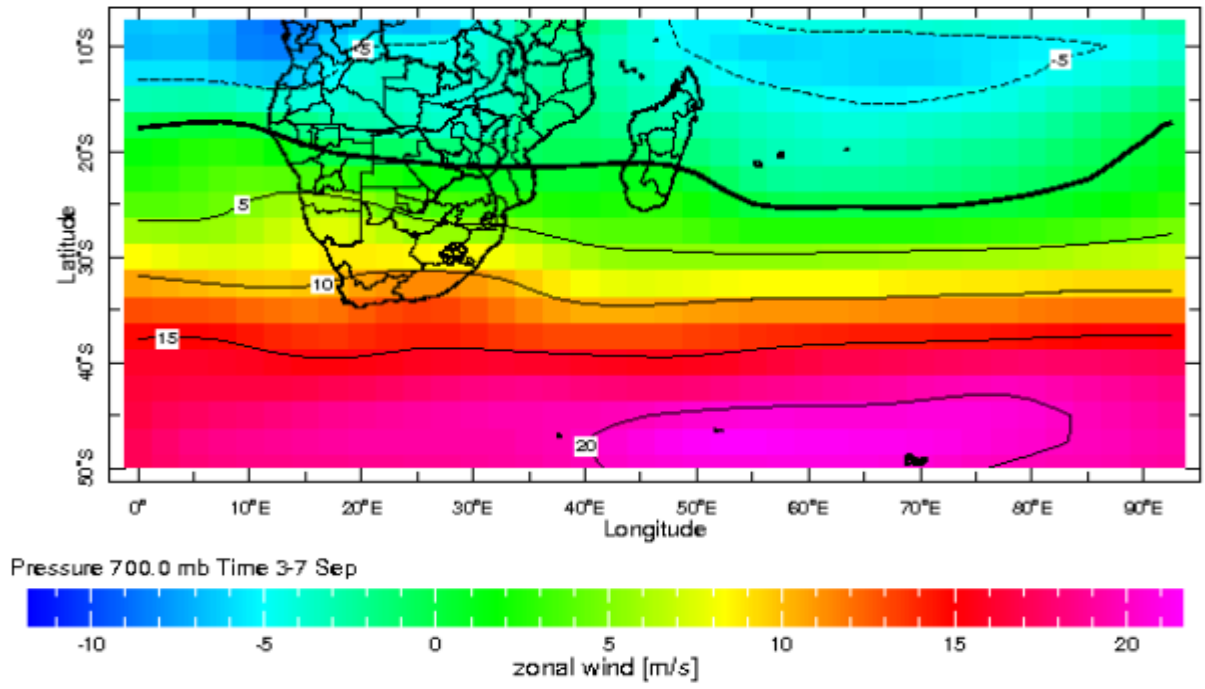


Figure 12 c: Climatological mean wind circulation during SON for 700 mb pressure level.

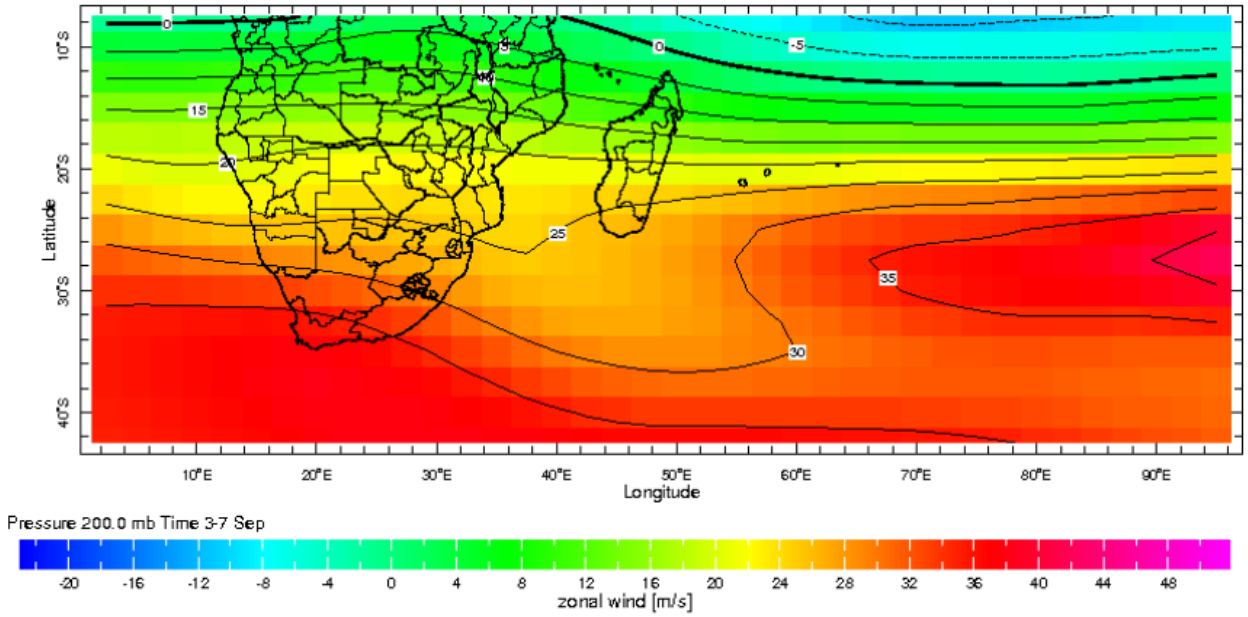


Figure 12 d: Climatological mean wind circulation during SON for 200 mb pressure level.

4.4 Predictability of DJF rainfall

Skill maps for correlations between observed rainfall anomalies and cross-validated predictions for southern Africa rainfall based on GPCC rainfall data set are shown in figure 13. A large part of southern Africa's rainfall as revealed in figure 13 shows positive correlations to SON averaged sea surface temperatures. One easy to identify area of negative correlations is southwestern tip of South Africa, a region that receives most of its rainfall during the austral winter months. Eswatini which was used as a focus of the study is dominated by positive correlations, which exceed 0.6 as evident in Figure 13. For predictive purposes it is important that the correlations be positive and significant and this will be further assessed in the coming sections.

4.4.1 Correlation between December to February (DJF) GPCC rainfall and September to November (SON) averaged SSTs

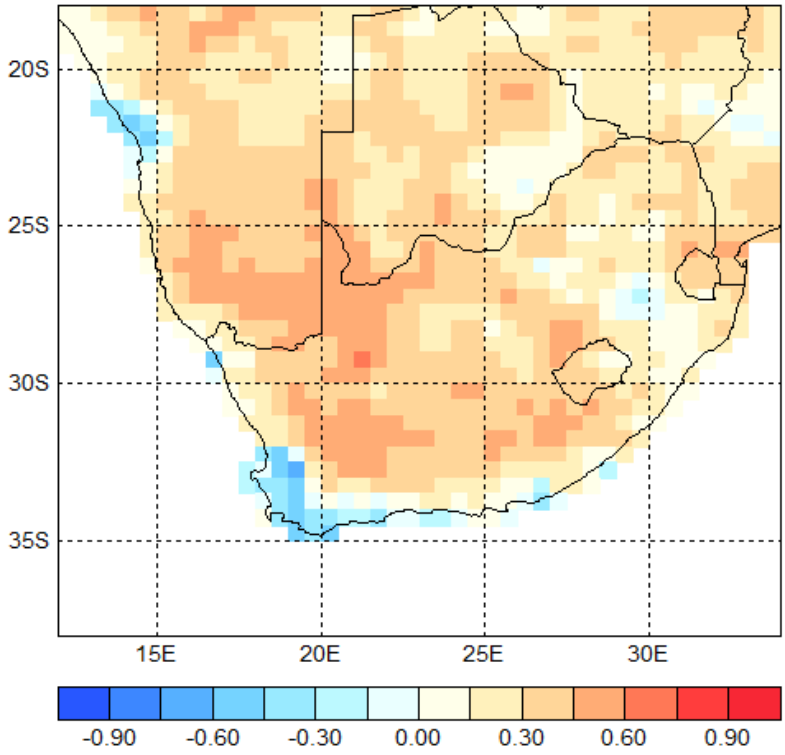


Figure 13: Correlation between cross-validated predictions using SON SSTs and GPCC rainfall.

4.4.2 Development of model to predict DJF rainfall using SON predictors

The first few principal components of each of the six variables as shown in table 3 were later used to develop multiple linear regression models for predicting rainfall anomalies during the DJF season. As displayed in table 3, the number of principal components retained for each ranged between 1 and 5. It can be observed that the first few MSLP and SST modes account for larger variance as compared to the first few wind modes. This can be attributed to the slow evolving state of the SSTs whereas the winds are more variable over time.

Table 3: Optimum number of modes retained for each SON predictor variable with variance explained.

Predictor (SON)	**Optimum Number of modes**	**Mode variance**					**Total variance**
		1	**2**	**3**	**4**	**5**	
SST	3	25.6	17.3	5.0			47.9
MSLP	3	23.2	9.5	8.4			41.1
U1000	5	10.9	9.1	6.6	5.4	4.8	37.0
U850	3	12.5	10.0	6.8			29.3
U700	5	13.0	9.0	6.8	5.8	5.2	39.8
U200	1	20.4					20.4

4.4.3 Correlations between cross-validated predictions and rainfall anomalies

The cross-validated predictions for each SON predictor variable were correlated with DJF station rainfall anomalies and table 4 shows the results. From the correlation results in table 4, figures in red represent significant correlations at 0.05 significance level and three stations have significant positive correlation with each of the six variables used in the study. These stations are namely Mbabane, Malkerns and Matsapha. However, a usable skill is taken to be 0.3 or greater so as to account for at least 10% predictability. SSTs and Zonal wind (U200) at 200 mb had the largest

number of significant correlations with the stations' rainfall used. Zonal wind at 1000 mb pressure level had the least number of significant correlations with Eswatini's rainfall. MSLP at 0.457 correlation with rainfall in Matsapha had the greatest significant correlation coefficient.

Table 4: Correlation between each station's rainfall and the six predictors.

Station	SSTV4	U1000	U850	U700	U200	MSLP
Mbabane	0.398	0.377	0.450	0.402	0.451	0.388
PiggsPeak	-0.018	0.194	0.062	0.242	0.053	-0.159
Malkerns	0.329	0.273	0.331	0.310	0.375	0.285
Mananga	0.275	-0.096	0.107	0.055	0.233	0.079
BigBend	0.372	0.142	0.287	0.194	0.343	0.137
Mpisi	0.156	0.084	0.090	0.091	0.186	0.254
Khubutha	0.030	0.177	0.210	0.244	0.157	-0.056
Matsapha	0.252	0.427	0.274	0.409	0.395	0.457
Nhlangano	0.225	0.222	0.200	0.212	0.264	0.070

4.4.4 Models constructed using SON predictors.

After correlating the rainfall to the six predictor variables, multiple linear regression equations for predicting DJF seasonal rainfall anomalies were developed for each of the three agro-ecological zones representative stations. The major principal components found in table 3 were used as the predictor variables. The forward selection technique was used to select candidate predictors for use in the final equations that were constructed. For the stations of Mbabane and Matsapha the predictability of rainfall is based on MSLP and zonal winds at 1000 and 200 mb pressure levels. In Big Bend, on the other hand most of the predictability came from SSTs.

The predicted DJF rainfall anomalies plotted against observed rainfall anomalies graphs for Mbabane, Matsapha and Big Bend are shown in figures 14 (a), (b) and (c) respectively.

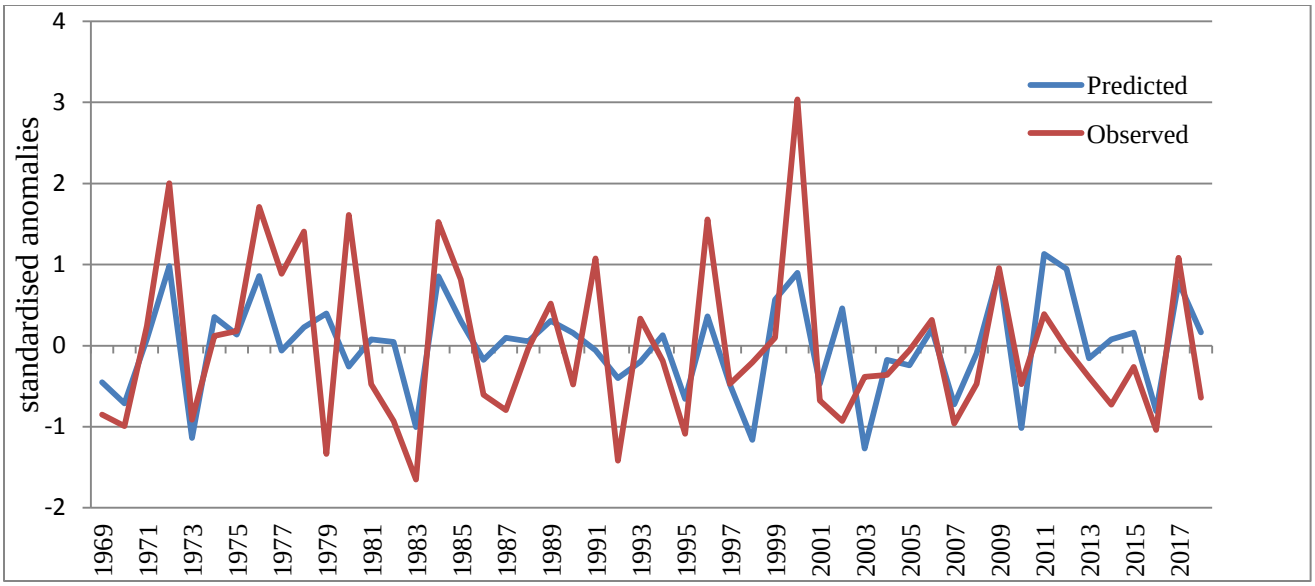


Figure 14 a: Predicted DJF rainfall anomalies versus observed rainfall anomalies for Mbabane.

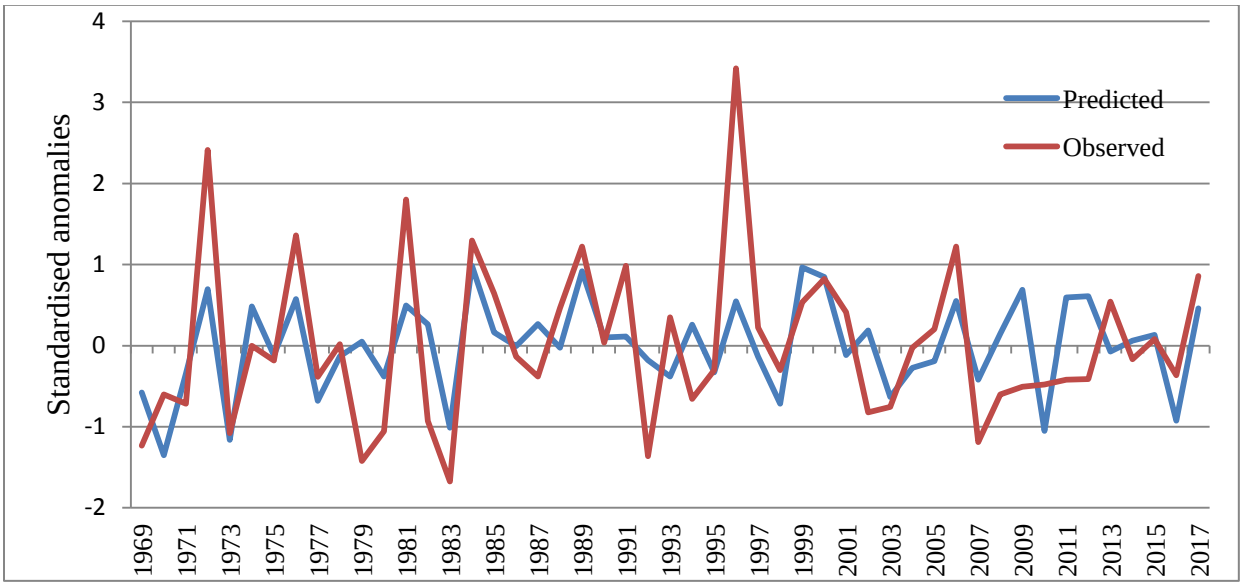


Figure 14 b: Predicted DJF rainfall anomalies total versus observed rainfall anomalies for Matsapha.

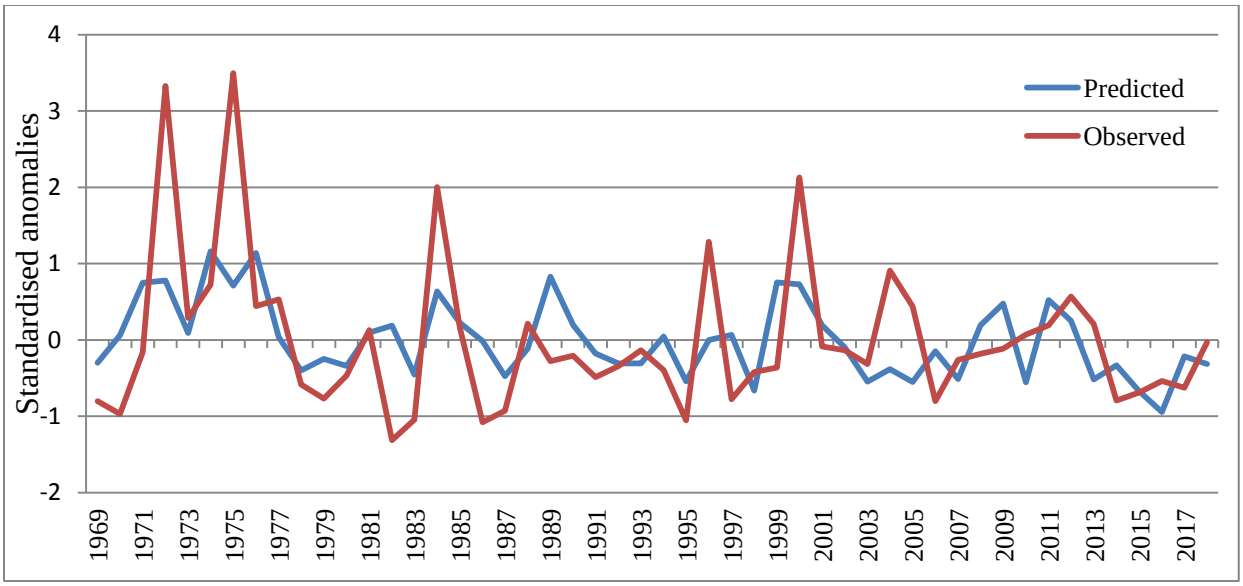


Figure 14 c: Predicted DJF rainfall total anomalies versus observed rainfall anomalies for Big Bend.

Validation of (SON) models

Table 5 shows ROC scores as obtained from the multiple linear regressions. The ROC scores are significant at 0.05 level of significance. Models for Mbabane and Matsapha distinguish better the above-normal category than the below-normal category while for Big Bend the converse is true. The ROC scores for all the rainfall categories were above 0.5 which shows that the models are skilful as they have the ability to differentiate between the extreme categories (below-normal and above-normal).

Table 5: ROC scores of the representative stations for models developed using SON predictors.

Station	**ROC (below-normal)**	**ROC (above-normal)**
Mbabane	0.695	0.772
Matsapha	0.715	0.784
Big Bend	0.715	0.654

4.5 Assessment of the potential for long-lead forecasts

The averaged predictor values for the period of June-July-August (JJA) were found to have a potential to predict the DJF seasonal rainfall total anomalies, giving a forecast lead time of 3 months. The optimum number of modes for JJA predictors and corresponding variances are shown in table 6. Similarly, to the SON predictors the optimum number of modes is identified by producing cross-validated predictions for each of the series in the Y data file, and correlating these with the observed values. Most of the major spatial and temporal modes of variability in the predictor variables averaged over the JJA season were found to be similar to those obtained for the SON season. However, the third major mode of SSTs (for JJA) showed a reverse structure to that of the SON season. Similarly, the first major modes of variability in zonal winds (at 1000 mb, 850, 700, and 200 pressure levels) are a reverse to the corresponding ones obtained for SON season.

4.5.1 Development of model to predict DJF rainfall using JJA predictors

Similarly, to the development of multiple linear regression equations using SON predictors, the first few principal components of the six predictor variables in table 6 were later used to develop multiple linear regression models for predicting rainfall anomalies during the DJF season. The SST modes retained a larger total variance as before mostly due to the low-frequency evolution of sea surface temperatures. As shown in table 6, only one major mode was retained for Mean Sea Level Pressure. Zonal wind at 850 mb and at 700 mb levels both had almost similar total variances from the 3 modes retained for each of them.

Table 6: Optimum number of modes retained for each JJA predictor variable together with variance explained.

Predictor (JJA)	**Optimum Number of modes**	**Mode variance**					**Total variance**
		1	**2**	**3**	**4**	**5**	
SST	5	26.0	13.4	5.7	5.4	4.2	54.7
MSLP	1	21.4					21.4
U1000	5	10.5	8.6	6.5	5.1	4.9	35.6
U850	3	12.1	9.4	6.6			28.1
U700	3	12.9	9.6	5.9			28.4
U200	4	17.3	8.7	6.8	6.4		39.2

4.5.2 Models constructed using JJA predictors.

The major modes of variability for each of the JJA predictors as given in table 6 were used to construct regression models to predict DJF rainfall anomalies giving a lead time of 3 months. The forward selection multiple regression approach was used for this purpose.

The predicted anomalies plotted against observed rainfall anomalies for each of the representative stations are shown in figures 15 (a) to (c), albeit with decrease in skill as compared to models developed using SON predictors.

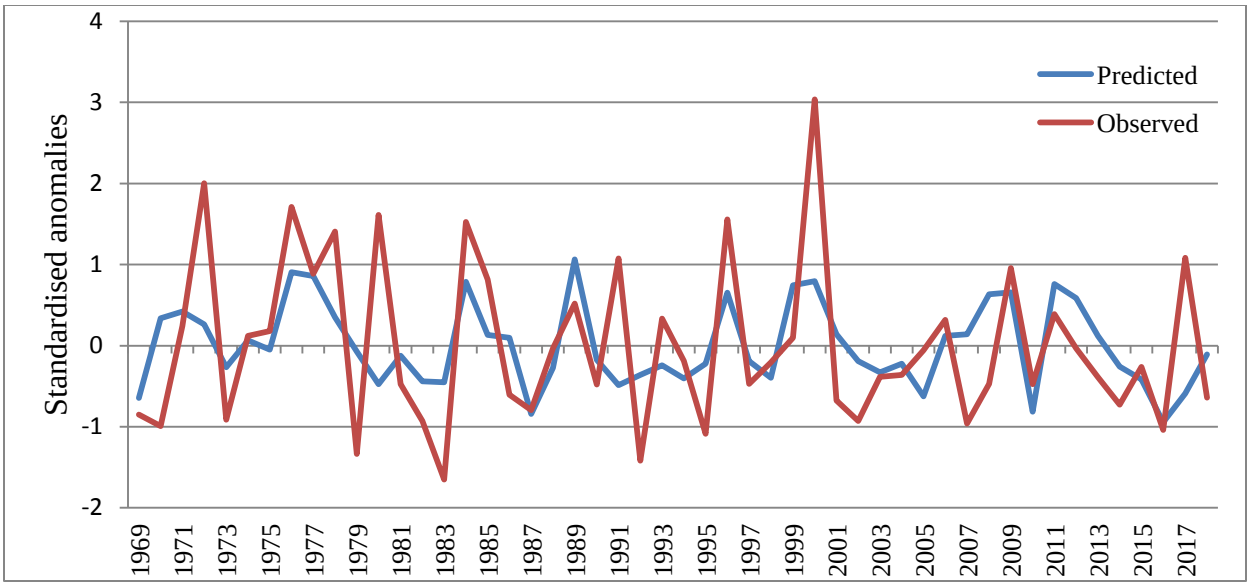


Figure 15 a: Predicted anomalies against observed rainfall anomalies for Mbabane.

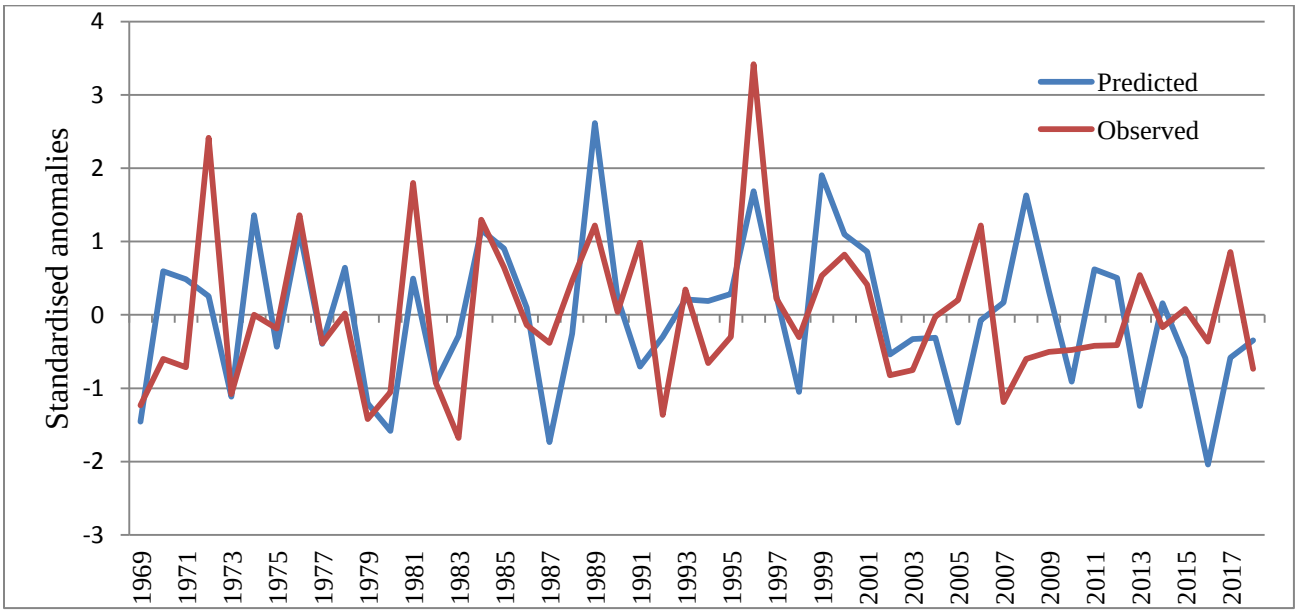


Figure 15 b: Predicted anomalies against observed rainfall anomalies for Matsapha.

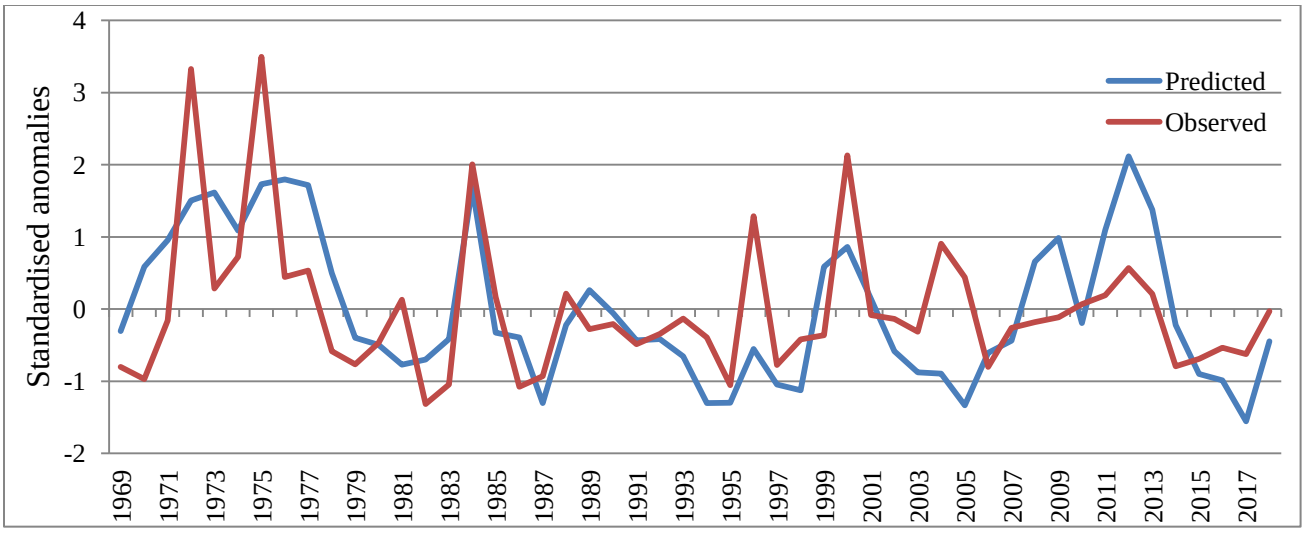


Figure 15 c: Predicted anomalies against observed rainfall anomalies for Big Bend.

Verification of JJA models

After the models were constructed, once again verification was performed using the relative operating characteristic (ROC) scores. The summary of the scores is given in table 7. Once again the ROC scores are above 0.5 meaning that the models are skilful and can be used for forecasting. The ROC scores are significant at the 0.05 level of significance. The highest ROC scores are obtained for the above-normal category implying that the models segregate the above-normal category much better than the other.

Table 7: ROC scores of the representative stations for models developed using JJA predictors.

Station	**ROC (below-normal)**	**ROC (above-normal)**
Mbabane	0.658	0.720
Matsapha	0.628	0.736
Big Bend	0.704	0.759

CHAPTER FIVE

5.0 CONCLUSIONS AND RECOMMENDATIONS

This chapter outlines conclusions and recommendations drawn from the study.

5.1 Conclusions

In the study, the predictive potential of seasonal rainfall (DJF season) using dominant ocean and atmospheric patterns was investigated over southern Africa with focus on Eswatini. Major patterns of ocean variability influencing rainfall included ENSO and positive IOD as revealed by principal component 2 in the SST modes of spatial variability. At most five major spatial modes with corresponding time modes of variability in each of the six variables utilised, were determined to influence rainfall over Eswatini. From the climatological mean circulation during SON, winds can be seen blowing westerly from South Atlantic and easterly from Indian Ocean at the lower tropospheric levels (1000 mb and 850 mb). However, the easterly flow appears to be stronger in the 850 mb level suggesting that the bulk of the moisture comes from the Indian Ocean. The higher levels 700 mb and 200 mb are dominated by westerly flow over a greater part of the study area. Most of the predictability in two of the stations used was dominated by MSLP and zonal winds which shows the importance of wind circulation. Hence, a notable outcome is that circulation at upper levels of the atmosphere have an important contribution to the predictability of rainfall in Eswatini.

Using SON predictors it was established that DJF rainfall is predictable with usable skill. Three of the stations namely Mbabane, Malkerns and Matsapha showed significant positive correlation with each of the six predictor variables. Predictive models were however developed for the three representative stations Mbabane, Matsapha and Big Bend, taken from agro-ecological zones shown in table 1. The predictive models developed over Eswatini were evaluated whether they are skilful or not using ROC scores. ROC scores above 0.5 were found, confirming that the models are skilful. This means that the models can be used to give a forecast since they effectively distinguish between the extreme rainfall categories which are of greatest concern to human livelihood and economy.

After establishing the predictability potential, an investigation was conducted to assess the possibility of long-lead forecasts. JJA predictors gave the longest forecast lead time of 3 months albeit with reduced but still usable skill. The models can therefore be used to give guidance on what is most likely to happen in the DJF season three months ahead of time. This is important for planning purposes and resource mobilisation especially for disaster management agencies. The formulated models also revealed the importance of atmospheric circulation which directly brings in the mechanism of moisture source and distribution over southern Africa region. The DJF season was chosen because of its importance to agricultural activities, especially as it is around that time when the widely grown maize crop in the region starts tasseling and also because this is the peak of the rainfall season hence, important for water resources.

5.2 Recommendations

Using Eswatini's rainfall stations the study revealed that seasonal rainfall forecasts can be made up to 3 months prior to the DJF season in the country. Hence, disaster management agencies can make use of the information in planning and preparing for any potential climate related disaster like drought or flooding. Farmers can also use the information to make important decisions on their farming practices concerning anticipated unpleasant climate conditions. Water resources managers can use forecasts information to plan ahead of time and implement water conservation practices depending on the forecasted conditions.

The study only focused on a period around the peak of the seasonal rainfall hence future research should consider the start of the season as it determines the time of planting and possibly sub-seasons for example the peak month of the season. Efforts should also be made to consider other potential predictor variables like geopotential heights and relative humidity among others. The predictors can also be from global dynamical models, and this will be done in the future extension of this study.

REFERENCES

Abiodun, B., Artan, G., Hart, N., James, R., Kay, G., Mutemi, J., Pokam,W., Senior, C., and Washington, R., 2018: Bull. Amer. Meteor. Soc., **BAMS** February 2018, 313-336: https://journals.ametsoc.org/doi/full/10.1175/BAMS-D-16-0090.1

Adamsonf, G. C. D., Endfield, G. H., Klein, J., Knivetone, D. R., and Nash, D. J., 2015: Tropical cyclone activity over Madagascar during the late nineteenth century. *International Journal of Climatology* **35**: 3249-3261.

Alfano, E., Avarado, L., Lopez, B. F. and Maldonado, T., 2013: Seasonal prediction of extreme precipitation events and frequency of rainy days over Costa Rica, Central America, Using Canonical Correlation Analysis. *Advances in Geosciences* **33**: 41-52.

Ambrosino, C., Chandler, R. E., and Todd, M. C., 2011: Southern African Monthly Rainfall Variability: An Analysis Based on Generalized Linear Models. *Journal of Climate* **24**: 4600-4617.

Ash, K. D., and Matyas, C. J., 2012: The influences of ENSO and the subtropical Indian Ocean Dipole on tropical cyclone trajectories in the southwestern Indian Ocean. *International Journal of Climatology* **32**: 41-56.

Ashok, K., Preethi, B., and Sabin, P., 2015: Impacts of the ENSO Modoki and other Tropical Indo-Pacific Climate-Drivers on African Rainfall. *Sci. Rep.* **5**, 16653; doi: 10.1038/srep16653.

Barnston, A. G., and Korecha, D., 2006: Predictability of June-September Rainfall in Ethiopia. *Monthly Weather Review* **135**: 628-650.

Barnston, A. G., Thiao, W., and Kumar, V., 1996: Long-Lead Forecasts of Seasonal Precipitation in Africa using CCA. *Weather and Forecasting* **11**: 506-520

Bartman, A., Beraki, A., Kgatuke, M., Landman, W. A., Mbedzi, M., and du Piesanie, A., 2009: Performance comparison of dynamical and empirical downscaling methods for South Africa from a seasonal climate modelling perspective. *International Journal of Climatology* **29**: 1535-1549.

Basher, R., Cane, M. A., Goddard, L., Mason S. J., Ropelewski, C. F., and Zebiak, S. E., 2001: Current approaches to seasonal-to-interannual climate predictions. *International Journal of Climatology* **21**: 1111–1152.

Becker, A., et al (2013): A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901-present. Earth System Science Data, http://dx.doi.org/10.5194/essd-5-71-2013.

Behera, S., Ratna, S. B., Ratnam, J. V., Takahashi, K., and Yamagata, T., 2013: An index for tropical temperate troughs over southern Africa. *Climate Dynamics* **41**: 421-441.

Beraki, A., and Landman, W. A., 2012: Multi-model forecast skill for mid-summer rainfall over southern Africa. *International Journal of Climatology* **32**: 303-314.

Beraki, A., DeWitt, D., Dong-Eun, L., Lotter, D., and Landman, W. A., 2012: Seasonal Rainfall Prediction Skill over South Africa: One-versus Two-Tiered Forecasting Systems. *Weather Forecasting* **27**: 489-501.

Blamey, R. C., Kolusu, S. R., Mahlalela, P., Reason, C. J. C., 2018: The role of regional circulation features in regulating El Niño climate impacts over southern Africa: A comparison of the 2015/2016 drought with previous events. *International Journal of Climatology* **38**: 4176-4295.

Botes, S., Goddard, L., Landman, W. A., and Shongwe, M., 2005: Assessing the predictability of extreme rainfall seasons over southern Africa. *Geophysics Research Letters* **32**: L23818, doi: 10.1029/ 2005GL023965.

Camberlin, P., Janicot, S., and Poccard, I., 2001: Seasonality and Atmospheric Dynamics of the Teleconnection between African Rainfall and Tropical Sea-Surface Temperature: Atlantic vs. ENSO. *International Journal of Climatology* **21**: 973-1005.

CartoGIS Services Coordinator, College of Asia and the Pacific, The Australian National University. http://asiapacific.anu.edu.au/mapsonline/base-maps/southern-africa [Accessed 31 March 2019].

Colberg, F., Reason, C. J. C. and Rodgers, K., 2004: South Atlantic response to El Niño-Southern Oscillation induced climate variability in an ocean general circulation model. *Journal of Geophysical Research*: Oceans **109**: 1978-2012.

Cook, C., Hewitson, B. C., and Reason, C. J. C., 2004: Wet and dry spells within particularly wet and dry summers in the South African summer rainfall region. *Climate Research* **26**: 17-31.

Cooper, G. R. J., McCarthy, S., and Tyson, P. D., 2002: Millennial to Multi-Decadal Variability in the Climate of Southern Africa. *International Journal of Climatology* **22**: 1105-1117.

Crétat, J., Fauchereau, N., Pohl, B., Reason, C., Richard, Y., and Rouault, M., 2012: Recurrent daily rainfall patterns over South Africa and associated dynamics during the core of the austral summer. *International Journal of Climatology* **32**: 261-273.

Engelbrecht, C. J., Engelbrecht, F. A., Landman, W. A., and Malherbe, J., 2012: Tropical systems from the southwest Indian Ocean making land fall over the Limpopo River Basin, southern Africa: a historical perspective. *International Journal of Climatology* **32**: 1018-1032.

Erdelyi, R., Hall, R. J., Hanna, E., Jones, J. M., and Scaife, A. A., 2017: Simple Statistical Probabilistic Forecasts of the Winter NAO. *Weather and Forecasting* **32**: 1585-1601.

Fauchereau, N., Hart, N. C. G. and Reason, C. J. C., 2013: Cloud bands over southern Africa: seasonality, contribution to rainfall variability and modulation by the MJO. *Climate Dynamics* **41**: 1199-1212.

Fauchereau, N., Pohl, B., and Richard, Y., 2007: Influence of Madden-Julian Oscillation on southern African Summer Rainfall. *Journal of Climate* **20**: 4227-4242.

Fauchereau, N., Richard, Y., Rouault, M., and Trzaska, S., 2003: Rainfall Variability and Changes in Southern Africa during the 20th Century in the Global Warming Context. *Natural Hazards* **29**: 139-154.

Food and Agriculture Organisation of the United Nations: Eswatini. *Water Report* 29, 2005. http://www.fao.org/nr/water/aquastat/countries_regions/SWZ/ [Accessed 11 March 2019].

Girmes, R., Hense, A., Menz, G., and Paeth, H., 2006: Improving Seasonal Forecasting in the Low Latitudes. *Monthly Weather Review* **134**: 1859-1879.

Gitau, W., Camberlin, P., Ogallo, L., and Okoola, R., 2014: Oceanic and atmospheric linkages with short rainfall season intraseasonal statistics over Equatorial Eastern Africa and their predictive potential. *International Journal of Climatology* DOI: 10.1002/joc.4131.

Goddard, L., and Landman, W. A., 2002: Statistical Recalibration of GCM Forecasts over Southern Africa Using Model Output Statistics. *Journal of Climate* **15**: 2038-2055.

Graham, N. E., and Mason, S. J., 1999: Conditional Probabilities, Relative Operating Characteristics, and Relative Operating Levels. *Weather and Forecasting* **14**: 713-725

Halpert, M. S., and Ropelewksi, C. F., 1987: Global and Regional Scale Precipitation Patterns Associated with the El Niño/Southern Oscillation. *Monthly Weather Review* **115**: 1606-1626.

Hart, N. C. G., Reason, C. J. C., and Washington, R., 2018: On the Likelihood of Tropical-Extratropical Cloud Bands in South Indian Convergence Zone during ENSO Events. *Journal of Climate* **31**: 2797-2817.

Hegerl, G. C., and Kenyon, J., 2010: Influence of the modes of climate variability on Global precipitation extremes. *Journal of Climate* **23**: 6248-6262.

Ininda, J. M., 1994: Numerical Simulation of the Influence of the Sea Surface Temperature Anomalies on the East African Seasonal Rainfall. Doctor of Philosophy in Meteorology submitted to the University of Nairobi.

Jury, M. R., 1996: Regional teleconnection patterns associated with summer rainfall over South Africa, Namibia and Zimbabwe. *International Journal of Climatology* **16**: 135-153.

Jury, M. R., Mason, S. J., and Mulenga, H. M., 1999: Exploratory Long-Range Models to Estimate Summer Climate Variability over Southern Africa. *Journal of Climate* **12**: 1892-1899.

Kaiser, H. F., 1960: The application of electronic computers to factor analysis. *Educ. and Psycol. Meas.* **20**: 141-151.

Kalnay et al.,The NCEP/NCAR 40-year reanalysis project, Bull. Amer. Meteor. Soc., 77, 437-470, 1996.

Kim, J. and Nicholson, S. E., 1997: The relationship of the El Niño-Southern oscillation to African rainfall. *International Journal of Climatology* **17**: 117-135.

Kottek, M. and Rubel, F., 2010: Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen-Geiger climate classification. *Meteorologische Zeitschrift* **19**: 135-141.

Layberry, R., Kniveton, D. R., and Williams, C. J. R., 2008: Influence of South Atlantic Sea Surface Temperatures on Rainfall Variability and Extremes over Southern Africa. *Journal of Climate* **21**: 6498-6520.

Lee, S., Mechoso, C. R., Neelin, J. D., and Wang, C., 2013: Inter-hemispheric Influence of the Northern Summer Monsoons on Subtropical Anticyclones. *Journal of Climate* **26**: 10193-10204.

Lutjeharms, J. R. E., Mavume, A. F., Rouault, M., and Rydberg, L., 2009: Climatology and Landfall of Tropical Cyclones in the South-West Indian Ocean. *Journal of Marine Science* **8**: 15-36.

Makarau, A., and Jury, M. R., 1997: Predictability of Zimbabwe Summer Rainfall. *International Journal of Climatology* **17**: 1421-1432.

Matyas, C. J., 2015: Tropical cyclone formation and motion in the Mozambique Channel. *International Journal of Climatology* **35**: 375-390.

Mulenga, H., and Reason, C. J. C., 1999: Relationships between South African rainfall and SST anomalies in the Southwest Indian Ocean. *International Journal of Climatology* **19**: 1651-1673.

Mwafulirwa, N. D. and Jury, M. R., 2002: Climate Variability in Malawi, Part 1: Dry Summers, Statistical Associations and Predictability. *International Journal of Climatology*, **22**: 1289-1302.

Nakamura, M., 2011: Impacts of SST Anomalies in the Agulhas Current System on the Regional Climate Variability. *JOURNAL OF THE ATMOSPHERIC SCIENCES* **58**: 2146-2162

Nicholson, S. E., 2014: The Predictability of Rainfall over the Greater Horn of Africa. Part I: Prediction of Seasonal Rainfall. *Journal of Hydrometeorology* **15**: 1011-1027.

Nyakwada, W., 2009: Predictability of East African Seasonal Rainfall with Sea Surface Temperature Gradient Modes. Doctor of Philosophy in Meteorology submitted to the University of Nairobi.

Palmer, P. I., Todd, M. C., and Washington, R., 2004: Water vapour transport associated with tropical-temperate trough systems over southern Africa and the southwest Indian Ocean. *International Journal of Climatology* **24**: 555-568.

Phakula, S., 2016: Modelling Seasonal Rainfall Characteristics over South Africa. Master of Science dissertation submitted to the University of Pretoria.

Reason, C. J. C., 2017: Climate of Southern Africa. *Oxford Research Encyclopedia of Climate Science*. DOI: 10.1093/acrefore/9780190228620.013.513

Schneider et al (2017): Evaluating the Hydrological Cycle over Land Using the Newly-Corrected Precipitation Climatology from the Global Precipitation Climatology Centre (GPCC). Atmosphere 8(3), 52; doi:10.3390/atmos8030052

Sharon, S. E., 1987: Rainfall Predictability in Equatorial and Southern Africa: Relationships with Surface Sea Temperatures along the Southwestern Coast of Africa. *Journal of Climate and Applied Meteorology* **26**: 561-578.

Wilks D. S., 2006: Statistical Methods in the Atmospheric Sciences. Second Edition, ISBN 13: 978-0-12-751966-1.

WMO 1953: World distribution of thunderstorm days: WMO/OMM-NO. 21 TP. 6

Worldatlas.com,2019: https://www.worldatlas.com/webimage/countrys/africa/lgcolor/szcolor.htm [Accessed 25 March 2019].