DIAGNOSIS AND PREDICTABILITY OF INTRASEASONAL CHARACTERISTICS OF WET AND DRY SPELLS OVER EQUATORIAL EAST AFRICA

By

GITAU, WILSON

Department of Meteorology,

School of Physical Sciences

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in Meteorology, University of Nairobi, Kenya

and

Doctor of Philosophy in Climatology, Université de Bourgogne, France.

March 2011

University of NAIROBI Library
0378899 9

ACKNOWLEDGEMENTS

Thanks to the **Almighty Lord** for seeing me through my university education this far and successive completion of this Doctorate thesis. May His Holy Name be glorified forever.

The author would like to take this earliest opportunity to thank the supervisors, **Prof Laban OGALLO,** Director of IGAD Climate Prediction and Applications Centre (ICPAC), **Prof Pierre CAMBERLIN** of Centre of Research in Climatology / University of Bourgogne (CRC/UB) in Dijon, France and **Dr Raphael E. OKOOLA** of Department of Meteorology at University of Nairobi (UoN) in Nairobi, Kenya for their continuous guidance, constructive criticisms, encouragement and their tireless review of the various manuscripts that culminated in successive and timely completion of this final volume.

The author is deeply indebted to the **Government of France** through its **Embassy in Nairobi, Kenya** for offering a fully-funded scholarship to undertake the study in France and three return-air tickets to France to carry out the research work written in this thesis. Thanks to the **Vice Chancellor of UoN, Prof G A O MAGOHA** and the **President of UB, Mr Jean-Claude FORTIER** for establishing collaboration for joint-supervision (cotutelle) of this research work.

Special thanks go to the University of Nairobi especially the staff in the Department of Meteorology. Thanks to the Chairman, **Prof J N MUTHAMA** and his academic staff for training me since my first year as an undergraduate in October 1997 up to the completion of this PhD work, to the Dean, School of Physical Sciences and Principal, College of Biological and Physical Sciences, and the University of Nairobi as a whole for investing so much resources to train me as a part of academic staff.

Special thanks are due to the **institutions** and **staff members** of **CRC/UB** and **ICPAC** for providing me with computational facilities and acquisition of the data sets as well as providing me with every opportunity to complete the work. I am equally indebted my fellow students at CRC/UB (**Amoussou, Cretat, Oettli, Lofti, Louvet, Pohl, Romain, Vivianne** and **others**). My words of appreciation to **Dr Nathalie PHILIPPON** (CRC/UB) for her efforts in Matlab programming. Special thanks to **Mr Zachary ATHERU** (ICPAC) and **Dr Faith GITHUI** (Victoria-Australia) for the availing time to discuss this work. My appreciation equally goes to the **Vice Chancellor, UoN** and **Director, ICPAC** for granting

me with fully-paid study leave to carry out this research at CRC/UB in Dijon, France.

I am sincerely and greatly indebted to my beloved wife, **Ms Esther NYAMBURA** for her patience, moral support and understanding during the long periods of absence while carrying out this research as well as taking utmost care of our beloved son, **Alex KINYUA**. I am equally indebted to my son **Alex,** for remaining a nice and well-behaved boy in my absence. I do hereby acknowledge **my parents, brothers** and **sister** for the support both in words and deeds that they provided to my family during my long periods of absence.

And to all those who provided voluntary assistance of all kinds and for everybody who takes time to read this thesis, it was appreciated and is hereby acknowledged.

ABSTRACT

Most of Eastern Africa has arid and semi-arid climate with high space-time variability in rainfall. The droughts are very common in this region, and often persist for several years, preceded or followed by extreme floods. Most of the livelihoods and socio-economic activities however remain rain-dependent leading to severe negative impacts during the periods of occurrence of climate extremes. It has been noted that one extreme event was capable of reversing national economic growth made over a period of several years. Thus no sustainable development can be attained in eastern Africa without effective mainstreaming of climate information in the development policies, plans and programmes.

Many past studies in the region have focused on rainfall variability at seasonal, annual and decadal scales. Very little work has been done at intraseasonal timescale that is paramount to most agricultural applications. This study aims at filling this research gap, by investigating the structure of rainfall season in terms of the distribution of wet and dry spells and how this distribution varies in space and time at interannual time scale over Equatorial Eastern Africa. Prediction models for use in the early warning systems aimed at climate risk reduction were finally developed. The specific objectives of the study include to; delineate and diagnose some aspects of the distribution of the wet and dry spells at interannual timescale; investigate the linkages between the aspects of the distribution of wet and dry spells identified and dominant large scale climate fields that drive the global climate; and assess the predictability of the various aspects of wet and dry spells for the improvement of the use in the early warning systems of the region.

Several datasets spanning a period of 40 years (1961 – 2000) were used. The data included gauged daily rainfall amount for the three Eastern Africa countries namely Kenya, Uganda, and Tanzania; Hadley Centre Sea Surface Temperature (SST); re-analysis data and radiosonde observations from Nairobi (Kenya) and Bangui (Central Africa Republic) upper air stations. The indices of El Niño-Southern Oscillation (ENSO), Indian Ocean Dipole and SST gradients which constituted the predefined predictors were also used.

Missing data gaps were initially filled and the quality of rainfall data assessed. Less than seven percent of the data were estimated in all cases. The study region was then classified into few near-homogeneous spatial and temporal rainfall regimes using empirical orthogonal function approach. Several intraseasonal statistics of the wet / dry spells were computed at

both local (station) and sub-regional (near-homogeneous zone) levels to provide baseline information on the various aspects of rainfall distribution during March-May (long rains) and October-December (short rains) rainfall seasons. The interannual variation in the above intraseasonal statistics at local and sub-regional levels was also assessed for any significant trend using the non-parametric Spearman rank correlation test. The linkages between the various intraseasonal statistics of the wet / dry spells including seasonal rainfall totals and large scale climate fields were assessed using the total and partial Pearson correlation analysis. Last but not least, the stepwise regression technique was used to develop multivariate linear regression models for predicting the various intraseasonal statistics of wet / dry spells. The skill of these models was finally assessed using various statistical techniques.

The results obtained indicated that the gap-filled and quality controlled daily rainfall observations were of good quality and formed the foundation of all the analyses that were undertaken in this study. For the first time, this study delineated daily rainfall over Equatorial Eastern Africa into six near-homogeneous sub-regions for both the long and the short rainfall seasons. They are however significant spatial differences in the patterns of daily rainfall occurrences for the individual seasons which may be attributed to different climate mechanisms and systems which are in play during the specific rainfall seasons.

At interannual scale, positive (negative) relationship existed between the intraseasonal statistics of wet (dry) spells and the seasonal rainfall totals over most locations and sub-regions. The relationship with the intraseasonal statistics of the wet spells was mainly significant (at 95% confidence level) while those of the dry spells were generally not statistically significant. The mean frequency of dry spells of 5 days or more (the number of wet days within the season) had the least (strongest) association with the seasonal rainfall totals. The relationships were stronger during the short rainfall season compared to the long rainfall season.

For the first time, the study showed significant trends in all the intraseasonal statistics of the wet / dry spells though at few isolated locations. However, significant increasing trend in the occurrence of dry spells of 5 days or more showed organised patterns for the two seasons. Climate change is becoming a major development concern not only over the region but the world over. Further studies are therefore required to examine whether the trends observed in the daily rainfall spells in this study reflects any regional climate change signals.

Results from total and partial Pearson correlation analysis identified several large scale oceanic and atmospheric signals with robust physical/dynamical linkages with the sub-regional intraseasonal statistics of wet / dry spells (SRISS). The results further showed that the linkages between sub-regional intraseasonal statistics of wet spells and large scale signals were mainly from atmospheric fields of zonal and meridional components of wind and the specific humidity during the long rainfall season. For the short rainfall season, stronger linkages with oceanic variables especially SST were noted. The atmosphere has less climatic memory when compared with the oceans. Past studies have indicated stronger predictability potentials for the short rainfall season. By identifying stronger linkages between intraseasonal characteristics of wet spells for long (short) rainfall season and the atmospheric (oceanic) variables, the study has for the first time provided some insights to the prediction challenges for the specific seasons. Thus future predictability efforts for the long rainfall season should ensure the inclusion of atmospheric variables in the prediction models.

The study has produced cross-validated multivariate linear regression (MLR) models for predicting some intraseasonal characteristics of wet spells that can be used to support the current generation of models being used by the IGAD Climate Prediction and Applications Centre and National Meteorological and Hydrological Services.

The results from this study have for the first time provided an in-depth knowledge on the intraseasonal modes of rainfall variability and improvement in the forecasting and early warning tools for the wet spells over the Equatorial Eastern Africa region. Better understanding and accurate prediction of rainfall totals and intraseasonal statistics of wet / dry spells is of paramount importance in the planning, development and management of all rainfall-sensitive socio-economic sectors of the economy such as agricultural and water resources; and further contribute to national efforts towards achievements of the Millennium Development Goals.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ANGCO	Zonal wind index at 925mb level located over Angola and its coast averaged for December-January period
BoBEN	SST index over Bay of Bengal averaged for July-August period
BoBEN-1	SST index with slight location variation from BoBEN averaged for December-January period
CINDO	Zonal wind index at 700mb level over equatorial central Indian Ocean averaged for December-January period
CSINDO	Meridional wind index at 200mb level over central parts of the southern Indian Ocean averaged for December-January period
EALLJ	East Africa Low Level Jet
EBBEN	Specific humidity index at 925mb level over southern Asia slightly to the east of Bay of Bengal averaged for December-January period
ECMAD	SST index over east coast of Madagascar averaged for July-August period
ECMAD-1	SST index with slight location variation from ECMAD averaged for January-February period
ECMWF	European Centre of Medium range Weather Forecasting
EEA	Equatorial Eastern Africa
ENSO	El Niño / Southern Oscillation
EOF	Empirical Orthogonal Analysis
EQAFR	Zonal wind index at 200mb level extending from Equatorial Eastern Africa averaged for July-August period
EQAFR-1	Zonal wind index at 200mb level with slight location variation from EQAFR averaged for December-January period
EQATO	Meridional wind index at 200mb level over equatorial Atlantic Ocean averaged for December-January period
EQIND	Specific humidity index at 700mb level extending through equatorial Indian Ocean into eastern Africa region averaged for July-August period
ERA40	ECMWF Re-Analysis of 40 years
GCM	General Circulation Model / Global Climate Model
GHA	Greater Horn of Africa

IOD	Indian Ocean Dipole
ISO	Intraseasonal Oscillations
ISS	Intraseasonal statistics of wet and dry spells
ITCZ	Inter-Tropical Convergence Zone
LISS	Local intraseasonal statistics of wet and dry spells
MAM	March-April-May
MARCON	Zonal wind index at 200mb level over the maritime continent averaged for July-August period
MDG	Millennium Development Goals
MJO	Madden Julian Oscillation
MLR	Multivariate Linear Regression
NCAR	National Centre for Atmospheric Research
NCEP	National Centre for Environmental Prediction
NE	North Easterlies
NEATO	Meridional wind index at 700mb level extending slightly over northern Atlantic Ocean averaged for December-January period
NEGHA	Meridional wind index at 925mb level over northeastern parts of Greater Horn of Africa averaged for December-January period
NINDS	Zonal wind index at 200mb level over northern India sub-continent averaged for December-January period
OLR	Outgoing Longwave Radiation
OND	October-November-December
PCA	Principal Component Analysis
QBO	Quasi-Biennal Oscillations
RPCA	Rotated Principal Component Analysis
SAFR	Meridional wind index at 925mb level over southern Africa averaged for December-January period
SCEINDO	Zonal wind index at 700mb level, south of central equatorial Indian Ocean averaged for December-January period
SE	South Easterlies

SINDS	**Zonal wind index at 925mb level near the southern tip of India sub-continent averaged for July-August period**
SINDS-1	**Zonal wind index at 925mb level with slight location variation from SINDS averaged for December-January period**
SINDS-2	**Zonal wind index at 925mb level with slight location variation from SINDS averaged for December-January period**
SMESEA	**Specific humidity index at 925mb level south of Mediterranean Sea averaged for December-January period**
SOI	**Southern Oscillation Index**
SRISS	**Sub-regional intraseasonal statistics of wet and dry spells**
SSA	**Meridional wind index at 200mb level to the south of the study area averaged for December-January period**
SST	**Sea Surface Temperature**
STAFR	**Geopotential height index at 700mb level over the southern tip of Africa continent averaged for December-January period**
SWAFRC	**Specific humidity index at 700mb level located at Angola coast on south-western Africa averaged for July-August period**
SWHAW	**SST index over southwestern of Hawaii averaged for July-August period**
WAFR	**Zonal wind index at 925mb level extending from Atlantic Ocean to western Africa averaged for December-January period**
WCAUS	**SST index over western coast of Australia averaged for July-August period**
WCAUS-1	**SST index with slight location variation from WCAUS averaged for December-January period**
WCAUS-2	**SST index with slight location variation from WCAUS averaged for December-January period**
WCSOA	**Specific humidity index at 925mb level on the western coast of southern Africa averaged for December-January period**
WINDO	**Meridional wind index at 925mb level over equatorial western Indian Ocean and equatorial Africa averaged for December-January period**
ZIND	**Zonal SST gradient over Indian Ocean**
ZPAC	**Zonal SST gradient over Pacific Ocean**

CHAPTER ONE
INTRODUCTION

1.1 Background

The economies of East African countries largely depend on rain-fed agriculture. Over Kenya for example, the agricultural sector forms the main socio-economic activity accounting for up to 30% of the country's gross domestic product, 60% of the export earnings and the largest source of employment (ICPAC, 2006). Variation in the yields of many crops to a large extent is dependent on rainfall amounts and their distribution in space and time. Rainfall is therefore the most important weather factor in the region. There are however large variability of rainfall in the region in both spatio-temporal distribution and magnitudes. This has been witnessed by the recent droughts (1999 - 2001 and 2005 – 2006) that affected many parts of the Horn of Africa. Localized floods were however recorded at the onset of rains in some locations. The spatio-temporal variability of rainfall over Eastern Africa at different time scales are due to complex topographical features and existence of large water bodies (Kongoti, 1989; Ogallo, 1989; 1993; Mukabana and Pielke, 1996; Indeje *et al.,* 2001; Oettli and Camberlin, 2005; Nyakwada, 2009).

The cummulation of the specific spatio-temporal variability of rainfall in both magnitudes and distribution is often having devastating socio-economic impacts. Impacts associated with climate extremes include floods and droughts resulting in loss of life and property, food insecurity, water scarcity, power and communication interruptions, poor infrastructure and other socio-economic disruptions. Detailed spatio-temporal information of rainfall on different temporal scales is therefore essential for effectively managing of all rainfall dependent socio-economic systems and for disaster risk reduction.

Many studies in the past have focused on understanding the rainfall variability at monthly, seasonal, and interannual time scales. These studies have included predictability studies using linkages between rainfall and large scale phenomena such as El Niño-Southern Oscillation.

Recent studies over the region that have concentrated on the understanding of atmospheric processes and prediction of rainfall at different timescales, especially at seasonal timescale based on SST and SST-derived variables include Mutai, 2000; Mutemi, 2003; Owiti, 2005; Owiti *et al.,* 2008; Nyakwada, 2009. Upper tropospheric temperature and geopotential

variables have also been used (Njau, 2006). These studies showed that over the Eastern Africa region, the short rainfall season (October to December) has higher predictability as compared to the long rainfall season (March to May). The long rainfall season has been associated with complex interactions between many regional and large scale mechanisms which generally induce large heterogeneities in the spatial rainfall distribution (Ogallo, 1982; Semazzi *et al.*, 1996; Okoola, 1998; Indeje *et al.*, 2000) and virtually negligible correlation with ENSO (Ogallo, 1988).

The higher predictability of rainfall during the October to December season is attributed to the strong linkage with the regional and global teleconnections (Mutemi, 2003; Black *et al.*, 2003; Black, 2005; Owiti, 2005; Owiti *et al.*, 2008). However, studies to improve the understanding on the nature and characteristics of rainfall on intraseasonal timescales, particularly daily timescale are still lacking. Notwithstanding, a number of studies have investigated intraseasonal convective variability and pentad mean rainfall characteristics (Okoola, 1998; Mutai and Ward, 2000; Camberlin and Okoola, 2003).

The occurrence of wet and dry spells within the rainfall season determines the water availability for the rain-fed agriculture. Very limited efforts have been made in the region to understand their characteristics well and predict the interannual variability of the intraseasonal characteristics of the wet and dry spells in the region. There are many previous studies on the interannual rainfall variability at monthly, seasonal and annual timescales and few studies on the intraseasonal variability. However the linkage between the interannual rainfall variability and the intraseasonal wet and dry spells is still missing. This will be the focus of this research as outline in the objective of the study in **section 1.3**. Detailed justification for this study is provided later in this chapter.

1.2 Statement of the problem

East Africa is characterized by limited natural resources especially water, minerals and agricultural land. High population growth rate, poor agricultural practices, deforestation, abject poverty and high levels of unemployment are but some of the socio-economic challenges that face the region.

The high population growth rate has led to people migrating into the arid and semi-arid land (ASAL) areas thereby affecting the ecosystems of the region and rendering them more vulnerable to hazards such as drought (Bryan and Southerland, 1989). The high population

variables have also been used (Njau, 2006). These studies showed that over the Eastern Africa region, the short rainfall season (October to December) has higher predictability as compared to the long rainfall season (March to May). The long rainfall season has been associated with complex interactions between many regional and large scale mechanisms which generally induce large heterogeneities in the spatial rainfall distribution (Ogallo, 1982; Semazzi *et al.*, 1996; Okoola, 1998; Indeje *et al.*, 2000) and virtually negligible correlation with ENSO (Ogallo, 1988).

The higher predictability of rainfall during the October to December season is attributed to the strong linkage with the regional and global teleconnections (Mutemi, 2003; Black *et al.*, 2003; Black, 2005; Owiti, 2005; Owiti *et al.*, 2008). However, studies to improve the understanding on the nature and characteristics of rainfall on intraseasonal timescales, particularly daily timescale are still lacking. Notwithstanding, a number of studies have investigated intraseasonal convective variability and pentad mean rainfall characteristics (Okoola, 1998; Mutai and Ward, 2000; Camberlin and Okoola, 2003).

The occurrence of wet and dry spells within the rainfall season determines the water availability for the rain-fed agriculture. Very limited efforts have been made in the region to understand their characteristics well and predict the interannual variability of the intraseasonal characteristics of the wet and dry spells in the region. There are many previous studies on the interannual rainfall variability at monthly, seasonal and annual timescales and few studies on the intraseasonal variability. However the linkage between the interannual rainfall variability and the intraseasonal wet and dry spells is still missing. This will be the focus of this research as outline in the objective of the study in **section 1.3**. Detailed justification for this study is provided later in this chapter.

1.2 Statement of the problem

East Africa is characterized by limited natural resources especially water, minerals and agricultural land. High population growth rate, poor agricultural practices, deforestation, abject poverty and high levels of unemployment are but some of the socio-economic challenges that face the region.

The high population growth rate has led to people migrating into the arid and semi-arid land (ASAL) areas thereby affecting the ecosystems of the region and rendering them more vulnerable to hazards such as drought (Bryan and Southerland, 1989). The high population

c) Assess the predictability of the aspects of wet and dry spells under (a) based on results from (b) for the improvement of the use in the early warning systems of the region.

1.4 Justification of the study

Most studies have addressed rainfall variability at seasonal, annual and decadal scales, but little work has been done at intraseasonal timescale. Although the total seasonal anomalies of rainfall and related variables indicate wet or dry seasons, there is often a demand from users (for example from the agricultural sector) for information about variability on intraseasonal timescales such as the active and break phases within the season (Jadadheesha *et al.*, 2003). It is well known that a season with above average rainfall may not be better than a below average season over an agricultural region if the rainfall are not well distributed in space and time (Usman and Reason, 2004). Crops are likely to do well with evenly distributed 'light' rains than a few isolated 'heavy' rainfall interrupted by prolonged dry periods. For crop cultivation, the consistency with which minimally required rainfall is received is more important than the total rainfall received.

The rainfall time series during the wet seasons is marked by periods of wetness and dryness, which are often called the wet (rainy) spells and dry spells respectively. The transitions from the wet to dry periods and vice versa evolve slowly such that there are typically three or so wet/dry episodes in the course of the wet season (Mpeta and Jury, 2001).

Ogallo *et al.* (2000) have reviewed the potential applications of seasonal to inter-annual climate predictions in agricultural planning operations. Information and knowledge of wet and dry spells would enrich these applications and improve the general adaptations of ecosystems and land-use activities. Clear understanding of the key intraseasonal rainfall variations over East Africa is crucial for planning and management purposes especially to farmers and water managers. Such advance information of forthcoming wet/dry spells could be used to strategize on agricultural and water management policies as well as mitigating the adverse effects of recurring extreme climate events while fully capitalizing when more abundant and evenly spread rainfall occurs.

This study was further motivated by previous studies done within and outside East Africa that have corroborated or revealed significant associations between rainfall season onsets, cessations and wet/dry spells on one hand and end-of-season agricultural yields on the other hand (Stewart and Harsh, 1982; Sivakumar, 1992; Oladipo and Kyari, 1993; Barrow *et al.*,

2003; Barrow, 2004; Komutunga, 2006). For example, a 20-days delay in the onset of the long rainfall season at Katumani in Eastern Kenya whose mean seasonal rainfall is 300mm would reduce the maximum expected maize yield by 25 to 30% (Stewart and Harsh, 1982) while occurrence of a prolonged dry spell during the flowering phase has been shown to cause an estimated 72 – 75% reduction on maximum expected maize yield (Barron *et al.*, 2003). A major application of dry spell analysis is to predict extended drought durations during the growing season, which forms a basis for planning the crop production strategies (Sharma, 1996).

Better understanding and accurate prediction of rainfall totals and intraseasonal statistics of wet and dry spells is of paramount importance in the policy planning and implementation of early warning systems as well as development and management of agricultural, water resources and other rainfall-dependent sectors of the economy. This is in line with the Millennium Development Goals (MDGs) that were formulated in the year 2000 by the United Nations. One of the millennium goals aimed at ensuring environmental sustainability through improved and sustainable access to safe drinking water most of which can be harvested from the rainfall. Timely availability of information on the distribution of wet and dry spells during the wet seasons which this study aims to derive may contribute significantly towards the achievement of this millennium development goal.

In summary, the key in understanding the rainfall variability lies in the acquisition of information on intraseasonal characteristics of rainfall. Such intraseasonal characteristics of rainfall are the onset, duration and cessation of the wet season, seasonal rainfall totals, mean rainfall intensity, mean duration of the spell and others as summarized in **Figure 1.1.** It should be clarified that the various aspects in **Figure 1.1** do not follow any order of importance whatever. The onset, cessation and duration of the seasonal rainfall have been discussed by Alusa and Mushi (1974), Okoola (1998) and Camberlin and Okoola (2003). The rest of the intraseasonal aspects have rarely been studied over East Africa and formed the scope of this study. Better understanding of the behaviour of the wet and dry spells could improve management of the excess water and promote more effective agricultural and environmental management activities by users of climate information.

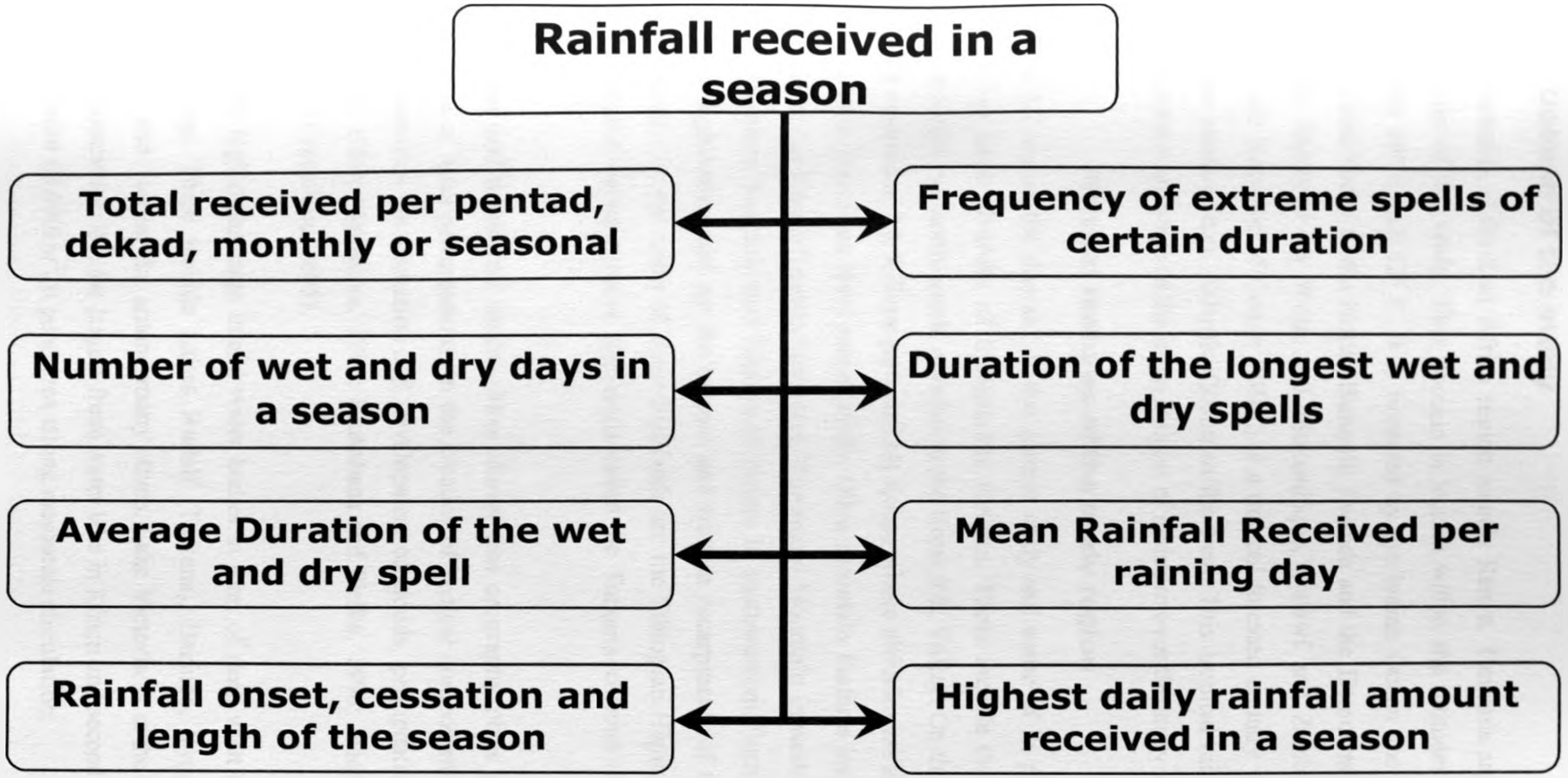


Figure 1.1: Various aspects of the rainfall received in a season

1.5 Domain of the study

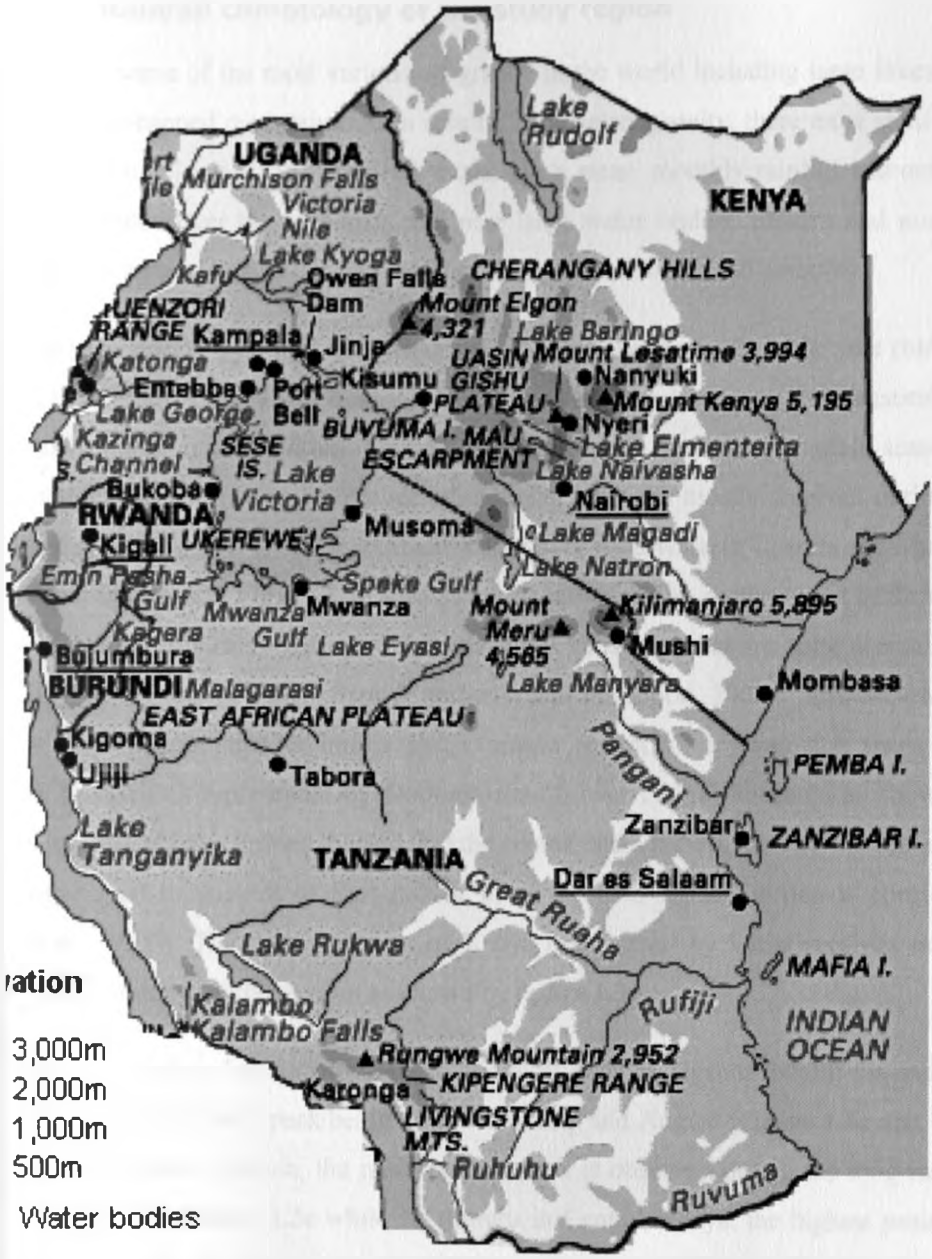
Three countries of the East Africa region namely Kenya, Tanzania and Uganda constituted the domain of this study. This domain is located within the latitudes 5° N and 12° S and longitudes 29° E and 42° E. It is bounded by the Indian Ocean and Somalia to the East, Ethiopia and Sudan to the North, Burundi, Rwanda and the Democratic Republic of Congo (formerly Zaire) to the West, and Mozambique, Malawi and Zambia to the South. The Democratic Republic of Congo (DRC) is a tropical forested country with a small coastline along the south-eastern Atlantic Ocean to the west. This tropical rain forest, Atlantic and Indian Oceans are some of the main sources of moisture over the study region.

1.5.1 Physical features of the study region

Figure 1.2 shows the domain of the current study and some of its physical features. East Africa has large diversity of topographic features. These include the eastern and western highlands that run north-south, parallel to the Great Rift Valley. On the highlands are snow-capped mountains; Mt Kilimanjaro and Mt Kenya whose altitudes are about 5892 metres and 5202 metres above sea level respectively. Other mountain features include Mt Elgon (4321 metres) on the Kenya/Uganda boundary, Ruwenzori Mountain in western Uganda, Mt Meru in northeastern Tanzania and Kipengere ranges in southwestern Tanzania. The eastern and western highlands make up the eastern and western escarpments of the Great Rift Valley respectively. To the north of these highlands are the Ethiopian Highlands with a low level valley region between these highlands called the Turkana channel (Kinuthia and Asnani, 1982).

Empirical and theoretical studies have shown that orography plays a leading role in the formation of local perturbations, in the creation of vertical components of wind speeds, etc, which promotes the formation and development of clouds, precipitation and thunderstorms (Kongoti, 1989; Mukabana, 1992; Mukabana and Pielke, 1996; Indeje *et al.,* 2000, 2001; Oettli and Camberlin, 2005).

The study region has large inland water bodies in form of deep vault lakes along the Great Rift Valley. These include Lakes Rudolf (Turkana), Baringo, Kyoga, Naivasha, Eyasi, Manyara and Tanganyika among many others. Lake Victoria is at the centre and shared by the three countries. It is the largest fresh water lake in Africa and second in the world, with an area of about 68,000km^2. It generates strong mesoscale circulation.



e 1.2: Domain of the study region showing the main physical features ***(Source: lopedia Britannica Online*** [§]***)***

media.web.britannica.com/eb-media//69/5369-004-167C12FF.gif

1.5.2 Rainfall climatology of the study region

East Africa has some of the most varied topography in the world including large lakes, Rift Valley and snow-capped mountains. As a result of this heterogeneity, there exist significant variations in climatological mean rainfall totals. High mean monthly rainfall amounts are mainly concentrated over the highlands and near large water bodies. Eastern and northern Kenya, parts of eastern Uganda and central Tanzania receive low rainfall amounts.

Nearer to the equator, two rainfall and two dry seasons are observed within the year (bimodal regime). The rainfall seasons are locally referred to as long and short rainfall seasons. The long rainfall period occurs within March-April-May while the short rainfall season is concentrated within October to December, with higher amounts mostly received during the long rainfall season as represented by Kabale station over southwestern Uganda and Musoma over northern Tanzania in **Figure 1.3a and 1.3b** respectively. The southern part of Tanzania was excluded from this study since it exhibits rainfall variations that are quite dissimilar to those of the other parts of East Africa (Camberlin and Philippon, 2002). The two rainfall seasons tend to merge together into a single season (unimodal regime) that spans from November to April as represented by Dodoma station over central Tanzania as shown by **Figure 1.3c**. Studies have further showed that the central and southern parts of Tanzania have an opposite signal to the rest of East Africa when the ENSO phenomenon is considered (Indeje *et al.*, 2000). The northern coast of Kenya represented by Lamu receives rainfall mainly during the long rainfall season as shown by **figure 1.3d**.

Parts of the Rift Valley, Lake Victoria basin and most parts of Uganda exhibit the trimodal regime with the third rainfall peak being observed in July and August **(Figure 1.3e and 1.3f)**. Over Soroti in western Uganda, the main rainfall peak is observed during the long rainfall season as shown by **Figure 1.3e** while Nyahururu in Central Kenya, the highest peak was observed during the July-August period **(Figure 1.3f)**. It is worthy to note from **Figures 1.3a–f** that though different locations may have unimodal, bimodal or trimodal nature of rainfall distribution, the time of occurrence and its peak are observed at different times of the year. This alludes to the complexity of the systems that influence rainfall over the location in question which are discussed in **section 2.3**. Detailed discussion on rainfall distribution over the East Africa region can be found in Ogallo (1980) and Ininda (1995) among others. A brief outline on the organization of this research thesis is provided in the next section.

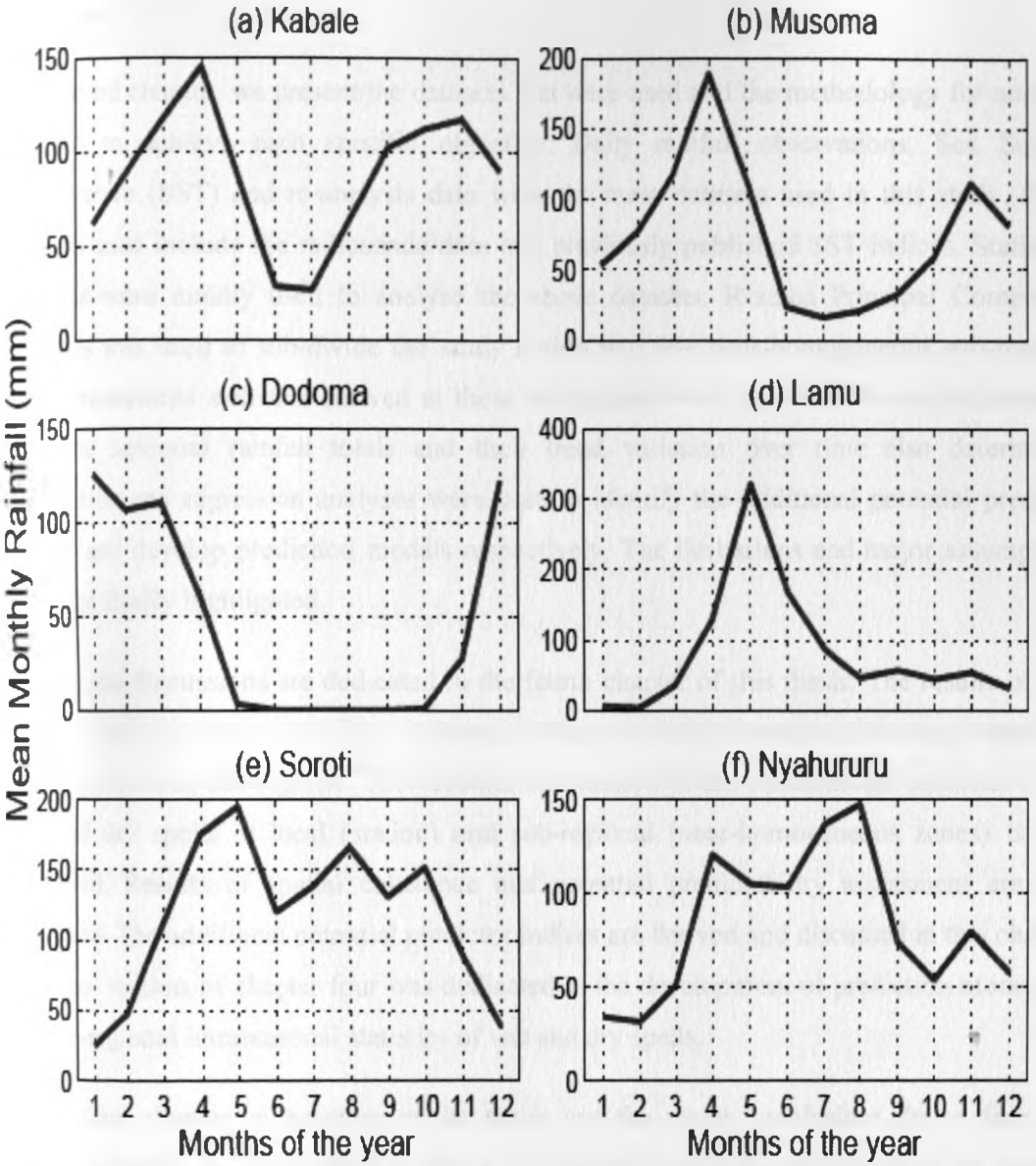


Figure 1.3: Patterns of annual cycle of rainfall distribution (1961 – 1990 average) for some selected stations over East Africa. Details of these stations are provided in Figure 3.1 and Table 3.1

1.6 Overview of the thesis

This thesis is organized into five major chapters, which are briefly outlined below. Chapter one provides the general introduction as well as the key objectives that were pursued in this study. The problem statement and justification of the study are also given. Also discussed are the physical features and rainfall climatology of the study domain. In the second chapter, all literatures that were relevant for the study are reviewed. The chapter also elaborates on the key climatic systems that influence the spatio-temporal distribution of the rainfall over the

study area.

In the third chapter, we present the datasets that were used and the methodology for analysis adopted to achieve each specific objective. Daily rainfall observations, Sea Surface Temperature (SST) and re-analysis data were the main datasets used in this study. Other datasets used include the radiosonde data and previously published SST indices. Statistical methods were mainly used to analyse the above datasets. Rotated Principal Component Analysis was used to sub-divide the study region into few near-homogeneous sub-regions. The intraseasonal statistics derived at these sub-regions were assessed for any relationship with the seasonal rainfall totals and their trend variation over time also determined. Correlation and regression analyses were used to identify the additional potential predictor indices and develop prediction models respectively. The limitations and major assumptions made are finally highlighted

Results and discussions are dedicated to the fourth chapter of this thesis. The results of data quality control are presented first, followed by those of the delineation of the study area into near-homogenous sub-regions. The baseline information of the intraseasonal statistics of the wet and dry spells at local (station) and sub-regional (near-homogeneous zones) is then presented. Results of spatial coherence and potential predictability assessment are then presented. The additional potential predictor indices are derived and discussed in this chapter. The final section of chapter four was dedicated to the development of prediction models for the sub-regional intraseasonal statistics of wet and dry spells.

In the final chapter, a summary of the thesis and the major conclusions drawn from the various analyses are highlighted, together with the recommendations that could be adopted and possibly explored further in future.

CHAPTER TWO
LITERATURE REVIEW

2.0 Introduction

Several studies have been carried out in an effort to understand the processes and systems associated with the spatio-temporal variability of rainfall at different timescales over the East Africa region. The recent past has seen an upsurge in studies aimed at assessment of the potential predictability of rainfall variability at different timescales. This literature review considered the above two aspects from previous studies dedicated to East Africa as well as other studies that are relevant to the current study.

As stated above, most of the studies have addressed monthly, seasonal, annual and longer timescales, with very little work at intraseasonal timescale.

2.1 Studies to understand the processes and systems

In this section, the literature highlighting studies dedicated to the intraseasonal variability of the rains over the region and their organization into wet / dry spells are reviewed first, followed by those at the interannual timescales. Those studies which analyses how intraseasonal characteristics of the rainfall vary at interannual timescales and how they have evolved over time are finally reviewed.

Washington and Todd (1999) have studied the variability of daily rainfall derived from satellite over Southern African-Southwest Indian Ocean from November to March. This study showed the leading mode of daily rainfall variability to be a tropical-temperate link spanning the latitudinal domain of the study. The study further indicated that these links have a parallel structure such that enhanced (suppressed) activity over Southern Africa in bands off the east coast are associated with suppressed (enhanced) activity over Southern Africa.

Mutai and Ward (2000) have indicated that the wet spells in East Africa are often associated with synoptic disturbances that migrate eastwards into Eastern Africa region in association with westerly near-surface wind anomalies.

Numerous studies have also used the Outgoing Longwave Radiation (OLR) as a surrogate for tropical rainfall (Nyakwada, 1991; Nogues-Paele and Mo, 1997; Okoola, 1998; Jagadheesha *et al.*, 2003; Okoola and Camberlin, 2003). This is based on an average of single morning and

evening passes of the satellite (Washington and Todd, 1999). Over East Africa which is within the tropics, the observed rainfall is dominantly from deep convective clouds. Some of these clouds extend as high as the tropopause levels and can therefore be seen by satellites as regions of cold temperatures and low OLR. The fact that spatial variations of temperature in the tropics are small makes it easier to interpret OLR data in the tropics.

Nyakwada (1991) studied the relationship between satellite derived outgoing longwave radiation (OLR) and some meteorological parameters. The study showed significant correlation between OLR and rainfall, with areal records giving better results as compared to the point records. Results from Principal Component Analysis (PCA) showed some similarities in the spatial and temporal characteristics of OLR and rainfall. Though the study confirmed that there exists a significant association between the OLR and rainfall and further developed regression equations, no attempt was made to forecast the rainfall using the developed regression equations.

The pattern and evolution of intraseasonal rainfall over East Africa and its teleconnections with the regional circulation have been studied by Mpeta and Jury (2001). Time-longitude Hovmoller plots of filtered anomalies of OLR and zonal winds at 850hpa level in the 7.5° to 10° S latitude band was used to reveal the nature of propagation and coupling of local circulation and convection. Time-longitude diagrams revealed eastward propagating and quasi-stationary features in the 7.5° to 10° S latitude band. Westward propagating features were found to be generally weak and short-lived. Many intraseasonal convective systems were found to pass across the Africa continent with small amplitude and propagate eastward into the Indian Ocean with increasing amplitude. Stronger equatorial convection and MJO activity were found to favour rainy conditions over East Africa and the adjacent west India Ocean, yet there was drier weather over much of sub-tropical Africa.

Okoola and Camberlin (2003) studied the intraseasonal oscillations associated with March - May rainfall in East Africa using pentad rainfall, OLR and NCEP global re-analysis datasets. The study depicted intraseasonal oscillations across equatorial East Africa with a 40 - 50 day periodicity that had large interannual variability. The cross-sectional analyses of the raw OLR showed eastward moving perturbations across equatorial Africa. The 20 - 75 day filtered OLR anomalies showed clearer eastward propagation. The study further noted that two or more active convection events were observed for most seasons while seasons with deficit rainfall had only one event. Space-lagged relationships in the convection between Gulf of

Guinea and Equatorial East Africa showed that convection over the Gulf of Guinea leads that over the Equatorial East Africa by 1 to 2 pentads, indicating that convection over Gulf of Guinea may be used in monitoring the start and subsequent performance in the Equatorial East Africa wet/dry events, especially when above normal seasonal rainfall are anticipated.

A study by Ngigi *et al.* (2005) over Laikipia district in upper Ewaso Ng'iro river basin of Kenya revealed that there is 80% probability of occurrence of dry spells exceeding 10 and 12 days during the long and short rainfall seasons respectively. The off-season dry spells, which occur after rainfall cessation, were longer and more severe than intraseasonal dry spells. The occurrence of off-season dry spells coincides with the critical crop growth stage especially the flowering and grain-filling stages.

Gitau (2005) studied the characteristics of wet and dry spells during the wet seasons over Kenya. The study using the wavelet method of analysis identified three wavelet bands in the occurrence of daily rainfall events. The wavelet bands identified were less than 10 days, 10 to 20 days and 20 to 32 days. The latter was associated with the lower modes of Madden-Julian Oscillation which have been noted in other parts of the world (Krishnamurti and Ardunay, 1980; Sikka and Gadgil, 1980; Kripalani *et al.*, 2004).

Other studies on the occurrence of the wet / dry spells over Eastern Africa include the works of Alusa and Gwage (1978), Ogallo and Chillambo (1982), Otengi and Ogallo (1984), Bazira and Ogallo (1985), Sharma (1996), Camberlin and Wairoto (1997) and Barrow *et al.* (2003) among others. A detailed review on other studies related with the occurrence of wet and dry spells over Kenya can be found in Gitau (2005), over Tanzania in Tilya (2006), and over Uganda in Bamanya (2007).

Besides the studies dedicated to the intraseasonal variability of the rains in the region and their organization into wet / dry spells, studies on the interannual timescale have been many. Studies by Ogallo (1988), Ogallo *et al.* (1988), Indeje (2000), Mutemi (2003), Owiti (2005) and Njau (2006) have clearly showed strong teleconnection between the seasonal rainfall totals on one hand and oceanic and atmospheric fields on the other hand.

Zorita and Tilya (2002) studied the rainfall variability in northern Tanzania in the March-May season and its links to large scale climate forcing. Monthly rainfall totals from 22 stations and spanning a period of 36 years (1963 – 1998) were used. The study used the sea-level pressure, air temperature, zonal and meridional wind components near the surface,

vertical velocity at 850mb level and winds at 200mb level all from the National Centre for Environmental prediction / National Centre for Atmospheric Research (NCEP/NCAR) re-analysis (Kalnay *et al.* 1996). Principal component analysis was undertaken on the monthly rainfall totals. Concurrent correlation analysis was used to analyse the association of the two leading principal components with the large scale climate forcing. The results indicated that the March and April rainfall anomalies are linked to zonal thermal contrast between the Indian Ocean and the Eastern African land mass, to zonal surface winds anomalies and to vertical velocity anomalies. On the other hand, May rainfall anomalies are associated with a meridional surface temperature contrast between the Indian Ocean and the Asian continent and meridional surface winds anomalies, indicating a relationship with the Indian Monsoon.

However, few studies considered the interannual variability of the characteristics of the rains apart from the seasonal rainfall totals. Ambenje *et al.* (2001) have analysed the trend in the seasonal precipitation and frequency of days with precipitation above some thresholds over 19 countries in eastern and southern Africa for the four standard seasons. The frequency of days with precipitation above 1mm, 12.5mm, 25.4mm, 50.8mm and 100mm were determined. Linear trend of the time series of the seasonal precipitation and frequency were then determined by linear regression. Results showed that there was a general tendency for trends of opposite signs to occur between the tropical (0° – 20° N/S) and subtropical latitudinal belts. Over equatorial eastern Africa, the results indicated that the seasonal precipitation and the associated frequency of days with precipitation above the various thresholds have decreased in the humid western parts and increased over the coastal and semi-arid regions to the east. The increase in seasonal precipitation over the coastal region and semi-arid zones were more pronounced during the September to November, and December to February seasons. This was associated with the warm phase of the El Niño / Southern Oscillation (ENSO) cycle which has occurred more frequently in the recent decades. Decreasing trend in the frequency of days with precipitation above 50.8mm was significant at 95% confidence levels during the March to May rainfall season over Equatorial Eastern Africa region. Trend in the frequency of days with precipitation above the moderate thresholds categories were however small in magnitude.

Moron *et al.* (2007) have examined the spatial coherence characteristics of daily station observations of rainfall over five tropical regions during the principal rainfall season(s). These regions were Senegal in West Africa, northern Queensland in Australia, northwestern

India, Nordeste in northern Brazil and Kenya in East Africa. This study considered three aspects of the rainfall which are seasonal rainfall total, daily rainfall frequency (number of wet days) and mean rainfall intensity (mean rainfall per rainy day). The study noted that mean rainfall frequency is the most coherent variable, followed closely by the seasonal total while the daily intensity was a distant third. Similar results had been obtained by Moron *et al.* (2006) over Senegal using 13 stations. It should be noted at this point that Kenya was represented by nine stations only during the two main seasons of long and short rains. Further, no attempt was made to identify regional subdivisions within the country.

2.2 Predictability studies and Forecast model development

Several studies have fitted the Markov chain models to the occurrence of the wet and dry spells over East Africa. These include the work of Ogallo and Chillambo (1982), Mungai (1984), Otengi and Ogallo (1984), Bazira and Ogallo (1985), Gitau (2005), Tilya (2006) and Bamanya (2007). These studies have shown that the first-order Markov chain models adequately describe the occurrence of the wet and dry spells over the eastern Africa region.

Ochola and Kerkides (2003) have used the concepts of conditional probability, Poisson probability distribution function and chi-square testing to develop a first-order Markov chain model that predicts the critical wet/dry spells over Kano plains in western Kenya. They found that the length of critical dry (wet) spell was 14 (12) days for the long rainfall season and 12 (8) days for the short rainfall season over Ahero Irrigation Scheme.

For India, Xavier (2002) showed that the evolution of intraseasonal oscillation of rainfall (dry and wet spells) is spatially and temporally coherent with that of circulation during the India summer monsoon. The study established potential predictability of the dry and wet spells from the 850mb relative vorticity. A forward stepwise multivariate linear regression model was developed and the skill of the predictions assessed at every step. The rainfall anomalies predicted by the empirical model were compared with the intraseasonally filtered rainfall anomalies and the model captured the extreme events with sufficient skill. Examination of these predictions indicated that predictions initiated from some initial conditions had more skill than others. It was found that 15-day predictions made from active or break conditions agreed much better with observations than those made from the transition initial conditions. Over Eastern Africa however, no such study on the predictability of the wet and dry spells is available.

At seasonal timescale and using ENSO index, an energy gradient from the East African highlands, 500-hPa geopotential height anomalies over the Near East and westerly winds from the Congo basin, Camberlin and Philippon (2002) developed seasonal multivariate linear regression prediction models for the March–May season over Kenya-Uganda with a multiple correlation coefficient of 0.66 in cross-validation mode. The multivariate linear regression (MLR) prediction model used the February predictors only due to the poor inter-monthly persistence of atmospheric and oceanic anomalies. The models main shortcoming was the absence of long lead-time for operational applications and practice.

Building on earlier results by Mutai *et al.* (1998) which identified SST predictors of the East Africa short rainfall season, Philippon *et al.* (2002) developed a prediction model for the seasonal rainfall totals during this season.

Hastenrath (2007) has shown strong concurrent correlation (-0.85) of short rains at the equatorial East Africa coast and the westerlies over the central equatorial Indian ocean. The equatorial westerlies drive the Wyrtki jet (Wyrtki, 1973) in the upper ocean and enhance the westward temperature gradient, a surface manifestation of powerful zonal–vertical circulation cell along the Equatorial Indian Ocean. Using the September values of a number of surface and upper air indices from equatorial zonal circulation cell as predictors, stepwise regression models were developed for the entire period (1958 – 1997) and separately for 1958 – 1977 (training period) and 1978 – 1997 (verification period). The evaluation of the results obtained showed that the correlations between the predictors and October-November rainfall series (the predictand) deteriorated although the equatorial zonal circulation cell remains strong throughout the entire period. The relation between the predictors and the predictand became very weak during the verification period.

Jury *et al.* (2009) found that the East African rainfall and zonal winds over the equatorial east Atlantic and West Indian Ocean which found an in-phase relationship. The strongest signal is a 2 to 2.3-year cycle from 1961 to 1968 and again in the late 1990s. The winds led rainfall by about 3 months from 1960 to 1970. However rainfall led wind by more than 3 months from 1970 to 1998. Further consideration of the East Indian Ocean zonal winds found a more robust teleconnection while cross-wavelet analysis revealed 2 to 4-year cycles and the time delay indicated that winds lead rainfall up to 8 months from 1982 to 1998. A model for OND seasonal rainfall developed using the central Indian Ocean zonal winds averaged over three months (JAS) was found to adequately hit 60% of the target categories but under-predicts the

intensity of big events.

From the foregoing discussions on previous studies, it has been observed that;

a. Most of the studies have concentrated on understanding the processes and systems based on the observed historical rainfall data. However other studies have used the outgoing long wave radiation as a surrogate of the observed rainfall. These studies cover both the interannual and intraseasonal timescales, but little work which combines the two timescales is available.

b. Some studies have made an effort to assess the predictability of the seasonal rainfall anomalies most based on the development of linear regression models. However, there is virtually no previous work available on the predictability of the intraseasonal statistics of the wet and dry spells. There is therefore the need to further our understanding on the intraseasonal statistics of the wet and dry spells in order to provide a more comprehensive picture on the evolution of rainfall activity within the season and assess its predictability.

c. The studies aiming at the prediction of seasonal rainfall anomalies have mainly concentrated on predictors with a one month lag which may be too soon for the users of such models. The monthly predictors that have been used are mostly released on 13/14 day of the next month which means that the models outputs will be available when the season have already started. There is therefore the need to consider predictors with longer time lags for the models outputs to be meaningful to the users. Alternatively, the variables/predictors which can be forecasted by the Global Circulation Models (GCMs) with a good skill could be used.

d. The few studies which have attempted to develop seasonal rainfall regression models have tended to concentrate mainly on the Indian Ocean and its circulation patterns without much consideration for other parts of the tropics. Other studies have also concentrated on the Central Pacific Ocean due to the influence of the ENSO phenomenon on the tropical climate. This study is aimed at considering the tropical region and parts of middle latitude in search for the predictors for seasonal rainfall and intraseasonal statistics of the wet and dry spells prediction. Apart from SSTs that are normally used in predictability studies, large-scale atmospheric predictors were also looked for. Despite the lower internal memory of the atmosphere as compared

to the ocean, previous studies have demonstrated the utility of these predictors, which also have the potential to be simulated by GCMs.

2.3 Systems that influence rainfall over the study domain

The spatio-temporal variability of rainfall over East Africa is controlled by a number of global, regional and local processes/systems. The variability results from complex interactions of these processes at various temporal scales. Observational studies have shown that the diurnal variation of precipitation in East Africa is largely determined by the mesoscale flows, the synoptic scale flows, and the interaction between the mesoscale and the synoptic scale flows (Asnani and Kinuthia, 1979; Mukabana and Pielke, 1996). The synoptic scale and higher scale circulations which affect weather and climate over East Africa include systems such as the monsoons, tropical cyclones, subtropical anticyclones, easterly and westerly wave perturbations, jet streams, global and regional modes of variability. These as well as the mesoscale systems are briefly discussed in the next sub-sections.

2.3.1 Inter-Tropical Convergence Zone

The Inter-Tropical Convergence Zone (ITCZ) may be defined as a narrow zone into which low-level tropical equatorward moving air masses from both hemispheres generally converge (Okoola, 1999a). It may be summarised as a zone marked with maximum cloudiness, humidity and precipitation; and minimum wind and pressure.

Over the East Africa region, the ITCZ has a rather complex structure consisting of zonal and meridional arms. The ITCZ is diffuse and thus difficult to locate at low levels but is detectable in the wind field near 700mb (Kiangi *et al.*, 1981). The structural complexity has been attributed to the geography of the Rift valley and the mountain chains of East Africa and the associated thermally-induced mesoscale circulations which makes the ITCZ patterns near the surface much diffused (Mukabana and Pielke, 1996). The zonal (conventional) arm has east-west orientation and oscillates in the north-south direction with the overhead sun. The double passage of the zonal arm of ITCZ over Eastern Africa region lagging behind the overhead sun is associated with the two rainfall seasons namely the long and the short rainfall seasons during which a large portion of the annual rainfall is received over Eastern Africa. The meridional arm which has a north-south orientation is formed by the convergence between the easterly winds from the Indian Ocean and moist westerlies from the Atlantic Ocean. This arm fluctuates from east to west and vice versa, with the easternmost extent

observed in July-August. The July/August rainfall received over most parts of Uganda, western Kenya and parts of Rift valley has been associated with the eastward extent of the westerlies from the Atlantic Ocean.

Over the East Africa, the ITCZ is the major synoptic-scale system that controls seasonal rainfall (Asnani, 1993; 2005). The fluctuations in the rainfall amounts and distribution have been attributed to the anomalies in the large scale factors that influence the characteristics of the ITCZ over East Africa region. The location of the ITCZ together with its overall horizontal and vertical structures largely depends on the intensity of the north-easterly and south-easterly winds which are in turn driven by the subtropical anticyclones. Comprehensive details of the ITCZ over East Africa region can be found in Ogallo (1993), Ininda (1995) and Okoola (1996) among others.

2.3.2 Monsoons

A monsoon is a wind in low-latitude climates that seasonally changes direction between winter and summer. Monsoons usually blow from the land in winter (called the dry phase, because the wind is composed of cool, dry air), and from water to the land in summer (called the wet phase, because the wind is composed of warm, moist air), causing a drastic change in the precipitation and temperature patterns on the area impacted by the monsoon.

The driving force for the monsoons is the differential heating of land and water surfaces by the solar radiation, which results in land-ocean pressure differences. The monsoonal winds are mostly confined to the tropics where the temperature contrast between the land and ocean is sufficiently high to generate the circulation. The monsoon is an important feature of atmospheric circulation, because large areas in the tropics and subtropics are under the influence of monsoons which bring humid air from over the oceans to produce rain over the land. The agricultural economies of impacted areas such as Asia and India frequently depend on the moisture provided by monsoon wind driven storm.

East Africa is subject to two monsoonal wind circulations, the Northeast (NE) and the Southeast (SE) monsoons. These monsoons coincide with the months of the year when the ITCZ is further from East Africa and thus are associated with relatively little rainfall (Okoola, 1999a). The northeast (NE) monsoon airstream occurs during the Northern Hemisphere winter (December to February) and emanates from the Arabian anticyclone which is situated on the Arabian Peninsula. It then recurves south of the equator to become a north-westerly

flow. The NE monsoonal winds have a sea trajectory of modest length thus they are warm and dry. The southeast (SE) monsoon current occurs during Northern Hemisphere Summer (June to August) and comes from the Mascarene highs over the southern Indian Ocean hence it is cool and moist. The flow then recurves north of the equator to become south-westerly.

Both monsoons are generally diffluent in the low levels and flow parallel to the coast. They are relatively shallow extending up to about 600hpa and capped aloft by an easterly flow resulting in a persistent inversion near 600hpa. The inversion inhibits cloud development, but it is occasionally broken by incursions of the westerlies (Okoola, 1982).

2.3.3 Tropical Cyclones

A tropical cyclone refers to an intense spiral storm that originates over warm tropical oceans and is characterized by low atmospheric pressure, strong winds and heavy rainfall. A characteristic feature of tropical cyclones is a warm centre with clear skies, light winds and low atmospheric pressure called the eye. Eye diameters are typically 40km but can range from under 10km to over 100km. The eye is surrounded by a dense ring of cloud about 16km high known as the eye wall which marks the belt of strongest winds and heavy rainfall. There is also a rapid variation of pressure across the storm which mostly occurs near the centre and resulting in very steep pressure gradient force, which is responsible for the strong winds present in the eye wall. Tropical cyclones derive their energy from the warm tropical oceans and do not form unless the Sea Surface Temperature (SST) is above 26.5°C, although once formed they can persist over lower SST.

Cyclones that affect the East Africa region (mostly southeastern coast of Tanzania) are those that form over Southwest Indian Ocean basin upto about 100° E. They generally occur from November to May but are more common during the months of January to March. On average, there are nine tropical disturbances a season, with about 50% of them reaching Tropical Cyclone (TC) status. However, their effect on East Africa weather may be indirect. Their formation during late March and early April often leads to delayed and below normal long rainfall over Eastern Africa region as was the case in 1984 (Okoola, 1999a). High frequency of the TC in the Mozambique Channel during 1984 resulted in winds being diverted to the Channel resulting into the non-establishment of the ITCZ over the Eastern Africa region during the long rainfall season. This led to loss of lives and livestock due to the drought that resulted.

2.3.4 Subtropical Anticyclones

These are synoptic-scale quasi-permanent pressure cells that form the descending arms of the tropical Hadley circulations. The pressure difference between the equatorial regions and the sub-tropical anticyclones drive the tropical trade winds. The four anticyclones affecting the synoptic flow over East Africa region are the Azores and Arabian anticyclones in the northern hemisphere (Griffiths and Solimani, 1972) and Mascarene and St. Helena anticyclones in the southern hemisphere (Van de Boogaard, 1977). The anticyclones are most intense during the winter season of each hemisphere and weaker during summer. The relative location, strength, structure and spatial orientation of these anticyclones determine whether they will pump in moist air or dry air over a region.

The Arabian anticyclone generates a stronger North Easterly (NE) flow during the short rainfall period than the South Easterly (SE) flow from the weaker Mascarene anticyclone. However, since the NE flow does not have long trajectory over the ocean as compared to the SE flow, it results in lesser rainfall during the September-November period.

The Mascarene and St. Helena are more pronounced during the southern hemisphere winter (June to August). The Mascarene anticyclone generally determines the characteristics of the moist SE monsoon flow over the Indian Ocean which influences rainfall over most of Eastern Africa. During the March-May season, the Mascarene anticyclone drives stronger and more moist SE flow into East Africa. Convergence of SE flow with the NE flow, both of which have stronger easterly component results into more rainfall in this season. The intensity and relative position of St. Helena anticyclone determines the position and depth of the quasi-permanent low pressure centre over central Africa, and therefore the intensity of the weather associated with it and how far to the east this weather will penetrate due to the strength of the meridional arm of the ITCZ.

2.3.5 Jet streams

A jet stream is a narrow, fast, upper atmospheric wind current, flowing at around 10 kilometers above the surface of the Earth. The jet stream may extend for thousands of kilometers around the world, but it is only a few hundred kilometers wide and usually less than 1.6 kilometers thick. A jet stream forms at the boundaries of adjacent air masses with significant differences in temperature. The jet stream is thus mainly found near the tropopause, at the transition between the troposphere (where temperature decreases with

height) and the stratosphere (where temperature increases with height).

The two jet streams that affect the weather and climate over the East Africa region are the Turkana Jet stream and the East Africa Low Level Jet stream (EALLJ). The Turkana jet stream is a strong SE low level jet in the Turkana Channel which separates the Ethiopian Highlands and the East Africa Highlands. This jet stream exists throughout the year, with the morning winds being stronger than the afternoon winds, mainly due to stronger vertical mixing and dilution of the jet maximum in the afternoon (Kinuthia and Asnani, 1982). Details of the Turkana Jet stream can be found in Kinuthia and Asnani (1982), Kinuthia (1992) and Indeje *et al.*, (2001) among others.

The East Africa Low Level Jet (EALLJ) stream occurs near the coast of East Africa. This jet stream is one of the major well-recognized cross-equatorial flows that have been studied through observational and numerical models (Findlater 1966; 1977; Krishnamurti et al. 1976; among others). The jet core is generally located between 1 and 1.6 km above the mean sea level and is associated with flows across the equator carrying Southern Hemisphere air northward up the African continent and ending at the Indian subcontinent. This jet stream induces strong currents and upwelling over the western equatorial Indian Ocean. It thus plays an integral role in the seasonal development of the Somali Current, an intense ocean current that flows northward only during the southwest monsoon. The jet builds during the months of April and May, becomes more pronounced in June to August and decays in September and October, during which the flow reverses to NE monsoons. Its horizontal divergence and vertical wind shear leads to dry conditions over East Africa.

2.3.6 Global and regional modes of climate variability

A mode of variability is a climate pattern with identifiable characteristics, specific regional effects, and often oscillatory behavior. Many modes of variability are used as indices to represent the general climatic state of a region affected by a given climate pattern. Such modes of variability may be found closer or far away from the target area, yet have an effect on the latter.

Climate dynamics research has demonstrated the existence of several modes of climate variability. The large scale modes of climate variability that relates to the East Africa rainfall include the El Niño/Southern Oscillation (ENSO), Indian Ocean Dipole (IOD) Mode, Quasi-Biennial Oscillations (QBO) and Intraseasonal Oscillations (ISO) among others.

2.3.6.1 Quasi-Biennial Oscillations

The Quasi-Biennial Oscillation (QBO) is a quasi-periodic reversal of the equatorial zonal wind between easterlies and westerlies in the tropical stratosphere with a mean period of 23 to 30 months averaging at about 26 months. The alternating wind regimes develop at the top of the lower stratosphere and propagate downwards at about 1.2 km per month until they are dissipated at the tropical tropopause. At the top of the vertical QBO domain, easterlies dominate, while at the bottom, westerlies are more likely to be found.

Several studies have confirmed the presence of the QBO in various atmospheric parameters. Some variables that have exhibited QBO include temperature, ozone, Indian monsoon and Africa rainfall (Ogallo *et al.,* 1994; Indeje and Semazzi, 2000). A study by Indeje and Semazzi (2000) has shown that about 36% of rainfall variability over Eastern Africa during the long rainfall season is associated with the QBO in the lower equatorial stratospheric zonal winds and further suggested that the relative role of QBO and rainfall over Eastern Africa is stronger in the time-lag sense than the simultaneous relationship.

2.3.6.2 El Niño / Southern Oscillation

El Niño / Southern Oscillation (ENSO) is a set of interacting parts of a single global system of coupled ocean-atmosphere climate fluctuations that come about as a consequence of oceanic and atmospheric circulations.

ENSO is the largest coupled ocean–atmosphere phenomenon resulting in climatic variability on interannual time scales (Godı´nez-Dominquez *et al.,* 2000). This wide ranging influence of ENSO has attracted the attention of the global climate community, particularly due to the well-documented economic and societal impacts, both today and throughout historical times, recorded locally and globally, within a wide latitudinal band about the equator.

El Niño which is the oceanic component of ENSO refers to the anomalous and sustained warming of the Sea Surface Temperature anomalies of magnitude greater than 0.5°C across the central and eastern tropical Pacific Ocean. The cooling phase is referred to as La Niña. When the anomaly is met for a period of less than five months, it is classified as El Niño or La Niña conditions; if the anomaly persists for five months or longer, it is classified as an El Niño or La Niña episode.

The atmospheric signature of ENSO, the Southern Oscillation (SO) reflects the monthly or

seasonal fluctuations in the air pressure difference between Tahiti and Darwin. In using the Southern Oscillation Index (SOI) based on just two stations, it must be recognized that there are many small-scale and high frequency phenomena in the atmosphere, such as the Madden–Julian Oscillation that can influence the pressures at stations involved in forming the SOI but that do not reflect the Southern Oscillation itself. As such, a 5-month running mean of SST anomalies and SOI is made in order to smooth out the possible intraseasonal variations in the tropical ocean.

While ENSO events show basically in phase variations between the Pacific and Indian Oceans, their signature in the Atlantic Ocean lag behind the Pacific events by 12 to 18 months. Many of the countries most affected by ENSO events are developing countries whose economies are largely dependent upon their agricultural and fishery sectors as a major source of food supply, employment and foreign exchange.

ENSO is the most prominent known source of interannual climate variability around the world including Eastern Africa with an irregular cyclicity of 3 to 8 years. Many studies have investigated the relationship between East African rainfall and ENSO. Mutemi (2003) found a strong relationship between rainfall over East Africa and evolutionary phases of ENSO. Shifts in the onset/cessation of rainfall patterns over some regions were observed while in others significant reduction in the seasonal peak was evidenced. Nicholson and Kim (1997) observed that ENSO has little influence on the long rainfall season but significantly modulates rainfall during the short rainfall season. Ogallo (1988) found significant instantaneous and time lagged negative correlations between East African seasonal rainfall and the Southern Oscillation Index (SOI). By correlating the global SST anomalies within the tropics (30° N/S) with the rotated principal component analyses (RPCA) modes of the autumn rainfall over Eastern Africa, Ogallo *et al.* (1988) found that 36% of the short rainfall variation in East Africa could be explained by SST variations in western Pacific and most of Indian Ocean.

Using an atmospheric General Circulation Model (GCM) forced with various combinations of Indian and Pacific SST anomalies, Goddard and Graham (1999) observed that while the SST variability of the tropical Pacific exerts some influence over the African region, it is the atmospheric response to the Indian Ocean variability that is essential for the model simulating robust rainfall response over eastern, central and southern Africa. This may point to the importance of the Indian Ocean Dipole (IOD) in climate studies which is discussed next.

Further details of the ENSO influence over East Africa can be found in Mutemi (2003), Ogallo (1988) and Ogallo *et al.* (1988) among others.

2.3.6.3 Indian Ocean Dipole

Previous studies have identified a unique ocean–atmosphere mode characterized by anomalously warm SSTs over the western Indian Ocean and anomalously cold SSTs in the eastern Indian Ocean (Saji *et al.*, 1999; Owiti, 2005; Owiti *et al.*, 2008). The evidence indicates that Indian Ocean SST anomalies have a significant impact on regional atmospheric circulation and rainfall anomalies that extend into Eastern and Southern Africa. As the wind flow entering East Africa mostly originates from the Indian Ocean, it would be reasonable to assume that Indian Ocean Dipole (IOD) SST anomalies would have a marked influence on the moisture supply to the adjacent landmasses (Reason, 2001).

Indian Ocean Dipole (IOD) refers to the occasional occurrences of see-saw SST anomalies over the southeastern and western parts of equatorial Indian Ocean (**Figure 2.1**). The difference between mean SST anomalies observed in tropical western Indian Ocean (50° E – 70° E, 10° S – 10° N) and tropical southeastern Indian Ocean (90° E – 110° E, 10° S – Equator) has been used to quantify the zonal temperature gradient representative of the IOD (Saji *et al.*, 1999).

Analysis on the evolutional phases of IOD index by Owiti (2005) and Owiti *et al*, (2008) indicate the significant SST anomalies begin to appear around April, attains maximum peak around October/November and starts decaying in January. Most cycles do not extend beyond one year. As such, the significant association between the IOD and Eastern Africa regional rainfall is stronger during the short (OND) rainfall season while the correlation values are generally not significant during the long (MAM) rainfall seasons.

Available records show that at times the strong positive (negative) IOD events co-occurred with El Niño (La Niña) episodes. This may be indicative of some possible interactions between ENSO and IOD. However, some strong IOD events were observed in non-ENSO events. A study by Trenberth (1997) indicate that warming over the western Indian Ocean during the ENSO events is associated with high moisture fluxes over the marine boundary layer. The increased tropospheric moisture associated with the warm El Niño events is advected into the Eastern Africa by the relatively strong easterly wind flow during the wet seasons. The advected moisture supports enhanced convection and orographic precipitation

through latent heat release thus sustaining wet conditions over the East African region. Comprehensive details of the IOD over East Africa region can be found in Saji *et al.* (1999); Black *et al.* (2003); Clark *et al.* (2003); Black (2005); Owiti (2005) and Owiti *et al.* (2008) among others.

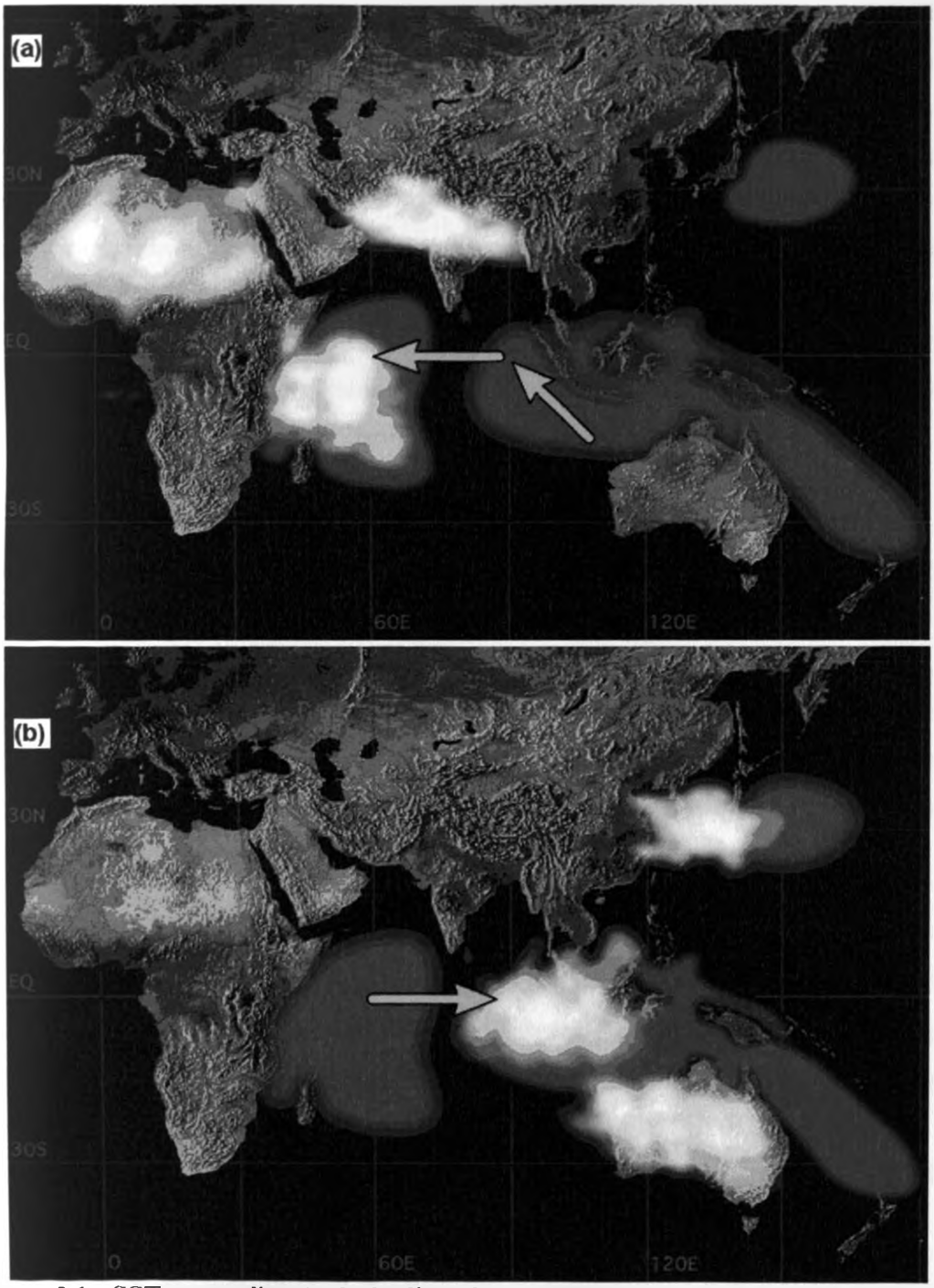


Figure 2.1: SST anomalies (red shading denotes warming; blue-cooling) during (a) positive and (b) negative Indian Ocean dipole (IOD) event. (*Source A. Suryachandra Rao, Institute for Global Change Research, Japan)*

2.3.6.4 Intraseasonal Oscillations

Studies have shown that intraseasonal oscillations (ISO) are present in the proxies of the rainfall such as outgoing longwave radiation over the tropical region (Anyamba, 1990; Soden and Fu, 1995; Barr-Kumarakulasinghe and Lwiza, 1998; Omeny, 2006). A study by Gitau (2005) over Kenya has suggested the existence of ISO in the occurrence of the daily rainfall events. A quasi biweekly oscillation with 10 to 20 days periodicity has been found in the occurrence of rainfall events (Okoola, 1989; Gitau, 2005). Another form of the intraseasonal oscillations that is most prominent in the tropical region is the Madden-Julian Oscillation (MJO). The Madden-Julian Oscillation plays an important role in climate variability and has a significant influence on medium-to-extended ranges of weather forecasting in the tropics (Jones *et al.* 2000; Pohl and Camberlin, 2006; Omeny, 2006; Omeny *et al.*, 2008). Goswami *et al.* (2003) have suggested that the slow evolution of the monsoon intraseasonal oscillations on account of the 30 - 60 days dominant periodicity could make it potentially predictable by up to about three weeks in advance during the Indian summer monsoon.

2.3.7 Mesoscale systems/features

Mesoscale systems are small-scale weather systems with the horizontal dimension ranging from 5 to 500 km and typically possessing lifetimes of a day or less. They cannot therefore be observed on synoptic charts. For such systems, the vertical motion is as important as the horizontal ones and Coriolis force has little or no effect due to the short lifetime or the over-riding magnitude of other forces. Proximity to the ocean, varied topography and existence of large inland lakes induces vigorous mesoscale circulations with a strong diurnal cycle over several parts of the East Africa region.

2.3.7.1 Effects of orography

Spatial distribution of weather in East Africa is to some extent determined by the interactions between the quasi-stationary mesoscale circulations and the seasonally varying large scale flow. By modeling the interaction of the mesoscale circulation and synoptic scale circulations, Mukabana and Pielke (1996) and Indeje *et al.* (2001) demonstrated that orography plays a role in causing rainfall at nearly all places in Kenya and East Africa respectively.

Oettli and Camberlin (2005) have defined statistical models to explain the spatial distribution of rainfall in Eastern Africa (southern Kenya and NE Tanzania) based on various

topographical descriptors. The results indicated that the north–south exposure contrasts are the main factor of rainfall variation, except in the northern summer (June to September). South-facing stations are wetter, especially during the long rainfall (March to May) season since southerly winds are slightly wetter than those with a northerly component. East-facing stations are wetter in the short rains season (October to December) and drier in the monsoon season. These variations coincide with seasonal atmospheric circulation changes over the study region. The study finally concluded that mean elevation had little effect on the monthly rainfall while other factors especially north-south exposure describe the interaction between rainfall and topography more adequately.

2.3.7.2 Land and Sea/Lake Breezes

These are diurnal local winds that are generated as a result of the different specific heat capacities of the water and land near the shores.

The sea/lake breeze is one of the most frequently occurring mesoscale weather systems. The sea/lake breeze refers to a diurnal, thermally driven circulation in which a surface convergence zone often exists between airstreams having over-water versus over-land histories. It results from the unequal sensible heat flux of the lower atmosphere over adjacent solar-heated land and water masses. Because of the large specific heat capacity of a water body, the air temperature changes little over the water while over land, the air mass warms during daytime. Occurring during periods of fair skies and generally weak large scale winds, the sea/lake breeze is recognizable by a wind shift to onshore, generally several hours after sunrise.

The reverse occurs at night, the land cools off quicker than the ocean due to differences in their specific heat capacities, which forces the dying of the daytime sea/lake breeze. If the land surface temperature drops below that of the adjacent sea/lake, the pressure over the water will be lower than that of the land, setting up a land breeze. The colder air from the land then moves offshore. Typically, the land breeze circulation is much weaker and shallower than its daytime counterpart, the sea/lake breeze.

Breeze circulations are created within the vicinity of Lake Victoria and along the coast. Sea/Lake breeze dominates during the afternoon/evening. The katabatic (drainage) winds coupled with the land breeze, dominate during late night/early morning up to at least 100 km from the shore. This circulation interacts with the seasonal flow and forces convection up to a

distance of even 150-200 km from the Lake Victoria shore (Mukabana, 1992; Okeyo, 1987).

The occurrence and strength of the both sea/lake and land breezes is controlled by land-sea surface temperature differences, the synoptic wind and its orientation with respect to the shoreline; the thermal stability of the lower atmosphere and the geometry of the shoreline and the complexity of the surrounding terrain.

CHAPTER THREE
DATA AND METHODS

This chapter provides the description of the datasets that were used in the current study to achieve the objectives discussed in **section 1.3**. It also provides the methodology that was adopted.

3.1 Datasets

Several secondary datasets were used in this study. These are the observed daily rainfall amounts, Hadley centre Sea Surface Temperatures (SSTs), NCEP/NCAR and ERA40 re-analysis data, radiosonde observations, the indices of Niño, Indian Ocean Dipole (IOD), and Sea Surface Temperature (SST) gradients. These datasets covered about 40 years starting from 1958. The daily rainfall dataset covers the East Africa region while the SSTs and re-analysis data covered the tropical region and part of the mid-latitudes (50° N - 50° S). Radiosonde observations were obtained over Nairobi in East Africa and Bangui in Central Africa.

Like the rest of Africa, East Africa continues to experience some difficulties with the availability of long-time climate data (see Figure 1 in Camberlin and Philippon, 2002). The available surface observations are rather sparse and their number has tremendously reduced over time. Each of the three East Africa countries has one operational upper-air observation station (Njau, 2006) out of which two have a lot of missing data.

3.1.1 Rainfall data

The observed daily rainfall amounts for 36 stations across the three East Africa countries and extending from January 1962 to December 2000 was used in this study. The amount of missing data from each station is highly variable (at most 7%). At times, data are missing for all the days in a month since the report forms are filled and sent to the headquarters of the National Meteorological services on a monthly basis. In such a case, the report forms were sourced from the Headquarters of the National Meteorological services and used to fill the gaps. However such cases were quite few.

The spatial distribution of the stations with long un-interrupted time series was carefully selected in order to minimize the amount of the missing data. At the same time, an evenly distributed gauge network throughout the study region was required. **Figure 3.1** shows the

spatial distribution of the Eastern African stations used in the study. **Table 3.1** which gives the names of the station used, their location and elevation. Based on the requirement of this study for a long un-interrupted time series of daily rainfall observations with few missing data points, the network of the station was assumed to be the most representative of the daily rainfall climatology over the study area. This dataset was obtained from the archives of Kenya Meteorological Department, IGAD Climate Prediction and Applications Centre (ICPAC) both of which are in Nairobi, Kenya and the Centre de Recherches de Climatologie (CRC) at Université de Bourgogne in Dijon, France.

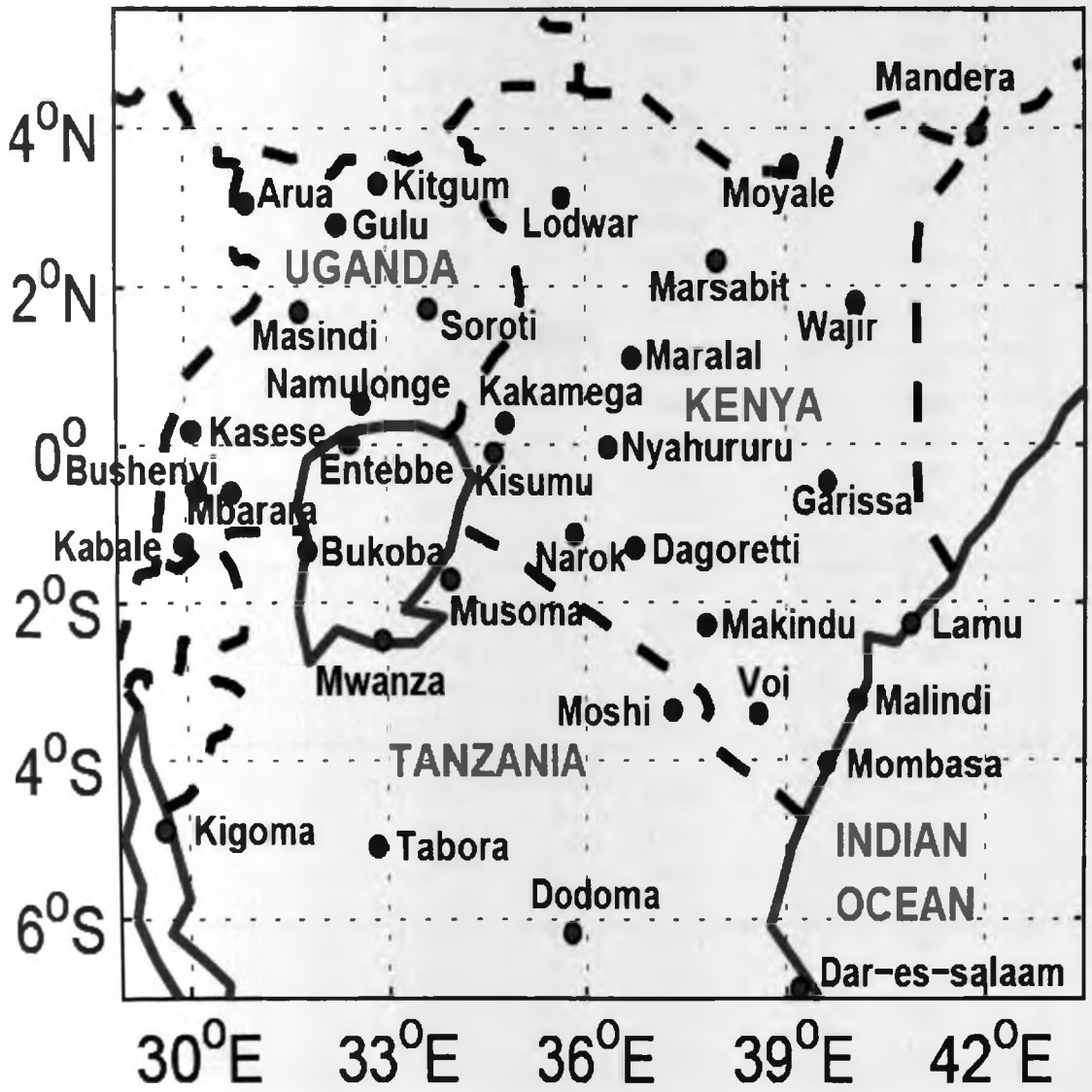


Figure 3.1: Network of the East African rainfall stations used

Table 3.1: Details of the East African rainfall stations used in the study

No	Stations	Latitudes	Longitudes (°E)	Elevation in M (AMSL)
		KENYA		
1	Mandera	3.93°N	41.87	230
2	Moyale	3.53°N	39.05	1113
3	Lodwar	3.12°N	35.62	566
4	Marsabit	2.32°N	37.98	1219
5	Maralal	1.10°N	36.70	1951
6	Wajir	1.80°N	40.07	244
7	Kakamega	0.28°N	34.75	1555
8	Nyahururu	0.03°S	36.35	2374
9	Kisumu	0.10°S	34.58	1146
10	Garissa	0.47°S	39.63	128
11	Dagoretti	1.30°S	36.75	1798
12	Narok	1.13°S	35.83	1890
13	Lamu	2.27°S	40.90	9
14	Makindu	2.28°S	37.83	1000
15	Malindi	3.23°S	40.10	3
16	Voi	3.40°S	38.57	579
17	Mombasa	4.03°S	39.62	57
		UGANDA		
18	Kitgum	3.30°N	32.88	940
19	Arua	3.05°N	30.92	1280
20	Gulu	2.78°N	32.28	1106
21	Soroti	1.72°N	33.62	1127
22	Masindi	1.68°N	31.72	1146
23	Namulonge	0.53°N	32.62	1150
24	Kasese	0.18°N	30.10	691
25	Entebbe	0.03°N	32.45	1183
26	Bushenyi	0.57°S	30.17	1590
27	Mbarara	0.60°S	30.68	1412
28	Kabale	1.25°S	29.98	1867
		TANZANIA		
29	Bukoba	1.33°S	31.82	1143
30	Musoma	1.70°S	33.93	1147
31	Mwanza	2.47°S	32.92	1139
32	Moshi	3.35°S	37.33	869
33	Kigoma	4.88°S	29.67	999
34	Tabora	5.08°S	32.83	1182
35	Dodoma	6.17°S	35.77	1120
36	Dar es salaam	6.87°S	39.20	53

3.1.2 Re-analysis data

Re-analysis data refers to a gridded dataset representing the state of the Earth's atmosphere, incorporating direct observations, remote-sensed observations and global climate model (GCM) output. Data from different sources such as surface, ship, aircraft, radiosonde, satellite and GCMs are quality controlled, merged and interpolated at grid points (assimilated) to obtain the re-analysis data.

Different types of re-analysis data have been developed over time by different climate centres. These include the re-analysis data from the National Centre for Environmental Prediction (NCEP) and the National Centre for Atmospheric Research (NCAR) as documented by Kalnay *et al.* (1996), ERA40 and ERA-interim from the European Centre for Medium-range Weather Forecast (ECMWF) as documented by Uppala *et al.*, (2005) and JRA25 from the Japanese Meteorological Agency among others. In the current study, the JRA dataset could not be considered since it spans a short duration of 26 years only from January 1979 to December 2004 for JRA (Kazutoshi *et al.*, 2005). The ERA-interim dataset could not be used for similar reason.

The NCEP/NCAR re-analysis and ERA40 datasets are both gridded to a horizontal resolution of 2.5° latitude by 2.5° longitude. Though the data assimilation system remained unchanged over the re-analysis periods to help eliminate perceived climate jumps associated with changes in the real time data assimilation system, the NCEP/NCAR re-analysis data is still affected by changes in the observing systems (Kalnay *et al.* 1996; Kanamitsu *et al.*, 2002). These two re-analysis datasets have been used in Eastern Africa with satisfactory results in Mutai and Ward (2000), Zorita and Tilya (2002), Camberlin and Philippon (2002), Black *et al.* (2003), Pohl *et al.* (2005) for the NCEP/NCAR and in Mukabana and Pielke (1996) and Okoola (1999b; 1999c) for the ERA40.

The NCEP/NCAR re-analysis dataset was downloaded from the database of the National Oceanic and Atmospheric Administration (NOAA) website while the ERA40 dataset was downloaded from the database of the European Centre for Medium-range Weather Forecasts (ECMWF) website. The two re-analysis datasets were compared with the radiosonde data over the study area and the surrounding regions. The re-analysis dataset that mostly replicated the radiosonde data was thus adopted and used for further analysis to accomplish the objectives of the study. It should however be noted that re-analysis dataset are not simple interpolation of the observed data to the grid-points and therefore we do not expect a perfect

match between the observed radiosonde and re-analysis data.

From the re-analysis dataset that mostly replicate the radiosonde data, four variables were extracted at 925mb, 700mb and 200mb levels representing the lower, middle and upper atmospheric levels. The four variables extracted were the zonal (u) and meridional (v) components of wind vector, the specific humidity (q) and the geopotential heights (φ).

Due to the importance attributed to the circulation patterns while studying the rainfall patterns, it was necessary to ascertain the re-analysis dataset that closely replicated the measured radiosonde data.

3.1.3 Radiosonde data

Radiosonde sounding systems use *in situ* sensors carried aloft by a small, balloon-borne instrument package, the radiosonde, to measure vertical profiles of atmospheric pressure, temperature, and moisture (relative humidity or wet bulb temperature) as the balloon ascends, and transmit the data to a ground-based receiver and data acquisition system. A rawinsonde is a radiosonde that is designed to also measure wind speed and direction. Rawinsondes are commonly referred to as radiosondes. The radiosonde electronic subsystems sample each sensor at regular intervals.

Upper-air winds (horizontal wind speed and direction) are determined during radiosonde ascents by measuring the position of the radiosonde relative to the earth's surface as the balloon ascends. By measuring the position of the balloon with respect to time and altitude, wind vectors can be computed and represent the layer-averaged horizontal wind speed and wind direction for successive layers.

An upper-air station exists in each of the three countries considered in this study. They are located at Nairobi (01° 18'S, 36° 45'E), Entebbe (00° 03'N, 32° 27'E) and Dar-es-Salaam (06° 50'S, 39° 12'E). Njau (2006) observed that the later two stations had a lot of data missing and hence unsuitable for analysis. In view of this, Nairobi upper–air station was chosen to represent the East Africa region.

The zonal and meridional wind components of radiosonde wind data at different standard pressure levels from two locations (Nairobi in Kenya and Bangui in Central Africa Republic) spanning a period of 30 years and extending from January 1959 to December 1988 was used to assess the quality of the re-analysis datasets. Bangui radiosonde station is located at 04°

24'N, 18° 31'E at an altitude of 366 M above Sea Level (Duree *et al.*, 2006). The radiosonde data for Bangui upper-air observational station was downloaded from the Research Data Archive at the National Centre for Atmospheric Research (NCAR) website[**]. The Bangui upper-air station was used because it is located over a region (Central Africa) where NCEP/NCAR and ERA40 re-analyses shows large discrepancies, yet it is important source of moisture advection from Congo Basin which significantly affect rainfall variability in East Africa. Another reason for the choice of Bangui was that the observed data has a common time overlay with the Nairobi upper-air data yet it is outside the study area but within the equatorial region.

3.1.4 Hadley centre sea surface temperature

Sea-Surface Temperatures (SSTs) play an important role in modulating rainfall variability. Idealised SST anomalies have been used to force global and regional circulation models to simulate rainfall variability and study the physical mechanisms behind the variability over various regions, such as the Sahel (Moron *et al.*, 2003), northwest Africa (Li *et al.*, 2003), Africa as a whole (Paeth and Friederichs, 2004), and southern Africa (Reason, 2002; Misra, 2003).

Many of these studies have suggested the importance of SST in modulating rainfall variability, either indirectly such as an alteration in the position of the ITCZ (Biasutti *et al.*, 2004) or by more direct 'local' mechanisms (Janowiak, 1988; Walker, 1990; Jury and Pathack, 1993; Jury *et al.*, 1993; Mason, 1995; Shinoda and Kawamura, 1996; Reason and Lutjeharms, 1998). Over the Eastern Africa region, SSTs and SST-derived indices have regularly been used for various studies including the seasonal rainfall prediction (Ogallo *et al.*, 1988; Nicholson and Kim, 1997; Mutai *et al.*, 1998; Latif *et al.*, 1999; Indeje *et al.*, 2000; Black *et al.*, 2003; Mutemi, 2003; Owiti, 2005; Nyakwada, 2009).

The Hadley centre SSTs used in this study are gridded to a horizontal resolution of 1° latitude by 1° longitude and covered 45° N/S latitude but spanned all longitudes (Rayner *et al.*, 2003). The monthly Hadley Centre SST covered a period of 40 years from January 1961 to December 2000.

3.1.5 Other datasets used

[**] www.dss.ucar.edu/datasets/ds430.0

Several other datasets, in the form of climatic indices, were used in this study. These include the Niño indices which were downloaded from the Climate Prediction Centre (CPC) website[††], the IOD indices which have been documented by Owiti (2005) and the SST-gradients developed by Nyakwada (2009). These indices are used over the Greater Horn of Africa countries (of which the study region forms part of) for seasonal rainfall prediction by IGAD Climate Prediction and Application Centre. These datasets are all at monthly timescale and cover the period 1961 to 2000. The locations where the IOD index is derived from were highlighted in **sub-section 2.3.6.3**. The locations where Niño and SST-gradient indices are derived are discussed next.

Figure 3.2 and **Table 3.2** shows the locations used to compute different Niño indices. The Niño indices have significant association with the seasonal rainfall over the eastern Africa region especially during the short rainfall season (Ogallo, 1988; Ogallo *et al.*, 1988; Indeje *et al.*, 2000; Mutemi, 2003).

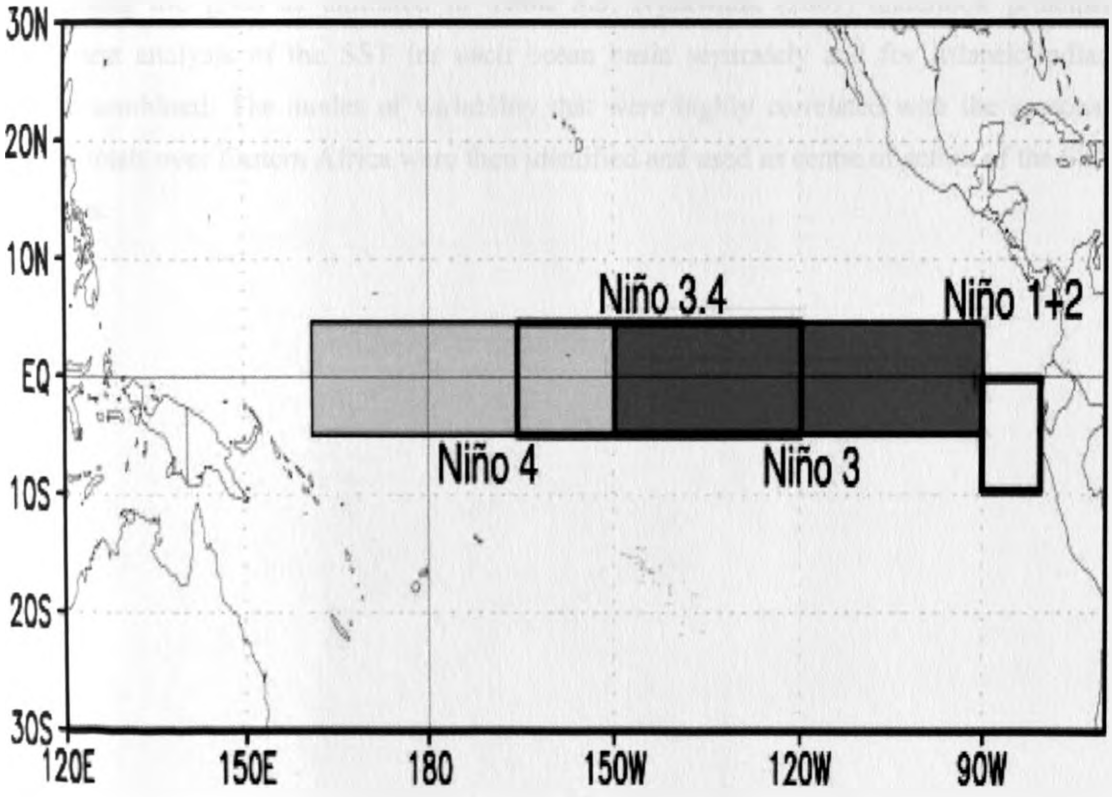


Figure 3.2: Graphical depiction of the four Niño regions (source: CPC)

[††] www.cpc.noaa.gov/data/indices/sstoi.indices

Table 3.2: The coordinates used to compute the various Niño indices (source: CPC)

	Longitudes (°)	Latitudes (°)
NIÑO 1	80 – 90 W	5 – 10 S
NIÑO 2	80 – 90 W	0 – 5 S
NIÑO 1+2	80 – 90 W	0 – 10 S
NIÑO 3	90 – 150 W	5 N – 5 S
NIÑO 4	150W – 160 E	5 N – 5 S
NIÑO 3.4	120 – 170 W	5 N – 5 S

Figure 3.3 gives the locations used to compute the zonal and meridional SST-gradient modes with the highest relationships with seasonal rainfall over East Africa (Nyakwada, 2009). The gradient modes are computed in the direction of the arrows from the SST anomalies representing the grids as indicated in **Table 3.3**. Nyakwada (2009) undertook principal component analysis of the SST for each ocean basin separately and for Atlantic-Indian Oceans combined. The modes of variability that were highly correlated with the seasonal rainfall totals over Eastern Africa were then identified and used as centre of action of the SST gradients.

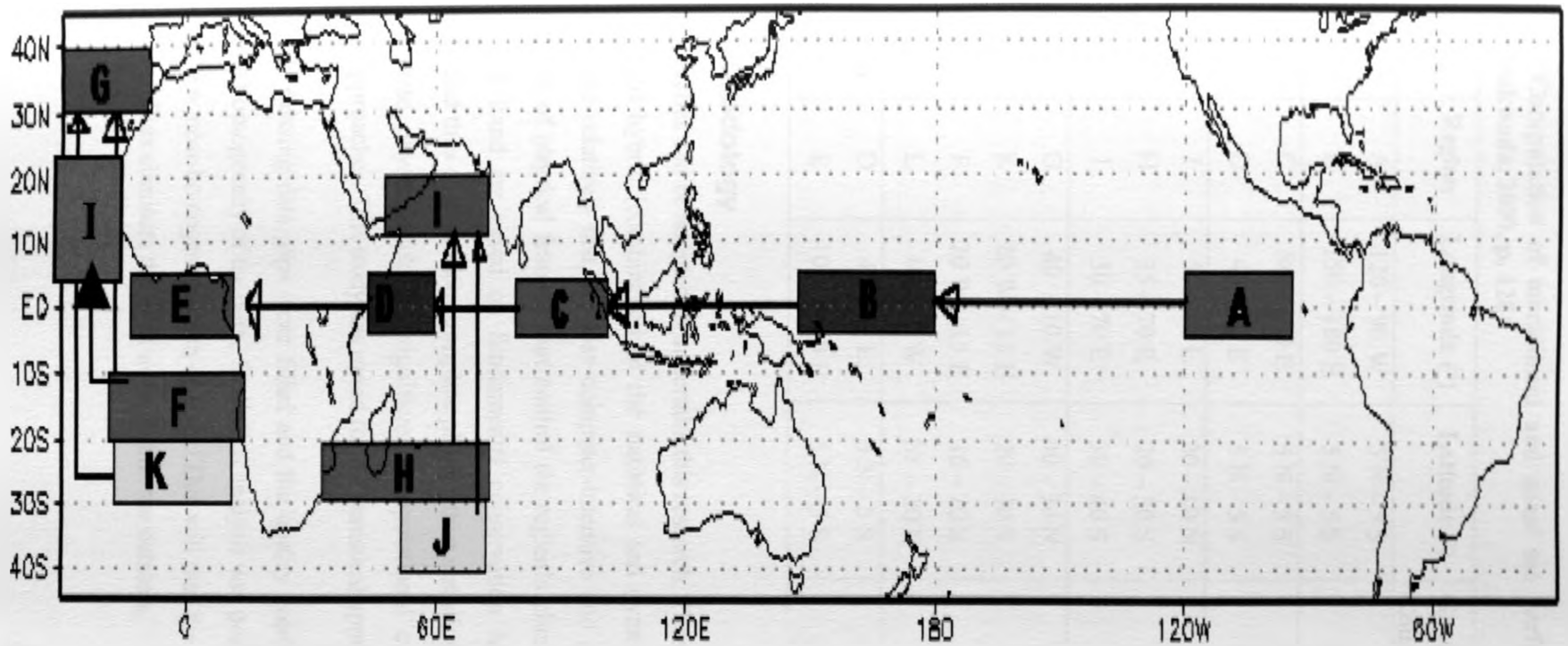


Figure 3.3: The locations used to compute the sea surface temperature gradients (Source: Nyakwada, 2009, p. 127)

Table 3.3: Computation of meridional and zonal sea surface temperature gradient (Source: Nyakwada, 2009, p. 128)

Ocean	Region	Longitude (°)	Latitude (°)	Gradient Computation	Gradient Name
Pacific	A	120 – 90 W	5 N – 5 S		
	B	150 – 180 E	5 N – 5 S	B-A	ZPAC
Indian	C	80 – 100 E	5 N – 5 S		
	D	40 – 60 E	5 N – 5 S	D-C	ZIND
	I	50 – 75 E	20 – 10 N		
	H	35 – 70 E	20 – 30 S	I-H	MIB1
	J	30 – 70 E	30 – 40 S	I-J	MIB3
Atlantic	G	40 – 10 W	40 – 30 N		
	K	20 W – 15 E	20 – 30 S	G-K	MAB3
	F	20 W – 15 E	10 – 20 S	G-F	MAB6
	L	40 – 15 W	10 – 20 N	L-F	MAB4
Atlantic and Indian	D	40 – 60 E	5 N – 5 S		
	E	10 W – 10 E	5 N – 5 S	E-D	ZAF

3.2 Methodology

Three approaches can be adopted to undertake this research. These approaches are statistical, dynamical and hybrid (combination of the statistical and dynamical approach). Statistical approaches are relatively fast, are less computer-intensive and generally strive for concise representation of physical features that control the region's climate. Dynamical approaches on the other hand are based on fundamental conservation laws for mass, energy and momentum and thus contain more complete physics than statistical approaches. However, the more complete physics evokes a significant computational cost that limits the use of dynamical approaches. This study has opted for the statistical approach as a research design.

Initially, the missing data gaps were filled and the quality control measures undertaken to ascertain the homogeneity of the data. Statistical analysis was performed to classify the study region into few near-homogeneous sub-regions. This will ease the interpretation of the results obtained as well as eliminate the local noise within the datasets.

3.2.1 Missing data and Quality control

The biggest drawback in long-term meteorological time series analysis is that recorded data available must be gap-filled and quality controlled to provide a reliable continuous homogeneous reference series in which divergences are only caused by variations in weather and climate (Lucio *et al.*, 2006). Several interpolation methods can be used to approximate the missing rainfall amounts. These include the correlation and regression methods, distance weighted method (Inverse distance and Shepard methods), Schafer method, Thiessen polygon method and Krigging method, among others (Ogallo, 1982, 1988; Basalirwa, 1991; Schafer, 1991; Lynch and Schulze, 1995).

In this study, missing data were estimated using the correlation and regression techniques. The station that was highly correlated with the one with missing data was initially identified. Regression equations were then derived for the two stations for the period during which both stations have the data. The regression equation was later used to estimate the missing data. Upon filling in the missing data gaps, the quality of the data was assessed before any analysis was undertaken. It is worthy to mention that less than seven percent of the daily rainfall was estimated. Continuous missing data were not estimated and such stations were excluded from this study.

The double mass curve analysis was used to test the consistency of the rainfall data. The method involves the comparison of the accumulated seasonal rainfall record at a station with that of the accumulated seasonal rainfall of the nearby station. For homogeneous records, the double mass curve appears as a single straight line.

The zonal and meridional components of the re-analysis wind at the closest grid points to the two radiosonde stations were extracted. A simple correlation analysis between the time series of zonal and meridional components of the radiosonde wind data at both Nairobi and Bangui on one side and the closest grid points on the other hand was undertaken. The re-analysis dataset which had the highest correlation coefficient over most of the upper air levels was used. It has been observed that both NCEP/NCAR and ERA40 datasets are high-quality data for application in climate related research. In the next section, the methodology used to delineate the study area into near-homogeneous sub-regions is discussed.

3.2.2 Regionalization of the study area into near-homogeneous sub-regions

The classification of locations into spatial rainfall regimes with similar temporal rainfall characteristics not only reduced the number of locations that are used in the study but also reduced the local noise associated with observation from an individual location while extracting spatially coherent signal. This in return made the interpretation of the results easier. The principal component analysis (PCA) approach was used to attain this.

PCA analysis is one of the most efficient ways of compressing geophysical data both in space and time, as well as separating noise from meaningful data. The technique aims at finding a new set of variables that capture most of the observed variance from the data through linear combinations of the original variables. PCA is in essence a non-model orientated tool, which allows a time display and a space display of a space-time field such as temperature and rainfall.

PCA is an orthogonal linear transformation that converts the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. Each variable is transformed into a linear combination of orthogonal (perpendicular) common components with decreasing variation. Each component carries different information, which is not related with other components. PCA is used for dimensionality reduction in a dataset while retaining those characteristics of the dataset that contribute most to its variance, by keeping lower-order principal components and ignoring higher-order ones. Such low-order components often contain the "most important" aspects of the dataset.

PCA produces a visual representation of the relative positions of the data in a space/time of reduced dimensions, thus indicating spatial/temporal relationships among the variables. The position of each of the data points is defined by a series of axes, each of which represents separate uncorrelated information.

The output is a covariance/correlation matrix denoting the transformation coefficients (eigenvectors) listed in decreasing order of variation. The total variance accounted for by each component is the eigenvalue.

The EOFs are generally plotted as contour or vector maps, from which one can assess which

regions are closely related, inversely related or unrelated, as well as identify centres of activity, regions of strong gradients among others. The relative importance of each mode is determined by its associated eigenvalue, which is used to calculate the variance attributable to that mode.

Let P be an m x n matrix of daily rainfall data, where m is the number of days and n is the number of stations. This matrix can be decomposed into linear functions of m temporal and n spatial vectors so that the rainfall observation P_{ij} on day i at station j is

$$P_{ij} = \sum_{k=1}^{n} a_{ik} e_{kj} \Leftrightarrow P = a\, e \qquad \textbf{(3.1)}$$

where a_{ik} is the element for day i in the kth time vector, and e_{kj} is the element for station j in the kth space vector.

The strength of the analysis is that often a large part of the spatial variability of the original data can be reproduced using only a few of the space vectors. These may be interpreted in relation to the underlying physical rainfall producing processes. The time vector may be seen as a time series of weights, giving more or less weight to a particular space vector (spatial rainfall pattern) each day. To recreate the original daily spatial rainfall pattern, the weighted spatial patterns are superimposed.

The space vectors may be found using either the correlation or the covariance matrix of the rainfall time series. The correlation matrix was used in this study. Svensson (1999) has indicated that large scale rainfall patterns become less clear when the covariance matrix was used over mountainous areas with the larger rainfall variances as compared to those on the plain. The inhomogeneous terrains over the study region thus justify the use of the correlation matrix.

Since the daily rainfall distribution at each station was skewed, the daily rainfall totals had to be transformed. Two approaches of the transformation that can be used are the square-root and logarithm transformations. Square-root transformation was used in this study since it is easily applied unlike the logarithm one which gives some difficulties when zero rainfall amounts are considered. The square-root transformed daily rainfall series works well over the East Africa region (Bärring, 1988; Camberlin and Okoola, 2003). Stephenson *et al.* (1999) have indicated that square root transformation is the optimal variance stabilizing

transformation for a Poisson process and thus is beneficial in stabilizing the variance of sporadic rainfall time series.

The square-root transformed rainfall data P_{ij}, were standardized by subtracting the mean $\overline{P}_j$, of the time series for each station and then dividing by the standard deviation σ_j, so that the new standardized rainfall A_{ij}, is

$$A_{ij} = \frac{P_{ij} - \overline{P}_j}{\sigma_j} \qquad \textbf{(3.2)}$$

The symmetric n x n correlation matrix C, calculated with regard to the time series (i.e. column-wise for the matrix A) is given by

$$C = \frac{A^{'} A}{(m-1)} \qquad \textbf{(3.3)}$$

The correlation matrix can be decomposed into eigenvectors e, and associated eigenvalues λ (Svensson, 1999). The eigenvectors are the space vectors described by **Equation 3.1**, and the corresponding eigenvalues are measures of the explained variance accounted for by each eigenvector. The eigenvalues are obtained by solving **Equation 3.4** while the eigenvectors are obtained by solving **Equation 3.5**.

$$|C - \lambda I| = 0 \qquad \textbf{(3.4)}$$

$$(C - \lambda I)e = 0 \qquad \textbf{(3.5)}$$

Comprehensive details on the application of PCA in atmospheric science studies can be found in von Storch and Zwiers (1999), Wilks (2006), Hannachi *et al.*, (2007) and in Jolliffe (2002) for the general application of EOF analysis.

There are at times some difficulties in interpretation of the obtained patterns (Ambaum *et al.*, 2001; 2002; Dommenget and Latif, 2002) because the physical modes are not necessarily orthogonal. Spatial orthogonality and temporal uncorrelation of the PCs impose limits on physical interpretability of loading patterns (Hannachi *et al.*, 2007). This is because physical processes are not independent and therefore physical modes are generally expected to be non-orthogonal. Horel (1981) pointed out that if the first EOF has a constant sign over its domain, then the second one will generally have both signs with the zero line going through the

maxima of the first EOF.

The difficulties associated with interpreting PCAs have led to the development of more tools to overcome this problem. The linear transformation of PCs, based on rotation is one such tool which have been introduced and yielded the concept of Rotated Principle Component Analysis (RPCA) as discussed by Richman (1986).

The main purposes of RPCA are to;

i) Alleviate the strong constraints of PCA mainly orthogonality/uncorrelation of outputs and domain dependence of spatial patterns (Dommenget and Latif, 2002);
ii) Obtain simple structures; and
iii) Ease in the interpretation of the obtained patterns.

Rotation of the EOF patterns can systematically alter the structures of EOFs. Rotation of EOF has the effect of redistributing the variance within the eigenvectors and therefore removing the ambiguities while conserving the variance extracted by the selected subset of non-rotated eigenvectors (Indeje, 2000). By constraining the rotation to maximise a simplicity criterion, the REOF patterns can be made simple.

Given a p x m matrix, $U_m = (\mathbf{u}1, \mathbf{u}2, \mathbf{u}3 \ldots, \mathbf{u}m)$ of the leading m PCA loadings, the rotation is achieved by seeking an m x m rotation matrix R to construct the REOFs K according to;

$$K = U_m\beta \qquad \textbf{(3.6)}$$

where β is either R or $(R^T)^{-1}$ depending on the type of rotation desired. The simplicity criterion for choosing the rotation matrix for maximisation problem is expressed by;

$$\max f\,(U_m\beta) \qquad \textbf{(3.7)}$$

over a specified subset or class of m x m square rotation matrices R.

Various rotation criteria exist (Richman, 1986). However, they can be broadly classified into two families (Jennrich, 2001; 2002) namely;-

1. Orthogonal in which the rotation matrix is chosen to be orthogonal and $\beta = R$.
2. Oblique in which the rotation matrix is chosen to be non-orthogonal and $\beta = (R^T)^{-1}$.

Varimax orthogonal rotation was used in this study as oppose to the Quartimax oblique

rotation for three reasons, namely:

i) The rotated PCA yield components that are easier to interpret physically.

ii) Studies done over the East Africa region using this method have obtained satisfactory results (Ogallo, 1980; Indeje, 2000; Ouma, 2000; Okoola and Camberlin, 2003; Komutunga, 2006).

iii) Thirdly and most important, by normalising the spatial eigenvectors to unity, Varimax rotation produces uncorrelated components that satisfy the assumptions of cluster analysis (Phillips and Denning, 2007).

Different authors have suggested different methods that can be used to determine the number of principal components that should be retained for rotation (Kaiser, 1959; Anderson, 1963; Castell, 1966; North *et al.*, 1982; Overland and Preisendorfer 1982). The method used in this study to determine the number of the principal components to be retained and rotated is the Monte Carlo simulation method.

The Monte Carlo method is used to simulate a statistical model under the assumption that a given null hypothesis H_o is true (von Storch and Zwiers, 1999). A matrix of random values of the size of the observed data is generated, in which the temporal auto-correlation found in the observed times-series is reserved. PCA is computed on this matrix, and the eigenvalues stored. This procedure was repeated 500 times. All the eigenvalues are ranked and the 95^{th} percentile considered as the 95% confidence threshold, to which the actual eigenvalues of the observed data set are compared. All eigenvalues higher than the threshold are judged significant.

Rotated Empirical Orthogonal Function (REOF) and simple correlation analyses were used to delineate the near-homogeneous rainfall sub-regions in East Africa using the quality-controlled daily gauge rainfall. The approach used in this analysis is similar to the one employed by Indeje (2000). Each Rotated Principal Component (RPC) time series obtained from REOF analysis was correlated with the stations' rainfall data and stations with significant correlation coefficient identified. Delineation of a near-homogeneous sub-region was accomplished by identifying the stations with the largest correlation with the RPC time series associated with the eigenvector of the daily rainfall in a season.

3.2.3 Baseline information of wet and dry spells

In order to successfully achieve the first specific objective of this study as outlined in **section 1.3**, several intraseasonal statistics (ISS) on the wet/dry spells were defined and computed at both local (station) and sub-regional (near-homogeneous) levels as described in **Table 3.4**. This provides the baseline information on the various aspects of daily rainfall performance as supplied by the alternating wet and dry spells.

Initially, a threshold for separation of wet and dry days and the definition of wet and dry spells was adopted. The frequency distribution of the wet and dry spells is determined, from which the various intraseasonal statistics of wet and dry spells are computed. Simple correlation analysis was undertaken to determine the association between the seasonal rainfall totals and the various intraseasonal statistics of the wet and dry spells. The variation of the seasonal rainfall totals and intraseasonal statistics of wet and dry spells with time was finally assessed using the non-parametric Spearman rank correlation analysis. These steps are elaborated in the subsequent sections.

3.2.3.1 Threshold used and definition of wet and dry spells

The occurrence of a wet or dry day is a mutually exclusive event (Chapman, 1998; Dobi-Wantuch et al., 2000). A threshold for delineating wet and dry days is required when analysing spells of rainfall since the frequency distribution of the length of the wet/dry spells is highly skewed and depends on the selected threshold (Bärring *et al.*, 2006).

Different authors have used different thresholds based on the aspect of the spells that they need to consider. Dobi-Wantuch *et al.* (2000) have indicated that the standard observational threshold of 0.1 mm provides a good representation of precipitation conditions for some observational records. Moon *et al.* (1994) and Matrin-Vide and Gomez (1999) have used 0.1 mm since it is the usual precision of rain-gauges. Frei *et al.* (2003) have used a higher threshold of 1.0 mm since it is more resistant to measurement errors related to light rainfall. Douguedroit (1987) and Lazaro *et al.* (2001) employed a threshold of 1.0 mm and argued that rainfall less than this amount evaporated off directly.

Perzyna (1994) used a threshold of 2.0 mm in order to remove any events featuring less rainfall and with very little significance in the river flow due to losses by interception and evaporation. Ceballos *et al.* (2004) have used a threshold of 10 mm since rainfall below this amount have only small effect on the soil water-content at a depth greater than 5cm from the

surface (Ceballos *et al.*, 2002). Such rainfall remains at the surface of the soil or on its plant cover, from where it readily returns to the atmosphere through evapo-transpiration.

Recent studies have mainly used more than one threshold for delineating the dry/wet days in observational records. Gitau *et al.* (2008) have used 1, 3 and 5 mm thresholds based on the average evapo-transpiration at the different locations in Kenya. Ambenje *et al.* (2001) have used five thresholds of 1, 12.5, 25.4, 50.8 and 100 mm to study the frequency distribution of days with precipitation above these thresholds over 19 countries in eastern and southern Africa for the four standard seasons. Ceballos *et al.* (2004) have used two thresholds of 0.1 and 10 mm to study dry spells on Duero basin in Spain.

In this study, a threshold of 1.0mm was used to delineate wet days from dry days. This was mainly because of two reasons. First lower thresholds (less than 1.0mm) are more vulnerable to measurement errors associated with light rainfall and readily evaporate given the higher evapo-transpiration rate at the study region. Secondly, higher thresholds (more than 1.0mm) substantially reduce the sample size of the data to be used for further analysis since they greatly reduce the number of wet days. Other studies over East Africa that have used a threshold of 1.0 mm include the work of Mungai (1984) and Ogallo and Chillambo (1982).

A dry day was therefore defined as any day that received rainfall less than 1.0 mm, while a wet day was any day that received rain equal to or in excess of 1.0 mm.

Once the threshold for wet days was fixed, the next aspect was the definition of wet and dry spells. Different authors have considered different definitions of the wet/dry spells. Two main examples of such definitions are given here. Peña and Douglas (2002) defined wet (dry) spells as days when 75% or more (35% or less) of the stations along the Pacific side of Nicaragua, Costa Rica, and Panama records rainfall.

However, most authors define wet and dry spells locally. Ogallo and Chillambo (1982) have defined a wet (dry) spell of length i as a sequence of i wet (dry) days preceded and followed by a dry (wet) day. It is this definition that was used in this study.

The intraseasonal statistics of wet and dry spells (ISS) can be defined at station (local), sub-regional (near-homogeneous zone) or regional (the whole of study area) levels. This study considers the former two levels and yields the concept of local intraseasonal statistics of wet and dry spells (LISS) at station level and sub-regional intraseasonal statistics of wet and dry

spells (SRISS) at near-homogeneous zone level. The 1.0mm threshold value was used to derive both the LISS and SRISS.

3.2.3.2 Local intraseasonal statistics of wet and dry spells

Based on the threshold value and definition of wet and dry spells discussed above, the wet and dry spells of varying lengths at local (station) level were tallied and organised into a frequency distribution table as described in Gitau (2005), Tilya (2006) and Bamanya (2007). From the frequency distribution tables, various intraseasonal statistics of the wet and dry spells were computed. **Table 3.4** gives a description of the various intraseasonal statistics that were computed from the frequency distribution table for each year and at each station. In addition, the seasonal rainfall totals was computed by summing up the daily rainfall observations for each individual season.

It is worth clarifying at this point that in order to determine the above intraseasonal statistics of the wet and dry spells, the dry periods before the first and after the last rainfall/wet spells were excluded. This was in order to avoid the long dry spells that occur at the beginning and at the end of the rainfall period, and which belongs to the preceeding and following dry seasons respectively. To accomplish this and since the date of onset and cessation of the rainfall period were not predetermined, the dry spells before (after) the first (last) wet spells for each rainfall season were excluded from the dry spell analysis.

The significance of the calculated Pearson correlation coefficient was tested by computing the *t* statistic given by **Equation 3.9**. The computed *t* statistic obtained was compared against the tabulated critical values and appropriate conclusions drawn.

$$t = \rho\sqrt{\frac{(N-2)}{1-\rho^2}} \qquad \textbf{(3.9)}$$

where the symbols have the same meanings as in **Equation 3.8**.

3.2.3.2.2 Trend analysis

The time series of meteorological parameters are often generated by a complex interaction of a number of weather/climate systems. The temporal behaviour of such weather/climate systems, in terms of trends and modes of variability is therefore of paramount importance for understanding of climate and the assessment of its potential impacts.

The interannual variation of the local intraseasonal statistics of wet and dry spells (LISS) was assessed for any significant trend using the Spearman rank correlation analysis. This is a non-parametric approach based on ranks and is used here since there are very few underlying assumptions about the structure of the data. This method produces robust results especially when the observations are thought not to satisfy the normal distribution (Helsel and Hirsch, 1992). In addition, the use of ranks rather than actual values makes it insensitive to outliers and missing values.

The sample $\{(X_i, Y_i)$ for $i = 1, 2, 3$ ---- n$\}$ is replaced by the corresponding sample of ranks $\{(\mathbf{R}_{Xi}, \mathbf{R}_{Yi})$ for $i = 1, 2, 3$ ---- n$\}$ where $\mathbf{R}_{Xi}$ is the rank of X_i amongst the **Xs** and $\mathbf{R}_{Yi}$ is similarly defined. The differences *d* between the ranks of each observation on the two variables are then calculated. If there are ties, the tied observations are assigned the corresponding average rank. The dependence between **X** and **Y** is then estimated with the Spearman rank correlation coefficient ρ which is given by **Equation 3.10**;

$$\rho = 1 - \frac{6\sum d_i^2}{n\left(n^2-1\right)} \qquad \textbf{(3.10)}$$

where:

d_i = the difference between each rank of corresponding values of X and Y, and n = the number of pairs of values. The significance of the computed spearman rank correlation coefficient was tested by computing the statistic t given by **Equation 3.9**.

3.2.3.3 Sub-regional intraseasonal statistics of wet and dry spells

Different methods can be used to calculate the sub-regional intraseasonal statistics (SRISS) based on daily rainfall from several locations. At sub-regional level, the local noise associated with observations from individual locations is minimized. The extraction of a spatial coherent signal (if any) is an important step towards the assessment of the predictability potential of a given climate variable.

Figure 3.4 shows a schematic diagram of three different approaches that can be used to compute the SRISS. The first method involves computing the local intraseasonal statistics of wet and dry spells (LISS) at individual locations, which are then averaged for a specific near-homogeneous zone to obtain the SRISS. In the second method, the observed daily rainfall amounts for several stations constituting a given sub-region (near-homogeneous zone) are first averaged and the SRISS derived from the sub-regional areal-average rainfall. The final method involves using the Principal Component Analysis (PCA) scores as obtained from regionalization for each sub-region (near-homogeneous zone). In the case of the Principal Component (PC) score, the threshold that could correspond to the 1.0 mm threshold which was used at the station level was chosen. The daily rainfall data for all the n stations constituting a near-homogeneous sub-region are grouped together into a single column and sorted in ascending order (starting with the smallest). The percentile p, corresponding to the value of 1.0 mm was then obtained. The Principal Component (PC) score was also sorted in ascending order and the position (p/n) obtained. The PC score threshold used is the value that corresponds to the p/n position rounded off upwards. This PC score value corresponds to 1.0 mm threshold used for station data.

Out of the three approaches available for computing the SRISS (**Figure 3.4**), the box-plot of the correlation coefficients between the LISS and the SRISS was used to determine the best approach for deriving SRISS. A comparative assessment carried out with SRISS obtained using the three approaches shown that the second approach gives results that are inconsistent with the other two.

The SRISS obtained were assessed for their association with seasonal rainfall totals at sub-

region level as well as their variation with time in a similar manner to those of the local intraseasonal statistics of wet and dry spells (LISS). This completed the analyses necessary to achieve the first specific objective of this study.

The second specific objective as outlined in **section 1.3** was investigating the linkages between the various sub-regional intraseasonal statistics of the wet and dry spells (SRISS) including seasonal rainfall totals and dominant large scale climate fields that drive the global climate during specific seasons. Initially the spatial coherence of the various intraseasonal statistics of the wet and dry spells was assessed by undertaking inter-station correlation analysis. This provides an indirect measure of potential predictability. The indices of Niño, IOD and SST gradients were used as the predefined predictor indices. Total and partial correlation analyses were then used to quantify the relationship between various SRISS and large scale climate fields. This enabled the identification of locations from which additional potential predictor indices were extracted. Plausible physical/dynamical explanation and comparative location assessment helped to reduce the number of the additional potential predictor indices. The various steps which were undertaken are discussed in details in the following sections.

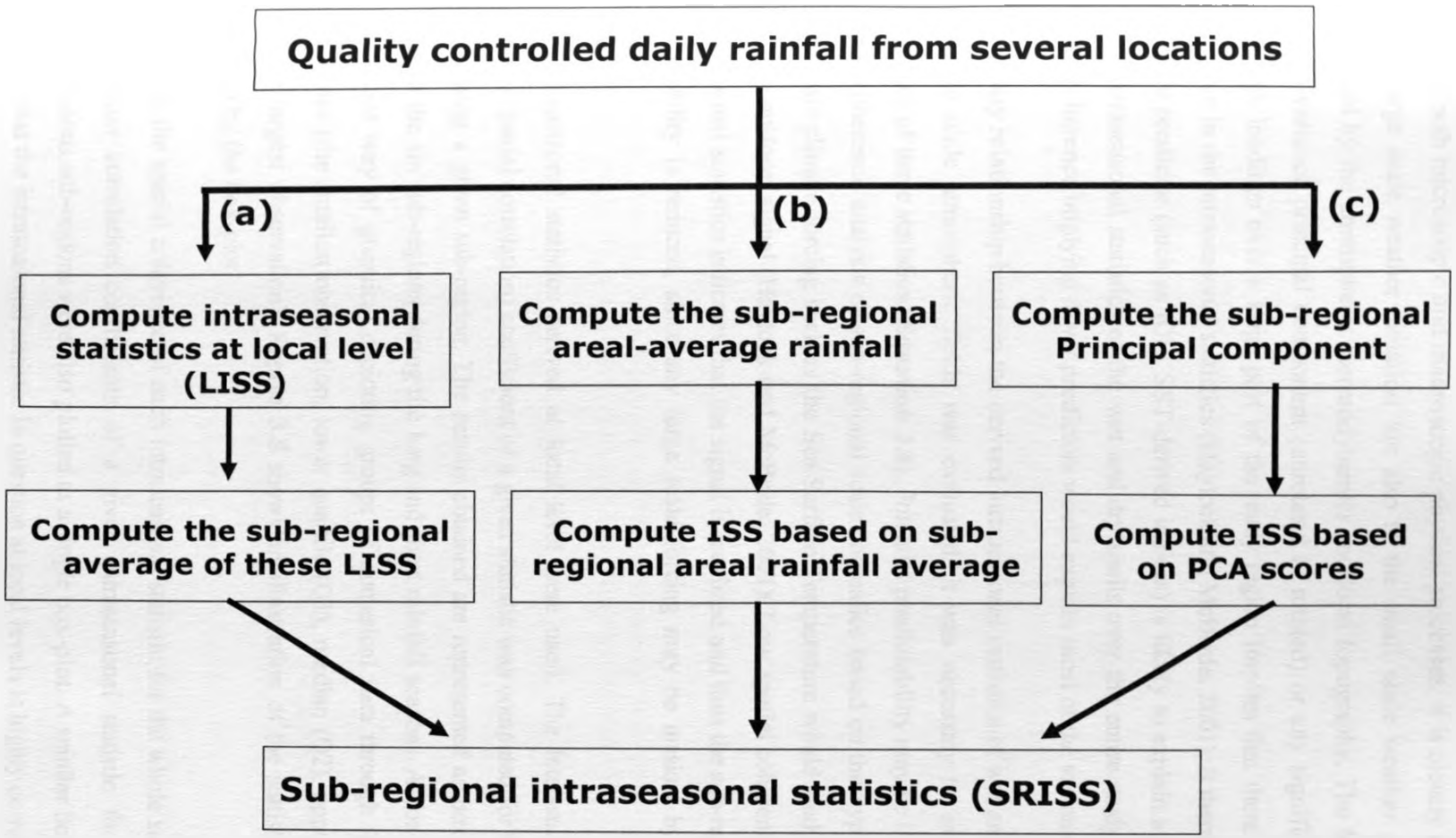


Figure 3.4: Schematic diagram on different approaches of calculating sub-regional intraseasonal statistics of wet and dry spells

3.2.4 Spatial coherence and potential predictability

Precipitation is the most difficult climate component for models to forecast, because it involves both microscopic and macroscopic physical processes, it is closely related not only to the large scale weather situation but also to the small scale weather systems, and is influenced by the atmospheric thermodynamics and local topography. The fact that there is no high-variance principal component (unrotated or rotated) or any significant component with high loadings over a large part of the study region implies that there is only limited coherence in the intraseasonal statistics (Haylock and McBride, 2001). It therefore means that no single predictor (such as IOD, SST-derived indices) is likely to explain a high proportion of the intraseasonal statistics of the wet and dry spells over the entire study region. Higher spatial coherence implying fewer predictors would explain most of the variance.

Before any relationship between the derived intraseasonal statistics of wet and dry spells and the large scale atmospheric fields was evaluated, it was necessary to assess the spatial coherence of these statistics (**Equation 3.8**). Potential predictability may be inferred from the spatial coherence analysis of sub-regional scale anomalies based on the hypothesis that any large scale climate forcing such as the Sea Surface Temperature would tend to give a rather spatially uniform signal (Haylock and McBride, 2001). Low spatial coherence of any of the intraseasonal statistics indicates that the signal is localized and thus the sub-regional potential predictability is reduced, since any large scale forcing may be masked by stronger local effects.

The intraseasonal statistics derived at local level were used. The inter-station correlation (same as spatial correlation) coefficient of a given statistic was computed for all the locations constituting a given sub-region. The results obtained are represented as box-plot. This was done for the six sub-regions during the long and short rainfall seasons. A box-plot provides a convenient way of graphical depicting groups of numerical data through the five-number summaries (the smallest observation, lower quartile (Q1), median (Q2), upper quartile (Q3) and the largest observation). **Figure 3.5** shows an illustration of the statistical summaries provided by the box-plot.

To derive the spatial coherence of each intraseasonal statistic for the whole study region, the inter-station correlation coefficients of a given intraseasonal statistic for all the near-homogeneous sub-regions were also plotted as a single box-plot. A smaller box-length would indicate that the intraseasonal statistic in question at local levels is highly correlated. A higher

The three atmospheric variables were not nested as such. However, the predictor search was confined to region (80°W – 120°E, 45°S – 45°N). The choice of this region was based on the fact that it includes the sub-tropical anticyclones which control moisture fluxes towards East Africa. It also enables the depiction of the wind features which directly affect East African climate, such as the Indian Ocean monsoon, the Indian and Atlantic Ocean Walker-type circulation cells, the Tropical Easterly Jet, the Subtropical Westerly Jets among others. It is worthy to mention that there was an assumption that higher latitude (latitudes beyond 45°N or 45°S) oceanic and atmospheric systems, at seasonal scale do not influence the rainfall characteristics over equatorial eastern Africa.

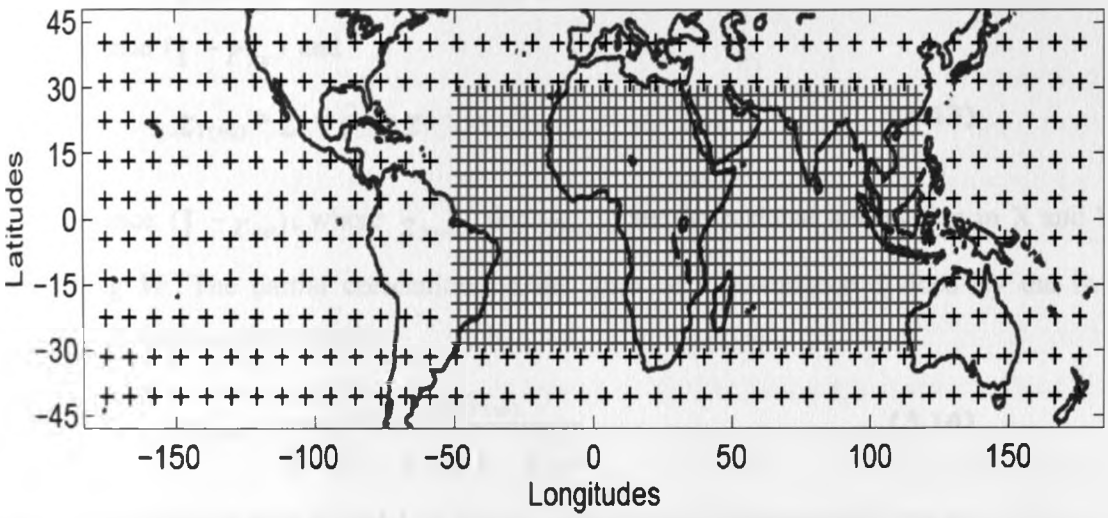


Figure 3.6: Map showing the nesting of the SST grid points. Red plus (+) are the fine grid resolution while the black plus (+) are for the coarse grid resolution.

The partial correlation between the predictands (rainfall totals as well as intraseasonal statistics) and the Hadley centre SSTs, atmospheric variables of zonal and meridional winds, specific humidity and geopotential height were then calculated while controlling the influence of the predefined predictor indices (**in section 3.2.5.1**) that were significant at 95% confidence level. This provided a list of additional potential predictors for the rainfall totals and the different intraseasonal statistics. The rationale behind the partial correlation analysis was that many large scale climate fields are influenced by major modes of variability such as ENSO (already described by the indices used in step 1) hence full correlation with East Africa rainfall may at times only reflect co-variations induced by the common forcing rather than a physical relationship.

Partial correlation $r_{YX\,W}$ allows us to determine what the correlation between any two

variables say X and Y would be, if the third variable W is held constant. This ensures that no variance predictable from W enters the relationship between Y and X. In z-score form, we can predict both X and Y from W, then subtract those predictions leaving only information in X and Y that is independent of W, as follows.

$$\hat{z}_X = r_{XW} z_W \text{ and } \hat{z}_Y = r_{YW} z_W \quad \textbf{(3.11)}$$

where $\hat{z}_X$ and $\hat{z}_Y$ are the predicted z-scores for X and Y respectively. Subtracting these predicted scores, we obtain

$$z_{X(res)} = z_X - \hat{z}_X = z_X - r_{XW} z_W \quad \textbf{(3.12)}$$

with variance $(1 - r^2_{XW})$ and

$$z_{Y(res)} = z_Y - \hat{z}_Y = z_Y - r_{YW} z_W \quad \textbf{(3.13)}$$

with variance $(1 - r^2_{YW})$, where $z_{X(res)}$ and $z_{Y(res)}$ are the residual information in X and Y controlling W. The partial correlation, in the form of a covariance divided by the two standard deviations, then equals

$$r_{YX.W} = \frac{\sum z_{X(res)} z_{Y(res)}}{N\sqrt{(1 - r^2_{XW})(1 - r^2_{YW})}} \quad \textbf{(3.14)}$$

Substituting **Equations 3.12** and **3.13** into the numerator of **Equation 3.14**, we get

$$r_{YX.W} = \frac{\sum (z_X - r_{XW} z_W)(z_Y - r_{YW} z_W)}{N\sqrt{(1 - r^2_{XW})(1 - r^2_{YW})}} \quad \textbf{(3.15)}$$

which gives

$$r_{YX.W} = \frac{\sum (z_X z_Y)/N - r_{YW}\sum (z_X z_W)/N - r_{XW}\sum (z_Y z_W) + r_{XW} r_{YW}\sum z^2_W / N}{\sqrt{(1 - r^2_{XW})(1 - r^2_{YW})}} \quad \textbf{(3.16)}$$

But **Equation 3.8** in z-score form becomes $r_{XY} = \sum (z_X z_Y)/N$. Thus **Equation 3.16** reduces to

$$r_{YX.W} = \frac{r_{XY} - r_{YW} r_{XW} - r_{XW} r_{YW} + r_{XW} r_{YW}}{\sqrt{(1 - r^2_{XW})(1 - r^2_{YW})}} \quad \textbf{(3.17)}$$

which finally becomes

$$r_{YX.W} = \frac{r_{XY} - r_{YW} r_{XW}}{\sqrt{(1 - r_{XW}^2)(1 - r_{YW}^2)}} \qquad (3.18)$$

This is the equation for a partial correlation that was used to search and identify predictors from the oceanic and atmospheric variables. The partial correlation approach has been successfully used by Behera *et al.* (2005) in determining the effect of IOD (ENSO) on short rainfall over Eastern Africa while the effect of ENSO (IOD) is removed.

Partial correlation maps were then produced. It was from these maps that the highly correlated regions were identified and used to compute the new indices. It is worthy to mention that the correlation box identified were at least 5° by 5° for the atmospheric variables and 6° by 6° for the oceanic variable. This means that at least four grid points were averaged for the atmospheric variable predictors (since they are gridded at 2.5° by 2.5°) and oceanic variable predictors (fine grid nested at 3° by 3°). This was to ensure that the predictors have less noise, remain stable and do not vary too fast from the time the forecast is made until the time the observations are made. Mutai *et al.* (1998) have combined the UK Met. Office SST version 4 (MOHSST4) which are initially at 1° by 1° to form a 10° by 10° grid boxes to improve data coverage and reduce noise. Gong *et al.* (2003) have further demonstrated that spatial aggregation increases the skill of seasonal total precipitation forecasts.

At times, none of the predefined indices were significantly correlated with the rainfall totals and intraseasonal statistics. In such situations, concurrent and lagged simple correlation analyses were first undertaken with Hadley Centre SSTs. The significant SST predictors identified were then used to undertake partial correlation with the atmospheric variables. There were also cases where two or more predefined predictors were significantly related to the same intraseasonal statistics but highly dependent on each other. The predictor that was most frequently picked was used. In case both predictors are equally frequently occurring, the predictor with the highest correlation coefficient was retained. In case two or more predefined predictors which are not significant related to each other were identified, there were all retained.

3.2.5.3 Selection of robust potential predictors

The foregoing procedure yielded quite a large number of oceanic and atmospheric predictors. There was therefore the need to reduce the high number of predictors. In this study, apart from the use of standard statistical methods, the selection of the potential predictors was also

based on the physical interpretation of the relationship with East Africa rainfall. Only those predictors with a plausible physical/dynamical relationship were retained and later used to generate the regression model. The interpretation of the lag-relationship was based on the persistence of the predictor within its geographical location or its modulation on other climate variables especially the Sea Surface Temperature (for atmospheric variables).

Similarly, upon identification of the several predictors for the different intraseasonal statistics, comparative analysis was undertaken to identify the predictors that were more or less located at the same position with a shift of a few degrees of latitude or longitude. This not only reduced the number of predictors further but also ensured that only robust predictors were retained. The number of predictors had to be reduced since we need to include only those predictors that have significant association with our predictants. Robust predictors are those predictors with strong and consistent association with the predictants and are further supported by logical physical or dynamical linkage with the majority of the predictants. Small shifts in the location of the predictors from one predictand to the next are likely to reflect sampling errors rather than real climatic features. The main misgiving with these steps was that the variance explained by the regression models developed from these few predictors was likely to be slightly reduced. However as observed in Philippon (2002), it is desirable that physical consistency outweighs statistical skill in empirical climate prediction.

3.2.6 Development of regression models

The final specific objective as outlined in **section 1.3** was predictability assessment of the various intraseasonal rainfall variables through statistical models based on the linkages already identified, for the improvement of early warning systems. Forward stepwise multivariate linear regression (MLR) analysis was used to develop empirical statistical prediction models with sufficient lead time. The concept of the adjusted correlation coefficient was used to determine the number of predictors to be retained in the model. The cross validation method and calculation of the Linear Error in Probability Space (LEPS) skill score were used to assess the performance of developed MLR models. The residuals from the models developed were finally evaluated using the Durbin-Watson statistics and Kolmogorov-Smirnov test. The intraseasonal statistics with correlation coefficient of less than 0.5 between the observed and the cross-validated model output time series were classified as unpredictable.

3.2.6.1 Multivariate Linear Regression

The multivariate linear regression (MLR) approach is a common method in seasonal climate prediction. It is the most frequently used method over the East Africa region and has provided seasonal forecasts with useful skills (Mutai *et al.*, 1998; Camberlin and Philippon, 2002; Korecha and Barnston, 2007; Nyakwada, 2009). Statistical relationships between various wet/dry spells statistics and oceanic/atmospheric predictors were developed using forward stepwise MLR approach at sub-regional scale.

In the forward stepwise MLR approach, each predictor variable is entered into the regression model in an order determined by the strength of their correlation with the predictand. The effect of adding each predictor is assessed and the predictor retained if it contributes significantly to the variance explained by the model. This procedure is repeated until all the predictors that contribute to the variance of the model are retained. Those predictors that do not significantly contribute to the explained variance of the model are thus left out.

A MLR model which expresses intraseasonal statistics at any specific time t (Y_t) as a function of atmospheric and oceanic predictors at time lag k (X_{it+k}) may be expressed in **Equation 3.19**;

$$Y_t = a + b_1 X_{1t+k} + b_2 X_{2t+k} + \ldots\ldots\ldots + b_n X_{nt+k} \qquad \textbf{(3.19)}$$

For zero lagged relationship, **Equation 3.19** becomes;

$$Y_t = a + b_1 X_{1t} + b_2 X_{2t} + \ldots\ldots\ldots + b_n X_{nt} \qquad \textbf{(3.20)}$$

where a is the regression constant and b_i are regression coefficients. Both the regression constant and coefficients were estimated from available records.

The strong inter-correlation between the predictors leads to multi-collinearity which means that the predictors are non-orthogonal. This results to lacks of the model's accuracy and may lead to unclear interpretation of the regression coefficients as measures of original effects (Mc Cuen, 1985). It further imposes the problem of redundancy and unnecessary loss of degrees of freedom especially when large numbers of correlated predictors are used (Krishna Kumar *et al.*, 1995). To increase the reliability of regression models while using the multi-collinearity predictors, the variance inflation factor, VIF (Fox, 1991) should first be determined. The VIF measures how much the variance of the estimated regression

coefficients are inflated by multi-collinear predictors compared to when the independent variables are uncorrelated. In the current study, only independent variables that are uncorrelated were used to generate the regression model since the variance inflation factor was not calculated.

The cross validation method was used to test the developed MLR models for the various sub-regional intraseasonal statistics of the wet and dry spells (SRISS). This method involves temporarily discarding observations from the dataset and then estimating the discarded observations. The estimated values are then compared with the discarded value (Isaaks and Srivastaka, 1989). In this study, three values were left out each time and regression models developed with the remaining values. The regression model developed was then used to estimate the discarded values. The method was used since the available time series of SRISS was not long enough to enable the subdivision of the time series into training and verification periods. More details of cross validation method can be obtained from Issaks and Srivastaka (1989), Barnston *et al.* (1996) and Wilks (2006).

3.2.6.2 Number of predictors to be retained

A popular measure of the strength of association in linear regression between the observation and the model output is the coefficient of determination R^2, defined as the proportion of variability in the outcome variable explained by the model. However, a serious problem with this measure is that it can substantially overestimate the strength of association when the number of predictors p, is not small relative to the number of observations n. It can attain its maximum value of 1 for any saturated model even when the predictors and outcome are independent of each other. The adjusted coefficient of determination overcomes this problem (Liao and McGee, 2003). The adjusted coefficient of determination, in the forward stepwise MLR analysis, discourages incorporating additional predictors that will make little marginal changes in the unexplained variance. The adjusted R^2 accounts for the number of the predictors in the model and only increases if the new predictor improves the model more than would be expected by chance. The number of predictors to be retained in the final MLR model was thus determined from the adjusted R^2 of the cross-validated model. When the addition of a new predictor results in a decrease of the adjusted R^2 or remains unchanged, the new predictor was excluded and the model was developed with the previous predictors only.

The adjusted R^2 is defined as

$$R^2_{ADJ} = 1 - \left(\frac{n-1}{n-p-1} \right) \left(1 - R^2\right) \qquad \textbf{(3.21)}$$

where p is the number of predictors, n is the number of observations and R is the correlation coefficient.

Equation 3.21 means that the adjusted R^2 incorporates only the unexplained (from the denominator term) and total variance (from the numerator term). Delsole and Shukla (2002) and Nyakwada (2009) have observed that fewer predictors tend to produce better models than those developed using large numbers of predictors.

In addition to the adjusted R^2, the Analysis of Variance (ANOVA) was used to test the statistical significance of the regression constants, together with the variance accounted for by oceanic and atmospheric predictors. Details of ANOVA test and other regression principles can be obtained from Kendall and Stuart (1961), Kendall (1976), and Wilks (2006) among other authors.

3.2.6.3 Assessment of the model performance

Several methods can be used to assess the performance/skill of prediction models. Zhang and Casey (2000) have broadly grouped them into four categories and highlighted their advantages and disadvantages using the Australian winter and summer seasonal rainfall forecast model hindcasts for a period of 96 years. The Linear Error in Probability Space (LEPS) score that was developed by Ward and Folland (1991) and later refined by Potts *et al.* (1996) was used in this study.

LEPS is defined as the mean absolute difference between the cumulative frequency of the model forecast and the cumulative frequency of the observations (Jolliffe and Stephenson, 2003). It evaluates the model skill by penalizing errors in terms of the distance between forecasts and observations in cumulative probability space. It gives relatively more penalty when forecasting events around average values but gives relatively higher scores and less penalty for forecasts of extreme events (Zhang and Casey, 2000).

The normalized linear error in probability space score is given by

$$S^{*} = 3(1 - |P_F - P_O| + P_F^2 - P_F + P_O^2 - P_O) - 1 \quad \textbf{(3.22)}$$

Where P_O is the cumulative probability distribution of the observations and P_F is the cumulative probability distribution of the regression model forecasts. A maximum score of 2 is achieved when $P_O = P_F = 0$ or $P_O = P_F = 1$ while a minimum score of -1 is attained when $P_O = 0$ and $P_F = 1$ or $P_O = 1$ and $P_F = 0$. It is often desirable to have a measure of overall skill over a range of -100% to 100%. For a sufficiently large ensemble of forecast being assessed together, a method has been developed. To achieve the skill range from -100% to 100%, the average skill (SK) for continuous, categorical and probability forecasts is defined by **equation 3.23.**

$$SK = \frac{\sum 100 S^{*}}{\sum S_m^{*}} \quad \textbf{(3.23)}$$

where the summation is over all pairs of forecasts and observations. The definition of S_m^{*} depends on whether the number is positive or negative. If positive, S_m^{*} is the sum of the maximum possible scores given by the observations. If the numerator is negative, S_m^{*} is the sum of the modulli of the worst possible scores given the observations. That in short means that negative values of SK score indicate that the models developed are worse off than climatology while positive values indicate that the models are better off than climatology. A value of zero means that the model is as good as the climatology. More details of its derivation can be found in Potts *et al.* (1996). Camberlin and Philippon (2002) have previously used this skill score measure over Eastern Africa.

3.2.6.4 Residual analysis from the regression models

A good multivariate linear regression model requires that the residuals (the difference between the actual observations and the forecasted values) are independent and have a normal distribution (Nayagam *et al.*, 2008). The Durbin–Watson statistic checks the significance of the assumption that the residuals for successive observations are uncorrelated / independent. Its value ranges from zero to four. Values more than two indicate that there

exists some negative autocorrelation and values less than two, a positive autocorrelation. The Durbin–Watson (DW) statistic is defined as

$$DW = \frac{\sum_{T=2}^{N}(E_T - E_{T-1})^2}{\sum_{T=1}^{N} E_T^2} \qquad \textbf{(3.24)}$$

where N is the number of residuals, E_T is the residual at the time T and E_{T-1} is the residual at time$T$-1.

The values of the Durbin-Watson statistic are compared with the critical values tabulated by Farebrother (1980) since the regression models generated did not have the constant term. If there exists any kind of significant lag one autocorrelation, then the assumption of independence of residuals is violated and the model can be improved further (Makridakis *et al.*, 1998).

One sample Kolmogorov-Smirnov test was used to ascertain that the residuals were normally distributed. Kolmogorov-Smirnov test determines whether an underlying probability distribution from a finite sample differs from a hypothesized distribution by comparing the empirical distribution function with the cumulative distribution function specified by the null hypothesis. Minor improvements made by Lilliefors leads to the Lilliefors test (Lilliefors, 1967).

The null hypothesis is that the residuals from the multivariate linear regression (MLR) models are normally distributed. The alternative hypothesis is that the residuals have a distribution different from the normal distribution function.

3.3 Limitations and assumptions of the study

In the scientific studies including climatology and meteorology, there are limitations that one comes across and assumptions that have to be made in order for the study to move forward. The current study was not an exception.

The first major limitation was that the many rainfall stations that were established in the colonial period have been stopped due to the high cost of operations. Only a few stations established in the colonial period still exist today which means that stations/locations with long time series of the daily rainfall series are limited. This had an effect on the network of the stations used. Another limitation was the slow pace of data digitization especially for the

non-synoptic stations. This has an effect of reducing the length of the data records for the stations used.

Based on the foregoing limitations, several assumptions were made. The first assumption made was that the station network and study period used in this study is representative of the study region based on availability of long uninterrupted time series of daily rainfall series. The results obtained and conclusions made may therefore have slight differences with similar studies made over the study region at a different time especially in the context of the climate change aspect.

Another assumption made was that higher latitude (latitudes beyond 45°N or 45°S) oceanic and atmospheric systems, at seasonal scale, do not distinctly influence the rainfall characteristics over the equatorial eastern Africa. The search of the linkages between the intraseasonal statistics of the wet and dry spells was therefore confined to the equatorial, tropical and mid-latitudes regions.

The nesting of the oceanic field was based on the assumption that SST fields with large spatial extent at far distance may be expected to influence the East Africa climate just like SST fields with small spatial extent at close distance. For atmospheric fields, the lower, middle and upper atmospheric levels can be adequately represented by the 925mb, 700mb and 200mb. The search for linkages with atmospheric variables from re-analysis was therefore restricted to these levels with the exception of the specific humidity which excluded the upper atmospheric level.

Small shifts in the location of the predictors from one predictand to the next were assumed to reflect sampling errors rather than real climatic features. This tends to slightly reduce the variance explained by the multivariate linear regression models developed from these few predictors. Philippon (2002) has indicated that it is desirable that physical consistency outweighs statistical skill in empirical climate prediction.

The identification of linkages between the large–scale climate fields and interannual variability of the sub-regional intraseasonal statistics of the wet and dry spells (SRISS) was done by total and partial linear correlation analysis. The multivariate linear regression models that are developed to predict the SRISS were also linear. These two assumptions were made despite the fact that climatic processes are non-linear. Under certain circumstances, the predictive part may therefore be underestimated.

The results obtained and conclusions derived in the next chapter are thus based on these major assumptions, taking into account the limitations already stated.

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.0 Introduction

This chapter presents the results obtained from various methods that were used to achieve the overall and specific objectives of the study. The results from data quality control analysis are however presented first since the quality of the data used in any study form fundamental basis upon which the information is derived and conclusions drawn. The methods used to estimate the missing data and the quality control checks were presented in **section 3.2.1**.

4.1 Data management

4.1.1 Double mass curve homogeneity test

Results from the double mass curve analysis of the gap-filled daily rainfall data indicated that a single straight line could be fitted to the cumulative seasonal rainfall totals for any two chosen stations. These results were similar to those obtained by Gitau (2005) and Komutunga (2006) among others. **Figures 4.1** and **4.2** show typical examples of the double mass curve that were obtained for the long and short rainfall season respectively.

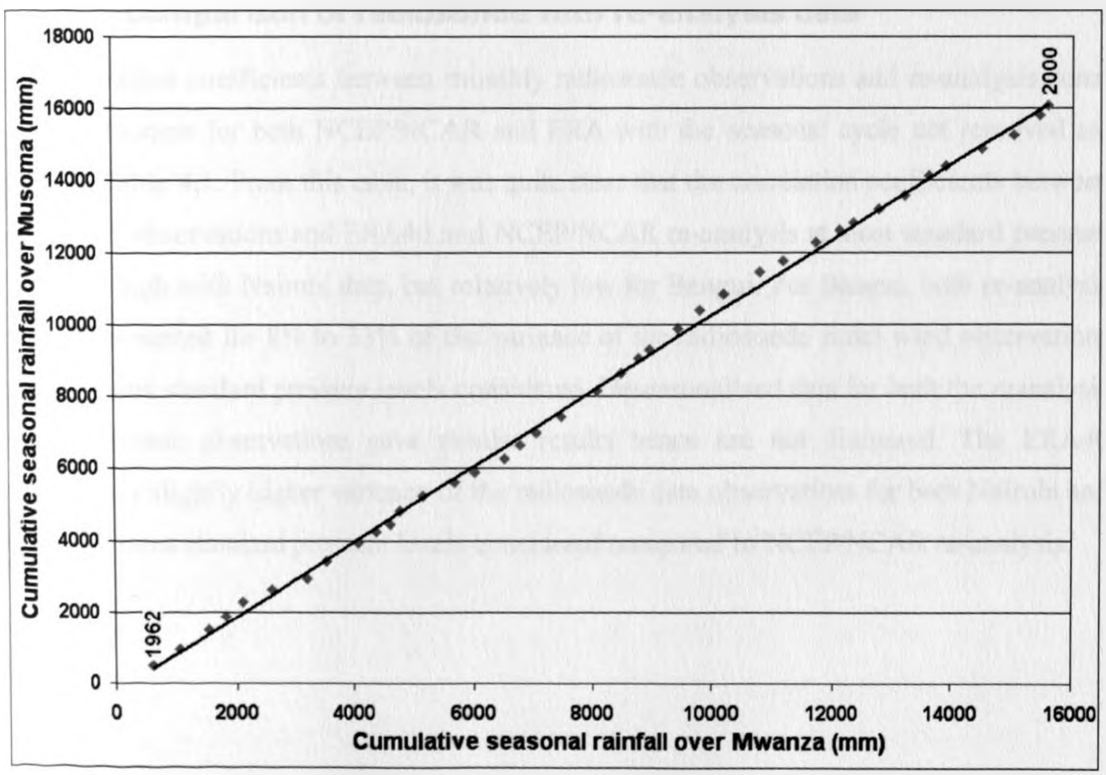


Figure 4.1: Double mass curve for Mwanza and Musoma during the long rainfall season

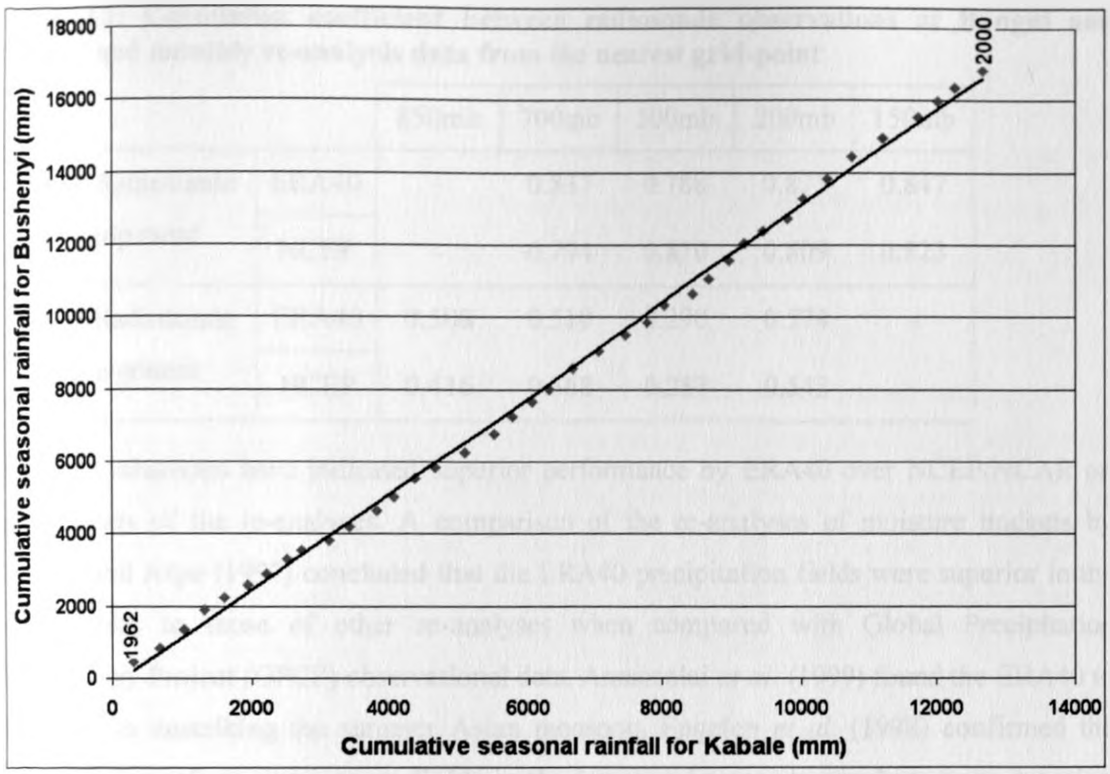


Figure 4.2: Double mass curve for Kabale and Bushenyi during the short rainfall season

4.1.2 Comparison of radiosonde with re-analysis data

The correlation coefficients between monthly radiosonde observations and re-analysis zonal wind component for both NCEP/NCAR and ERA with the seasonal cycle not removed are given in **Table 4.1**. From this table, it was quite clear that the correlation coefficients between radiosonde observations and ERA40 and NCEP/NCAR re-analysis at most standard pressure levels are high with Nairobi data, but relatively low for Bangui. For Bangui, both re-analysis records accounted for 8% to 33% of the variance of the radiosonde zonal wind observations at the various standard pressure levels considered. Deseasonalised data for both the reanalysis and radiosonde observations gave similar results hence are not discussed. The ERA40 accounts for slightly higher variance of the radiosonde data observations for both Nairobi and Bangui at most standard pressure levels considered compared to NCEP/NCAR re-analysis.

Table 4.1: Correlation coefficient between radiosonde observations at Bangui and Nairobi and monthly re-analysis data from the nearest grid-point

		850mb	700mb	500mb	200mb	150mb
Nairobi Radiosonde U-component	ERA40	-	0.837	0.788	0.877	0.847
	NCEP	-	0.794	0.810	0.809	0.823
Bangui Radiosonde U-component	ERA40	0.508	0.510	0.290	0.574	-
	NCEP	0.416	0.368	0.287	0.543	-

Several evaluations have indicated superior performance by ERA40 over NCEP/NCAR on some facets of the re-analyses. A comparison of the re-analyses of moisture budgets by Stendel and Arpe (1997) concluded that the ERA40 precipitation fields were superior in the extra-tropics to those of other re-analyses when compared with Global Precipitation Climatology Project (GPCP) observational data. Annamalai *et al.* (1999) found the ERA40 to be better in describing the summer Asian monsoon. Engelen *et al.* (1998) confirmed the ECMWF re-analysis water vapor fields in the lower and upper troposphere were superior. Newman *et al.* (2000) evaluation of the NCEP, National Aeronautics and Space Administration (NASA), and ERA40 which focused especially on the warm pool area of the Pacific from the standpoint of outgoing longwave radiation, precipitation and 200-mb divergence, found substantial problems with all re-analyses, although ERA40 gave the best estimates of the 200-mb divergence.

A study by Camberlin *et al.* (2001) over Africa, south of Sahara using split moving-windows dissimilarity analysis (Cornelius and Reynolds, 1991; Kemp *et al.*, 1994) has shown that major discontinuities exist in the time series of five NCEP/NCAR variables considered prior to 1968 at nearly all levels but more widespread for the lower troposphere. The five variables that were considered are the zonal (u) and meridional (v) components of the wind, geopotential height (φ), air temperature (T) and specific humidity (H).

The observed differences in the re-analysis datasets are due to the different observational databases, different analysis systems that may run at different resolutions, and different model dynamics and physics. In the tropics, the constraint of geostrophy on the divergent circulation is weak and thus there is considerable sensitivity to the diabatic heating field particularly that associated with moist processes (Annamalai *et al.*, 1999). In data sparse areas, the analysis heavily depends on the first guess supplied by the forecast model which in

return will be sensitive to the diabatic heating distribution produced by the physical parameterizations used in the model.

The consistency of the ERA40 in replicating the radiosonde observations at the two locations indicates the degree to which it represents the truth and hence its reliability when used for model development and validation. It was therefore used in the subsequent analysis in this study.

For the data management, the study concluded that the daily rainfall observations were of good quality and ERA40 re-analysis was most representative of the radiosonde observations. They could therefore be subjected to further analyses in order to attain the overall and specific objectives of the current study. These quality controlled data formed the foundation of all the analyses that were undertaken in this study. The results of zoning the study domain into few near-homogeneous sub-regions are discussed in the next section.

4.2 Near-homogeneous sub-regions for the study area

Most of the zoning of rainfall network into near-homogeneous rainfall sub-regions over the East Africa has been based on the Rotated Principal Component Analysis (RPCA). These include the studies by Ogallo (1980), Basalirwa (1991), Indeje *et al.* (2000) and Komutunga (2006) among others. However, none of these studies have zoned the Eastern Africa region into near-homogeneous rainfall sub-regions based on the observed gauged daily rainfall. The results obtained from this study were nevertheless compared with those of other studies that used observed gauged rainfall data at other timescales.

Application of the Rotated Empirical Orthogonal Functions (REOF) and simple correlation analysis to the gap-filled quality controlled daily rainfall data yielded 6 near-homogenous rainfall sub-regions for both the long and the short rainfall seasons as shown by **Figures 4.3** and **4.4** respectively. This simply means that only 6 Principal Components were found to be significant at 95% confidence level according to Monte Carlo testing. It should be stressed that REOF produces patterns of rainfall variability rather than patterns of actual rainfall since the data after square-root transformation were normalized before the procedure was carried out and the mean rainfall removed (Williams *et al.*, 2007). The regionalisation was therefore based on the occurrence and intensity of daily rainfall events, so that stations/locations that receive rainfall under related synoptic conditions fall within the same sub-region (Tennant and Hewitson, 2002).

The six near-homogeneous zones derived from the daily rainfall series over East Africa during the long rainfall season (**Figure 4.3**) are: 1. CK – Central and Western Kenya, 2. CS – Coastal strip of Kenya and Tanzania, 3. NK – Northeastern Kenya, 4. WT – Western Tanzania and Southern Uganda, 5. EH – Southeastern lowlands of Kenya and Northeastern Tanzania, and 6. WU – Western Uganda consisting of most parts of Uganda. However, Arua and Lodwar (located at the northern fringe of study region) could not unambiguously be attributed to any of the near-homogeneous sub-regions. They were thus grouped together with the nearest stations.

The six near-homogeneous zones delineated for the daily rainfall series during the short rainfall season (**Figure 4.4**) are: 1. CK – Central Kenya and southeastern lowlands, 2. WU – Western Kenya and most parts of Uganda, 3. NK – Northeastern Kenya, 4. CS – Coastal strip of Kenya and Tanzania, 5. CT – Central and Northern Tanzania, and 6. LV – Western of Lake Victoria and western Tanzania. The spatial patterns of the near-homogeneous zones for the two seasons are quite different with the exception of northeastern Kenya and the coastal strip of Kenya and Tanzania, which looks similar with some few modifications.

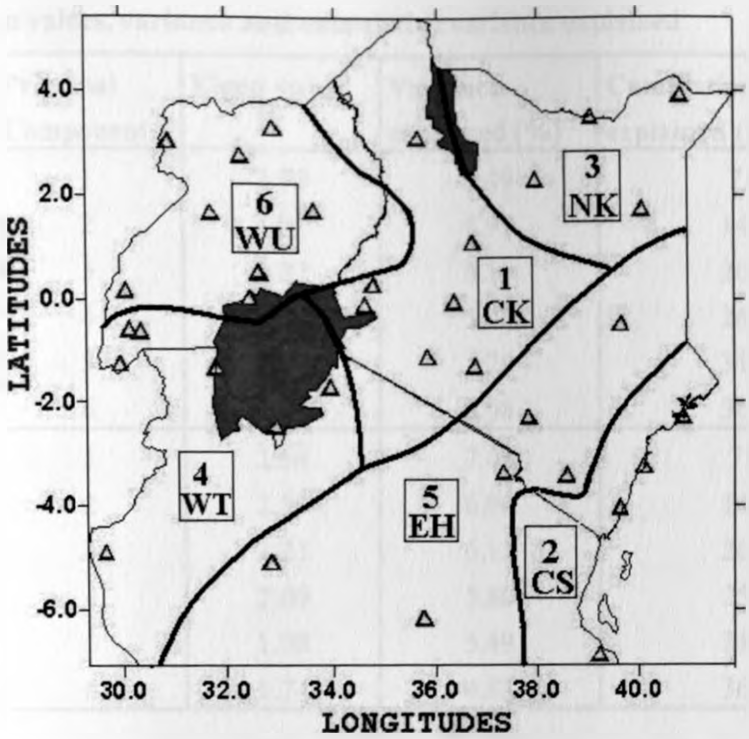


Figure 4.3: The six near-homogeneous sub-regions obtained from daily rainfall series during long rainfall season

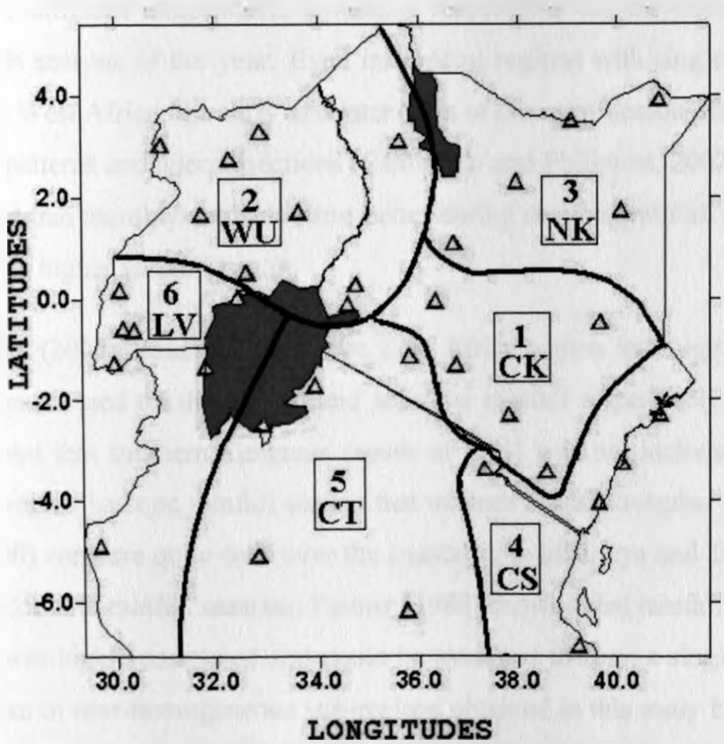


Figure 4.4: The six near-homogeneous sub-regions obtained from daily rainfall series during short rainfall season

Table 4.2: Eigen values, variance and cumulative variance explained

	Principal Component	Eigen value	Variance explained (%)	Cumulative variance explained (%)
Long Rainfall Season (March-April-May)	1	2.70	7.49	7.49
	2	2.51	6.97	14.46
	3	2.22	6.16	20.62
	4	2.05	5.70	26.32
	5	1.89	5.26	31.59
	6	1.83	5.08	36.67
Short Rainfall Season (October-November-December)	1	2.54	7.04	7.04
	2	2.50	6.96	14.00
	3	2.21	6.13	20.13
	4	2.09	5.80	25.93
	5	1.98	5.49	31.42
	6	1.74	4.83	36.26

Though the results indicate six significant PCs for both long and short rainfall seasons, the spatial patterns for the near-homogeneous rainfall sub-regions have slight variations. This may point to the different atmospheric dynamics responsible for the behavior of climate during the various seasons of the year. Even in tropical regions with single-season regimes such as India and West Africa, the early and later parts of the rainy season often exhibit slight distinct regional patterns and teleconnections (Camberlin and Philippon, 2002). Ininda (1995) has recommended that monthly analysis were better during the long rainfall season since the season experiences higher variability.

A study by Indeje (2000) classified the entire East Africa region into eight and nine near-homogeneous zones based on the annual and seasonal rainfall respectively. It is worthy to clarify at this point that southern Tanzania (south of 7° S) was not included in the current study since this region has one rainfall season that extends from November to May. Results from Indeje (2000) compare quite well over the coastal strip of Kenya and Tanzania (Zone 1 in Indeje, 2000) for both rainfall seasons. Farmer (1988) showed that rainfall over the coastal areas of Kenya was highly correlated and could be averaged to form a single rainfall index. The spatial pattern of near-homogeneous sub-regions obtained in this study is quite similar to those constructed by Ogallo (1988) and Indeje (2000) with some slight variations. The slight variations are attributed to the fact that the spatial coherence assessed from daily rainfall data arises from both interannual and intraseasonal variability, whereas the use of seasonal and

annual totals provides information on interannual variability only.

Collectively, the six significant Principal Components (PCs) accounted for 36.67% and 36.26% of the total variance in the daily rainfall over the Equatorial East Africa region during the long and short rainfall seasons respectively (**Table 4.2**). Compared to the total variance in the seasonal rainfall explained by the first six significant PCs of 57% and 73% for long and short rainfall seasons respectively (Indeje 2000), the current values are quite low. Camberlin and Philippon (2002) have indicated that for the Greater Horn of Africa, it takes 11 components to explain half of the total variance in the month-by-month analysis using the extended PCA. Further, the use of the scree test and examination of the spatial patterns retained only the first 6 PCs that are subjected to Varimax orthogonal rotation. These 6 significant PCs together have explained 35.5% of the variance (Camberlin and Philippon, 2002).

The low percentage of the total variance of daily rainfall explained in the current study may be attributed to the noise in the daily rainfall which is smoothed out when summed to obtain seasonal and annual totals and also the fact that the synoptic and mesoscale systems that mainly influence rainfall at a daily time scale, but only affect a few stations at a time cannot be captured at seasonal or higher timescales.

Station rainfall data are only representative of some area of variable size and shape surrounding the rain gauge (Huffman *et al.*, 1997). Such point measurements of a spatially variable parameter can be highly erratic relative to the rest of the area, and be a poor representation of the effect of the large scale processes. This is particularly true of daily data. Clustering the rainfall stations into near-homogeneous rainfall sub-regions based on occurrence of rainfall events overcame this problem. Hence stations that receive rain under similar synoptic weather conditions are placed in the same sub-region. Near-homogeneous sub-regions based on daily rainfall events captured the higher-order intraseasonal variations in rainfall better so that individual synoptic situations and atmospheric circulation anomalies plays a greater role.

In conclusion, the study has for the first time been able to classify occurrence of daily rainfall over the equatorial eastern Africa region into six near-homogeneous rainfall sub-regions for both the long and short rainfall seasons. There were however significant spatial differences in the patterns for the individual seasons. The total variance of the daily rainfall explained by

the six significant PCs are 36.67% and 36.26% for the long and short rainfall seasons respectively. Though the total variance explained is low, it can be seen that the sub-regions obtained by the REOF and simple correlation are climatologically reasonable (i.e. coherent sub-regions within which high covariance is expected) and they seem related to specific topographic contexts.

Upon the assessment of the data quality, identification of the significant RPCs and classification of the study domain into near-homogeneous rainfall sub-regions, the baseline characteristics of the wet/dry spells over the study area were studied next.

4.3 Baseline information of wet and dry spells

As indicated in **section 3.2.3.1**, only one threshold of 1.0 mm was used to delineate the wet days from dry ones. A higher threshold of 3.0mm was found to substantially reduce the sample size of the data required for further analysis and make comparisons with most other studies less straightforward. It was hence dropped. This section presents the various aspects of the intraseasonal statistics of the wet and dry spells.

4.3.1 Local intraseasonal statistics of wet and dry spells

Initially the various local (station level) intraseasonal statistics of wet and dry spells (LISS) for the long and short rainfall seasons are individually discussed and compared. Next, the association of the LISS with the seasonal rainfall was discussed and finally the trend result of the LISS with time was presented. The LISS for the long rainfall (MAM) season are presented and discussed first. For both **figures 4.5** and **4.6**, the kriging method of interpolation was used to extrapolate on the areas where the data was not available. Details on this interpolation method can be found in Cressie (1990, 1991).

4.3.1.1 Local intraseasonal statistics during long rainfall season

Figures 4.5a and **b** show the spatial pattern of the mean length of wet and dry spells respectively over the study area during the long rainfall season. **Figure 4.5a** shows that longer wet spells are reported over the coastal strip of East Africa, northern Tanzania closer to Mt Kilimanjaro and next to Lake Victoria, central and western Kenya. **Figure 4.5b** shows that longer dry spells are confined to northern and eastern Kenya, Central Tanzania and northeastern part of Uganda. The mean frequency of wet (dry) spells exceeding 3 (5) days and the longest wet (dry) spells had similar spatial patterns hence they are not shown. These patterns relatively conform to those of the mean seasonal rainfall amounts with drier (wetter)

areas tending to show longer (shorter) dry spells and shorter (longer) wet spells.

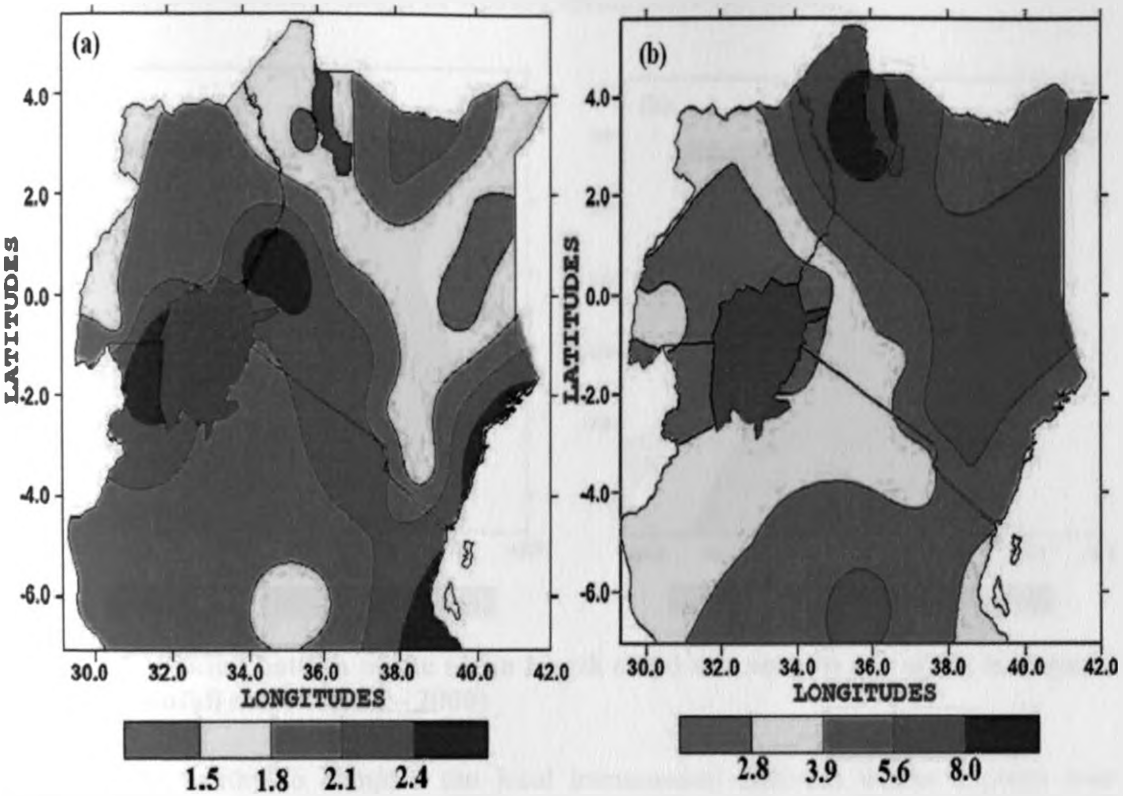


Figure 4.5: Spatial pattern of the mean length of (a) wet and (b) dry spells in days for the long rainfall season (1962 – 2000)

4.3.1.2 Local intraseasonal statistics during short rainfall season

Figures 4.6a and **b** show the local intraseasonal statistics of the wet and dry spells respectively during the short rainfall season. Notable differences were found when the mean lengths of the wet spells for the two wet seasons are compared. The longer wet spell that was observed over the coastal strip during the long rainfall season was notably absent. However for the short rainfall season, most parts along the coast have much less rainfall than during the long rainfall season. The longer wet spell that had been observed over the western parts of Lake Victoria extended into parts of western Uganda and most parts of western Tanzania (**Figure 4.6a**). It can also be seen that the mean length of a wet spell is slightly longer during the long rainfall season over most parts of study domain. This may help to explain why the long rainfall season receives slightly more rainfall than the short rainfall season over most parts.

In the case of the mean length of the dry spell (**Figure 4.6b**), the patterns are closer to those found during the long rainfall season (**Figures 4.5b**). The spatial patterns of the mean

frequency of wet (dry) spells of 3 (5) days or more and those of the longest wet/dry spells are similar to those of the mean length of wet/dry spells, hence not shown.

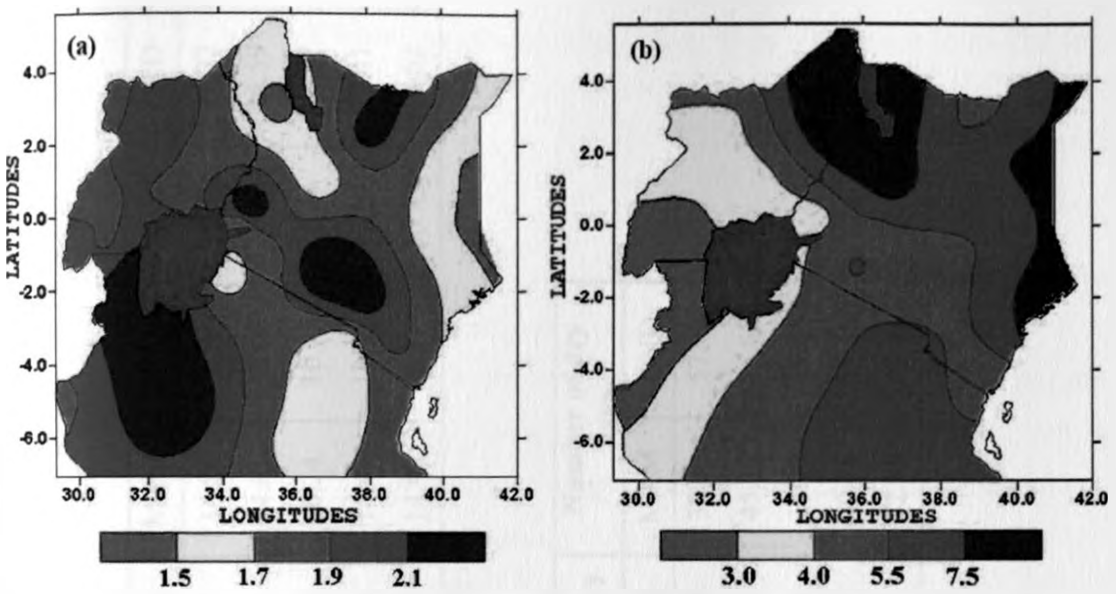


Figure 4.6: Spatial pattern of the mean length of (a) wet and (b) dry spells in days for the short rainfall season (1962 – 2000)

It is equally worthy to compare the local intraseasonal statistics within a given near-homogeneous sub-region as derived in **section 4.2** for the two rainfall seasons. Two such sub-regions remain relatively unchanged between the long and the short rainfall seasons. They are the coastal strip of Kenya and Tanzania and the northeastern parts of Kenya **(Figures 4.3** and **4.4)**. The coastal strip (zone 4 and zone 2 for the long and short rainfall season respectively with an additional station of Moshi during the short rainfall season) was selected in order to compare the local intraseasonal statistics during the two rainfall seasons. The results of this comparison are shown in **Table 4.3**.

The study concluded that during the long rainfall season, the mean lengths of the wet (dry) spells are longer (shorter) and records the longest wet spell. Higher (lower) mean frequency of wet (dry) spells of 3 (5) days or more were obtained in the long rainfall season as compared to the short rainfall at both local and sub-regional levels. There are more wet days during long rainfall season as compared to the short rainfall season which has more dry days. In terms of spatial patterns, the local intraseasonal statistics of the dry spells did not show marked differences between the long and the short rainfall seasons, unlike the case for the local intraseasonal statistics of wet spells that had marked differences between the two seasons.

Table 4.3: Local intraseasonal statistics of (a) wet and (b) dry spells over coastal strip of East Africa

(a) Wet spells	Mean Length		Longest Spell		3 days or more wet		Mean Intensity		Number of wet days	
	MAM	OND	MAM	OND	MAM	OND	MAM	OND	MAM	OND
Malindi	2.92	1.67	20	10	4.26	1.67	14.88	9.39	34.33	17.72
Mombasa	2.45	1.90	16	9	3.56	2.85	14.81	10.66	30.00	25.59
Lamu	2.78	1.56	23	10	3.38	1.23	16.24	10.16	28.54	12.97
Dares Salaam	2.85	1.82	17	11	6.10	2.59	14.41	13.71	42.64	22.41
Moshi*	2.44	1.53	21	7	5.13	1.31	14.93	8.49	36.82	15.62

* Indicates a station in a different sub-region during the long rainfall season

(b) Dry spells	Mean Length		Longest Spell		5 days or more dry		Number of dry days	
	MAM	OND	MAM	OND	MAM	OND	MAM	OND
Malindi	3.67	7.45	26	72	2.69	4.08	39.31	57.13
Mombasa	4.03	4.71	26	29	3.54	4.23	45.13	55.62
Lamu	4.09	10.04	28	64	2.92	3.08	38.39	49.05
Dares Salaam	2.93	5.96	26	39	2.49	4.26	40.28	57.46
Moshi	3.07	6.52	23	41	2.44	3.67	42.85	53.18

4.3.1.3 Relationship with the local seasonal rainfall totals

Figures 4.7a–i and **4.8a–i** indicate the spatial patterns of the Pearson correlation coefficient between seasonal rainfall totals and the various intraseasonal statistics of wet and dry spells for the long and short rainfall seasons respectively at local level. The results show that there is a significant positive (negative) correlation between the seasonal rainfall totals and the intraseasonal statistics of the wet (dry) spells over most locations during the two wet seasons considered.

During the long rainfall season, all individual locations indicated a strong significant positive correlation between the seasonal rainfall totals on one side and number of wet days in the season (**Figure 4.7a**) and the mean rainfall intensity per rain day (**Figure 4.7c**) on the other side. Most locations in Uganda and Tanzania had a weaker though significant positive correlation of seasonal rainfall and the mean length of a wet spell (**Figure 4.7b**), duration of the longest wet spell (**Figure 4.7f**) and the mean frequency of wet spell of 3 days or more (**Figure 4.7g**). However, they are a few exceptions. For instance, Bushenyi in southwestern Uganda had a negative (though insignificant) correlation between the seasonal rainfall totals and mean length of a wet spell (**Figure 4.7b**) and duration of the longest wet spell (**Figure 4.7f**).

The Pearson correlation coefficient between the intraseasonal statistics of dry spell and the seasonal rainfall totals during the long rainfall season are shown by **Figures 4.7 d–e, h–i**. The duration of the longest dry spell (**Figure 4.7 h**) was the least significantly correlated with the seasonal rainfall totals, followed by the mean frequency of the dry spell of 5 days or more (**Figure 4.7 i**) and the number of the dry days (**Figure 4.7 d**) in that order. The northern part of Kenya especially did not have statistically significant association (at 95% confidence level) with these three intraseasonal statistics of the dry spells. The mean length of the dry spell during the long rainfall season however had significant negative correlation with the seasonal rainfall totals over most locations (**Figure 4.7 e**).

The seasonal rainfall totals during the short rainfall season have a high correlation with the intraseasonal statistics of both wet and dry spells (**Figures 4.8a–i**). Only the number of wet days in a season (**Figure 4.8a**) had significant association (at 5% significant level) over all the locations. However, seasonal rainfall totals over Maralal in north Kenya for example have insignificant relationship with the mean length of the wet spell (**Figure 4.8b**), mean rainfall intensity per rain day (**Figure 4.8c**) and the duration of the longest wet spell (**Figure 4.8f**).

Some locations over southwestern Uganda and western Tanzania have insignificant relationship between the seasonal rainfall and mean frequency of wet spells of 3 days or more (**Figure 4.8g**). Similar to the long rainfall season, the mean length of the dry spell report the significant negative association with the seasonal rainfall totals over most locations (**Figure 4.8e**).

An interesting observation made during the short rainfall season was that the mean frequency of dry spells of 5 days (**Figures 4.8i**) and more and the number of dry days (**Figures 4.8d**) had a positive linear relationship with the seasonal rainfall totals over northern and northeastern parts of Kenya (Lodwar, Maralal, Marsabit, Mandera, and Wajir). Lamu and Dodoma were also noted to be in this group (**Figures 4.8d** and **i**). These locations are in the arid and semi-arid lands (ASALs) and receive little rainfall during this season. The above pattern was attributed to the fact that as the seasonal rainfall total increases, the seasonal length also increases and thus the number of dry days and mean frequency of dry spell of 5 days or more increases. Otherwise as the seasonal rainfall reduces, the number of the dry days reduces since the rest of the period within the season constitutes the dry season.

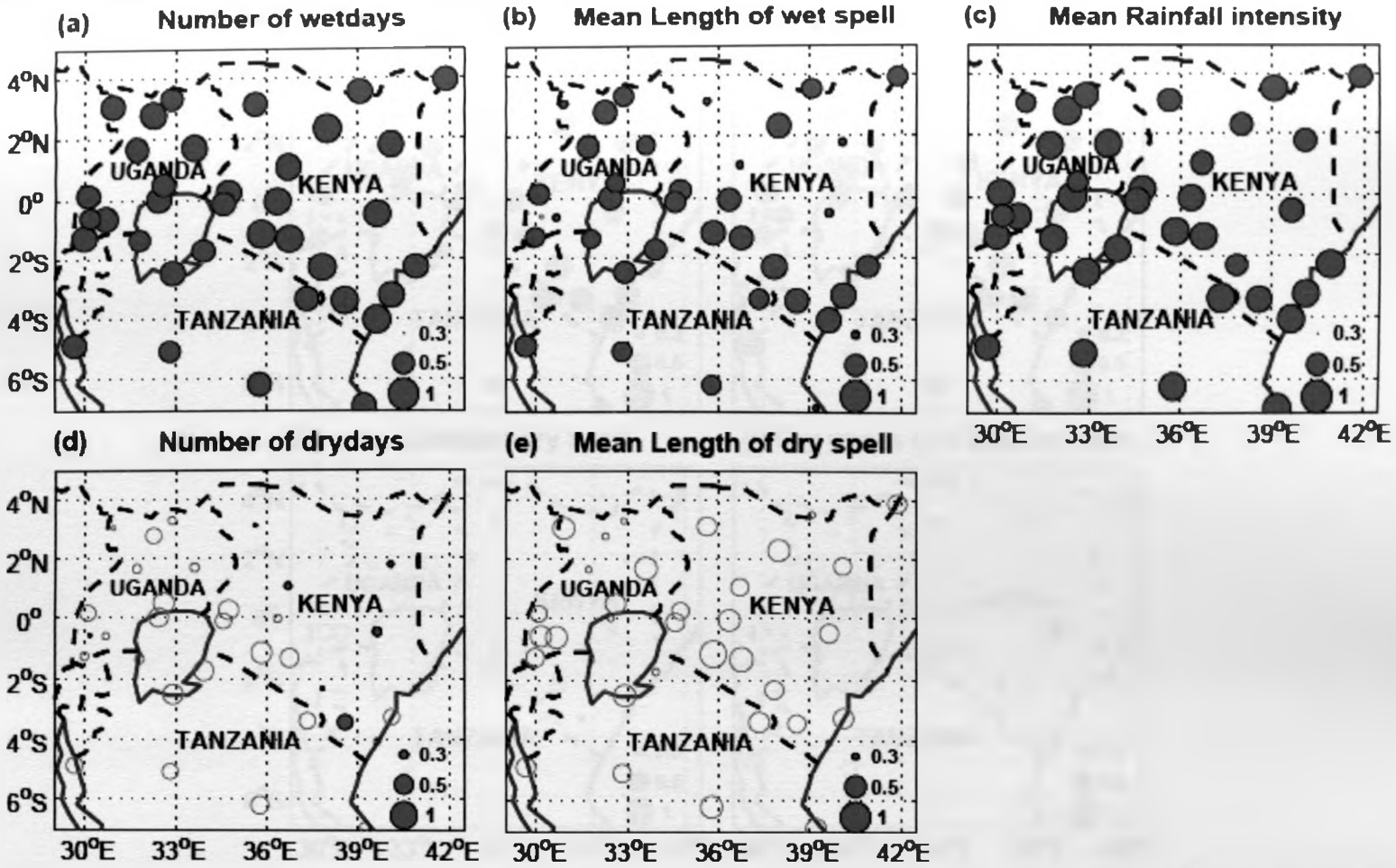


Figure 4.7: Maps of the Pearson correlation coefficient between seasonal rainfall totals and (a) number of wet days in the season, (b) mean length of wet spell, (c) mean rainfall intensity, (d) number of dry days in the season, and (e) mean length of dry spell during the long rainfall season. Closed (open) circles indicate positive (negative) correlation. Green (red) indicates the coefficient is significant (insignificant) at 95% confidence level

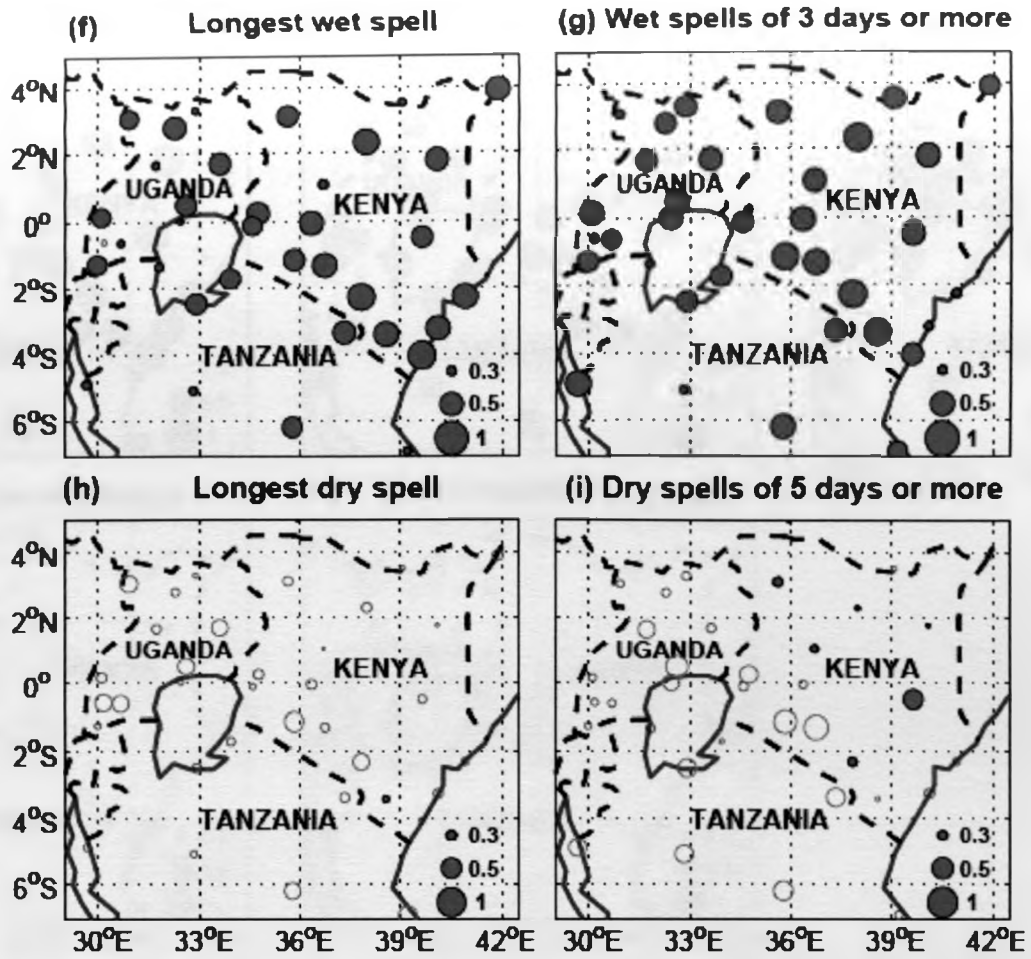


Figure 4.7 (*cont.*): Maps of the Pearson correlation coefficient between seasonal rainfall totals and (f) duration of the longest wet spell, (g) frequency of wet spells of 3 days or more, (h) duration of the longest dry spell, and (i) frequency of dry spells of 5 days or more during the long rainfall season. Closed (open) circles indicate positive (negative) correlation. Green (red) indicates the coefficient is significant (insignificant) at 95% confidence level

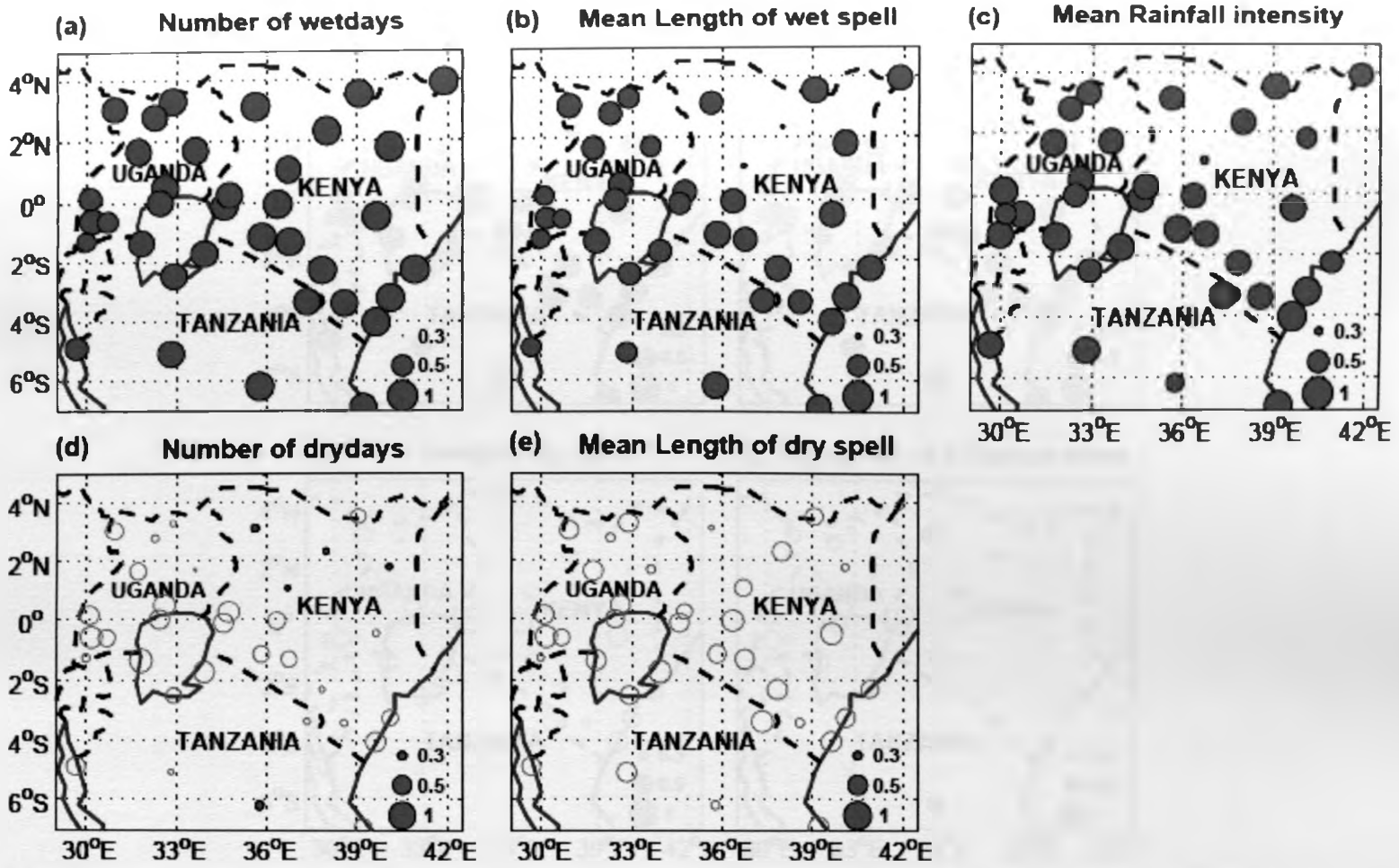


Figure 4.8: Maps of the Pearson correlation coefficient between seasonal rainfall totals and (a) number of wet days in the season, (b) mean length of wet spell, (c) mean rainfall intensity, (d) number of dry days in the season, and (e) mean length of dry spell during short rainfall season. Closed (open) circles indicate positive (negative) correlation. Green (red) indicates the coefficient is significant (insignificant) at 95% confidence level

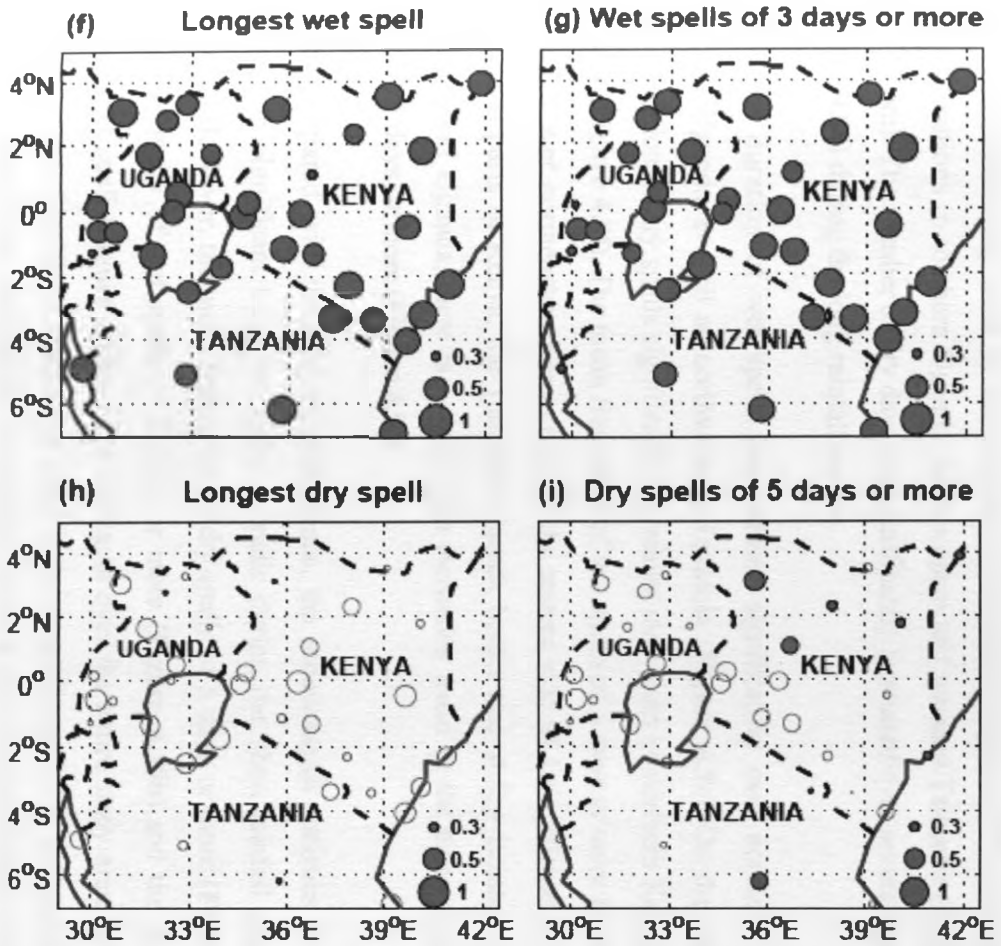


Figure 4.8 (*cont.*): Maps of the Pearson correlation coefficient between seasonal rainfall totals and (f) duration of the longest wet spell, (g) frequency of wet spells of 3 days or more, (h) duration of the longest dry spell, and (i) frequency of dry spells of 5 days or more during the short rainfall season. Closed (open) circles indicate positive (negative) correlation. Green (red) indicates the coefficient is significant (insignificant) at 95% confidence level

Trend results

for the trends of the seasonal rainfall and intraseasonal statistics of wet and dry :al scale are shown in **Figures 4.9a–j** and **4.10a–j** for the long and short rainfall spectively. It can be seen from **Figure 4.9a** that most locations have negative trend for the seasonal rainfall totals but were not significant during the ll season. However, Lodwar in northern Kenya and Bukoba in northwestern how significant decreasing trend for the seasonal rainfall total (**Figure 4.9a**). decreasing trend in the number of wet days (**Figure 4.9b**) was observed over ations in southern Uganda, northwestern and western Tanzania (Bukoba, Kigoma). The number of dry days has significantly increased over several parts of Uganda e) during the long rainfall season.

duration of wet spells has reduced significantly over northern, western and ganda as well as northwestern Tanzania (**Figure 4.9c**). On the other hand, the ion of dry spells significantly increased over parts of northern Kenya and western gure 4.9f**). The mean frequency of wet spells of 3 days or more have significantly ver northeastern Kenya and on the eastern side of Lake Victoria (**Figure 4.9h**). parts of Kenya, the Tanzania-Uganda border and at few isolated locations over nd Uganda, there is a significant increasing trend in the mean frequency of dry days or more (**Figure 4.9j**).

pared with the long rainfall season, the intraseasonal statistics of wet and dry significant trend are quite sporadic during the short rainfall season (**Figures** lowever, the mean frequency of dry spells of 5 days or more (**Figure 4.10j**), the uency of wet spells of 3 days or more (**Figure 4.10h**) and the duration of the t spell (**Figure 4.10g**) have increased over the entire study area during the short ison. Significant increasing trend in the mean frequency of dry spells of 5 days or noted over most parts of Uganda, western and coastal Kenya during the short ason (**Figure 4.10j**). In their study, Ambenje *et al.* (2001) had noted that most the tropics exhibited a reduction (though not significant) in both the seasonal als and associated frequency which is consistent with the current results.

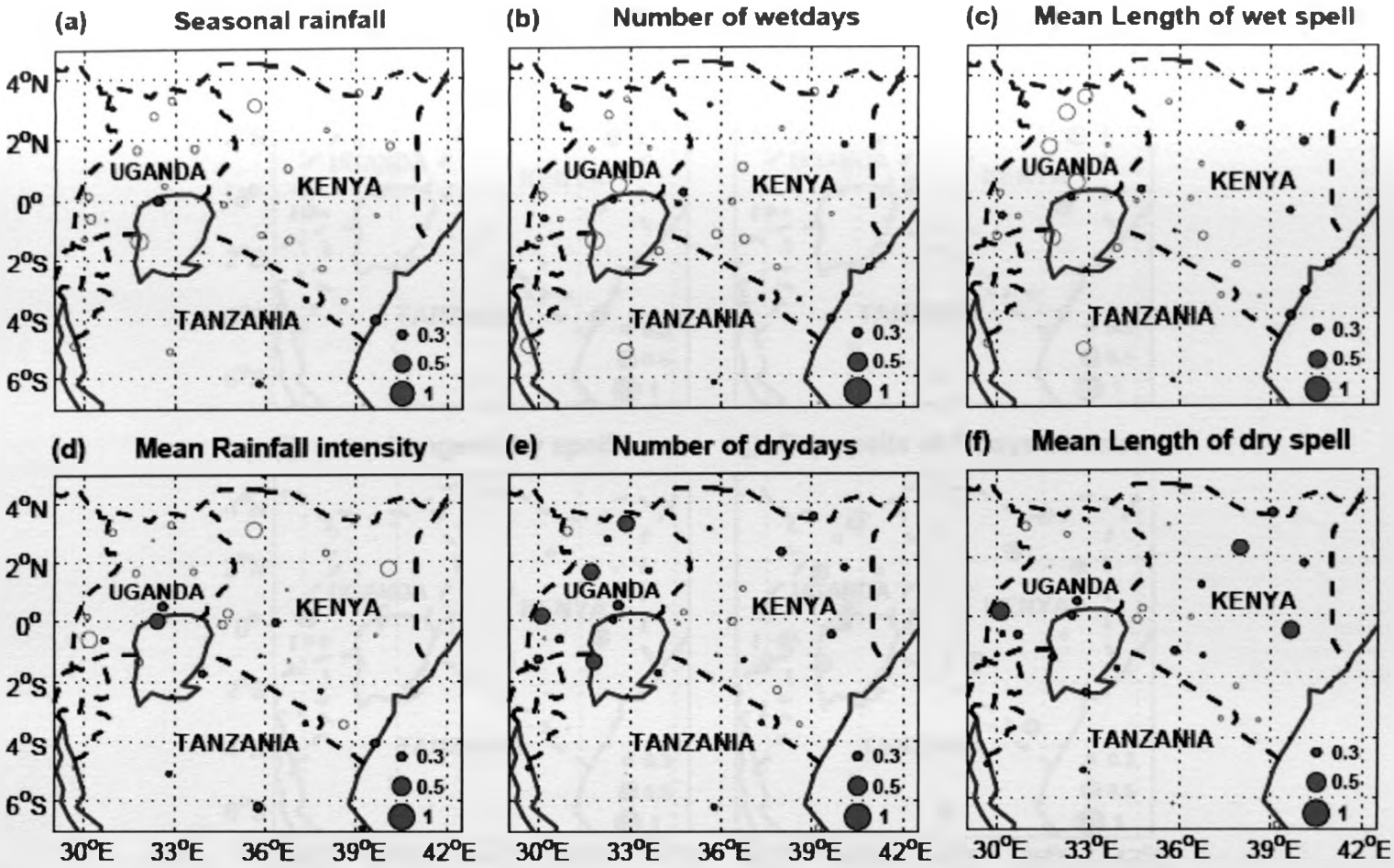


Figure 4.9: Maps of the Spearman rank correlation coefficient of (a) seasonal rainfall totals, (b) number of wet days in the season, (c) mean length of wet spell, (d) mean rainfall intensity, (e) number of dry days in the season, and (f) mean length of dry spell during the long rainfall season. Closed (open) circles indicate increasing (decreasing) trend. Green (red) indicates the trend is significant (insignificant) at 95% confidence level

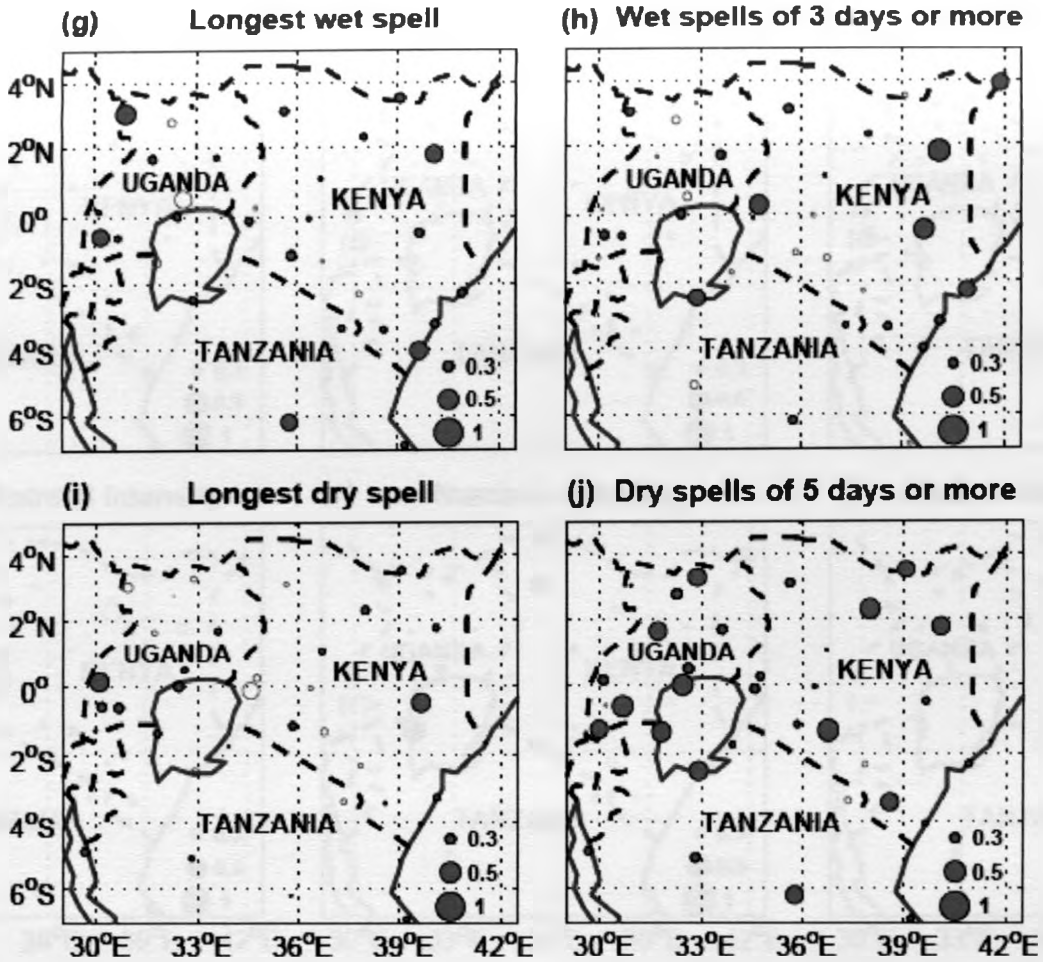


Figure 4.9 (*cont.*): Maps of the Spearman rank correlation coefficient of (g) duration of the longest wet spell, (h) frequency of wet spells of 3 days or more, (i) duration of the longest dry spell, and (j) frequency of dry spells of 5 days or more during the long rainfall season. Closed (open) circles indicate increasing (decreasing) trend. Green (red) indicates the trend is significant (insignificant) at 95% confidence level

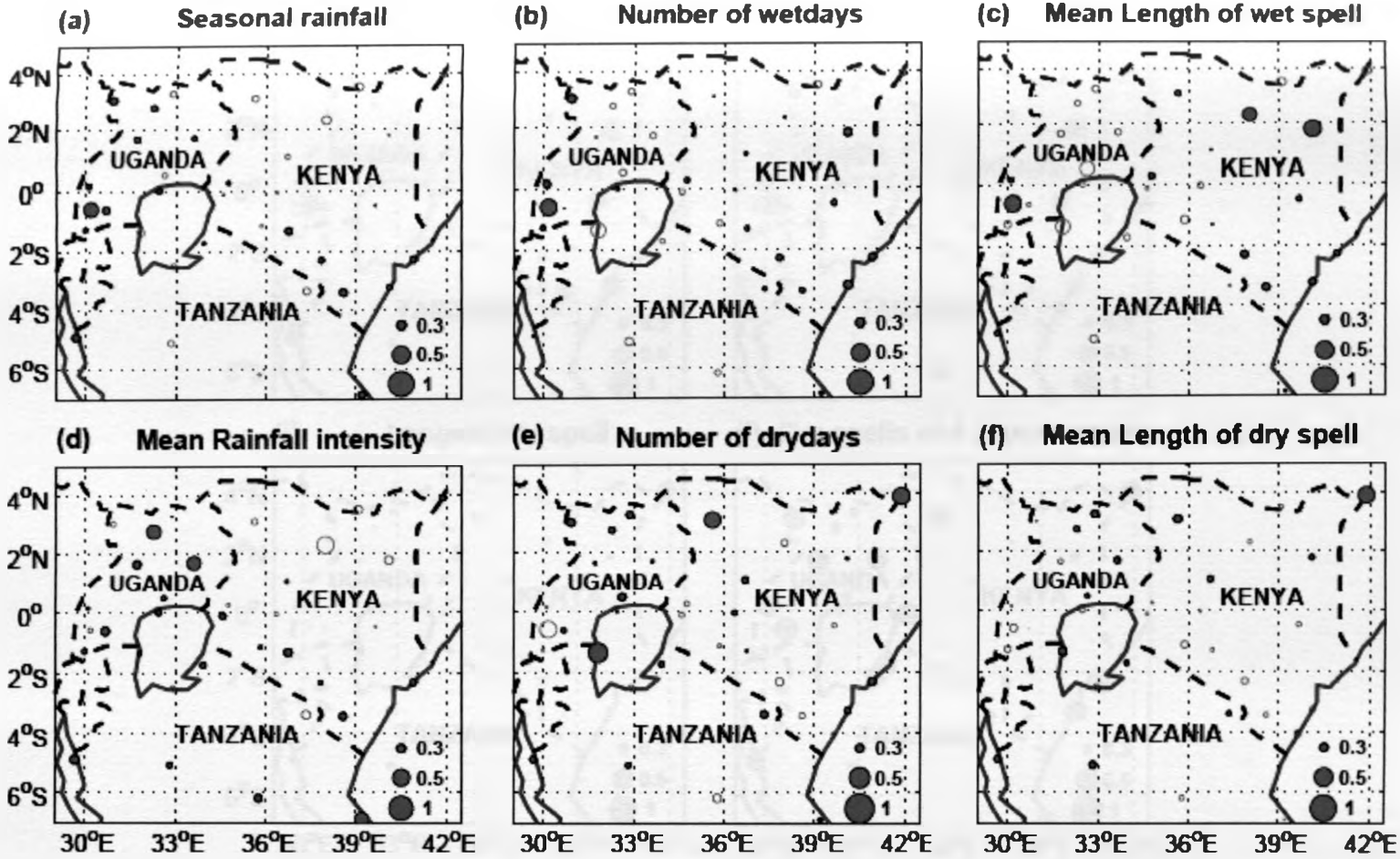


Figure 4.10: Maps of the Spearman rank correlation coefficient of (a) seasonal rainfall totals, (b) number of wet days in the season, (c) mean length of wet spell, (d) mean rainfall intensity, (e) number of dry days in the season, and (f) mean length of dry spell during the short rainfall season. Closed (open) circles indicate increasing (decreasing) trend. Green (red) indicates the trend is significant (insignificant) at 95% confidence level

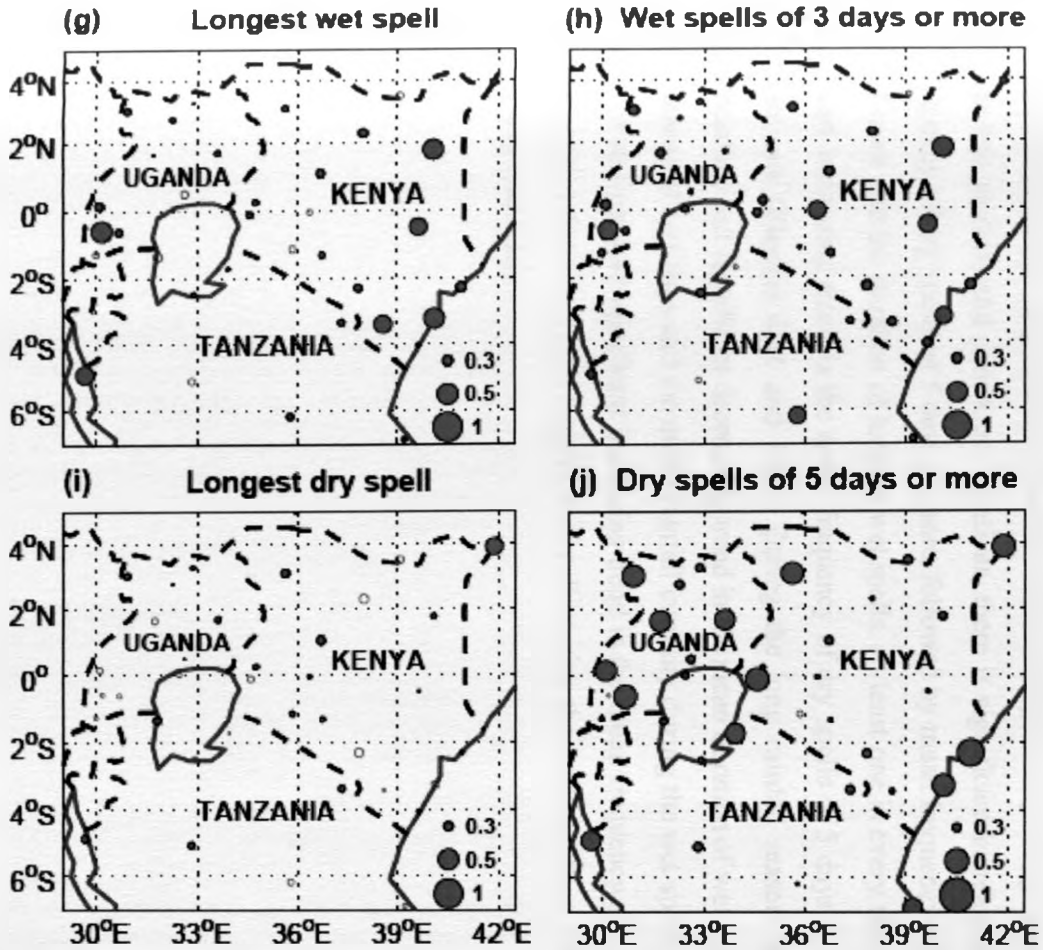


Figure 4.10 (*cont.*): Maps of the Spearman rank correlation coefficient of (g) duration of the longest wet spell, (h) frequency of wet spells of 3 days or more, (i) duration of the longest dry spell, and (j) frequency of dry spells of 5 days or more during short rainfall season. Closed (open) circles indicate increasing (decreasing) trend. Green (red) indicates the trend is significant (insignificant) at 95% confidence level

An interesting observation from **Figures 4.10a–j** is that, for the short rainfall season, the absence of any trend in seasonal totals masks out significant trends in the distribution of the rainfall (**Figures 4.10a**). The positive trends in both the 3-day wet spells (**Figures 4.9h** & **4.10h**) and 5-day dry spells (**Figures 4.9j** & **4.10j**) may reflect a change in the rainfall distribution, with longer spells becoming more common.

Figures 4.11 and **4.12** provide a summary on the percentage number of stations with significant trends for the long and short rainfall seasons respectively. They clearly show that during the two seasons and over most locations, there is significant increasing trend in the mean frequency of dry spells of 5 days or more, followed by mean frequency of wet spells of 3 days or more and the duration of longest wet spells. At least one in every three stations has a significant increasing trend in the mean frequency of dry spells of 5 days or more in both rainfall seasons (**Figures 4.11** and **4.12**). During the long rainfall season (**Figure 4.11**), several locations had significant decreasing trend in the mean duration of wet spells, followed by the number of wet days and the mean rainfall intensity during the wet spells. At least one in every six stations had significant increasing trend in the mean frequency of wet spells of 3 days or more and duration of the longest wet spells during the long and short rainfall seasons (**Figures 4.11** and **4.12**).

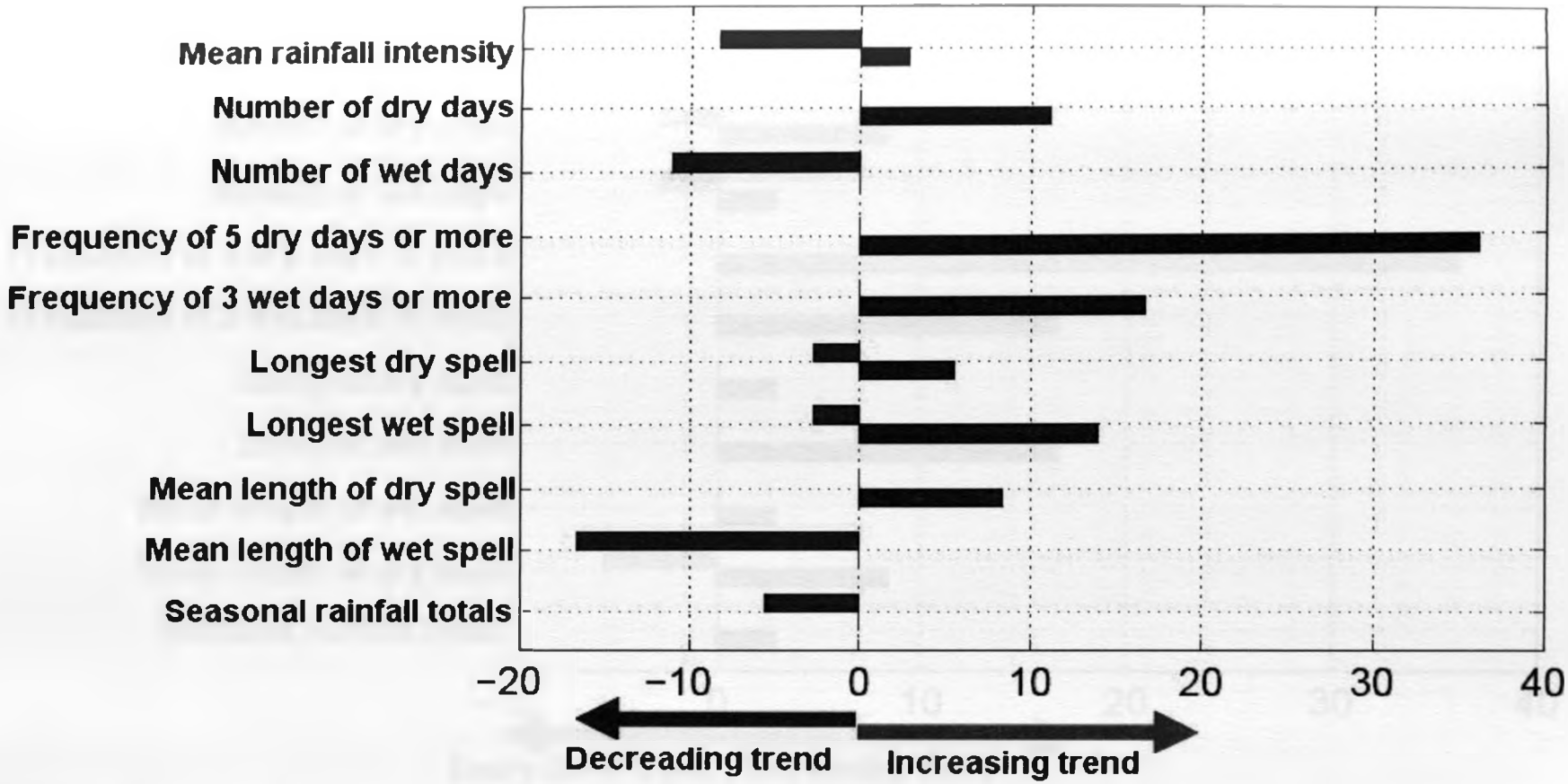


Figure 4.11: Percentage number of stations with significant decreasing (negative) and increasing (positive) trends for seasonal rainfall totals and the various intraseasonal statistics of wet and dry spells during the long rainfall (MAM) season

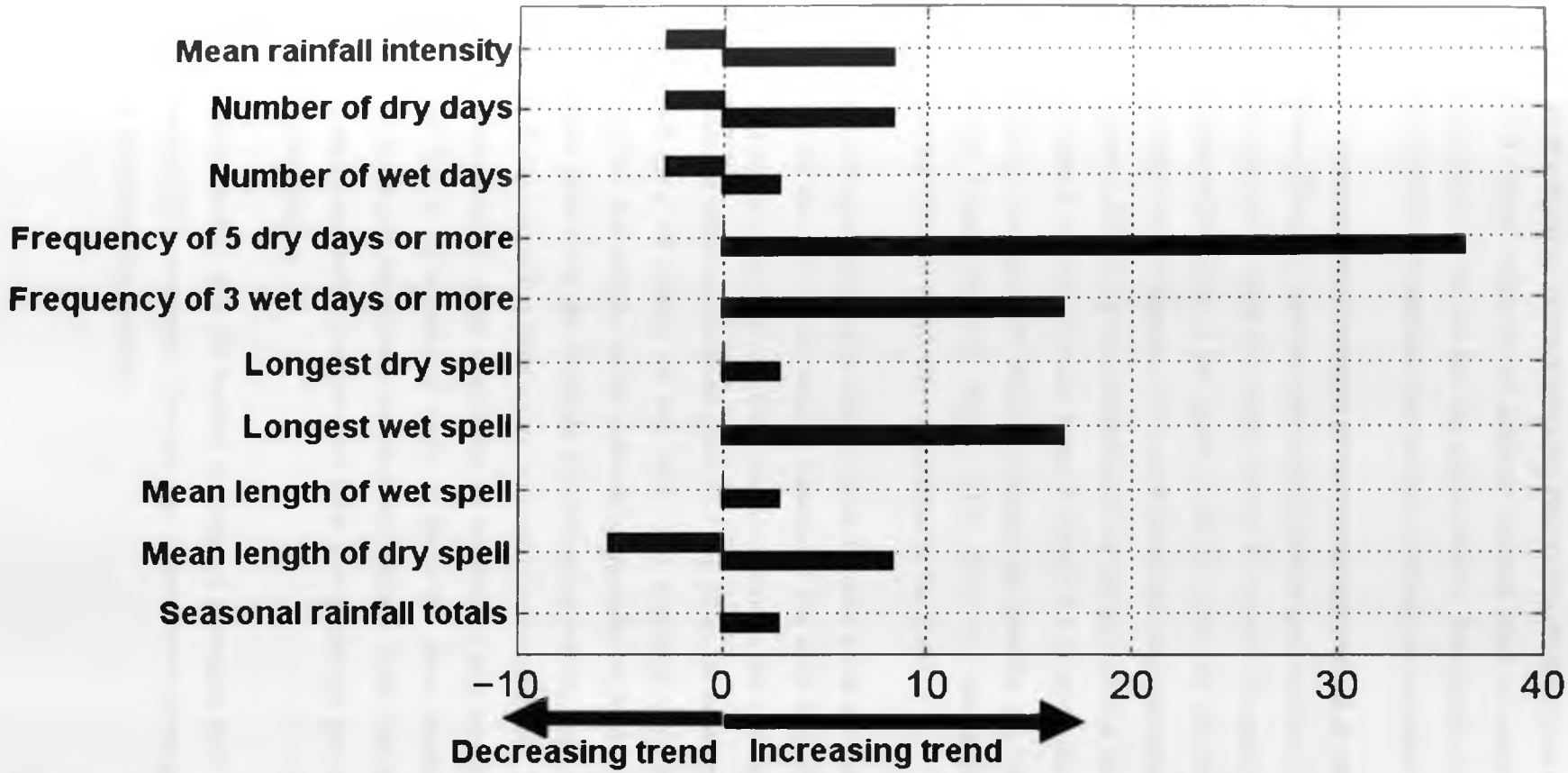


Figure 4.12: Percentage number of stations with significant decreasing (negative) and increasing (positive) trends for seasonal rainfall totals and the various intraseasonal statistics of wet and dry spells during the short rainfall (OND) season

4.3.2 Sub-regional intraseasonal statistics of wet and dry spells

Section 3.2.3.3 clearly indicated the different methods used to compute the sub-regional intraseasonal statistics of the wet and dry spells (SRISS). The results obtained are discussed in this section, but first the statistics from the three methods are compared.

4.3.2.1 Comparative analysis of the three definitions of SRISS

Figures 4.13a–c illustrates how the sub-regional intraseasonal statistics of wet and dry spells (SRISS) were derived by using the coastal strip of Kenya and Tanzania (sub-region 2) as an example. **Figure 4.13a** shows a line graph of the PC score for this sub-region during the MAM 1977 season. In this instance, the 1.0mm threshold used corresponds to -0.214 for the PC score. **Figure 4.13b** shows the distribution of the wet days at local (station) level where a red dot represents a wet day. The last graph in **Figure 4.13b** shows the distribution of wet days obtained by averaging the rainfall amounts and plotting the resultant series while maintaining the 1.0mm threshold. **Figure 4.13c** shows the distribution of the wet days obtained from the PC score which were represented as bar graph.

The local and sub-regional statistics obtained from **Figures 4.13a–c** are shown by **Table 4.4**. The table shows that there is an outright biasness if the daily rainfall amounts from the individual stations are averaged and then used to determine the sub-regional statistics. For instance, while the other two methods gives 31.5 and 29 as the number of wet days (NW), this approach gave 49 number of wet days. This approach tends to overestimate the components of the wet statistics while underestimating those of the dry statistics. Bärring *et al.* (2006) have shown that the threshold for delineating wet/dry days on area-average are quite different as compared to when using the point observational data. They found out that by using the threshold of 1.0mm to delineate the wet and dry days on the point observations, the threshold had to be adjusted in order to obtain the same results as those of point observations. Averaging the intraseasonal statistics obtained at the local level to obtain areal-averaged intraseasonal statistics on the other hand give results that are consistent with those of the PCA score analysis.

It is concluded therefore that the method of temporal averaging daily rainfall time-series before generating ISS is unsuitable. The next step is therefore to further assess the respective merits of the two remaining methods.

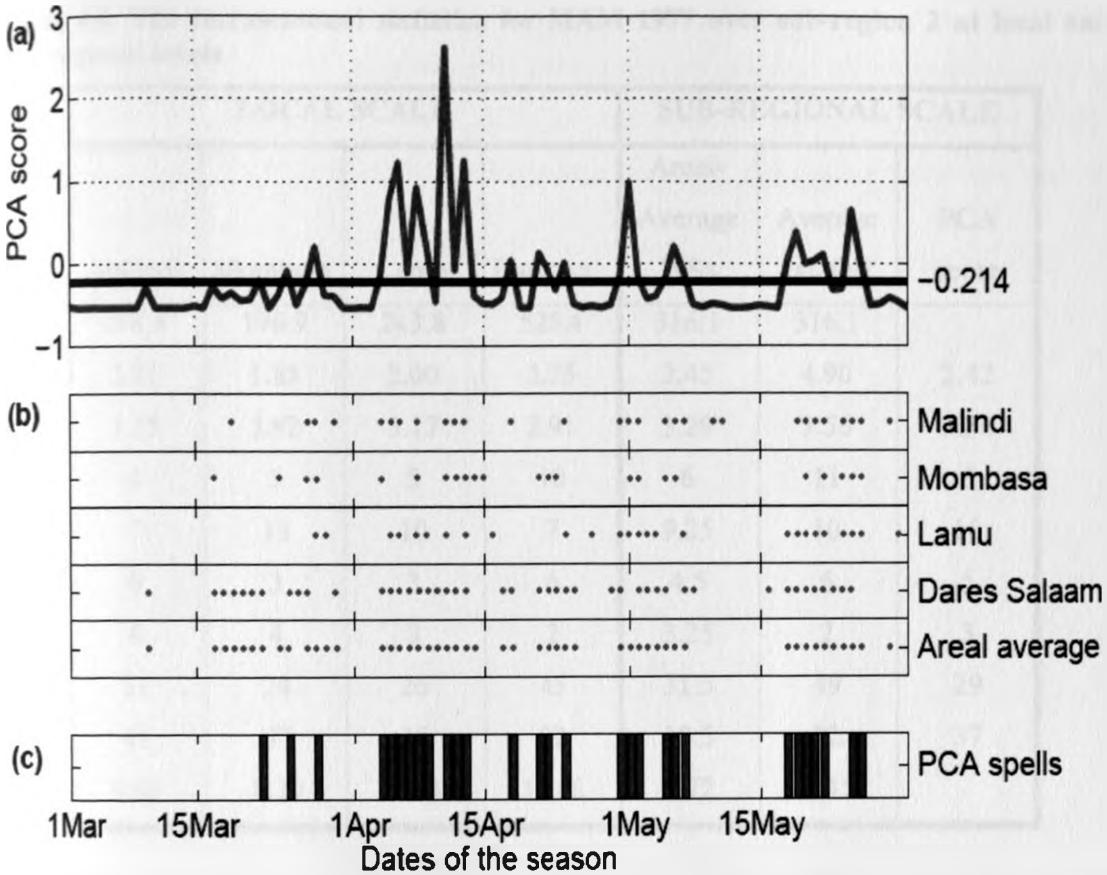


Figure 4.13: The temporal distribution of wet and dry spells during the MAM 1977 over the coastal strip of Kenya and Tanzania (sub-region 2) at local and sub-regional levels. (a) The PC score time series, (b) the distribution of wet and dry spells at four individual stations and from the areal-average rainfall for the four stations, and (c) the distribution of wet and dry spells from the PCA score time series. The x-axis is the dates of the season and is common to the three graphs.

Table 4.4: The intraseasonal statistics for MAM 1977 over sub-region 2 at local and sub-regional levels

ISS	LOCAL SCALE				SUB-REGIONAL SCALE		
	Malindi	Mombasa	Lamu	Dares sa	Areal-Average ISSs	Average rainfall	PCA score
SR	298.4	196.9	243.8	525.4	316.1	316.1	
MW	2.21	1.85	2.00	3.75	2.45	4.90	2.42
MD	3.15	3.92	3.17	2.91	3.29	3.56	3.36
LW	4	5	5	10	6	11	6
LD	7	13	10	7	9.25	10	10
3W	6	3	3	6	4.5	6	5
5D	4	4	3	2	3.25	2	3
NW	31	24	26	45	31.5	49	29
ND	41	47	38	32	39.5	32	37
MI	9.63	8.20	9.38	11.68	9.72	6.45	

The different intraseasonal statistics obtained at the local level (for each station) were correlated with those obtained for the PCA scores and those areal-averaged for each sub-region. The aim was to assess how representative of the local rainfall distribution were the types of the sub-region indices. The box-plots of the correlation coefficients during the long and short rainfall seasons are shown by **Figures 4.14** and **4.15** respectively. Both figures indicate that seasonal rainfall totals and number of wet days have the highest correlation coefficient in both cases while the mean frequency of dry spells of 5 days or more have the least coefficient. A closer look shows that during the long rainfall season there are no outliers unlike the short rainfall season (**Figures 4.14a** and **4.15a**). In the case of areal-average SRISS, both seasons show significant correlations for all the components considered (**Figure 4.14b** and **4.15b**). More outliers are also observed in these correlations as compared to the PCA-based SRISS, but on the whole the values obtained for the PCA-based data are lower, which simply means that the PCA-based data was less representative of the local rainfall distribution. Ogallo *et al.* (1988) have used both the PC-based and arithmetic areal-average seasonal rainfall totals indices at near-homogeneous zones over East Africa region to study their teleconnection with the global sea surface temperature anomalies.

It was concluded in this sub-section that the SRISS obtained from averaging the daily rainfall amounts from the individual stations are the most unrealistic and thus could not be used in the current study. The PCA-based SRISS is not as representative as the areal-average SRISS. However, it is free from outliers. The SRISS from these two approaches were thus kept for further analysis.

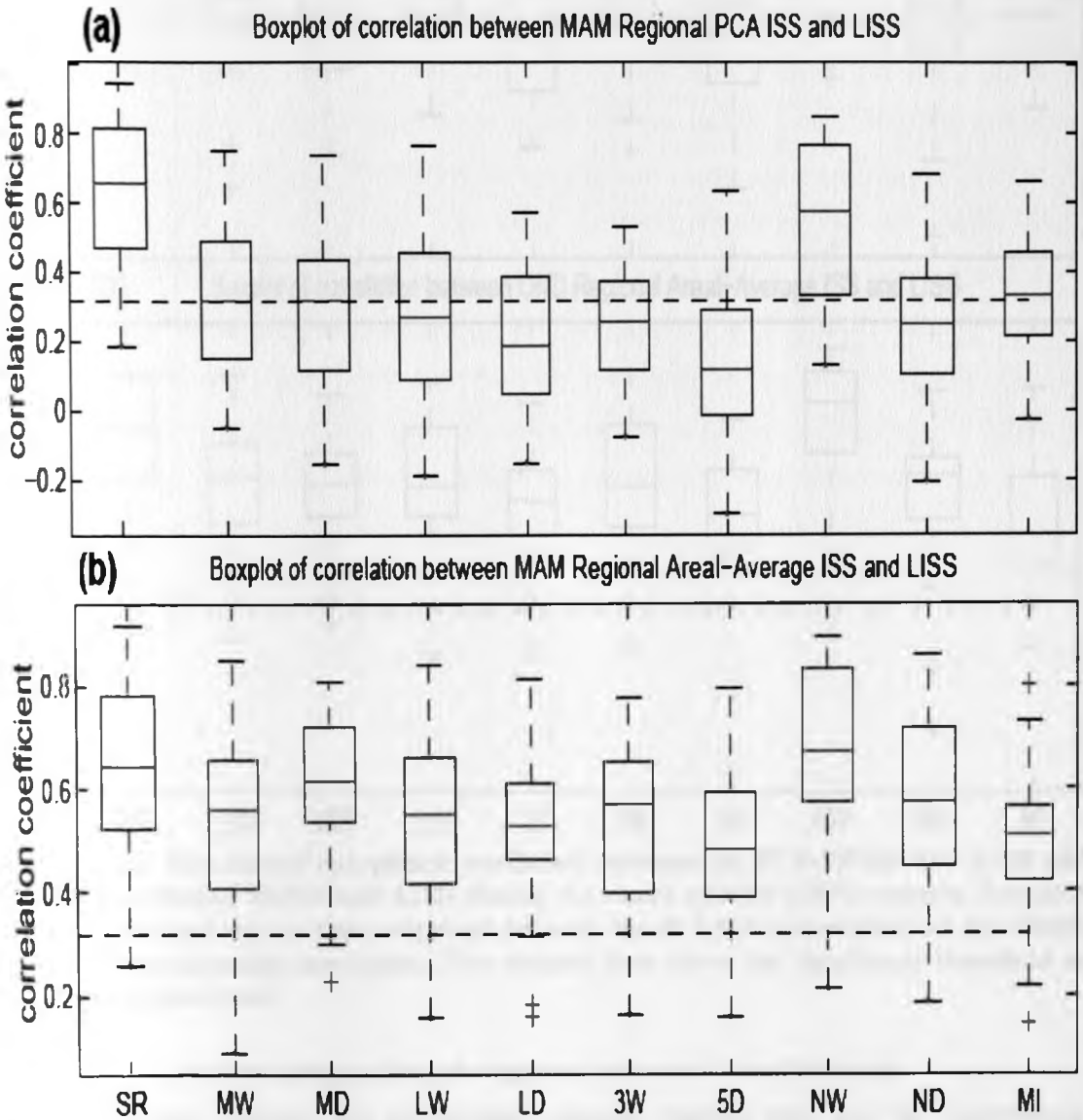


Figure 4.14: Box-plot of correlation coefficient (1962 - 2000) between (a) PCA-SRISS and LISS, and (b) areal-averaged SRISS and LISS during the long rainfall (MAM) season. For each ISS, the correlations are those obtained between the 36 LISS time-series and the SRISS for the corresponding sub-region. The dashed lines show the significant threshold at 95% confidence levels.

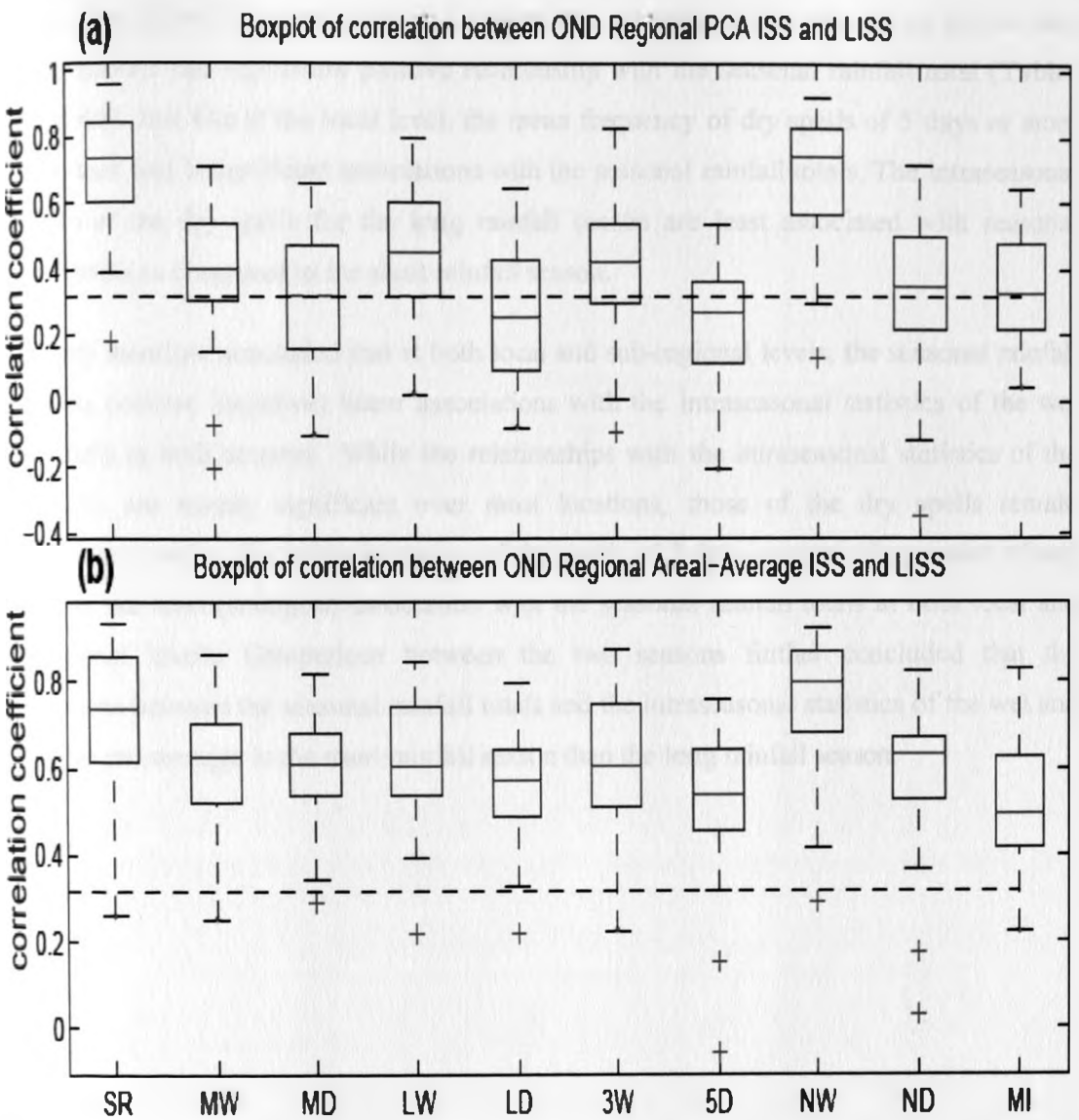


Figure 4.15: Box-plot of correlation coefficient between (a) PCA-SRISS and LISS and (b) areal-averaged SRISS and LISS during the short rainfall (OND) season. For each ISS, the correlations are those obtained between the 36 LISS time-series and the SRISS for the corresponding sub-region. The dashed lines show the significant threshold at 95% confidence levels

4.3.2.2 Relationship with sub-regional seasonal rainfall totals

The relationship between the sub-regional seasonal rainfall total and the sub-regional intraseasonal statistics of wet and dry spells (SRISS) obtained from the PCA scores and those obtained from areal-averaging of the intraseasonal statistics at locations forming a sub-region is presented in this part. The objective is to assess how dependent are the seasonal rainfall totals on the distributions of the rainfall within the rainfall season as supplied by wet and dry spells.

At the sub-regional level, most of the intraseasonal statistics of the wet spells for the two rainfall seasons had significant positive relationship with the seasonal rainfall total (**Tables 4.5** and **4.6**). Just like at the local level, the mean frequency of dry spells of 5 days or more had the least and insignificant associations with the seasonal rainfall totals. The intraseasonal statistics of the dry spells for the long rainfall season are least associated with seasonal rainfall totals as compared to the short rainfall season.

The study therefore concluded that at both local and sub-regional levels, the seasonal rainfall totals has positive (negative) linear associations with the intraseasonal statistics of the wet (dry) spells in both seasons. While the relationships with the intraseasonal statistics of the wet spells are mainly significant over most locations, those of the dry spells remain insignificant mostly. The mean frequency of dry spells of 5 days or more (the number of wet days) has the least (strongest) association with the seasonal rainfall totals at both local and sub-regional levels. Comparison between the two seasons further concluded that the associations between the seasonal rainfall totals and the intraseasonal statistics of the wet and dry spells are stronger in the short rainfall season than the long rainfall season.

Table 4.5: Pearson correlation coefficient between the seasonal rainfall totals and intraseasonal statistics during long rainfall season at sub-regional level for the period 1962 - 2000. Bold number indicates the coefficient is significant at 95% confidence level

Sub-region		Wet days	Dry days	Mean Wet	Mean Dry	Longest Wet	Longest Dry	3 Wet days	5 Dry days	Intensity
Central and western Kenya	PCA score	**0.94**	**-0.72**	**0.56**	**-0.73**	**0.71**	**-0.45**	**0.61**	**-0.41**	**0.92**
	Areal average	**0.92**	-0.23	**0.74**	**-0.61**	**0.71**	**-0.41**	**0.77**	**-0.33**	**0.63**
Coastal strip of Kenya and Tanzania	PCA score	**0.77**	**-0.35**	**0.61**	**-0.39**	**0.73**	**-0.36**	0.12	-0.11	**0.95**
	Areal average	**0.82**	**-0.33**	**0.65**	**-0.42**	**0.72**	**-0.40**	**0.48**	-0.31	**0.82**
Northeastern Kenya	PCA score	**0.92**	**0.08**	**0.73**	**-0.46**	**0.73**	-0.13	**0.70**	-0.02	**0.86**
	Areal average	**0.88**	**0.11**	**0.67**	**-0.74**	**0.59**	-0.28	**0.76**	-0.06	**0.51**
Western Tanzania and southern Uganda	PCA score	**0.80**	**-0.63**	**0.60**	**-0.33**	**0.41**	**-0.36**	**0.53**	**-0.34**	**0.99**
	Areal average	**0.72**	**-0.38**	**0.57**	**-0.52**	0.29	-0.22	**0.44**	-0.26	**0.68**
Southeastern lowlands of Kenya and northeastern Tanzania	PCA score	**0.92**	-0.24	**0.71**	**-0.55**	**0.70**	-0.15	**0.72**	**-0.38**	**0.86**
	Areal average	**0.87**	-0.12	**0.75**	**-0.54**	**0.70**	**-0.37**	**0.86**	-0.22	**0.64**
Most parts of Uganda	PCA score	**0.86**	-0.26	0.26	**-0.40**	0.25	**-0.40**	**0.75**	0.29	**0.98**
	Areal average	**0.69**	-0.10	**0.39**	**-0.49**	0.28	**-0.37**	**0.49**	-0.08	**0.70**

Table 4.6: Pearson correlation coefficient between the seasonal rainfall totals and intraseasonal statistics during short rainfall season at sub-regional level for the period 1962 - 2000. Bold number indicates the coefficient is significant at 95% confidence level

Sub-region		Wet days	Dry days	Mean Wet	Mean dry	Longest wet	Longest dry	3 Wet days	5 Dry days	Intensity
Central Kenya and southeastern lowlands	PCA score	**0.95**	-0.30	**0.71**	**-0.55**	**0.72**	**-0.40**	**0.76**	-0.20	**0.88**
	Areal average	**0.97**	**-0.42**	**0.86**	**-0.85**	**0.77**	**-0.68**	**0.93**	**-0.49**	**0.75**
Western Kenya and most parts of Uganda	PCA score	**0.90**	**-0.82**	**0.44**	**-0.57**	**0.59**	**-034**	**0.66**	-0.22	**0.96**
	Areal average	**0.94**	-0.28	**0.77**	-0.20	**0.81**	-0.28	**0.82**	**-0.36**	**0.67**
Northeastern Kenya	PCA score	**0.96**	-0.12	**0.77**	**-0.33**	**0.90**	**-0.36**	**0.68**	-0.03	**0.67**
	Areal average	**0.91**	0.08	**0.66**	**-0.45**	**0.76**	**-0.36**	**0.80**	0.23	**0.75**
Coastal strip of Kenya and Tanzania	PCA score	**0.96**	**-0.57**	**0.74**	**-0.67**	**0.75**	**-0.58**	**0.85**	-0.25	**0.70**
	Areal average	**0.96**	**-0.40**	**0.88**	**-0.64**	**0.88**	**-0.63**	**0.93**	-0.04	**0.86**
Central and northern Tanzania	PCA score	**0.93**	**-0.44**	**0.77**	**-0.62**	**0.72**	**-0.33**	**0.58**	**-0.42**	**0.89**
	Areal average	**0.93**	**-0.52**	**0.79**	**-0.59**	**0.73**	**-0.49**	**0.87**	**-0.65**	**0.57**
Western of Lake Victoria and western Tanzania	PCA score	**0.89**	**-0.79**	**0.65**	**-0.63**	**0.33**	**-0.41**	**0.74**	**-0.68**	**0.97**
	Areal average	**0.79**	**-0.69**	**0.76**	**-0.62**	**0.65**	**-0.45**	**0.75**	**-0.61**	**0.60**

4.3.2.3 Trend results

The spearman rank correlation coefficients of seasonal rainfall totals and SRISS over the period 1962 to 2000 for the long and short rainfall seasons are shown by **Tables 4.7** and **4.8** respectively. During the long rainfall season, most of the intraseasonal statistics did not have a significant trend at sub-regional scale apart from sub-region 4 covering the western parts of Lake Victoria, northwestern and western Tanzania and southern Uganda. This sub-region had significant decreasing trend in seasonal rainfall totals, number of wet days in a season and mean duration of wet spells for both the PCA score and areal-average derived statistics. Further, the PCA score derived mean rainfall intensity and duration of longest wet spell had significant decreasing trend while significant increasing trend was observed for number of dry days in a season and the mean frequency of dry spells of 5 days or more for same sub-region 4. It is also worthy to highlight that sub-regions 3, 4 and 5 covering Northeastern Kenya; Coastal strip of Kenya and Tanzania; and Central and northern Tanzania respectively had significant increasing trend in the occurrence of dry spells of 5 days or more derived from PCA score.

As shown in **Table 4.8,** most of the intraseasonal statistics did not have significant trends at sub-regional level during the short rainfall season, except for some positive trends in the mean frequency of dry spells of 5 days or more.

In conclusion, significant trends were noted in all the intraseasonal statistics of the wet and dry spells though at few isolated locations and sub-regions during the two rainfall seasons. However, significant increasing trend in the mean frequency of dry spells of 5 days or more shows an organized pattern for the two seasons at both local and sub-regional levels. Some crops such as maize are particularly sensitive to long dry spells around the flowering stage. The frequency of prolonged dry spells of various durations needs therefore to be studied.

Table 4.7: Spearman rank correlation coefficient of the seasonal rainfall totals and intraseasonal statistics at sub-regional scale during long rainfall season for the period 1962 - 2000. Bold number indicates significant trend at 95% confidence level

Sub-region		Seasonal Rainfall	Intensity	Wet days	Dry days	Mean Wet	Mean Dry	Longest Wet	Longest Dry	3 Wet days	5 Dry days
Central and western Kenya	PCA score	-0.19	-0.20	-0.18	0.25	-0.22	0.13	-0.07	0.11	0.06	0.23
	Areal average	-0.20	-0.24	-0.22	-0.08	-0.11	0.05	-0.01	-0.17	-0.27	0.09
Coastal strip of Kenya and Tanzania	PCA score	0.13	0.15	0.07	0.17	-0.01	-0.04	0.14	0.12	0.11	0.29
	Areal average	0.14	0.06	0.08	-0.13	0.19	-0.03	0.24	-0.01	-0.01	-0.07
Northeastern Kenya	PCA score	-0.24	-0.22	-0.13	0.28	-0.18	**0.32**	0.03	**0.35**	0.01	**0.39**
	Areal average	-0.23	**-0.42**	-0.08	0.19	-0.01	0.18	0.02	0.14	0.02	**0.36**
Western Tanzania and southern Uganda	PCA score	**-0.33**	**-0.34**	**-0.34**	**0.33**	**-0.40**	-0.06	**-0.41**	0.19	-0.10	**0.42**
	Areal average	**-0.42**	-0.14	**-0.40**	0.12	**-0.42**	0.19	-0.18	0.05	-0.04	0.13
Southeastern lowlands of Kenya and northeastern Tanzania	PCA score	-0.02	-0.05	-0.08	0.05	0.03	0.14	0.10	-0.06	0.24	**0.34**
	Areal average	-0.00	0.13	-0.09	0.00	-0.13	0.18	-0.02	0.12	-0.11	-0.01
Most parts of Uganda	PCA score	-0.15	-0.14	-0.04	0.11	-0.28	-0.16	0.08	0.04	-0.05	0.10
	Areal average	-0.09	-0.00	-0.28	**0.35**	**-0.51**	0.08	**-0.33**	-0.07	-0.20	0.20

Table 4.8: **Spearman rank correlation coefficient of the seasonal rainfall totals and intraseasonal statistics at sub-regional scale during short rainfall season for the period 1962 - 2000. Bold number indicates significant trend at 95% confidence level**

Sub-region		Seasonal Rainfall	Intensity	Wet days	Dry days	Mean Wet	Mean Dry	Longest Wet	Longest Dry	3-wet days	5-dry days
Central Kenya and southeastern lowlands	PCA score	0.10	0.12	0.22	-0.07	0.08	-0.30	0.06	0.02	**0.46**	0.18
	Areal average	0.13	0.13	0.08	-0.14	0.05	-0.12	0.03	-0.02	0.17	-0.10
Western Kenya and most parts of Uganda	PCA score	0.08	0.07	-0.06	0.13	-0.21	0.06	-0.13	0.10	0.22	0.16
	Areal average	0.07	0.12	-0.06	0.21	-0.12	0.20	-0.08	0.07	-0.07	**0.34**
Northeastern Kenya	PCA score	-0.23	-0.22	-0.20	0.12	-0.15	0.23	-0.10	0.18	0.05	0.29
	Areal average	-0.24	**-0.32**	-0.10	0.07	0.04	0.21	-0.02	0.06	-0.00	0.14
Coastal strip of Kenya and Tanzania	PCA score	0.08	0.10	0.12	0.16	0.01	-0.12	0.16	-0.06	0.12	**0.47**
	Areal average	0.13	0.16	0.11	0.17	0.01	-0.10	0.07	-0.08	0.06	0.28
Central and northern Tanzania	PCA score	-0.11	-0.11	-0.12	-0.04	0.06	0.30	0.01	0.11	0.15	0.28
	Areal average	-0.08	0.13	-0.19	-0.18	-0.18	-0.06	-0.24	0.01	-0.13	0.04
Western of Lake Victoria and western Tanzania	PCA score	0.09	0.10	0.10	-0.02	-0.23	-0.28	-0.28	-0.20	**0.39**	0.10
	Areal average	0.12	0.27	-0.01	0.07	-0.18	-0.05	-0.11	-0.22	0.11	0.11

4.4 Spatial coherence and potential predictability results

For relatively homogeneous sub-regions, the spatial coherence analysis provides a measure of potential predictability (Moron *et al.*, 2006). An illustration of the within-the-region (inter-station) differences in the interannual variability of the intraseasonal statistics is shown by **Figures 4.16a** and **b**. It is found that all the 5 stations making up the sub-region 1 (central highlands and southeastern lowlands of Kenya) during the short rainfall season display similar year-to-year variations in the standardized number of wet days (**Figure 4.16a**). Both SRISS indices (PCA and RIS) well replicate these variations. This reveals that the number of wet days is a spatially very coherent variable over sub-region.

For the same season and over the same sub-region, the duration of the longest dry spell between individual locations and at the sub-regional level are quite contrasted (**Figure 4.16b**). This simply suggests that during the short rainfall season, there is high potential to predict the number of wet days and lower potential predictability for the duration of the longest dry spells over central highland and southeastern lowlands of Kenya (sub-region 1).

The inter-station correlation coefficient was next used as an evaluation of spatial coherence for each sub-region. **Figures 4.17a** and **b** shows the inter-station correlation coefficients of intraseasonal statistics of wet and dry spells at two sub-regions during the long rainfall season. For sub-region 1 (central highlands and western Kenya), only the seasonal rainfall totals and the number of wet days have significant correlation coefficients between almost all the stations, though quite low. For the other variables, significant correlations are restricted to a few station couples (top whiskers and crosses on **Figure 4.17a**). Sub-region 6 which represents most parts of Uganda on the other hand has almost no significant correlation except for some couples of stations, as denoted by a few outliers in several of the intraseasonal statistics (**Figure 4.17b**).

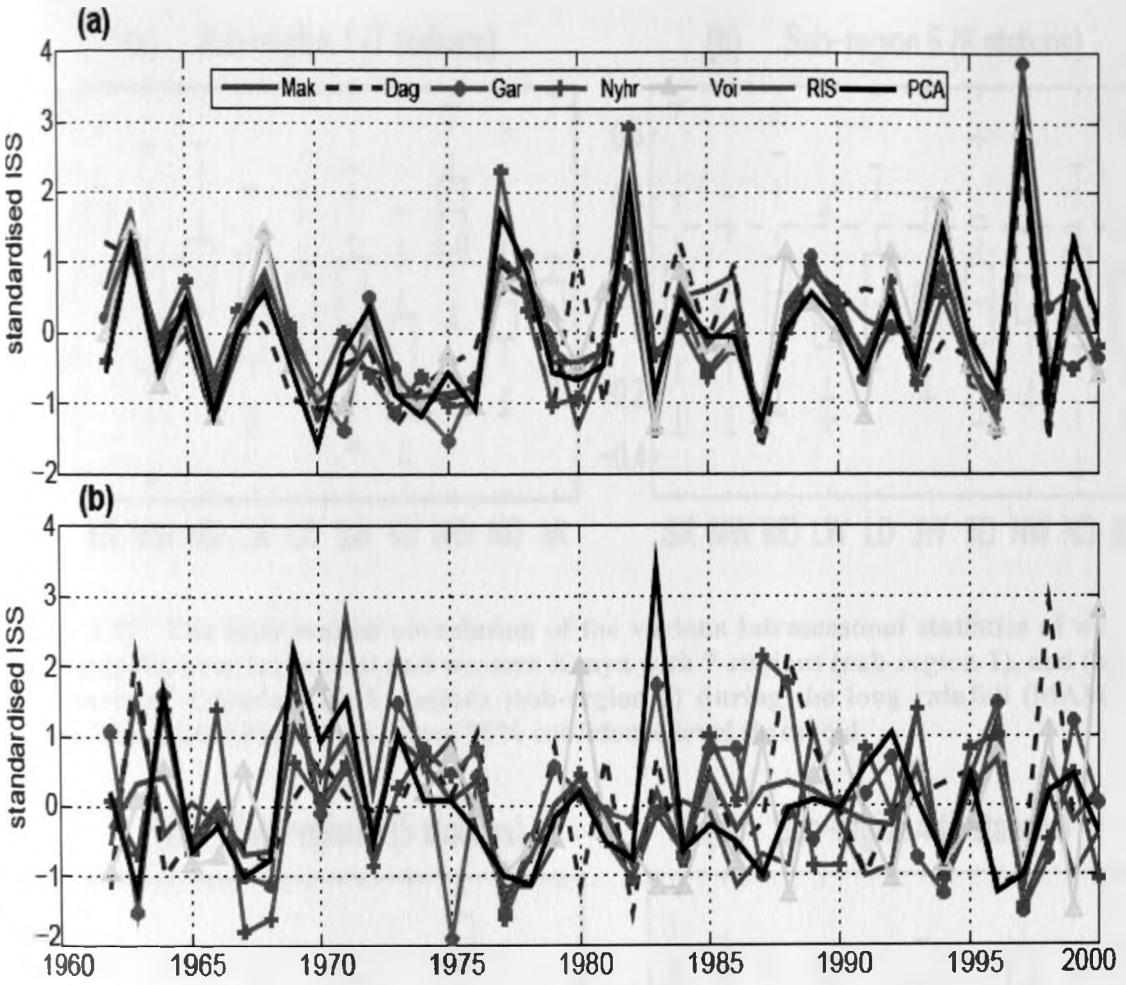


Figure 4.16: The standardized (a) number of wet days in a season and (b) duration of the longest dry spell over central highlands and southeastern lowlands of Kenya (sub-region 1) during the short rainfall (OND) season for the sub-region as a whole (RIS and PCA) and for the individual stations (Makindu, Dagoretti, Garissa, Nyahururu and Voi) which belongs to this sub-region

There were similar observations during the short rainfall season though the significance of correlation coefficient was slightly higher (**Figures 4.18a** and **b**). In addition to the two variables identified above, the mean frequency of wet spells of 3 days or more and the mean length of the dry spells for sub-region 1 (central highlands and southeastern lowlands of Kenya as shown in **Figure 4.18a**) and the duration of the longest wet spells for sub-region 4 (coastal strip of Kenya and Tanzania as shown in **Figure 4.18b**) were also significant.

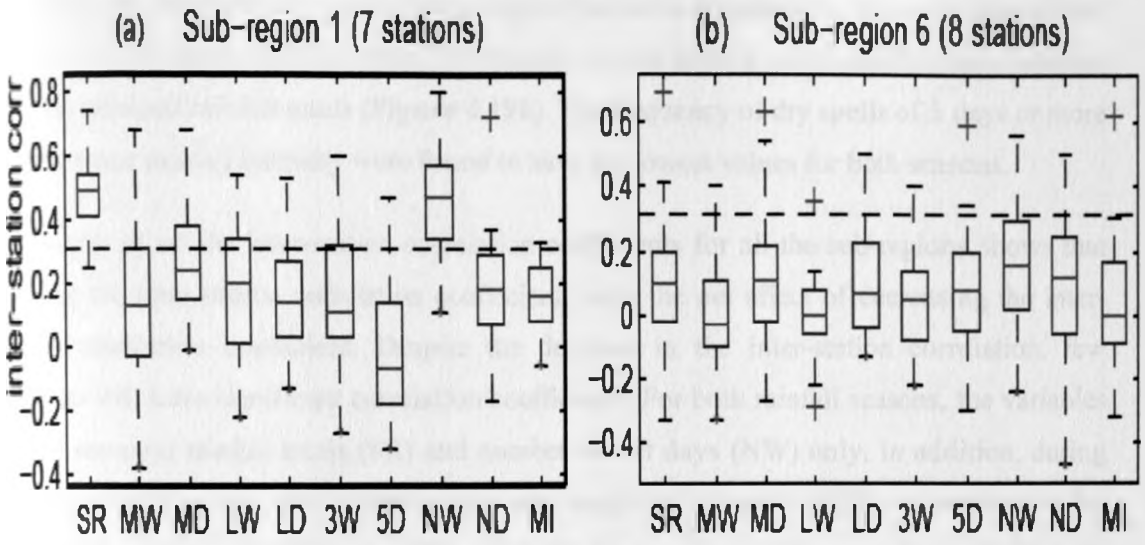


Figure 4.17: The inter-station correlation of the various intraseasonal statistics of wet and dry spells over (a) central and western Kenya with 7 stations (sub-region 1), and (b) most parts of Uganda with 8 stations (sub-region 6) during the long rainfall (MAM) season. The dotted line across shows 95% confidence level threshold

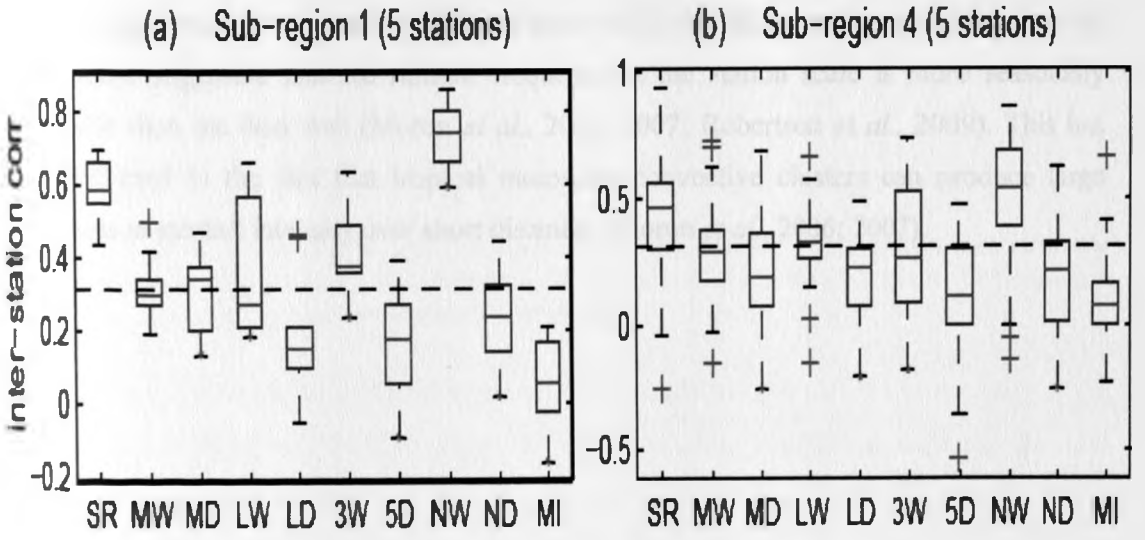


Figure 4.18: The inter-station correlation of the various intraseasonal statistics of wet and dry spells over (a) central highlands and southeastern lowlands of Kenya with 5 stations (sub-region 1), and (b) coastal strip of Kenya and Tanzania with 5 stations (sub-region 4) during the short rainfall (OND) season. The dotted line across shows 95% confidence level threshold

When all the inter-station correlation coefficients from different sub-regions are assembled together, the seasonal rainfall totals and number of wet days were found to have the greatest spatial coherence during the two rainfall seasons as shown by **Figures 4.19a** and **b**. The

values for the short rainfall season were slightly higher as compared to those of long rainfall season. For the short rainfall season, the number of wet days is even slightly more coherent than the seasonal rainfall totals (**Figure 4.19b**). The frequency of dry spells of 5 days or more and the mean rainfall intensity were found to have the lowest values for both seasons.

A box-plot of all the inter-station correlation coefficients for all the sub-regions shows that merging the inter-station correlation coefficients have the net effect of decreasing the inter-station correlation coefficient. Despite the decrease in the inter-station correlation, few variables still have significant correlation coefficients. For both rainfall seasons, the variables are the seasonal rainfall totals (SR) and number of wet days (NW) only. In addition, during the long rainfall season, the variables are mean length of dry spells (MD) and number of dry days (ND) as shown by **Figure 4.19a** while for the short rainfall season, the variables are mean length of wet spells (MW), duration of the longest wet spell (LW) and mean frequency of wet spells of 3 days or more (3W) as shown by **Figure 4.19b**. This means that the spatial coherence (hence potential predictability) is reasonably high in a few sub-regions for these variables. Given the relatively higher spatial coherence of inter-annual anomalies of rainfall frequency compared to seasonal rainfall and mean daily rainfall intensity, recent work in the tropics have suggested that the rainfall frequency at the station scale is more seasonally predictable than the later two (Moron *et al.*, 2006; 2007; Robertson *et al.*, 2009). This has been attributed to the fact that tropical mesoscale convective clusters can produce large differences in rainfall intensity over short distances (Moron *et al.*, 2006; 2007).

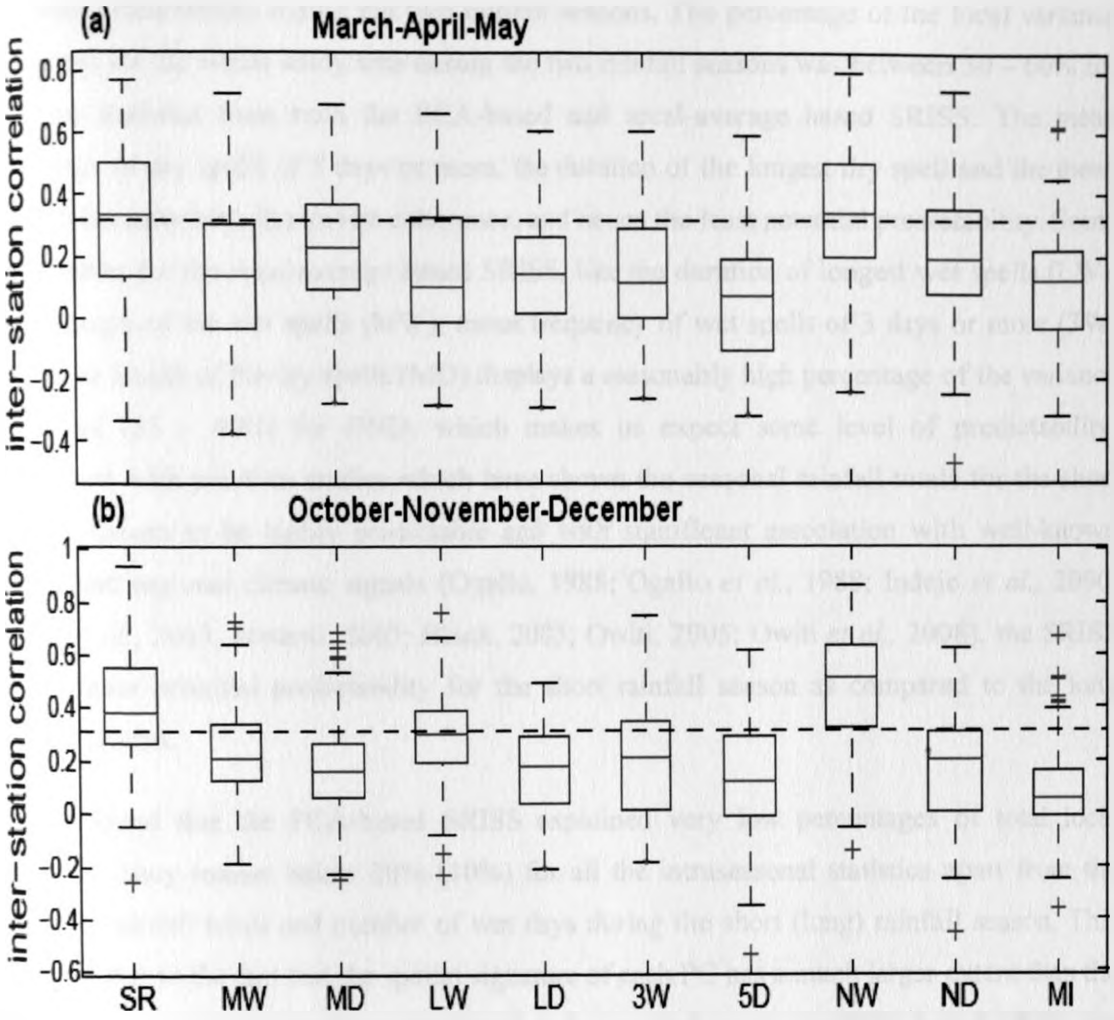


Figure 4.19: Box plot of inter-station correlation coefficients of all stations within the study region for the (a) long (MAM) and (b) short (OND) rainfall seasons. The dotted line across shows 95% confidence level threshold

Another way to characterize spatial coherence is to determine and plot the percentage of the local variance explained for each variable. The correlation coefficient between the sub-regional intraseasonal statistics of wet and dry spells (SRISS) including seasonal rainfall totals time series (both PC scores based and the areal average of the local intraseasonal statistics) and the intraseasonal statistics at local levels were averaged for the whole study region. The average correlation coefficient obtained is squared to obtain the percentage of the local total variance explained as described in **section 3.2.4**.

Figure 4.20 shows the percentage of the total local variance explained by the sub-regional intraseasonal statistics indices for the long and short rainfall seasons. The figure clearly shows that the seasonal rainfall totals and the number of wet days in a season have higher

potential predictability during the two rainfall seasons. The percentage of the local variance explained for the whole study area during the two rainfall seasons was between 30 – 60% for the two statistics from both the PCA-based and areal-average based SRISS. The mean frequency of dry spells of 5 days or more, the duration of the longest dry spell and the mean rainfall intensity have the lowest coherence, and hence the least potential predictability. Some of variables for the areal-average based SRISS, like the duration of longest wet spells (LW), mean length of the wet spells (MW), mean frequency of wet spells of 3 days or more (3W) and mean length of the dry spells (MD) displays a reasonably high percentage of the variance explained (35 – 40%) for OND, which makes us expect some level of predictability. Consistent with previous studies which have shown the seasonal rainfall totals for the short rainfall season to be highly predictable and with significant association with well-known global and regional climate signals (Ogallo, 1988; Ogallo *et al.*, 1988; Indeje *et al.*, 2000; Black *et al.*, 2003; Mutemi, 2003; Black, 2005; Owiti, 2005; Owiti *et al.*, 2008), the SRISS have higher potential predictability for the short rainfall season as compared to the long rainfall season.

It was found that the PCA-based SRISS explained very low percentages of total local variance. They remain below 20% (10%) for all the intraseasonal statistics apart from the seasonal rainfall totals and number of wet days during the short (long) rainfall season. This could be due to the fact that the spatial signature of each PC has a much larger extent than the sub-region to which it has been associated with. In other words, the PCA-based SRISS are not strictly sub-regional. The results further demonstrate that sub-regional indices of seasonal rainfall totals and intraseasonal statistics derived from areal-average are more representative than those derived from the PC scores.

The percentage of the variance of local random series explained by the area-average SRISS was also determined. This was accomplished by generating random Gaussian time series, and aggregating them by computing the average. The number of stations in each sub-region was maintained. The percentage of the local variance was then computed. This was repeated 500 times and the 95^{th} percentile extracted. It is the percentage local variance which is exceeded only 5 times out of 100 based on random time-series. This 95% confidence level is 17% for MAM and 19% for OND. All the RISS values computed from the real data (**Figure 4.20**) surpass these thresholds which mean that the spatial coherence in all cases is significant. In other words, there is a climate signal in all the variables. However, for some variables like the

mean frequency of dry spells of 5 days or more (5D) and mean rainfall intensity (MI), the percentage of local variance explained is only marginally higher than the 95% threshold. The slight difference in thresholds between MAM and OND is due to the fact that the number of stations in each sub-region is slightly different between the two seasons. Thus the SRISS including seasonal rainfall totals derived from the areal-average were investigated further for their association with large scale climate fields as discussed in the next section.

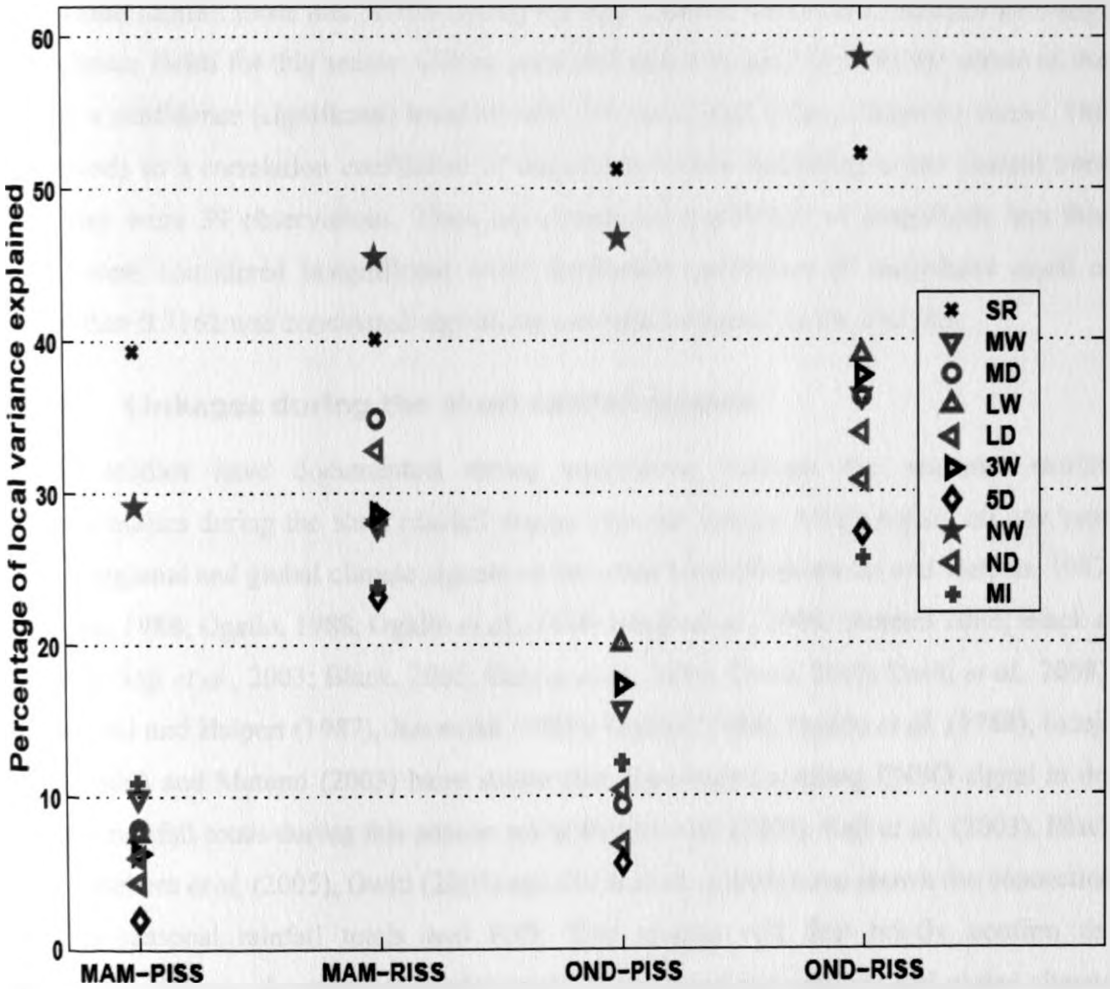


Figure 4.20: The local variance explained by sub-regional intraseasonal statistics of wet and dry spells derived from PCA scores (PISS) and from areal-averaging (RISS) during the long (MAM) and short (OND) rainfall season

4.5 Linkages between large scale climate fields and sub-regional intraseasonal statistics of wet and dry spells

The methodology of assessing the linkages between the large scale climate fields on one side and interannual variability of rainfall totals and sub-regional intraseasonal statistics of wet and dry spells (SRISS) on the other side was discussed in **section 3.2.5**. The results obtained will be discussed in this section. Due to the higher potential predictability identified above for the seasonal rainfall totals and SRISS during the short rainfall season, the linkages with large scale climate fields for this season will be presented and discussed first. In the whole of this section, a confidence (significant) level of 95% (5%) was used unless otherwise stated. This corresponds to a correlation coefficient of magnitude 0.3162 according to the student t-test since they were 39 observations. Thus any correlation coefficient of magnitude less than 0.3162 were considered insignificant while correlation coefficient of magnitude equal or greater than 0.3162 was considered significant and hence retained in the analysis.

4.5.1 Linkages during the short rainfall season

Several studies have documented strong association between the seasonal rainfall totals/anomalies during the short rainfall season over the eastern Africa region on one hand and the regional and global climate signals on the other hand (Ropelewski and Halpert, 1987; Janowiak, 1988; Ogallo, 1988; Ogallo *et al.*, 1988; Indeje *et al.*, 2000; Mutemi 2003; Black *et al.*, 2003; Saji *et al.*, 2003; Black, 2005; Behera *et al.*, 2005; Owiti, 2005; Owiti *et al.*, 2008). Ropelewski and Halpert (1987), Janowiak (1988), Ogallo (1988), Ogallo *et al.* (1988), Indeje *et al.* (2000) and Mutemi (2003) have shown that there exists a strong ENSO signal in the seasonal rainfall totals during this season while Black *et al.* (2003), Saji *et al.* (2003), Black (2005), Behera *et al.* (2005), Owiti (2005) and Owiti *et al.*, (2008) have shown the connection between seasonal rainfall totals and IOD. This section will first briefly confirm the relationship between the seasonal rainfall totals on one hand and regional and global climate signals on the other hand; and further sort whether the different SRISS are themselves associated with the regional and global climate signals, using previously defined climate indices (NINO, IOD, SST gradients) as discussed in **section 3.2.5.1.** This will be followed by a presentation of the additional potential predictors as derived in sections **3.2.5.2** and **3.2.5.3**. Along with that will be the discussion on how these indices influence the seasonal rainfall totals and SRISS.

4.5.1.1 Linkages with predefined SST predictors

Figures **4.21a–f** shows the concurrent and lagged correlation analysis results of the areal-average seasonal rainfall totals and number of dry days in a season over the six sub-regions (Z1 to Z6) with 1-month, 2-months average and 3-months average of Niño 3 index from the months of October to May. These two variables (seasonal rainfall totals and number of dry days) have been selected to illustrate the typical behavior of the relationship with ENSO for different lags, and different timescales. **Figure 4.21a** clearly indicates significant concurrent positive relationship between the seasonal rainfall totals and Niño 3 index over the six sub-regions. This relationship diminishes as lagged correlations are considered and several months averaged (**Figure 4.21b** and **c**). By the month of June, the correlation coefficient was less than 0.3 for both 1-month (**Figure 4.21a**) and 2-months average (**Figure 4.21b**) of Niño 3 index which is insignificant (at 0.95 confidence levels). Similar results were obtained for Niño 1+2, Niño 3, Niño 4, IOD indices, SST gradient across the equatorial Indian Ocean (ZIND) for both seasonal rainfall totals and number of wet days in a season. These results are in agreement with those of Mutemi (2003) on ENSO, Owiti (2005) on IOD and Nyakwada (2009) on SST gradients.

However, such a strong concurrent and lagged relationship is not always the case during the short rainfall season as illustrated by **Figures 4.21d–f**. In this case, the number of dry days in a season is not statistically related to the Niño 3 index at four out of the possible six sub-regions from both concurrent and lagged correlation results (**Figures 4.21d–f**).

Figure 4.21a further shows that there is some noise when the one month index is considered. It can be seen that there is a drop in the correlation coefficient with the Niño 3 September index which is again recovered by the Niño 3 August index. However, once the index is averaged for two or three months, the decrease in the lagged correlation coefficient is rather smooth (**Figures 4.21b** and **c**). With the consideration for a sufficient lead time in the development of prediction models and further noting that the lagged correlation coefficients beyond June are insignificant, the use of the July-August two-month average for the all predictors was seen as suitable in the current study. The indices for July-August are available by mid-September meaning there will be adequate time to update the indices before the start of the OND season.

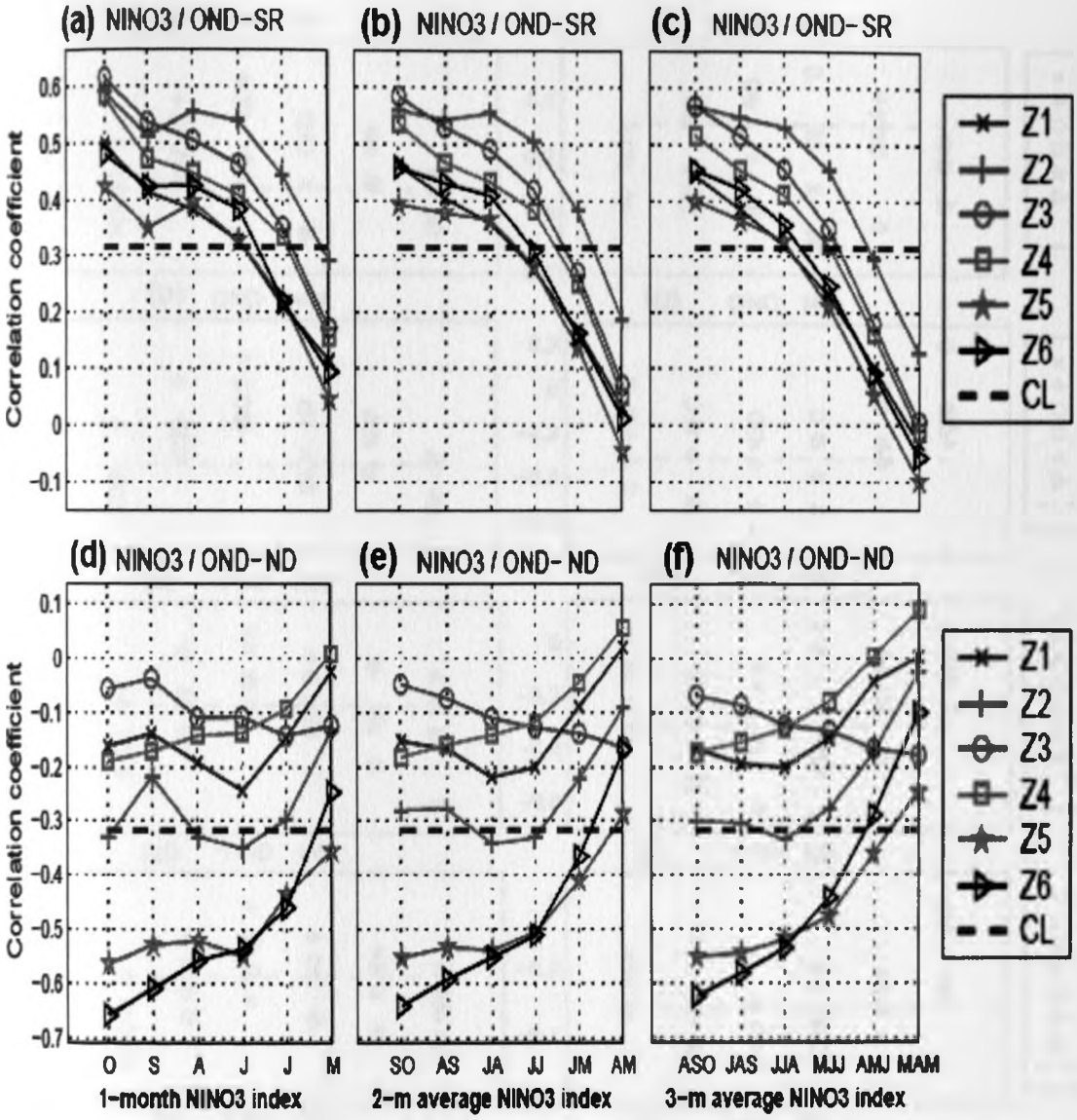


Figure 4.21: Correlation coefficient analysis between areal-averaged October-November-December (OND) seasonal rainfall totals (SR) and (a) 1 month (b) 2 months' average and (c) 3 months' average of Niño 3 index (October to May), for the six rainfall sub-regions Z1 to Z6; and areal-averaged OND number of dry days (ND) and (d) 1 month (e) 2 months' average and (f) 3 months' average of Nino 3 index. CL denotes 95% significant level threshold

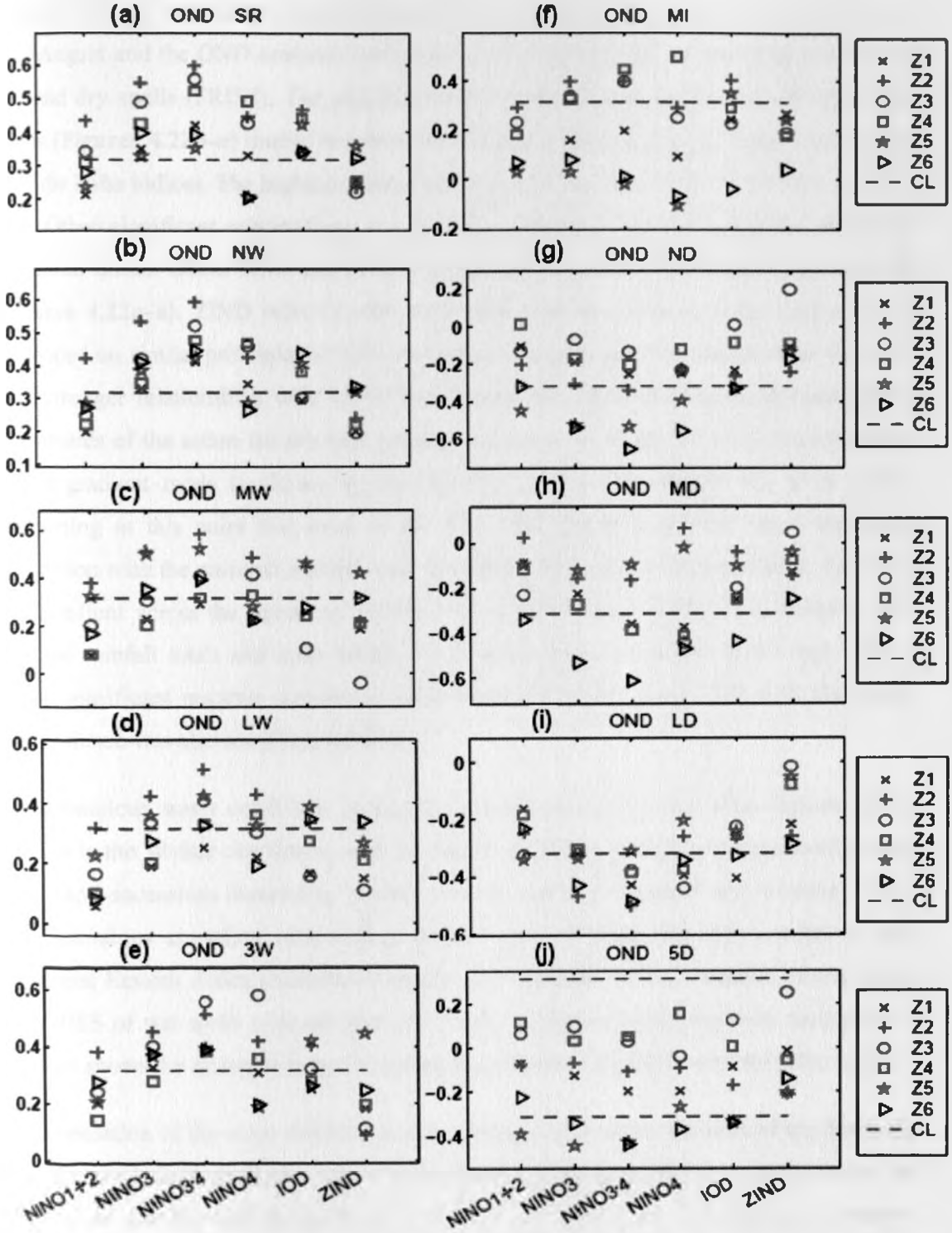


Figure 4.22: Correlation coefficient between predefined predictors averaged for July-August (x-axis) and areal-averaged October-November-December (a) seasonal rainfall totals, (b) number of wet days, (c) mean length of wet spells, (d) longest wet spell, (e) frequency of 3 wet days or more, (f) mean rainfall intensity, (g) number of dry days, (h) mean length of dry spells, (i) longest dry spell, and (j) frequency of 5 dry days or more, over the six rainfall sub-regions Z1 to Z6. CL shows the 95% confidence level threshold

Figures 4.22a–j summarize the relationship between the predefined predictors averaged for July-August and the OND seasonal rainfall totals and sub-regional intraseasonal statistics of wet and dry spells (SRISS). The seasonal rainfall totals (**Figure 4.22a**) and SRISS of wet spells (**Figures 4.22b–e**) during the short rainfall season have a positive lagged association with the Niño indices. The highest correlations are for Niño 3.4 and the lowest ones for Niño 1+2. Other significant relationships are generally obtained with SST gradients across the equatorial Indian Ocean (IOD and ZIND), though the correlations are often higher for IOD (**Figures 4.22a–e**). ZIND refers to the zonal SST gradient mode over the Indian Ocean developed on similar principles as IOD but centred along the equator, and has been shown to have stronger relationships with SOND rainfall than the classical IOD (Nyakwada, 2009). The centres of the action for this SST gradient are shown in **Table 3.3** while the calculation of the gradient mode is shown by the direction of arrow in **Figure 3.3**. It is worthy mentioning at this point that most of the other SST gradients did not show significant association with the seasonal rainfall totals and SRISS thus are not discussed here. The zonal SST gradient across the equatorial Pacific Ocean (ZPAC) had significant correlation with seasonal rainfall totals and other SRISS. Multi-collinearity assessment shows that it has a highly significant negative association (correlation coefficient about -0.7) with the ENSO indices hence was also not discussed here.

The anomalous warm conditions during the boreal autumn over the Niño regions induce changes in the Walker circulation, with anomalous ascending motion over Equatorial Eastern Africa and anomalous descending motion over the maritime continent and southern Africa. The anomalous ascending (descending) motions tend to bring wet (dry) conditions over Equatorial Eastern Africa (Maritime continent and southern Africa). Seasonal rainfall totals and SRISS of wet spells over sub-region 2 (which covers western Kenya and most parts of Uganda) shows the strongest lagged correlation coefficients especially with the Niño indices.

The association of the mean rainfall intensity and the intraseasonal statistics of the dry spells with the predefined predictors were rather diverse (**Figures 4.22f–j**). In many cases, the correlations are low and insignificant, but there are exceptions. Sub-region 6 (southern Uganda and western Tanzania) has strong lagged correlation coefficients between the SRISS of dry spells and the predefined SST predictors. For instance, a correlation of -0.6 was found between the Nino3.4 and the mean length of the dry spells, suggesting longer dry spells during La Niña years (**Figure 4.22h**).

It is interesting to observe that while the number of the wet days (NW) shows that strongest association with Niño indices (**Figure 4.22b**), the duration of the longest dry spell (LD) also shows a relatively strong and coherent/uniform response to Niño indices (**Figure 4.22i**). This means that a very long dry spell is expected to occur throughout East Africa during the short rainfall season with La Niña conditions, with potentially adverse effect on crops. The mean frequency of dry spells of 5 days or more shows the weakest control by Niño and other predefined indices (**Figure 4.22j**). This is closely followed by mean rainfall intensity (**Figure 4.22f**) and then the mean duration of the dry spells (**Figure 4.22h**) and number of dry days (**Figure 4.22g**) in that order. The weak control of the mean rainfall intensity by the ENSO indices and the low spatial coherence observed earlier may be attributed to the fact that tropical mesoscale convective clusters produce large differences in rainfall intensity over short distances (Moron *et al.*, 2006; 2007).

From the strong significant correlations with the predefined predictors, two indices with strong significant lagged correlations with the seasonal rainfall totals and SRISS were chosen. These were the Niño3.4 and ZIND indices whose average values for July-August are not related, yet they are associated with seasonal rainfall totals and most of the SRISS during the short rainfall season. They can be thought of as representing the climate signals from Pacific and Indian Ocean sea surface temperature in general terms for this study.

4.5.1.2 Linkages with additional potential predictors

Additional potential predictors were searched for in oceanic (Hadley centre SST) and atmospheric (ERA40) fields as described in **sub-section 3.2.5.2**. Concurrent and lagged correlation analysis of the seasonal rainfall totals and SRISS on one hand and the oceanic and atmospheric variables on the other hand while controlling the effects of significantly correlated predefined indices identified several common potential predictors. These oceanic and atmospheric predictors are briefly described in **Table 4.9**. **Table 4.10** shows the number of SRISS including the seasonal rainfall totals that have significant association with a given predictor at 95% confidence level. For example WCAUS is only associated with 2 intraseasonal statistics while BoBEN is associated with all the 9 SRISS and the seasonal rainfall totals.

The association of the nine additional (oceanic and atmospheric) predictors with the seasonal rainfall totals and SRISS during the short rainfall season is summarized by **Figures 4.23a–e** and **4.24a–e**. Just like with the predefined predictors, the seasonal rainfall totals and the

SRISS of wet spells are more coherent in their responses to these predictors (**Figures 4.23a–e**). Oceanic potential predictor, BoBEN and atmospheric potential predictor, SINDS for example have a significant negative and lagged relationship with the seasonal rainfall totals (**Figure 4.23a**) and all the SRISS of wet spells (**Figures 4.23b–e**). The various SRISS of the wet spells responds more or less uniformly to any predictor identified for all the sub-regions. For example, the duration of the longest wet spell (LW) had a correlation coefficient close to the 95% confidence level (+0.31) with SWHAW potential predictor over the six sub-regions (**Figure 4.23d**). The insignificant relationship found for some of the coefficients was attributed to the fact that for the sake of simplicity, total correlations are shown in **Table 4.10**, **Figures 4.23a–e** and **4.24a–e**, while the identification and selection of the additional potential predictors was based on partial correlation analysis, after the effect of significantly correlated predefined predictors has been removed.

The SRISS of dry spells were somehow diverged in their responses to the additional potential predictors identified just like was the case with the predefined predictors (**Figures 4.24a–e**). Considering the same predictor, SWHAW had varied correlation coefficients with mean duration of the dry spells (MD) at about +0.20 for two sub-regions (sub-region 2 and 5) and about -0.30 to -0.60 for the remaining four sub-regions (**Figure 4.24c**). Most of SRISS of dry spells have insignificant association (at 95% confidence level) with the additional potential predictors (**Figures 4.24a–e**). However, there are several exceptions. One such example is the atmospheric predictor SINDS that has a generally consistent response with the mean rainfall intensity (**Figure 4.24a**) and all the SRISS of dry spells with the exception of the mean frequency of dry spells of 5 days or more in 3 sub-regions (**Figures 4.24b–d**). These show that though the response of the intraseasonal statistics of dry spells may not be uniform for any given oceanic or atmospheric signals, there are a few exceptions.

In the next sections, each of these predictors is described in details and a physical interpretation on how the predictor influences the SRISS for which it is significantly correlated provided. The SST-based potential predictors will be discussed first followed by the atmospheric potential predictors derived from the ERA40.

Table 4.9: Brief description of the additional potential predictors for the short rainfall (OND) season and their location details

Index name	Description	Location details (°)	
		Longitude	Latitude
ECMAD	SST index on the east coast of Madagascar over south-western Indian Ocean	56 – 63 E	18 – 12 S
BoBEN	SST index over Bay of Bengal extending to west coast of Malaysia and Indonesia	83 – 90 E	12 – 17 N
SWHAW	SST index on the south-western of Hawaii in the Pacific Ocean	140 – 120 W	10 – 25 N
WCAUS	SST index on western coast of Australia over the south-eastern Indian Ocean	95 – 105 E	24 – 15 S
SINDS	Zonal wind component index at 925mb level to the south of the Bay of Bengal near the southern tip of India sub-continent	70 – 90 E	5 – 10 N
EQAFR	Zonal wind component index at 200mb level extending from Equatorial Africa into Equatorial Atlantic Ocean	0 – 45 E	10 – 5 S
MARCON	Zonal wind component index at 200mb level over the maritime continent and extending over the equatorial Indian Ocean	85 – 110 E	2.5 S – 2.5 N
SWAFRC	Specific humidity index at 700mb level located at Angola coast on south-western Africa and extending to Atlantic Ocean on the west and Zambia to the east	5 – 15 E	25 – 15 S
EQIND	Specific humidity index at 700mb level extending from the southern tip of India subcontinent, through equatorial Indian Ocean into the eastern Africa region	35 – 90 E	0 – 10 N

Table 4.10: **A summary of the association between the identified additional potential predictors (July-August) and the sub-regional intraseasonal statistics of wet and dry spells for the October-November-December rainfall season and the most strongly correlated intraseasonal statistic and sub-region**

Predictor	Atmospheric Level	Index Name	Number of SRISS associated with the predictor (out of 10)	Strongest total correlation		
				SRISS	Sub-region	Coefficient
SST	surface	ECMAD	3	LW	3	*-0.37*
		BoBEN	10	3W	6	*-0.55*
		SWHAW	9	MD	4	*-0.56*
		WCAUS	2	ND	4	0.46
u-wind	925mb	SINDS	10	SR	2	*-0.70*
	200mb	EQAFR	7	ND	6	0.46
		MARCON	10	ND	4	*-0.58*
Specific humidity	700mb	SWAFRC	9	3W	1	0.43
		EQIND	7	SR	2	0.42

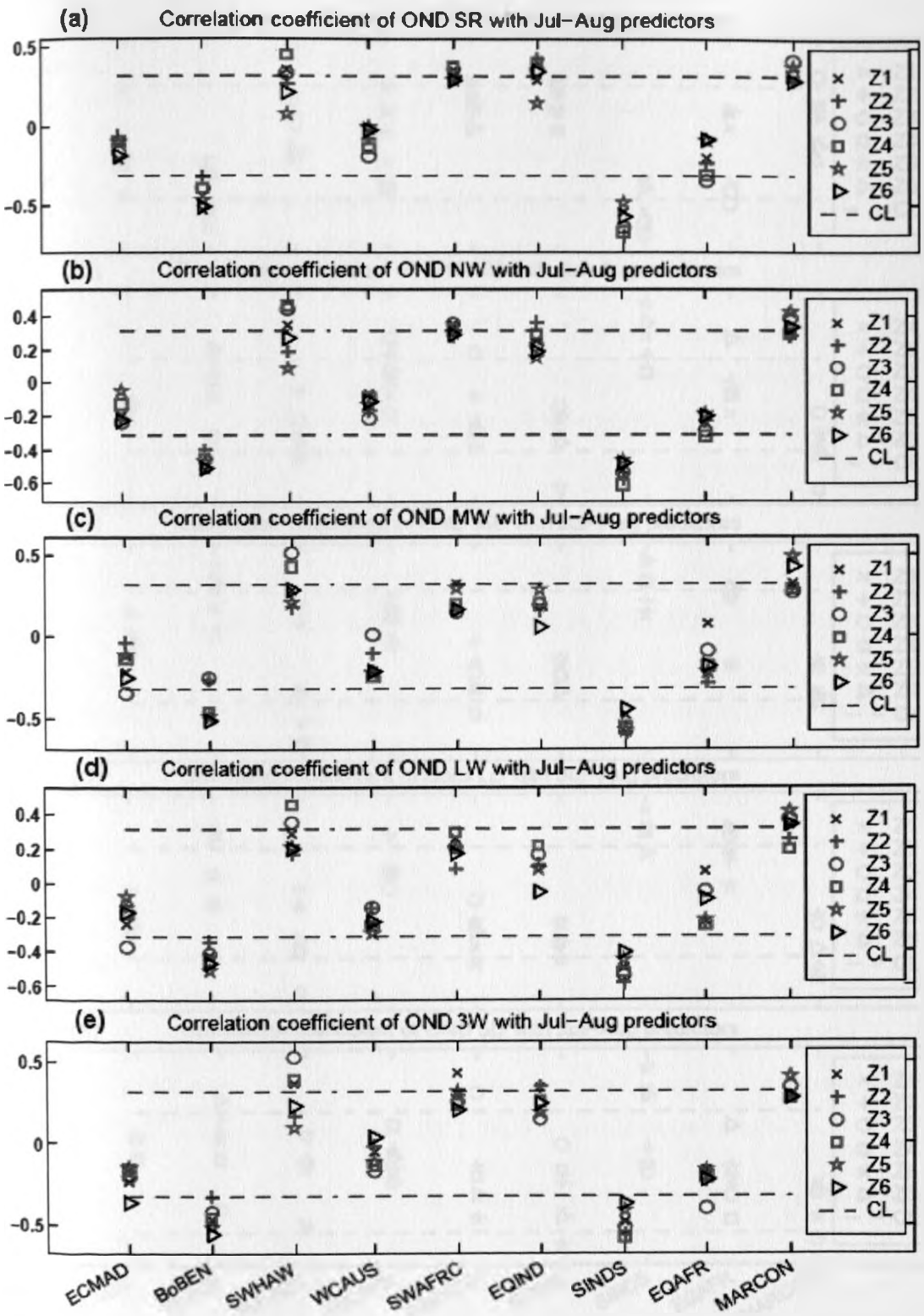


Figure 4.23: Correlation coefficient between the nine additional potential predictors identified averaged over July-August period and the areal-averaged October-November-December (a) seasonal rainfall totals, (b) number of wet days, (c) mean length of wet spell, (d) longest wet spell and (e) frequency of 3 wet days or more, over the six rainfall sub-regions Z1 to Z6. CL shows 95% confidence level threshold

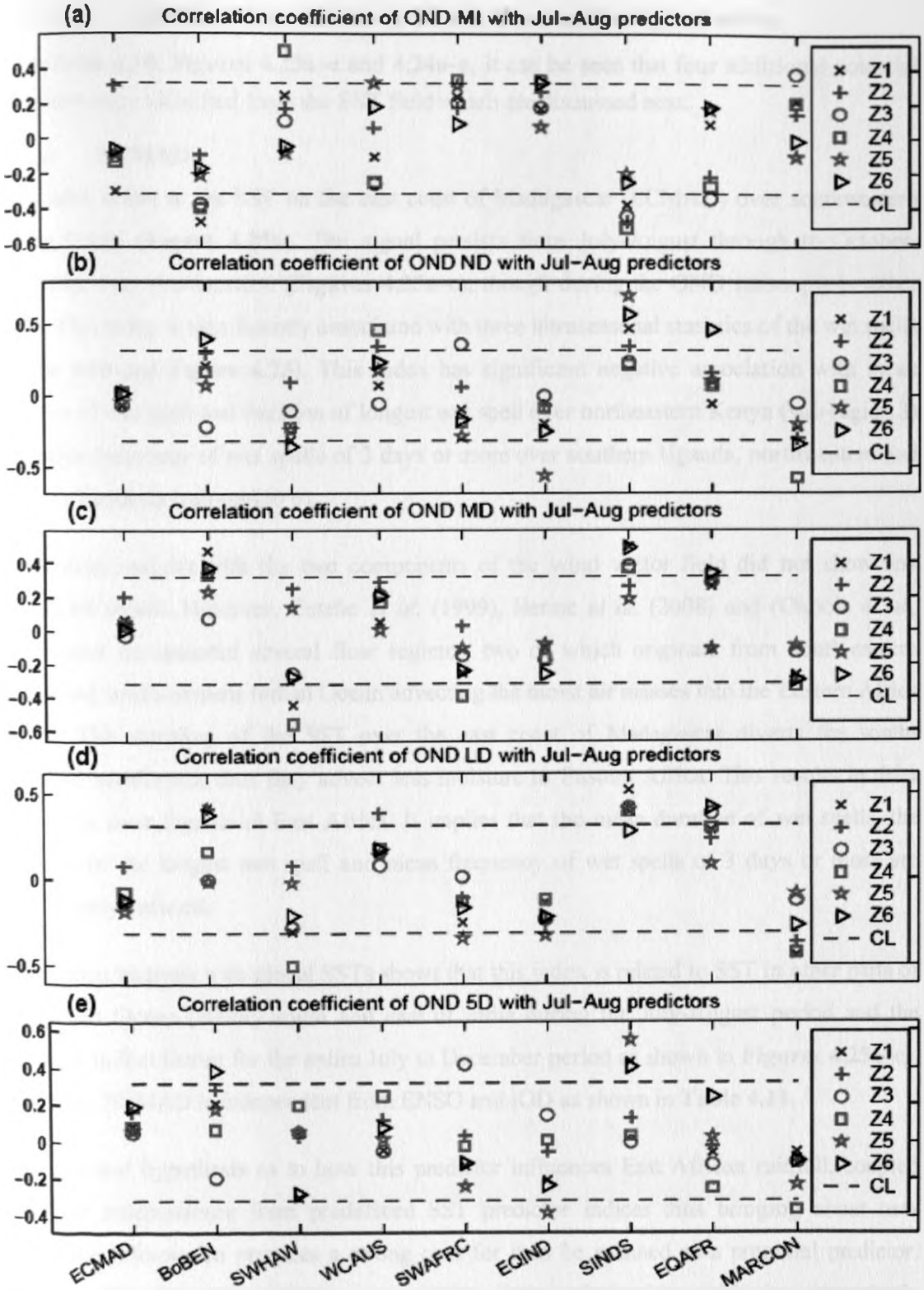


Figure 4.24: Correlation coefficient between the nine additional potential predictors identified averaged over July-August period and the areal-averaged October-November-December (a) mean rainfall intensity, (b) number of dry days, (c) mean length of dry spell, (d) longest dry spell and (e) frequency of 5 dry days or more, over the six rainfall sub-regions Z1 to Z6. CL shows the 95% confidence level threshold

4.5.1.2.1 Additional predictors from the sea surface temperature

From **Table 4.10, Figures 4.23a–e** and **4.24a–e**, it can be seen that four additional potential predictors were identified from the SST field which are discussed next.

(a) ECMAD

This index refers to the SST on the east coast of Madagascar (ECMAD) over southwestern Indian Ocean (**Figure 4.25a**). The signal persists from July-August through to October-December over this location (**Figures 4.25a–c**), though during the OND season, it is rather weak. This index is significantly associated with three intraseasonal statistics of the wet spells (**Table 4.10** and **Figure 4.23**). This index has significant negative association with mean duration of wet spell and duration of longest wet spell over northeastern Kenya (sub-region 3) and mean frequency of wet spells of 3 days or more over southern Uganda, northwestern and western Tanzania (sub-region 6).

Correlation analysis with the two components of the wind vector field did not show any significant signal. However, Gatebe *et al.* (1999), Henne *et al.* (2008) and (Okoola *et al.*, 2008) have documented several flow regimes, two of which originate from south-eastern Africa and south-western Indian Ocean advecting the moist air masses into the Eastern Africa region. The warming of the SST over the east coast of Madagascar diverts the south-easterlies southwards thus they advect less moisture to Eastern Africa. This results in drier conditions over Equatorial East Africa. It implies that the mean duration of wet spells, the duration of the longest wet spell and mean frequency of wet spells of 3 days or more are significantly reduced.

Correlation analysis with global SSTs shows that this index is related to SST in other parts of the Indian Ocean (mainly south and east of India during the July-August period and the southern Indian Ocean for the entire July to December period as shown in **Figures 4.25a–c**). However, ECMAD is independent from ENSO and IOD as shown in **Table 4.11.**

The physical hypothesis as to how this predictor influences East African rainfall, coupled with the independence from predefined SST predictor indices thus bringing about new predictive information provides a strong case for it to be retained as a potential predictor. However the fact that this index has weaker linear relationships with the atmospheric variables (zonal and meridional components of wind vector, specific humidity and geopotential height) over East Africa during the OND season marks its major weakness.

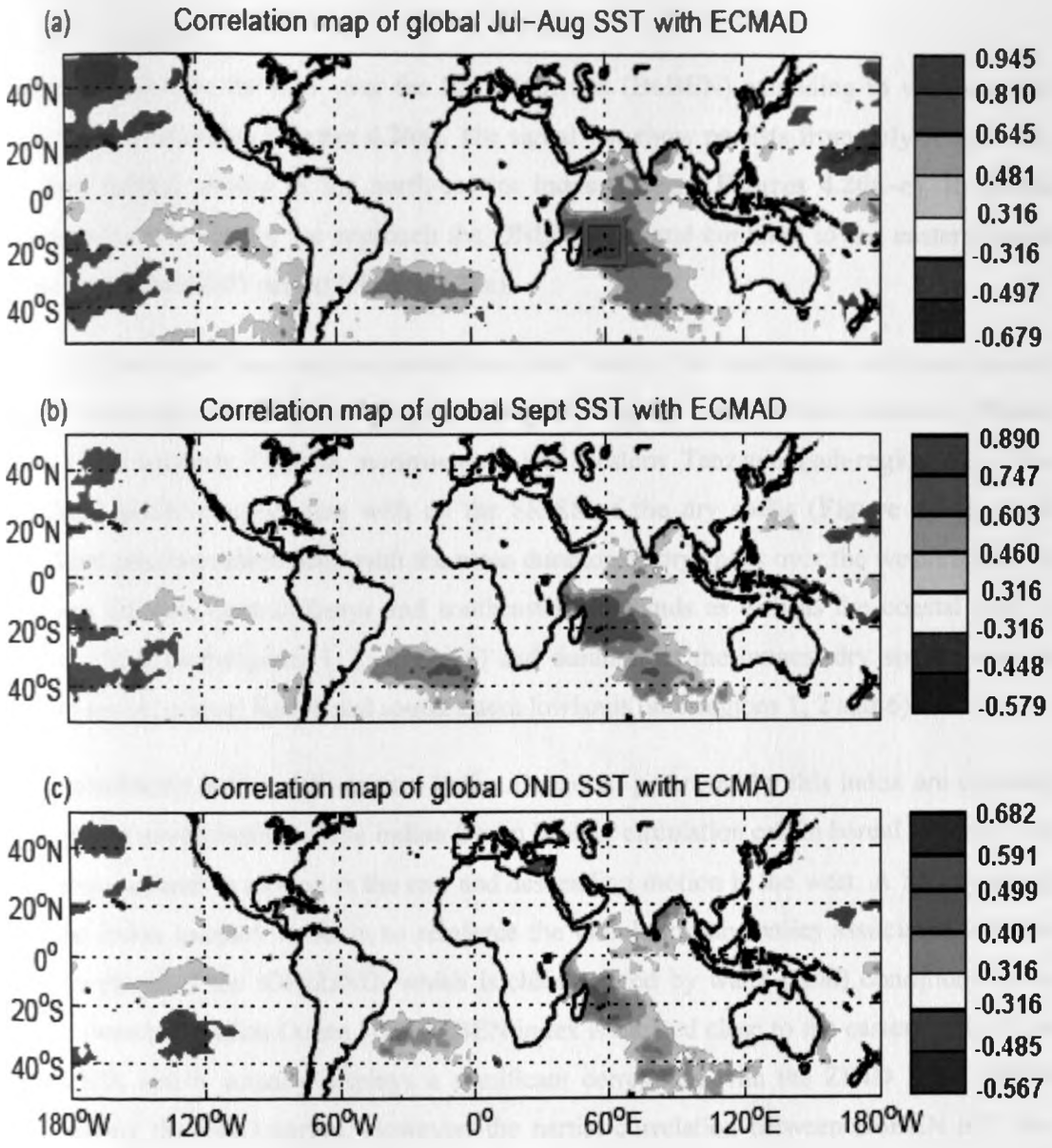


Figure 4.25: Map of significant correlation between East Coast of Madagascar (ECMAD) SST index and global SST for (a) July-August, (b) September and (c) October-December. The green rectangle in (a) shows the approximate location of ECMAD SST index computed for July-August period from 1962 to 2000

Table 4.11: Correlation coefficients between East Coast of Madagascar (ECMAD) SST index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	ZIND
Jul-Aug	0.26	0.19	0.25	0.14	0.28	-0.02
OND	0.23	0.26	0.19	0.22	0.17	0.01

(b) BoBEN

This index refers to the SST over the Bay of Bengal (BoBEN) extending to west coast of Malaysia and Indonesia (**Figure 4.26a**). The signal somehow persists from July-August into the short rainfall season in the north-eastern Indian Ocean (**Figures 4.26a–c**). Its spatial extent tends to reduce as we approach the OND season and confined to the eastern Indian Ocean during the OND period (**Figure 4.26c**).

This index has significant negative association over most of the sub-regions with the seasonal rainfall totals and all SRISS of the wet spells including the mean rainfall intensity (**Figure 4.23**). Over southern Uganda, northwestern and western Tanzania (sub-region 6), it has significant positive association with all the SRISS of the dry spells (**Figure 4.24**). It has significant positive relationship with the mean duration of dry spells over the western parts of the study domain, central Kenya and southeastern lowlands as well as the coastal strip of eastern Africa (sub-regions 1, 2, 4 and 6) and duration of the longest dry spells over the western sector, central Kenya and southeastern lowlands (sub-regions 1, 2 and 6).

Warm conditions in the north-eastern Indian Ocean, as portrayed by this index are expected to result in a strengthening of the Indian Ocean Walker circulation cell in boreal autumn, with anomalous ascending motion in the east and descending motion in the west. A SST warming over the index location is likely to reinforce the circulation anomalies associated with the negative phase of the IOD/ZIND, which is characterized by warm (cold) conditions in the eastern (western) Indian Ocean. The BoBEN index is located close to the eastern pole of the IOD/ZIND, and it actually displays a significant correlation with the ZIND index (**Table 4.12**) during the OND period. However, the partial correlation between BoBEN and East Africa rainfall, independent of ZIND is still significant, which means BoBEN brings independent predictive information. The strengthening of the Indian Ocean Walker circulation cell results in the reduction of seasonal rainfall totals, number of wet days, duration of the longest wet spells, and the mean frequency of the wet spells of 3 days or more over Eastern Africa. The mean duration of the dry spells and duration of the longest dry spells are also increased.

Correlation analysis with the global SST shows that this index has no signal over the tropical Pacific Ocean (**Figures 4.26a–c**), which is further confirmed by the insignificant correlation coefficients between the Niño indices and this index (**Table 4.12**). This independence from

ENSO and indices from Indian Ocean coupled with the fact that it shows significant relationship with different atmospheric variables (zonal and meridional components of wind vector, specific humidity and geopotential height) over East Africa during the OND season, justify its retention as a potential predictor.

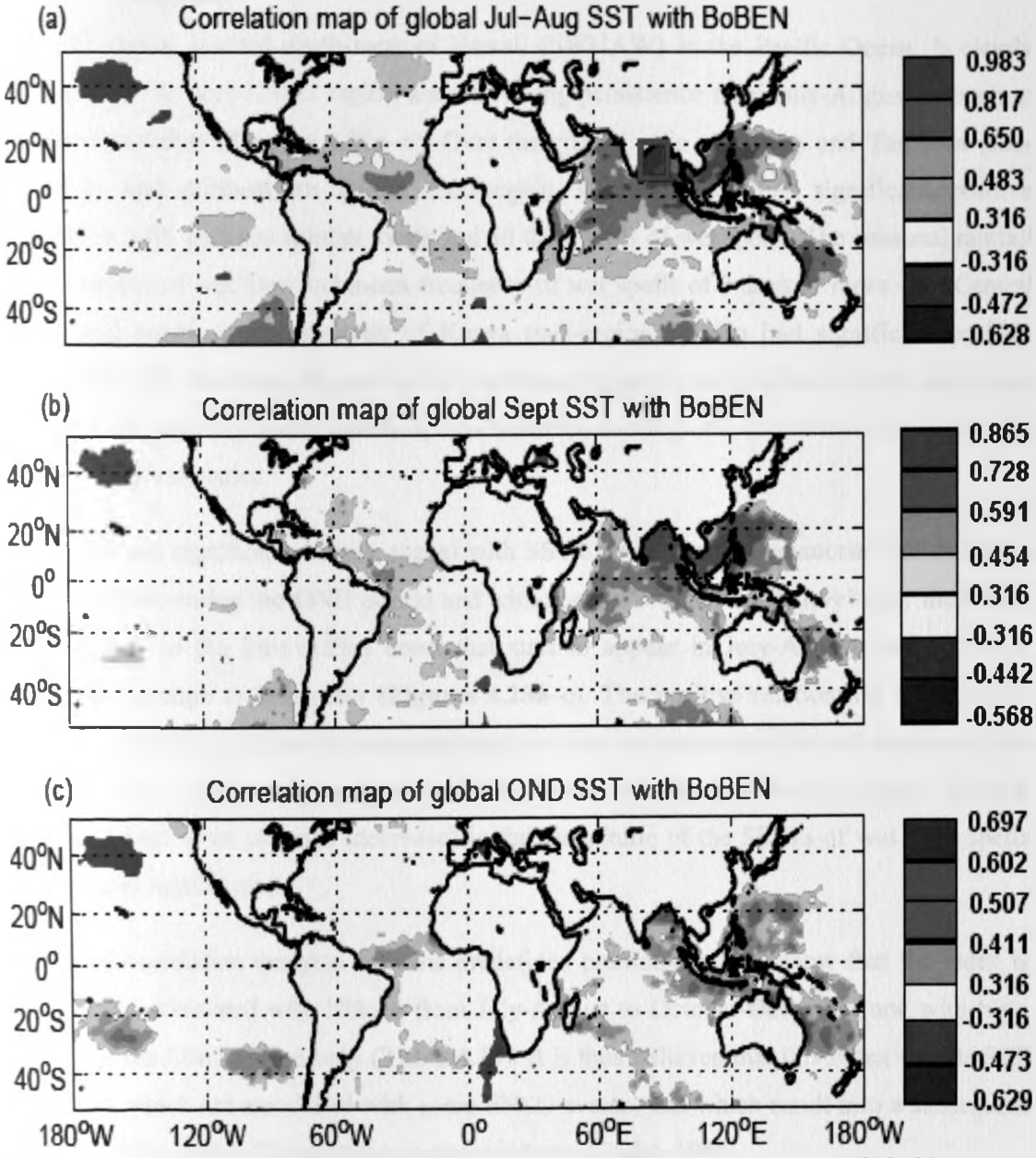


Figure 4.26: Map of significant correlation between Bay of Bengal (BoBEN) SST index and global SST for (a) July-August, (b) September and (c) October-December. The green rectangle in (a) shows the approximate location of BoBEN SST index computed for July-August period from 1962 to 2000

Table 4.12: Correlation coefficients between Bay of Bengal (BoBEN) SST index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	ZIND
Jul-Aug	0.28	-0.02	0.02	-0.16	-0.12	-0.27
OND	0.03	-0.15	-0.18	-0.23	-0.21	**-0.39**

(c) SWHAW

This SST index, located south-west of Hawaii (SWHAW) in the Pacific Ocean, is clearly distinct from the core ENSO region and has strong persistence from July-August through to October-December (**Figures 4.27a–c**). Over the coastal strip of Kenya and Tanzania (sub-region 4) and northeastern Kenya (sub-region 3), this index has significant positive association with seasonal rainfall totals and all the SRISS of wet spells. The seasonal rainfall totals, number of wet days and mean frequency of wet spells of 3 days or more over Central Kenya and southeastern lowlands of Kenya (sub-region 1) also had significant positive association with this index (**Figure 4.24**). Significant negative association with this index was observed for some dry spells statistics, over scattered sub-regions, and in less consistent way than wet spells statistics.

This index has significant positive signal with SST over the central equatorial Indian Ocean (**Figure 4.27c**) during the OND period and with zonal winds at 925mb level over the Indian Ocean closer to the East Africa coast that start to appear in July-August, and grows in September through to December (**Figures 4.28a–c**). The positive relationship of SRISS of wet spells with the zonal wind component implies that the south-easterlies are weakened thus depositing more moisture over Eastern Africa, which results in wetter conditions. The wet conditions lead to an increase (decrease) in the magnitude of the SRISS of wet (dry) spells and seasonal rainfall totals.

Results of correlation analysis with the predefined predictor indices show that the index is significantly correlated with Niño 4 from July-August to October-December and with Niño 3.4 during the OND season only (**Table 4.13**). It is thus believed that this index depicts SST conditions which are associated with some ENSO events, and which result into a subsequent warming of the Indian Ocean in the northern autumn (Cadet, 1985).

The independence of this predictor from most of the Niño indices and indices from Indian Ocean, and its relationship to several atmospheric variables around East Africa during the OND season, justify its retention as an additional potential predictor.

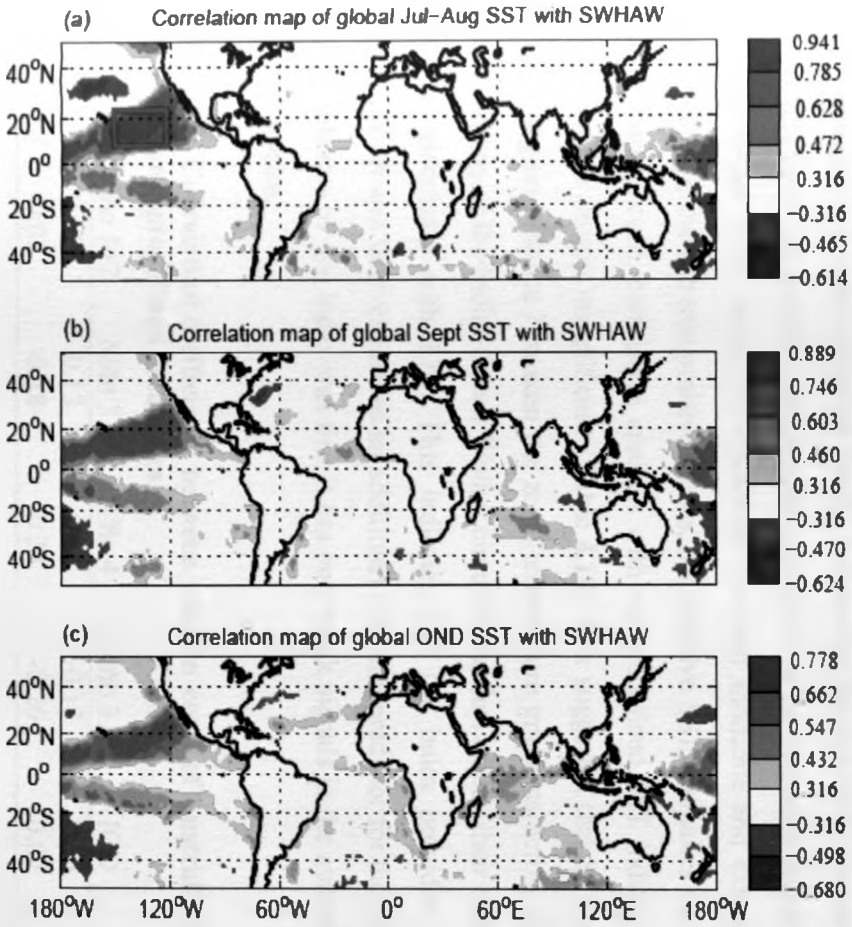


Figure 4.27: Map of significant correlation between South-West of Hawaii (SWHAW) SST index and global SST for (a) July-August, (b) September and (c) October-December. The green rectangle in (a) shows the approximate location of SWHAW SST index computed for July-August period from 1962 to 2000

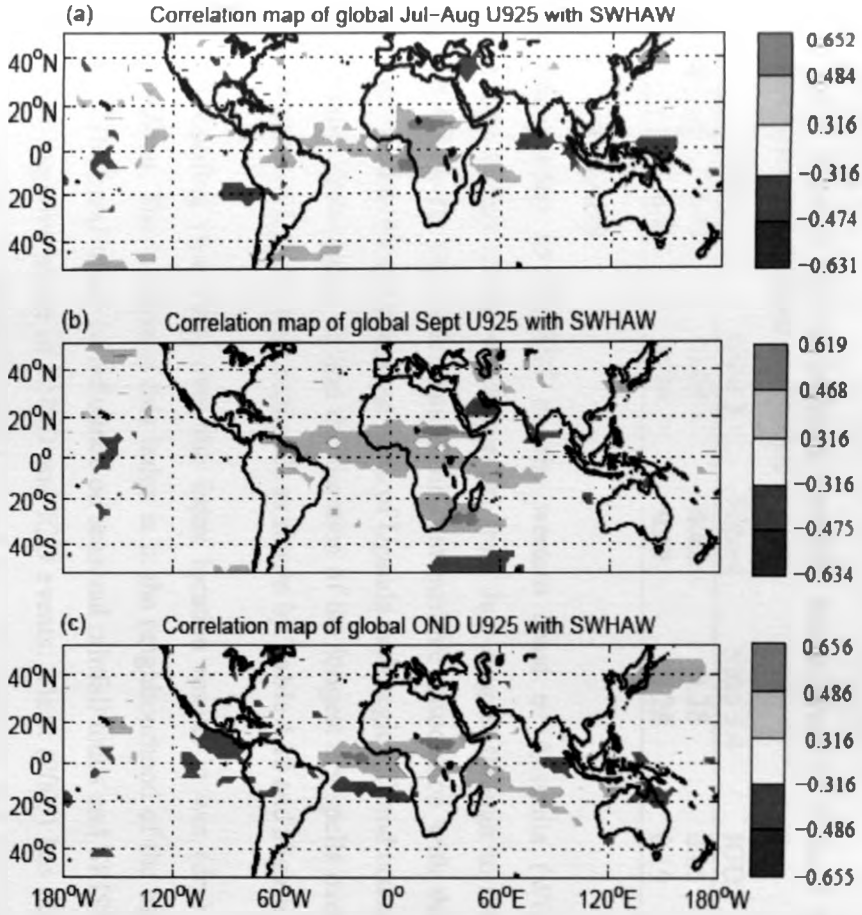


Figure 4.28: Map of significant correlation between South-West of Hawaii (SWHAW) SST index and global U925 for (a) July-August, (b) September and (c) October-December

Table 4.13: Correlation coefficients between South-West of Hawaii (SWHAW) SST index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	ZIND
Jul-Aug	0.14	0.09	**0.62**	0.28	0.05	-0.10
OND	**0.36**	0.24	**0.49**	0.28	0.19	-0.03

(d) WCAUS

This index refers to the SST on the western coast of Australia (WCAUS) over the southeastern Indian Ocean which persists from July-August through to October-December (**Figures 4.29a–c**). This index has significant positive association with the number of dry days over western Kenya and most parts of Uganda (sub-region 2) and coastal strip of Kenya and Tanzania (sub-region 4); and the duration of the longest dry spells over western Kenya and most parts of Uganda (sub-region 2) as shown by **Table 4.10** and **Figure 4.24**.

The SST cooling (warming) over the index location results in wet (dry) conditions over Eastern Africa. The location of this index is in the neighbourhood of the eastern pole of the IOD and ZIND and hence its influence on seasonal rainfall totals and SRISS may be similar to those of negative phase of ENSO and IOD events. Black (2005) has observed localized regions of cooling along the Sumatran coast and off Australia that was caused by ENSO-related anomalies in water transport via the Indonesian through-flow (Meyers, 1996). Black et al. (2003) have documented anomalously cold SSTs in the southeast Indian Ocean near the north Australian coast, which are significant for both moderate and extreme IOD events. However during the extreme positive events, the negative SST anomalies are centered on the equator. Furthermore, correlation analysis with the predefined SST indices shows that the coefficients are mostly insignificant (**Table 4.14**). It is suggested that SST variations in the WCAUS region enhance / weaken the zonal temperature gradients across the Indian Ocean thus impacting on the Indian Ocean Walker circulations, thus adding their effects or not to the gradients associated with IOD. This indicates that the index provides new predictive information. It was thus retained as an additional potential predictor for the OND season. The major weakness of this index was that it has only weak signals in the atmospheric variables over East Africa.

Table 4.14: Correlation coefficients between western coast of Australia (WCAUS) SST index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	ZIND
Jul-Aug	0.08	-0.13	-0.12	-0.18	0.09	-0.31
OND	-0.08	-0.28	-0.24	-0.29	0.02	-0.27

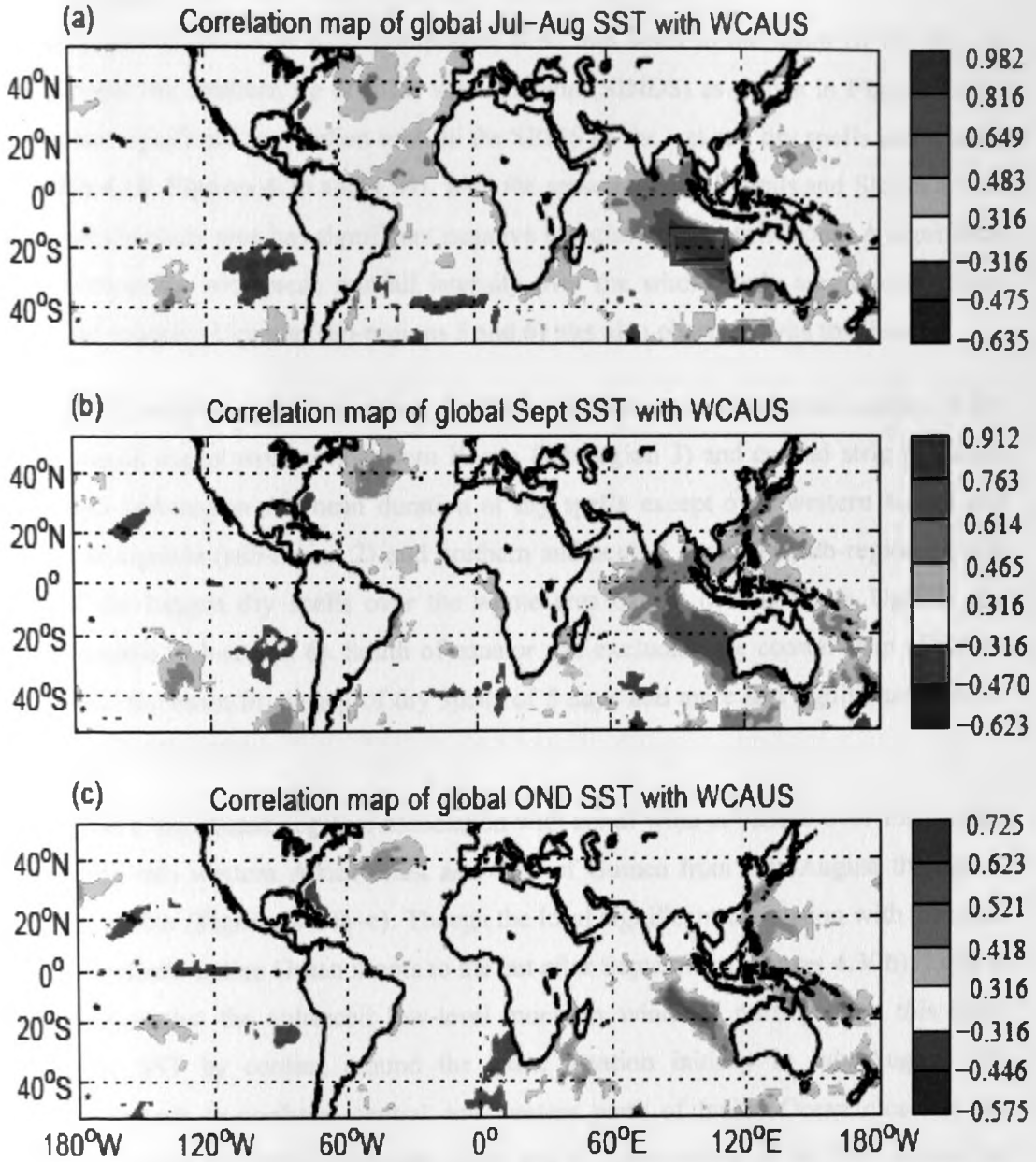


Figure 4.29: Map of significant correlation between western coast of Australia (WCAUS) SST index and global SST for (a) July-August, (b) September and (c) October-December. The green rectangle in (a) shows the approximate location of WCAUS SST index computed for July-August period from 1962 to 2000

4.5.1.2.2 Additional predictors from the wind field

Three additional predictors were identified from the zonal component of the wind field (**Tables 4.9 and 4.10**). None of the SRISS picked any predictor from the meridional component of the wind field. The associations with tropical zonal winds suggest that Walker circulation anomalies are involved in the teleconnections.

(a) SINDS

This index refers to the zonal wind component at 925mb level to the south of the Bay of Bengal and near the southern tip of India sub-continent (SINDS) as shown in **Figure 4.30a**. This index has significant association with all the SRISS of the wet and dry spells and rainfall totals (**Table 4.10, Figures 4.23** and **4.24**). With the seasonal rainfall totals and SRISS of wet spells, the whole study area has significant negative association with this index. A significant negative relationship with mean rainfall intensity over the whole study area except inland Tanzania and southern Uganda (sub-regions 5 and 6) was also observed with this index.

Over the whole study area, the index have significant positive association with number of dry days in a season except over northeastern Kenya (sub-region 3) and coastal strip of Kenya and Tanzania (sub-region 4); mean duration of dry spells except over western Kenya and most parts of Uganda (sub-region 2) and northern and central Tanzania (sub-region 5); and duration of the longest dry spells over the whole area except over southern Uganda and western Tanzania (sub-region 6). South of equator and excluding the coastal strip of Kenya and Tanzania, the mean frequency of dry spells of 5 days and more had significant positive association with this index.

The index has a significant negative association with zonal wind at 925mb over East Africa and extending into western Africa coast and Gulf of Guinea from July-August through to October-December (**Figures 4.30a–c**). Though the local significant correlation with the zonal wind over Northern Indian Ocean seems to die out after September (**Figures 4.30b**), **Figures 4.31a–c** suggest that the enhanced low-level monsoon winds as portrayed by this index modulate the SST by cooling around the index location initially in July-August. The modulation spreads to northern, central and western parts of Indian Ocean closer to the western pole of the IOD/ZIND (**Figures 4.31b** and **c**). The cooling of the SST around the western pole of the IOD/ZIND is, at times associated with the drier conditions over the eastern Africa. The drier conditions results in reduction in the magnitude of seasonal rainfall totals and SRISS of the wet spells as well as an increase in the magnitude of the SRISS of dry spells.

Correlation analysis with the global SST shows that this signal is significantly but negatively correlated with SST over the Niño regions from July-August through to October-December and over much of the northern and western Indian Ocean during the October-December period. This is further confirmed by the strong negative significant correlation coefficients

obtained with the Niño and Indian Ocean indices (**Table 4.15**).

The dependence of the SINDS wind index on Niño and Indian Ocean indices is discussed here. The SINDS wind index was identified when the partial correlation analysis was undertaken between the July-August zonal component of wind field at 925mb level and number of wet days during the OND season while controlling the effects of July-August predefined predictors (Niño3.4 and ZIND) and the five additional potential predictors earlier identified from the SST field.

Table 4.16 shows the concurrent total and partial correlation coefficients between the SINDS wind index and number of wet days over the six sub-regions while controlling the effects of predefined and additional potential predictors averaged for July-August period. The number of wet days has significant negative (positive) total correlation coefficient with SINDS (Niño 3.4) over the six sub-regions and significant positive total correlation coefficient with ZIND over two sub-regions.

Table 4.15: Correlation coefficients between southern tip of India sub-continent (SINDS) zonal wind index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	ZIND
Jul-Aug	**-0.57**	**-0.66**	**-0.61**	**-0.70**	**-0.45**	-0.07
OND	**-0.72**	**-0.69**	**-0.60**	**-0.69**	**-0.61**	**-0.32**

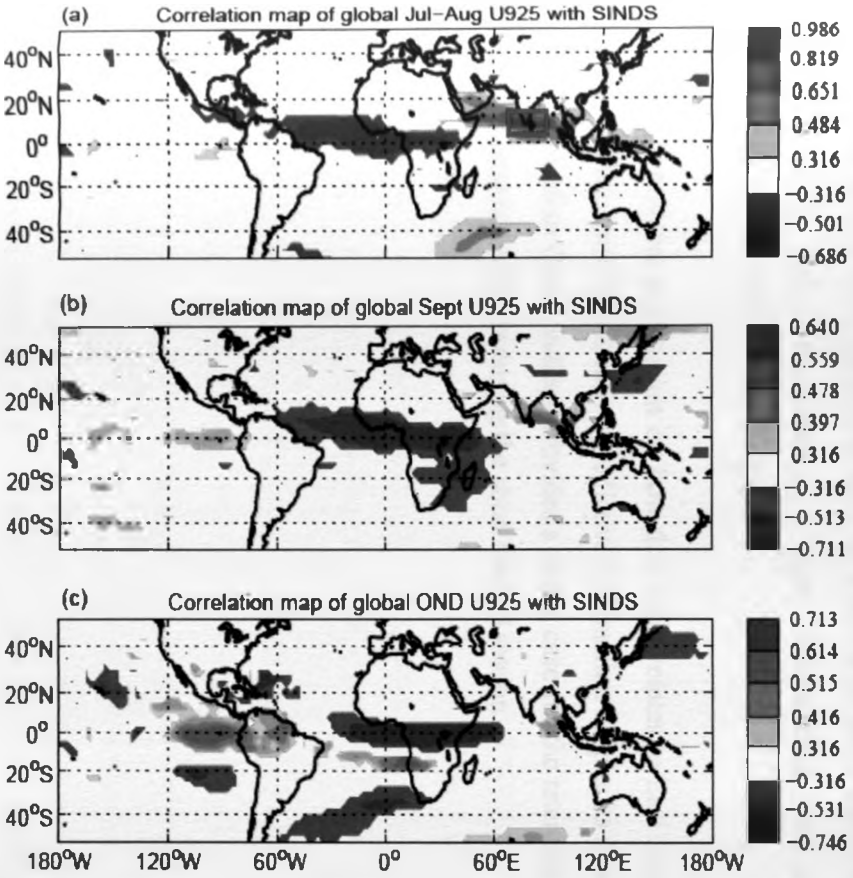


Figure 4.30: Map of significant correlation between southern tip of India sub-continent (SINDS) zonal wind index and global U925 for (a) July-August, (b) September and (c) October-December. The green rectangle in (a) shows the approximate location of SINDS zonal wind index computed for July-August period from 1962 to 2000

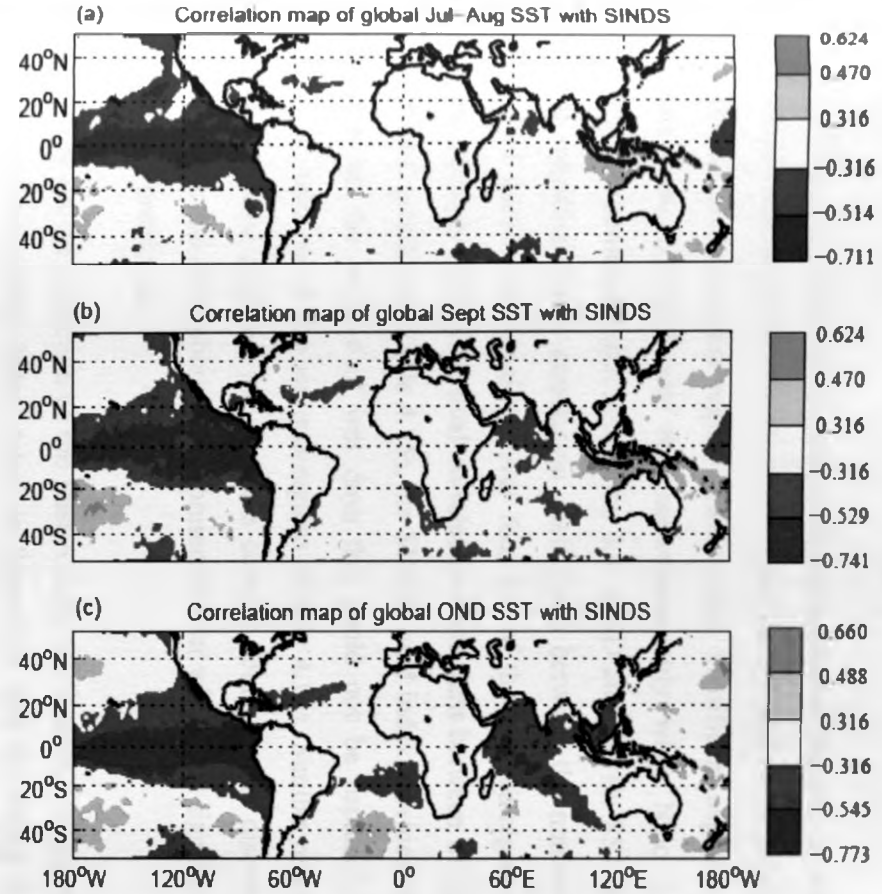


Figure 4.31: Map of significant correlation between southern tip of India sub-continent (SINDS) zonal wind index and global SST for (a) July-August, (b) September and (c) October-December

Significant negative partial correlation coefficients between number of wet days (RNW) and SINDS while controlling the effects of ZIND and the six additional potential SST predictors were obtained over the entire region (**Table 4.16**). Significant negative partial correlation coefficients were obtained over two sub-regions only when Niño 3.4 was controlled. Controlling the combined effects of Niño 3.4, ZIND and the six SST potential predictors, significant negative partial correlation coefficients between the number of wet days and SINDS wind index were obtained over four out of the six sub-regions. This means that despite the strongly significant total correlation coefficient between SINDS wind index and the predefined predictors (**Table 4.15**), the SINDS wind index provides additional predictive information on the number of wet days that could not be captured by the predefined predictors in two out of the six sub-regions (**Table 4.16**). Similar remarks apply to other additional potential predictor indices from oceanic fields already discussed earlier such as SWHAW index and atmospheric fields presented later which had significant correlation with the predefined predictors.

In their study on the prediction of the East African OND rains, Philippon *et al.* (2002) also found that an atmospheric index, taken in September and describing the Indian monsoon intensity, was having some predictive skill, in addition to more traditional SST predictors. The present study further demonstrates a partly independent predictive skill of Asian monsoon dynamics as early as July-August.

The plausible physical explanation on how this index relates to East Africa rainfall and the fact that this index provides additive predictive information despite its strong association with the Niño and Indian Ocean indices provides a strong case for the retention of this index as an additional potential predictor during the short rainfall season.

Table 4.16: **Total and partial correlation coefficients between areal-averaged number of wet days (NW) and southern tip of India sub-continent (SINDS) zonal wind index while controlling other predictors for July-August period. Bold numbers indicate that the coefficient is significant at 95% confidence level**

	Total correlation with NW			Partial correlation between NW & SINDS while controlling			
	SINDS	Niño3.4	ZIND	Niño3.4	ZIND	Other potential SST predictors	Niño3.4, ZIND & other potential SST predictors
1	**-0.53**	**0.41**	0.20	**-0.37**	**-0.52**	**-0.48**	**-0.42**
2	**-0.59**	**0.59**	0.24	-0.30	**-0.59**	**-0.56**	**-0.43**
3	**-0.54**	**0.52**	0.20	-0.29	**-0.54**	**-0.46**	-0.24
4	**-0.62**	**0.48**	0.22	**-0.45**	**-0.62**	**-0.55**	**-0.52**
5	**-0.47**	**0.42**	**0.32**	-0.27	**-0.47**	**-0.46**	**-0.43**
6	**-0.49**	**0.45**	**0.34**	-0.27	**-0.49**	**-0.45**	-0.28

(b) EQAFR

This index refers to the zonal wind component at 200mb level extending from Equatorial Africa (EQAFR) into Equatorial Atlantic Ocean (**Figure 4.32a**). This signal persists from July-August through to October-December (**Figures 4.32a–c**). Over the coastal strip of Kenya and Tanzania (sub-region 4), this index has significant inverse association with seasonal rainfall totals and number of wet days. Significant negative association were obtained between this index and seasonal rainfall totals, number of wet days in a season, mean frequency of wet spells of 3 days or more over northeastern Kenya (sub-region 3) as shown in **Figure 4.23**. Significant positive association are observed over southern Uganda, northwestern and western Tanzania (sub-region 6) between this index and the number of dry days, mean duration of the dry spells and the duration of the longest dry spell (**Figure 4.24**).

The negative (positive) association of this index with SRISS of wet (dry) spells implies that a weakening of the upper level easterlies over the Eastern Africa region tends to precede dry conditions over East Africa. Correlation analysis with global SST shows that this index is associated with the cooling of SST over northern and western Indian Ocean and most parts of the tropical eastern Pacific Ocean during the July to December period (**Figures 4.33a–c** and **Table 4.17**). It should be noted that this index has significant negative association with the

zonal component of wind and specific humidity both at 925mb level extending from equatorial eastern Atlantic Ocean, through equatorial Africa into the equatorial western Indian Ocean (not shown). The weakening of the upper level easterlies over Eastern Africa coupled with the cooling of the SST in the northern and western Indian Ocean and tropical eastern Pacific Ocean are typical of the negative phase of the IOD and ENSO events that results in dry conditions over the Eastern Africa.

This index also displays quite a strong persistence at 200mb level over Africa from July to December (**Figures 4.32a–c**). During the OND season, the index has a symmetrical (about equator) but negative association with zonal wind component at 200mb level at 20° N/S and extending from longitudes 0° to about 90° E. This persistence may be partly explained by the strong connection with ENSO, itself a persistent phenomenon. However, the wind signal also has an independent component anchored at African longitudes, out-of-phase between the upper (200mb) and the lower levels (925mb) as discussed above. It is therefore suggested to depict variations in the (zonal) Walker circulation above equatorial Africa, partly associated to SST anomalies and possibly to land surface conditions. These circulation anomalies have an evident connection to East African rainfall.

This justifies the retention of this index as an additional potential predictor, although the strong association with the Niño and Indian Ocean indices marks its major weakness.

Table 4.17: Correlation coefficients between Equatorial Africa (EQAFR) zonal wind index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	ZIND
Jul-Aug	**-0.46**	**-0.57**	**-0.75**	**-0.71**	-0.32	-0.00
OND	**-0.65**	**-0.70**	**-0.75**	**-0.73**	**-0.49**	-0.29

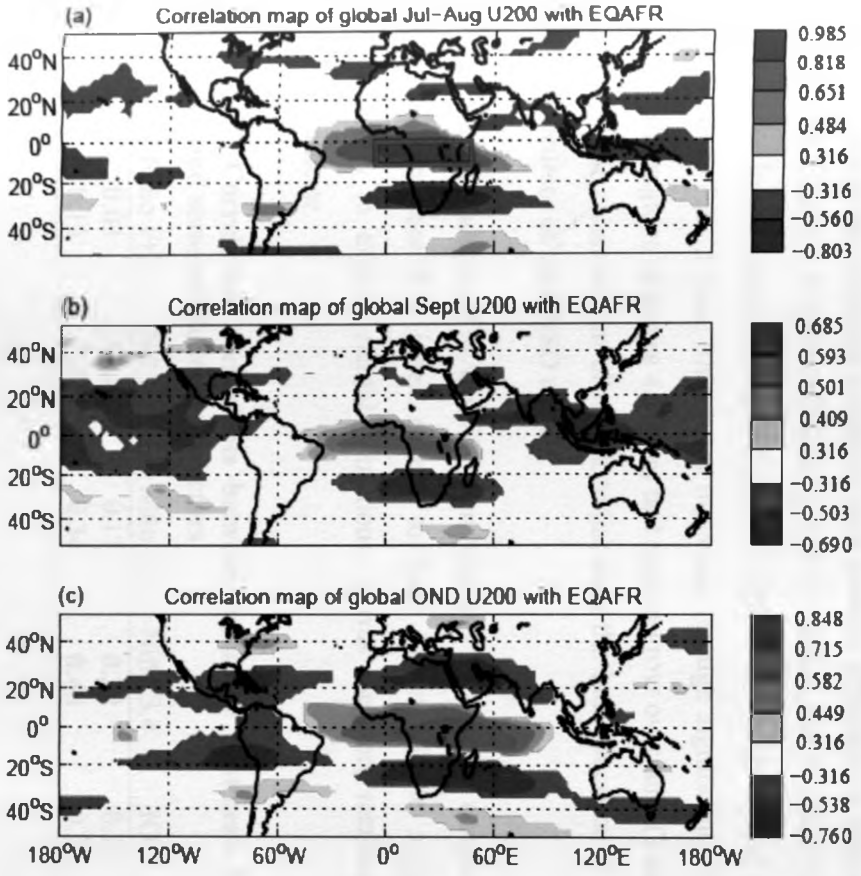


Figure 4.32: Map of significant correlation between Equatorial Africa (EQAFR) zonal wind index and global U200 for (a) July-August, (b) September and (c) October-December. The green rectangle in (a) shows the approximate location of EQAFR zonal wind index computed for July-August period from 1962 to 2000

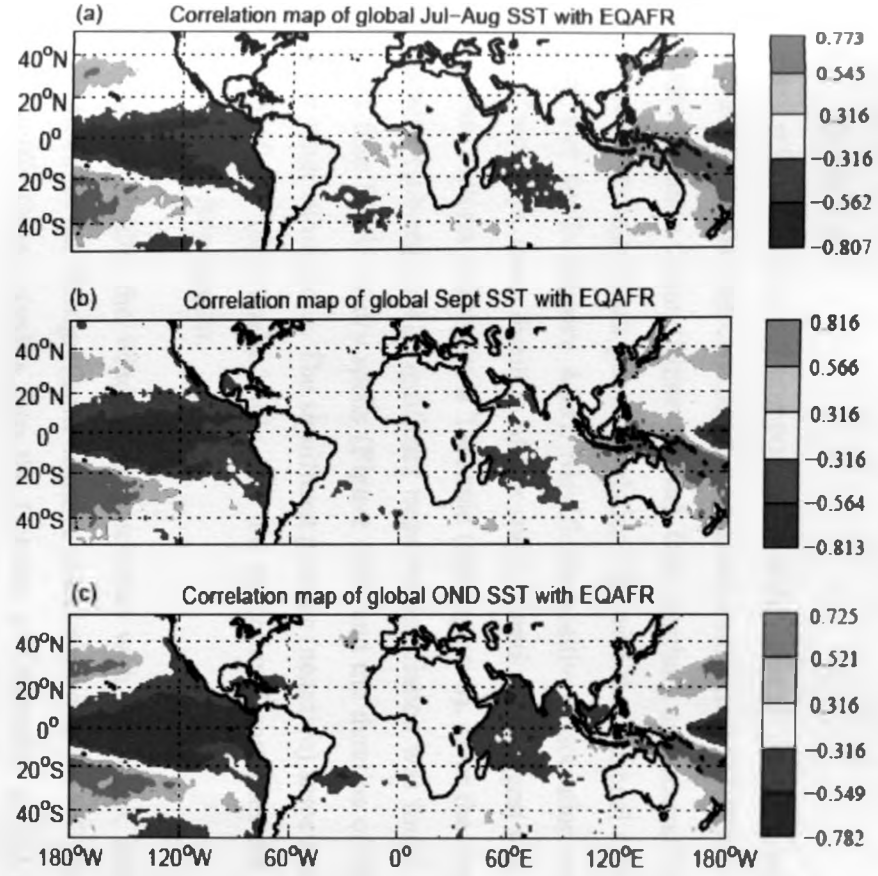


Figure 4.33: Map of significant correlation between Equatorial Africa (EQAFR) zonal wind index and global SST for (a) July-August, (b) September and (c) October-December

(c) MARCON

This index refers to the zonal component of wind field at 200mb level over the coast of Malaysia and Indonesia representing the maritime continent (MARCON) and extending over the equatorial Indian Ocean (**Figure 4.34a**). Over northern and central Tanzania (sub-region 5), this index has significant positive relationship with the seasonal rainfall totals and four SRISS of wet spells (**Figure 4.23**). Significant positive association was also noted with number of wet days, mean duration of wet spells, duration of longest wet spells over western part of Lake Victoria and western Tanzania (sub-region 6). Over the coastal strip of Kenya and Tanzania (sub-region 4), significant negative association was found between this index and all the four SRISS of dry spells (**Figure 4.24**) and the duration of the longest dry spells over a few other sub-regions. The significant positive (negative) association of this index and the SRISS of wet (dry) spells imply that the upper level easterlies are enhanced over the index location in July-August.

Similar to the SINDS, the wind signal associated with this index dies off in September (**Figures 4.34a–c**) but the significant negative association with the SST over this location and western Pacific Ocean persists into the October to December period (**Figures 4.35a–c**). Lower SSTs over the Maritime continent may produce atmospheric subsidence anomalies, a feature which weakens Walker circulations over the Indian and Pacific Oceans, thus resulting in an increase in the seasonal rainfall totals and SRISS of wet spells and a drop in SRISS of dry spells over East Africa. This index has weak (though significant) positive correlations over the Niño regions (**Figures 4.35a–c**) that tend to grow over time (**Table 4.18**). This may suggest that the enhancement of the easterlies over the index location in July to September maybe a precursor of the ENSO events.

The partial independence of MARCON from Niño and IOD indices coupled with the physical explanation on how this index relates to East Africa climate suggest this index is complementary to explain rainfall variations. It was therefore retained as an additional potential predictor.

Table 4.18: Correlation coefficients between maritime continent (MARCON) zonal wind index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	ZIND
Jul-Aug	0.03	**0.34**	0.17	**0.39**	0.22	0.20
OND	0.19	**0.43**	**0.34**	**0.48**	0.24	**0.44**

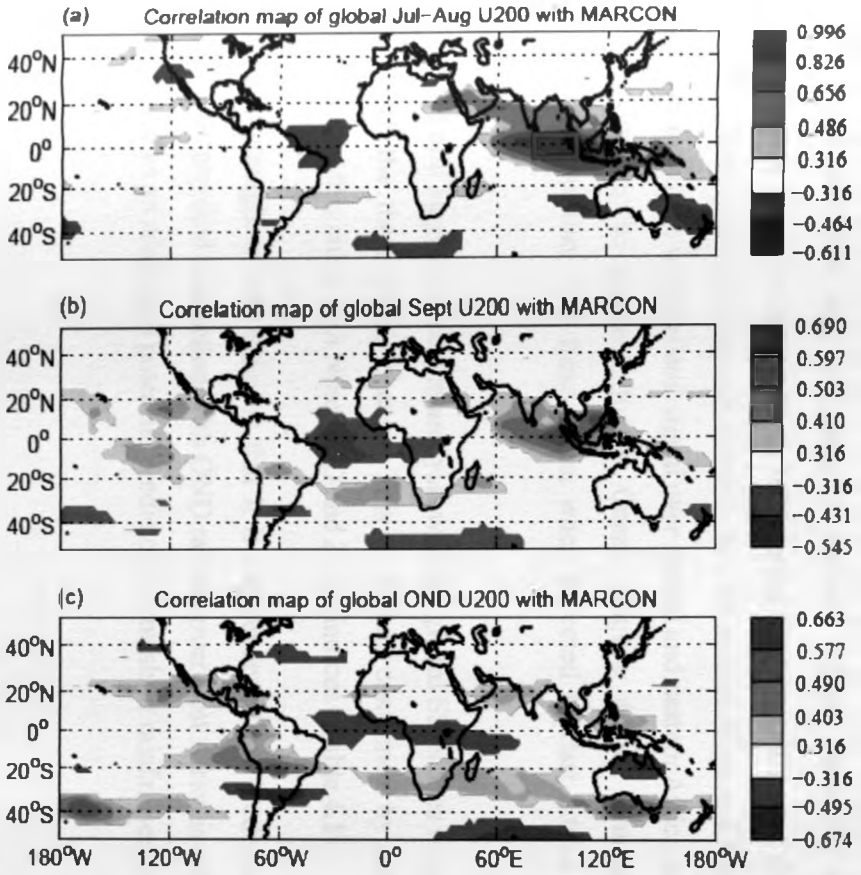


Figure 4.34: Map of significant correlation between maritime continent (MARCON) zonal wind index and global U200 for (a) July-August, (b) September and (c) October-December. The green rectangle in (a) shows the approximate location of MARCON zonal wind index computed for July-August period from 1962 to 2000

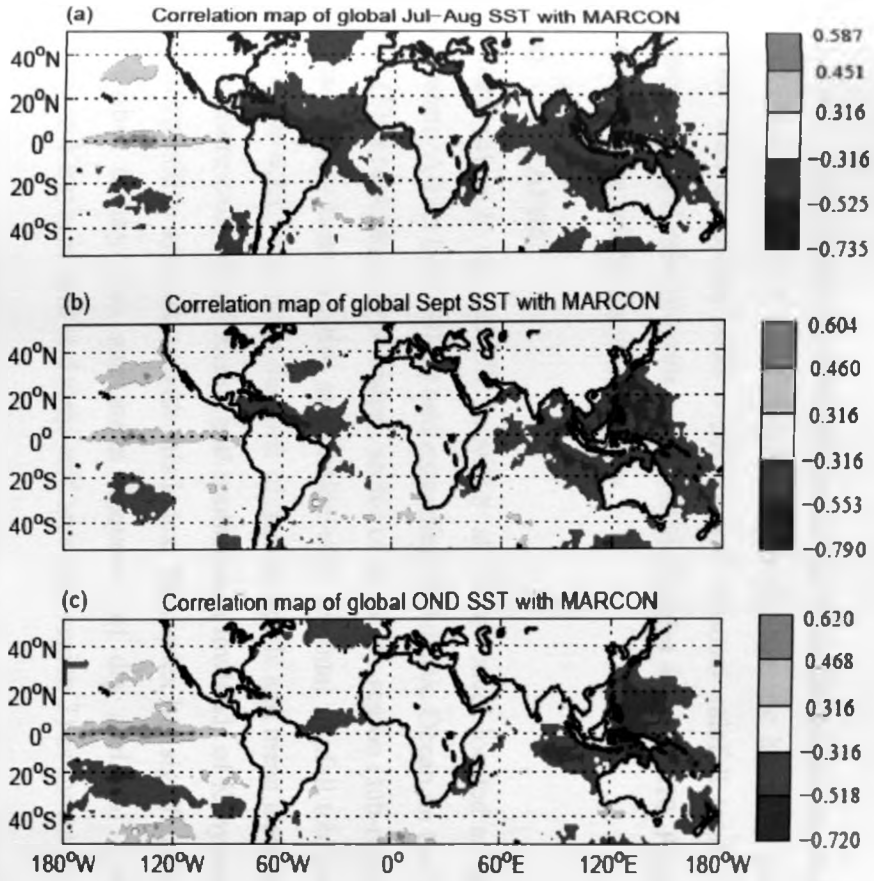


Figure 4.35: Map of significant correlation between maritime continent (MARCON) zonal wind index and global SST for (a) July-August, (b) September and (c) October-December

4.5.1.2.3 Additional predictors from the specific humidity field

Two additional predictors were identified from the specific humidity field at 700mb level. SWAFRC was associated with 8 SRISS and seasonal rainfall totals while EQIND was associated seasonal rainfall totals and 6 SRISS (**Tables 4.9** and **4.10, Figures 4.23** and **4.24**). These indices are briefly discussed below.

(a) SWAFRC

This index refers to the specific humidity at 700mb level located at the Angola coast on southwestern Africa (SWAFRC) and extending to Atlantic Ocean on the west and Zambia to the east (**Figure 4.36a**). Over eastern sector of the study region (sub-regions 1, 3 and 4), this index has a significant positive association with seasonal rainfall totals and number of wet days (**Figure 4.23**). The mean duration of the wet spells and mean frequency of wet spells of 3 days or more over central Kenya and southeastern lowlands of Kenya (sub-regions 1) have significant positive association with this index. This index has significant positive correlation with number of dry days and mean frequency of dry spells of 5 days or more over northeastern Kenya, an arid and semi-arid area (sub-region 3) as shown by **Figure 4.24**.

This index has a significant positive relationship with the specific humidity at 700mb level over Arabian Sea, Red sea, most parts of northern Africa, equatorial Atlantic Ocean and southern Indian Ocean around latitude 30°S during July-August period (**Figure 4.36a**). Over September, the signal weakens and seems to be a bit noisier (**Figure 4.36b**). During the OND period, the index has well-defined signal over central and eastern Africa and equatorial Indian Ocean extending to southern Indian Ocean (**Figure 4.36c**). Enhanced low- to mid-tropospheric moisture over these areas, when advected to East Africa may result in wet conditions.

Concurrent and lagged correlation analysis with the global SSTs does not show any persistent signal over the three global oceans. This is further confirmed by the weak correlation coefficients of this index with Nino, IOD and ZIND indices (**Table 4.19**).

The major strength of this index is that it is independent from predefined SST predictors. Although its physical connection with OND rainfall over East Africa is not straightforward, it was retained as an additional potential predictor for the short rainfall season.

Table 4.19: **Correlation coefficients between southwestern Africa (SWAFRC) specific humidity index and some predefined predictors**

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	ZIND
Jul-Aug	0.06	0.07	0.07	0.08	0.24	0.23
OND	0.20	0.22	0.01	0.14	**0.35**	**0.36**

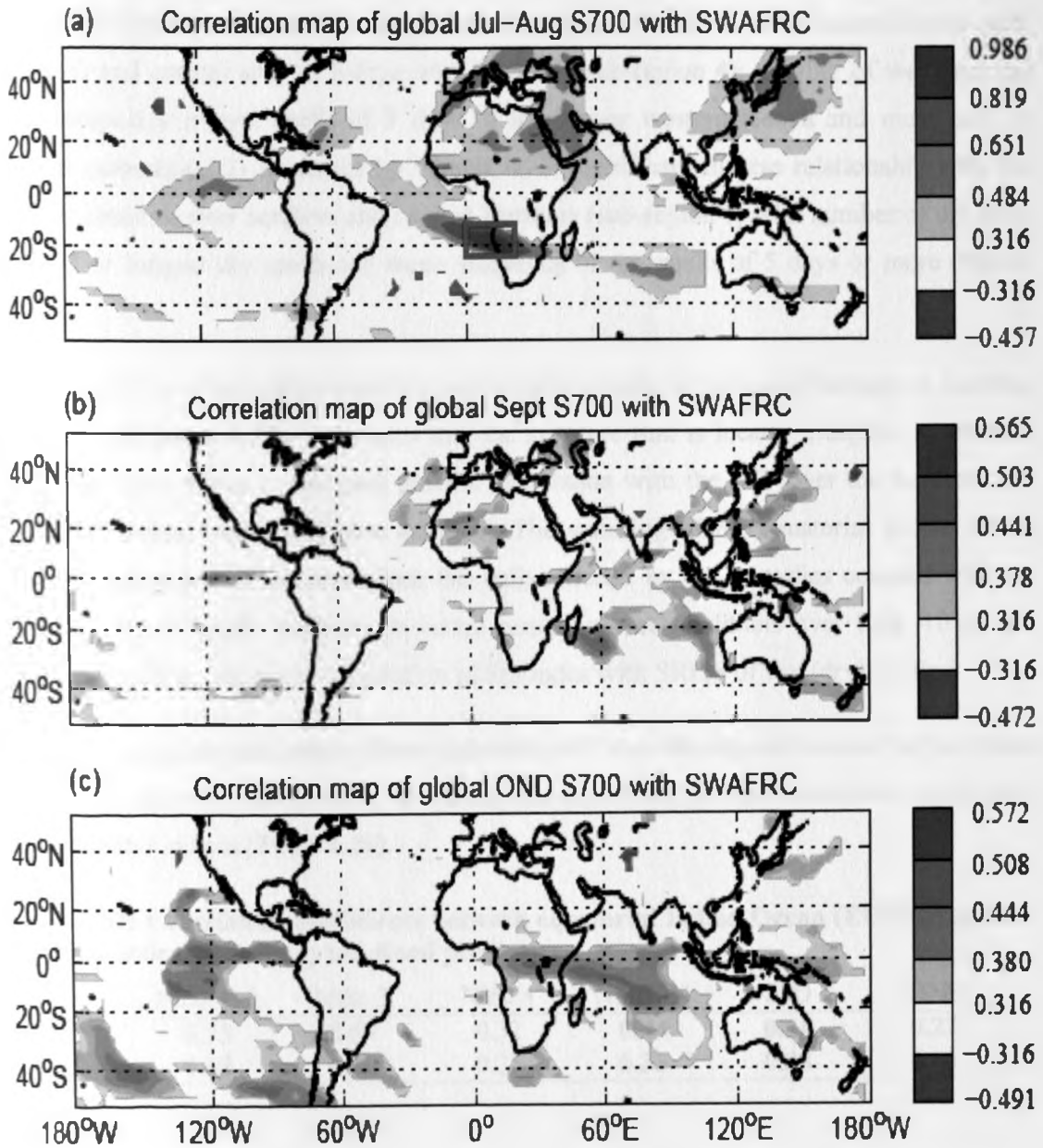


Figure 4.36: Map of significant correlation between southwestern Africa (SWAFRC) specific humidity index and global S700 for (a) July-August, (b) September and (c) October-December. The green rectangle in (a) shows the approximate location of SWAFRC specific humidity index computed for July-August period from 1962 to 2000

(b) EQIND

This index refers to the specific humidity at 700mb level extending from the southern tip of India subcontinent, through equatorial Indian Ocean (EQIND) into the eastern Africa region (**Figure 4.37a**). The index persists from July-August through to October-December (**Figures 4.37a–c**). Significant positive relationship exists between this index and seasonal rainfall totals over western sector of the study area (sub-regions 2 and 6), northeastern Kenya (sub-region 3) and coastal strip of Kenya and Tanzania (sub-region 4); number of wet days and mean frequency of wet spells of 3 days or more over western Kenya and most parts of Uganda (sub-region 2) as shown by **Figure 4.23**. Significant inverse relationship with this index is obtained over northern and central Tanzania (sub-region 5) with number of dry days, duration of longest dry spells and mean frequency of dry spells of 5 days or more (**Figure 4.24**).

The persistence of this signal over the eastern Africa from July-August through to October-December (**Figures 4.37a–c**) implies that the moisture that is locally available is retained. The index also shows a persistent positive correlation with the SST over the northern and equatorial Indian Ocean (**Figures 4.38a–c**). The warming of the equatorial Indian Ocean SST, the advection of moisture from the Indian Ocean by the easterlies coupled with the retention of the locally available moisture results in wet conditions over East Africa and hence the positive (negative) association of this index with SRISS of wet (dry) spells.

The strong positive association of this index with SST over the tropical eastern Pacific Ocean (**Figures 4.38a–c**) is confirmed by the strong and significant positive correlation coefficients with the Niño indices (**Table 4.20**).

Table 4.20: Correlation coefficients between equatorial Indian Ocean (EQIND) specific humidity index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	ZIND
Jul-Aug	**0.73**	**0.59**	0.32	**0.44**	**0.41**	0.27
OND	**0.62**	**0.46**	0.23	**0.33**	**0.38**	0.18

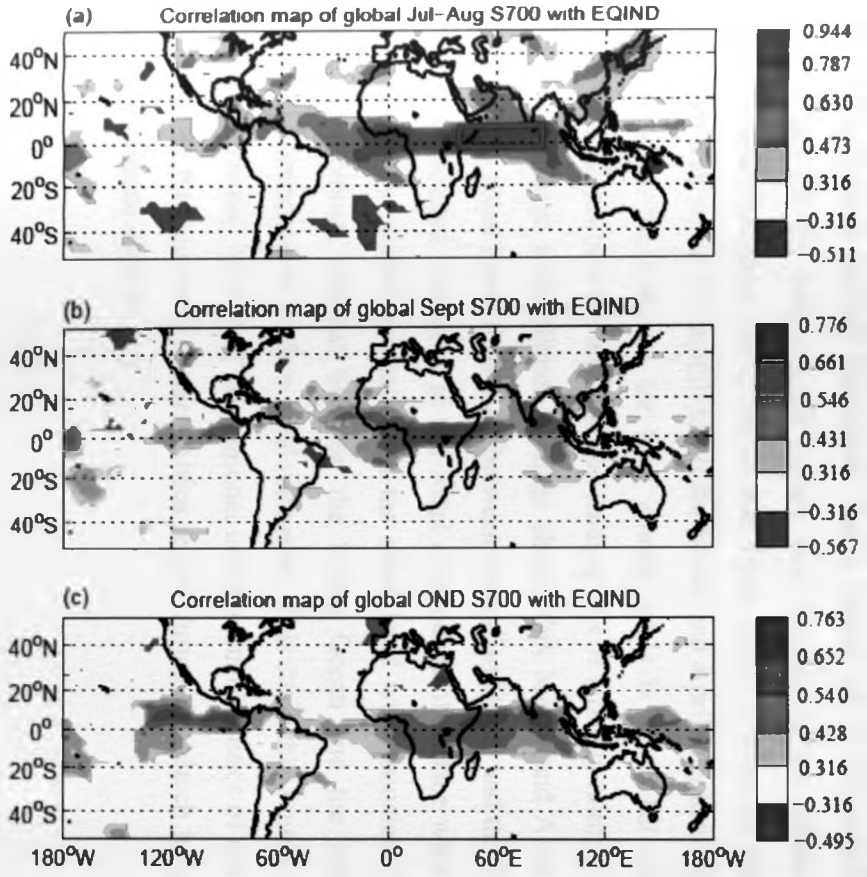


Figure 4.37: Map of significant correlation between equatorial Indian Ocean (EQIND) specific humidity index and global S700 for (a) July-August, (b) September and (c) October-December. The green rectangle in (a) shows the approximate location of EQIND specific humidity index computed for July-August period from 1962 to 2000

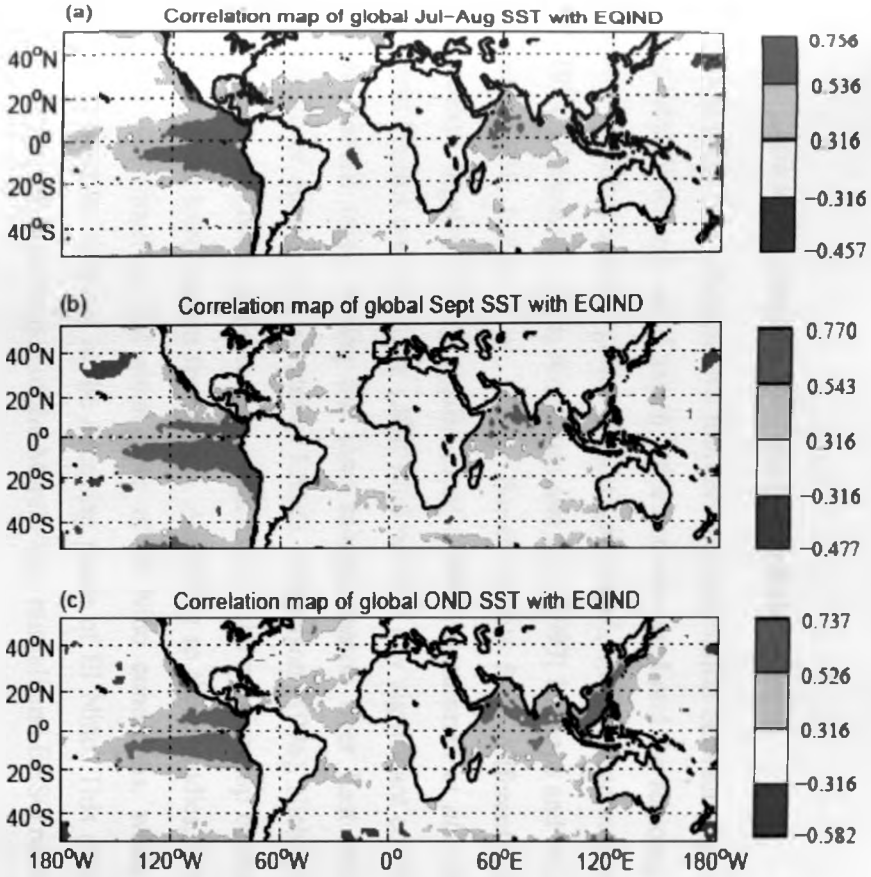


Figure 4.38: Map of significant correlation between equatorial Indian Ocean (EQIND) specific humidity index and global SST for (a) July-August, (b) September and (c) October-December

4.5.2 Linkages during the long rainfall season

Different studies have documented inconsistent association between the rainfall totals during the long rainfall season over Eastern Africa on one hand and the regional and global climate signals on the other hand (Ogallo, 1988; Ogallo *et al.*, 1988; Hastenrath *et al.*, 1993; Rowell *et al.*, 1994; Nicholson, 1996; Nicholson and Kim, 1997; Philipps and McIntyre, 2000; Mutai and Ward, 2000; Indeje *et al.*, 2000). Based on data from different periods and different spatial scales, Ogallo (1988), Ogallo *et al.* (1988), Hastenrath *et al.* (1993), Rowell *et al.* (1994) and Philips and McIntyre (2000) did not find any significant correlations between the seasonal rainfall totals during the long rainfall season over East Africa and either the atmospheric or oceanic component of ENSO. Nicholson (1996), Nicholson and Kim (1997) and Indeje *et al.* (2000) indicate that shifts exist in the relationship between Niño3 SST and seasonal rainfall totals across the season. According to these studies, weak positive rainfall anomalies are found on the onset year of El Niño conditions, while more pronounced negative anomalies develop in the decaying phase of El Niño. This shows that uncertainty still remains in the significance of the March–May rainfall and ENSO relationship.

There is however a general consensus that a month by month analysis provides better understanding on the long rainfall season over Eastern Africa (Mutai and Ward, 2000; Camberlin and Philippon, 2002; Zorita and Tilya, 2002). Camberlin and Philippon (2002) have shown that while the rainfall totals for March and April may have the same response to El Niño events, the response for May rainfall totals is somewhat different. Concurrent correlation analysis between the first two leading PCs of monthly rainfall totals over northern Tanzania and large scale climate forcings shown that March and April rainfall anomalies were linked to zonal thermal contrast between the Indian Ocean and the Eastern African land mass and associated anomalies in the zonal component of surface wind (Zorita and Tilya, 2002). The May rainfall anomalies on the other hand were associated with a meridional surface temperature contrast between the Indian Ocean and the Asian continent, and meridional component of surface wind anomalies. This study thus reassessed the relationship between the rainfall totals and sub-regional intraseasonal statistics of the wet and dry spells (SRISS) over the equatorial eastern Africa region during the March to May season and the large scale climate signals.

Based on these facts and the low spatial coherence of the seasonal rainfall totals and SRISS that was observed in the earlier parts of this study, the long rainfall season was split into two parts namely the March-April part and the May part. The linkages between the predefined and potential (both oceanic and atmospheric fields) on one hand and the rainfall totals and SRISS on the other hand were analysed separately for the two parts of the long rainfall season.

4.5.2.1 Linkages during the March-April period

As indicated in **section 4.5.1.1** earlier, all predictors are averaged for two-months with a lead time of one month to allow for the updating of the predictors identified before the start of the season. This means that the December-January values were averaged to obtain the predictor index for the March-April part of the season.

4.5.2.1.1 Linkages with the predefined SST predictors

The total correlation between some of the predefined predictor indices and the SRISS and rainfall totals during the March and April period of the long rainfall season are illustrated by **Figures 4.39a–j**. In general, the predefined predictors (including the ENSO indices) do not have significant association with rainfall totals and SRISS during the March-April period. However, occasionally some predefined predictors surpass the 95% significance threshold. Two meridional gradients are the most frequent in surpassing the threshold (**Figures 4.39a, b, d, i** and **j**). These are MIB1 (a meridional gradient across the Indian Ocean) and MAB3 (a meridional gradient across the Atlantic Ocean) as discussed by Nyakwada (2009). Their relationship with East Africa also remains tenuous, which justifies the search for additional potential predictors which is discussed in the next section.

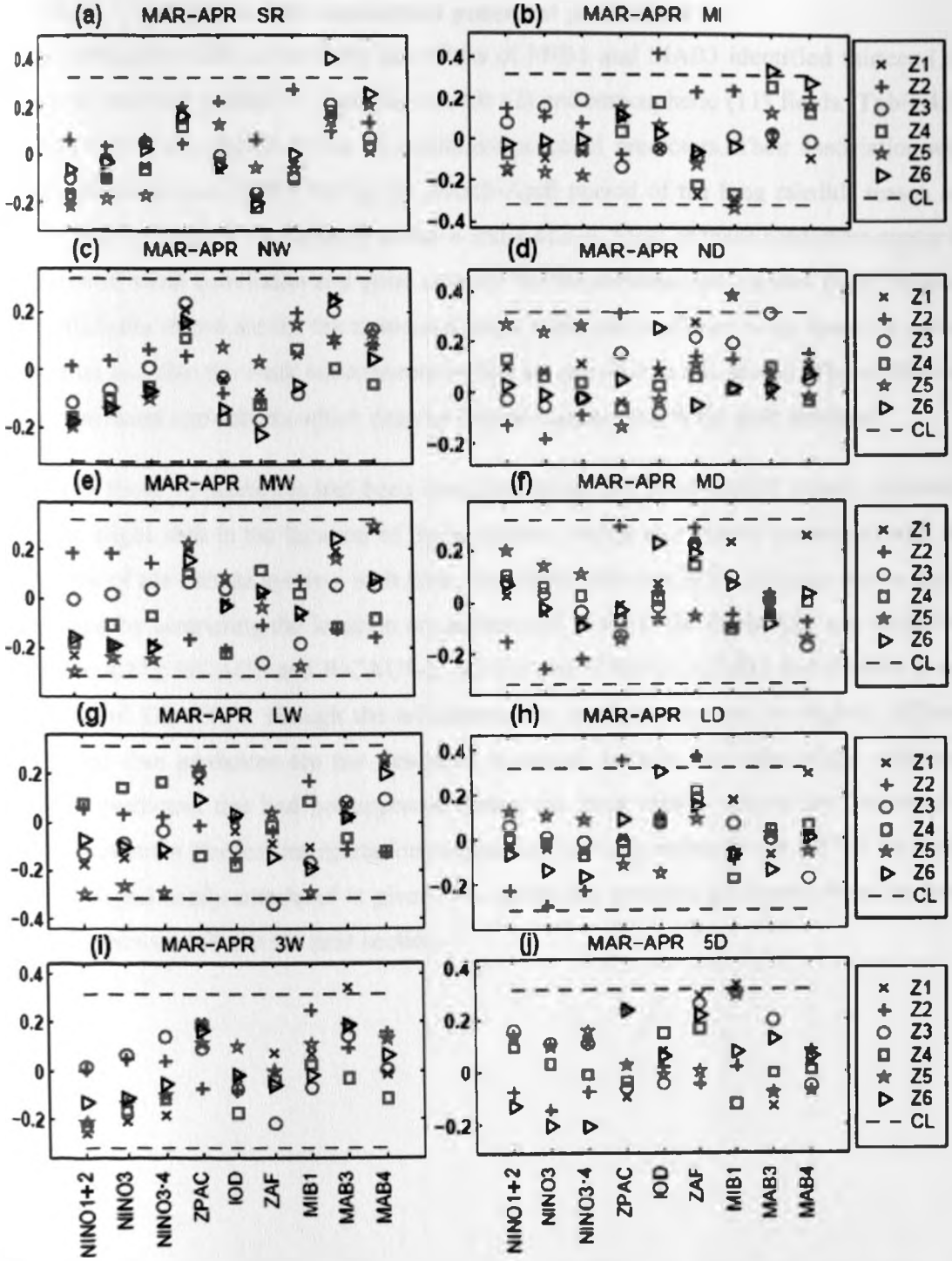


Figure 4.39: Correlation coefficient between predefined predictors averaged for December-January period (x-axis) and the areal-averaged March-April for (a) rainfall totals, (b) mean rainfall intensity, (c) number of wet days, (d) number of dry days, (e) mean length of wet spells, (f) mean length of dry spells, (g) longest wet spell, (h) longest dry spell, (i) frequency of 3 wet days or more, and (j) frequency of 5 dry days or more, over the six rainfall sub-regions Z1 to Z6. CL shows the 95% confidence level threshold

4.5.2.1.2 Linkages with additional potential predictors

Partial correlation while controlling the effects of MIB1 and MAB3 identified thirteen (13) additional potential predictors from the oceanic (2) and atmospheric (11) fields. **Table 4.21** provides a brief description of the 13 additional potential predictors. Their association with the rainfall totals and SRISS during the March-April period of the long rainfall season are summarised by **Table 4.22, Figures 4.40a–e** and **4.41a–e**. Most of these predictors appear to have general weak correlation and quite unstable for the different sub-regions partly because the coefficients shown are for the total correlation while the predictors were based on partial correlation and also the weak teleconnections that are peculiar to this season. There still exist some significant correlations which deserve further examination in the next sections.

Several of these 13 predictors had been identified during the short rainfall season. However, there are slight shift in the location of the predictors, which was mainly associated with the evolutions of the climate systems with time. The slight variation in the location can be easily recognized by comparing the location co-ordinates of BoBEN (in **Table 4.9**) and BoBEN-1 (in **Table 4.21**); WCAUS and WCAUS-1; SINDS and SINDS-1; SINDS and SINDS-2; and EQAFR and EQAFR-1. Though the teleconnections mechanisms may be slightly different also, these five predictors are not described in details further. The other eight additional potential predictors that had not appeared during the short rainfall season are discussed in details next and a physical interpretation on how they possibly influence the SRISS for which they are significantly correlated is given. The additional potential predictors from the wind field are discussed first in the next section.

4.5.2.1.2 Linkages with additional potential predictors

Partial correlation while controlling the effects of MIB1 and MAB3 identified thirteen (13) additional potential predictors from the oceanic (2) and atmospheric (11) fields. **Table 4.21** provides a brief description of the 13 additional potential predictors. Their association with the rainfall totals and SRISS during the March-April period of the long rainfall season are summarised by **Table 4.22, Figures 4.40a–e** and **4.41a–e**. Most of these predictors appear to have general weak correlation and quite unstable for the different sub-regions partly because the coefficients shown are for the total correlation while the predictors were based on partial correlation and also the weak teleconnections that are peculiar to this season. There still exist some significant correlations which deserve further examination in the next sections.

Several of these 13 predictors had been identified during the short rainfall season. However, there are slight shift in the location of the predictors, which was mainly associated with the evolutions of the climate systems with time. The slight variation in the location can be easily recognized by comparing the location co-ordinates of BoBEN (in **Table 4.9**) and BoBEN-1 (in **Table 4.21**); WCAUS and WCAUS-1; SINDS and SINDS-1; SINDS and SINDS-2; and EQAFR and EQAFR-1. Though the teleconnections mechanisms may be slightly different also, these five predictors are not described in details further. The other eight additional potential predictors that had not appeared during the short rainfall season are discussed in details next and a physical interpretation on how they possibly influence the SRISS for which they are significantly correlated is given. The additional potential predictors from the wind field are discussed first in the next section.

Table 4.21: Brief description of the additional potential predictors for March-April period of long rainfall season and their location details

Index name	Description	Location details (°)	
		Longitude	Latitude
BoBEN-1	A slight location variation of BoBEN (Bay of Bengal) SST index	68 – 73 E	12 – 17 N
WCAUS-1	A slight location variation of WCAUS (West coast of Australia) SST index	106 – 118 E	20 – 12 S
SINDS-1	A slight location variation of SINDS (southern tip of Indian sub-continent) zonal wind index at 925mb level	55 – 80 E	10 – 5 S
ANGCO	Zonal wind index at 925mb level located over Angola and its coast	10 – 20 E	25 – 20 S
WAFR	Zonal wind index at 925mb level from Atlantic Ocean into western Africa	35 – 15 W	15 – 25 N
SINDS-2	A slight location variation of SINDS (southern tip of Indian sub-continent) zonal wind index at 925mb level	55 – 65 E	5 – 10 N
CINDO	Zonal wind index at 700mb level over equatorial central Indian Ocean	70 – 80 E	2.5 S – 2.5 N
SCEINDO	Zonal wind index at 700mb level, south of central equatorial Indian Ocean	80 – 105 E	17.5 – 12.5 S
NINDS	Zonal wind index at 200mb level over northern India subcontinent	80 – 90 E	20 – 30 N
EQAFR-1	A slight location variation of EQAFR (Equatorial Africa) zonal wind index at 200mb level	10 – 20 E	10 – 5 S
NEGHA	Meridional wind index at 925mb level over north-eastern parts of Greater Horn of Africa in eastern Sudan, northern Ethiopia and parts of Djibouti	35 – 45 E	5 – 15 N
WINDO	Meridional wind index at 925mb level over equatorial western Indian Ocean and equatorial Africa	50 – 60 E	2.5 S – 2.5 N
EBBEN	Specific humidity index at 925mb level over southern Asia slightly to the east of Bay of Bengal	95 – 105 E	20 – 25 N

Table 4.22: A summary of the association between the identified additional potential predictors and the sub-regional intraseasonal statistics of wet and dry spells for the March-April period of the long rainfall season and the most strongly correlated intraseasonal statistic and sub-region

Predictor	Atmospheric Level	Index Name	Number of SRISS associated with the predictor (out of 10)	Strongest total correlation		
				SRISS	Sub-region	Coefficient
SST	Surface	BoBEN-1	1	MI	2	0.36
		WCAUS-1	2	MI	1	***-0.45***
u-wind	925mb	SINDS-1	8	MW	4	***-0.58***
		ANGCO	5	SR	3	***-0.39***
		WAFR	7	3W	2	***-0.50***
		SINDS-2	3	5D	6	0.42
	700mb	CINDO	6	NW	6	0.45
		SCEINDO	4	MD	3	***-0.47***
	200mb	NINDS	4	MI	2	0.42
		EQAFR-1	7	3W	5	0.52
v-wind	925mb	NEGHA	5	MD	3	***-0.44***
		WINDO	2	MD	1	0.39
Specific humidity	925mb	EBBEN	6	MW	5	***-0.48***

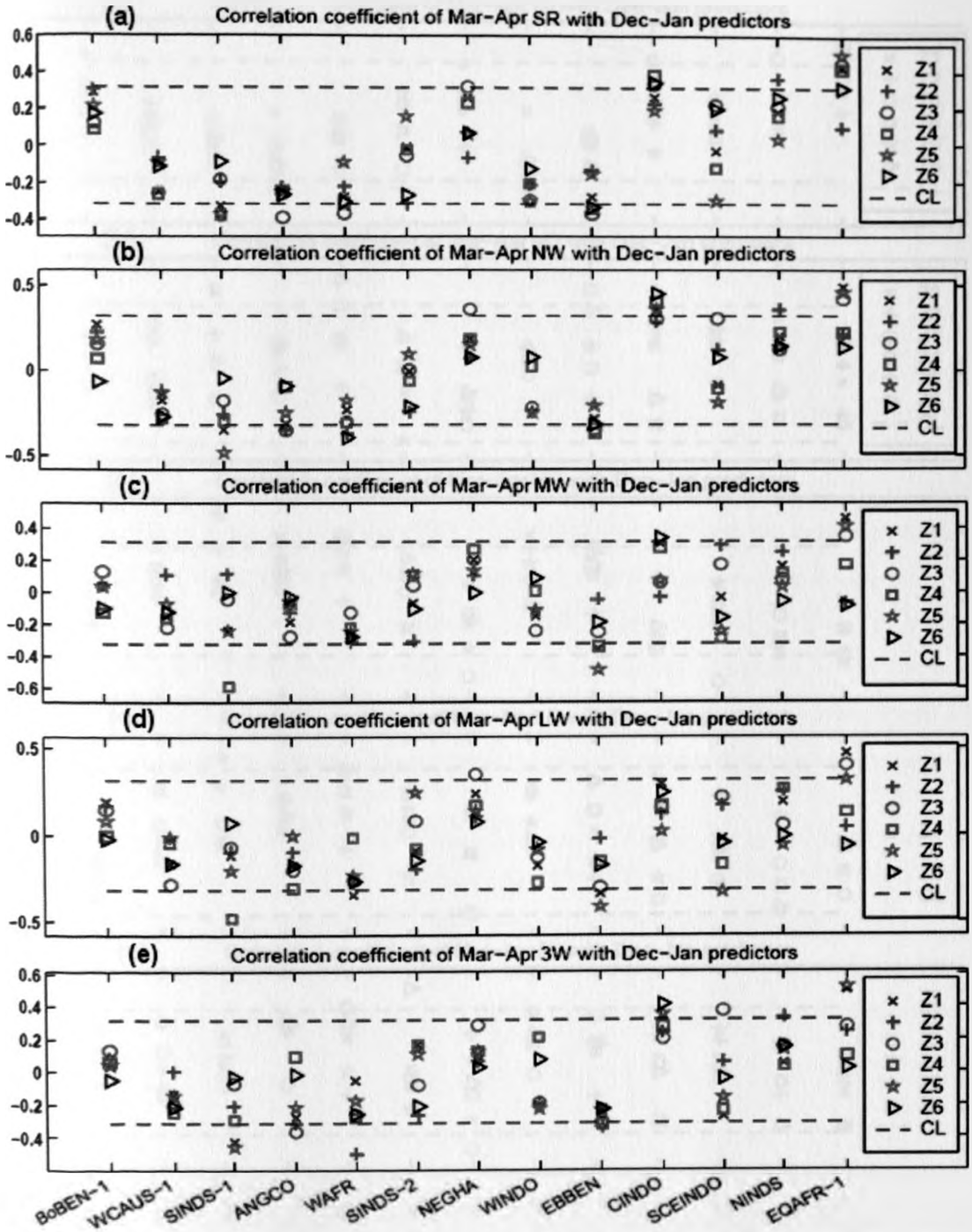


Figure 4.40: Correlation coefficient between the thirteen additional potential predictors identified averaged over December-January period and the areal-averaged March-April (a) rainfall totals, (b) number of wet days, (c) mean length of wet spell, (d) longest wet spell and (e) frequency of 3 wet days or more, over the six rainfall sub-regions Z1 to Z6. CL shows the 95% confidence level threshold

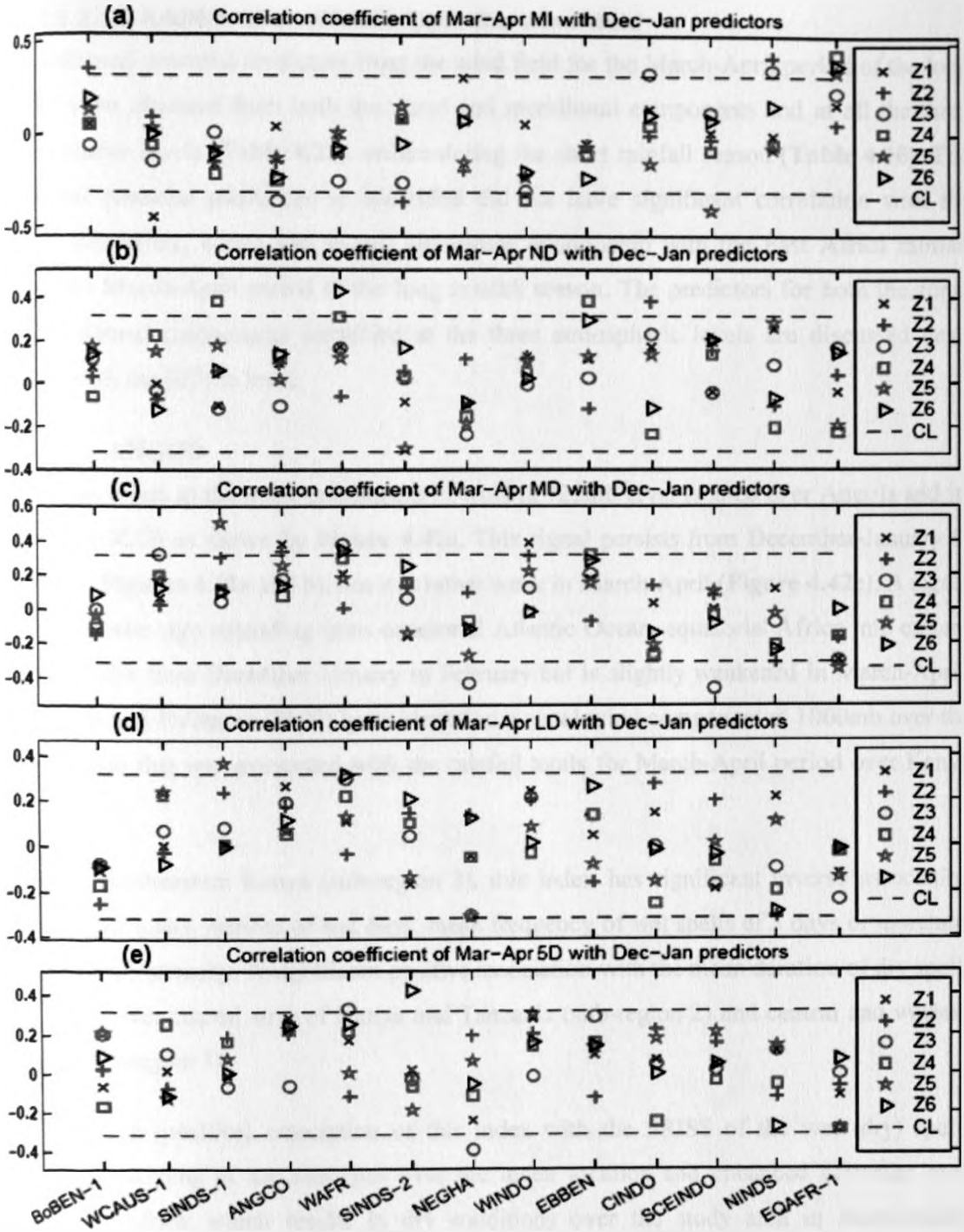


Figure 4.41: Correlation coefficient between the thirteen additional potential predictors identified averaged over December-January period and the areal-averaged March-April (a) mean rainfall intensity, (b) number of dry days, (c) mean length of dry spell, (d) longest dry spell and (e) frequency of 5 dry days or more, over the six rainfall sub-regions Z1 to Z6. CL shows the 95% confidence level threshold

4.5.2.1.2.1 Additional predictors from the wind field

The additional potential predictors from the wind field for the March-April period of the long rainfall were obtained from both the zonal and meridional components and at all the three representative levels (**Table 4.22**), unlike during the short rainfall season (**Table 4.10**). The additional potential predictors so identified did not have significant correlation with the MAB3 and MIB1, which had shown significant relationship with the East Africa rainfall during the March-April period of the long rainfall season. The predictors for both the zonal and meridional components identified at the three atmospheric levels are discussed next, starting with the 925mb level.

(a) ANGCO

This index refers to the zonal component of wind at 925mb level located over Angola and its coast (ANGCO) as shown by **Figure 4.42a.** This signal persists from December-January to February (**Figures 4.42a** and **b**), but it is rather weak in March-April (**Figure 4.42c**). A signal of the opposite sign extending from equatorial Atlantic Ocean, equatorial Africa into eastern Africa persist from December-January to February but is slightly weakened in March-April. Camberlin and Philippon (2002) have identified a zonal wind component at 1000mb over the Congo basin that was associated with the rainfall totals for March-April period over Kenya and Uganda.

Over the northeastern Kenya (sub-region 3), this index has significant inverse association with rainfall totals, number of wet days, mean frequency of wet spells of 3 days or more and mean rainfall intensity. A significant positive association with the mean duration of dry spells was noted over coastal strip of Kenya and Tanzania (sub-region 2) and central and western Kenya (sub-region 1).

The negative (positive) association of this index with the SRISS of the wet (dry) spells implies weakening of the easterlies over the index location and enhanced easterlies over equatorial Africa, which results in dry conditions over the study area in March-April. Correlation analysis with global SST (not shown) shows that this index is associated with SST cooling over the central Indian Ocean and the Niño regions in December-January. In February and March-April, the cooling is spread over the western Indian Ocean. The index is associated with moisture reduction over the equatorial Indian Ocean and equatorial Africa in December-January and to some extent in March-April. The weakening of the easterlies, the cooling of the SST over Indian Ocean and Niño regions coupled with the moisture reduction

over the Equatorial Indian Ocean and equatorial Africa result in dry conditions over the study area. Hence SRISS of dry spells (wet spells and rainfall totals) during the March-April period of long rainfall season have positive (negative) association with this index.

The plausible physical explanation on how this index relates to East Africa rainfall provides a strong case for its retention as an additional potential predictor.

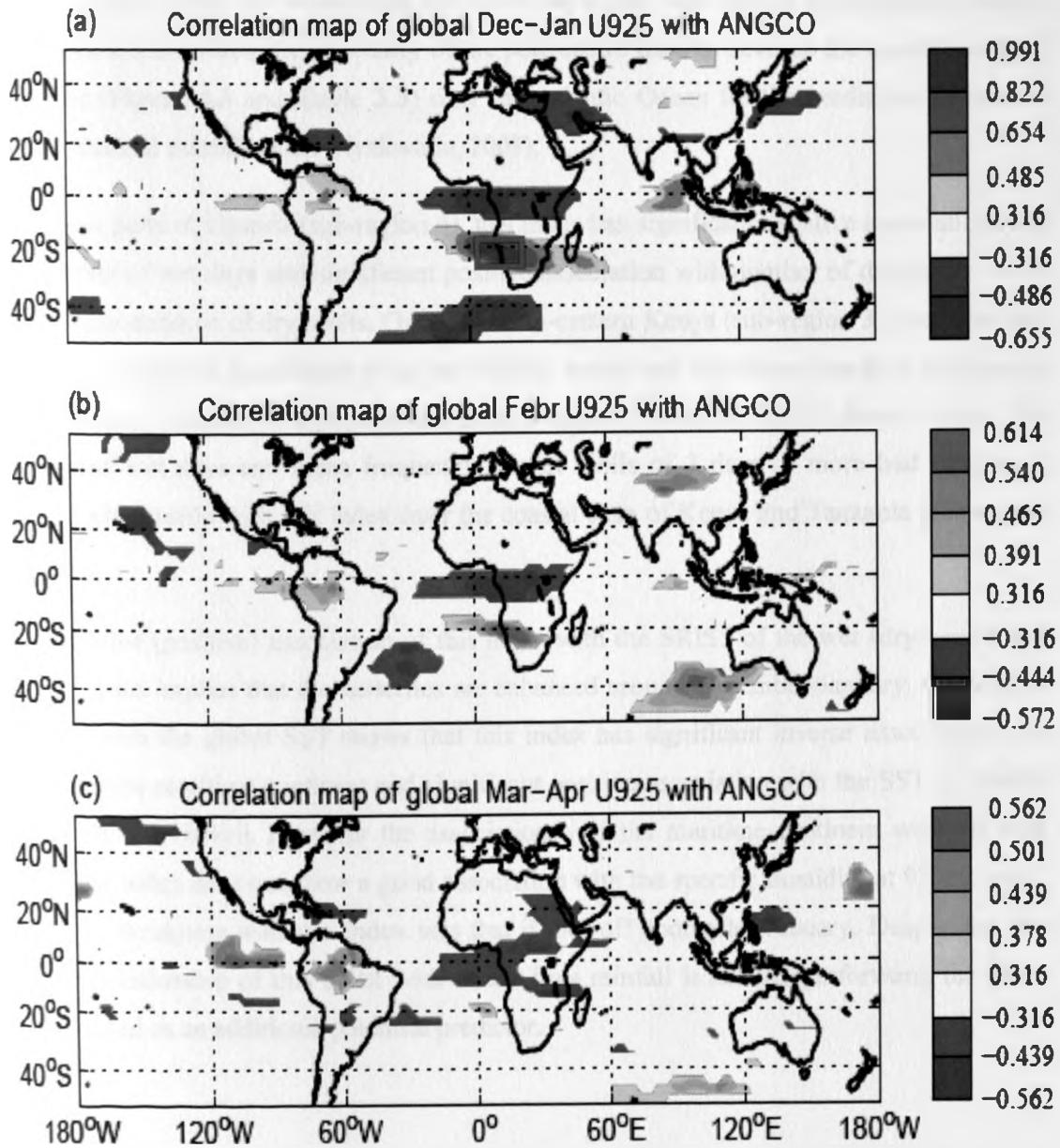


Figure 4.42: Map of significant correlation between Angola and its coast (ANGCO) zonal wind index and global U925 for (a) December-January, (b) February and (c) March-April. The green rectangle in (a) shows the approximate location of ANGCO zonal wind index computed for December-January period from 1961 to 2000

(b) WAFR

This index refers to the zonal component of wind at 925mb level extending from the Atlantic Ocean into western Africa (WAFR) around 20°N as shown by **Figure 4.43a**. However the local signature of this wind signal dies off in December-January (**Figures 4.43a–c**). Henne *et al.* (2008) have documented six flow regimes towards East Africa. One such regime which may be associated with WAFR index is the North Africa free tropospheric flow observed from January to May and accounting for 6% of the totals flow regime observations studied. The index is also within the proximity of the pole centre used to develop the meridional SST gradients (**Figure 4.3** and **Table 3.3**) over the Atlantic Ocean for the prediction of eastern Africa seasonal rainfall totals (Nyakwada, 2009).

Over most parts of Uganda (sub-region 6), this index has significant negative association with the number of wet days and significant positive association with number of dry days as well as the mean duration of dry spells. Over the north-eastern Kenya (sub-region 3), the index has significant negative association with the rainfall totals and significant positive association with the mean duration of dry spells and mean frequency of dry spells of 5 days or more. The number of wet days and mean frequency of wet spells of 3 days or more had significant inverse relationship with this index over the coastal strip of Kenya and Tanzania (sub-region 2).

The negative (positive) association of this index with the SRISS of the wet (dry) spells and rainfall totals implies that the easterlies are enhanced around December-January. Correlation analysis with the global SST shows that this index has significant inverse association with SST over the maritime continent and significant positive association with the SST on Pacific Ocean around Hawaii. However the association over the maritime continent weakens with time. This index does not show a good association with the specific humidity at 925mb level. The main weakness with this index was that it dies off soon after January. Despite the fact that the relationship of this index with East Africa rainfall is not straightforward, the index was retained as an additional potential predictor.

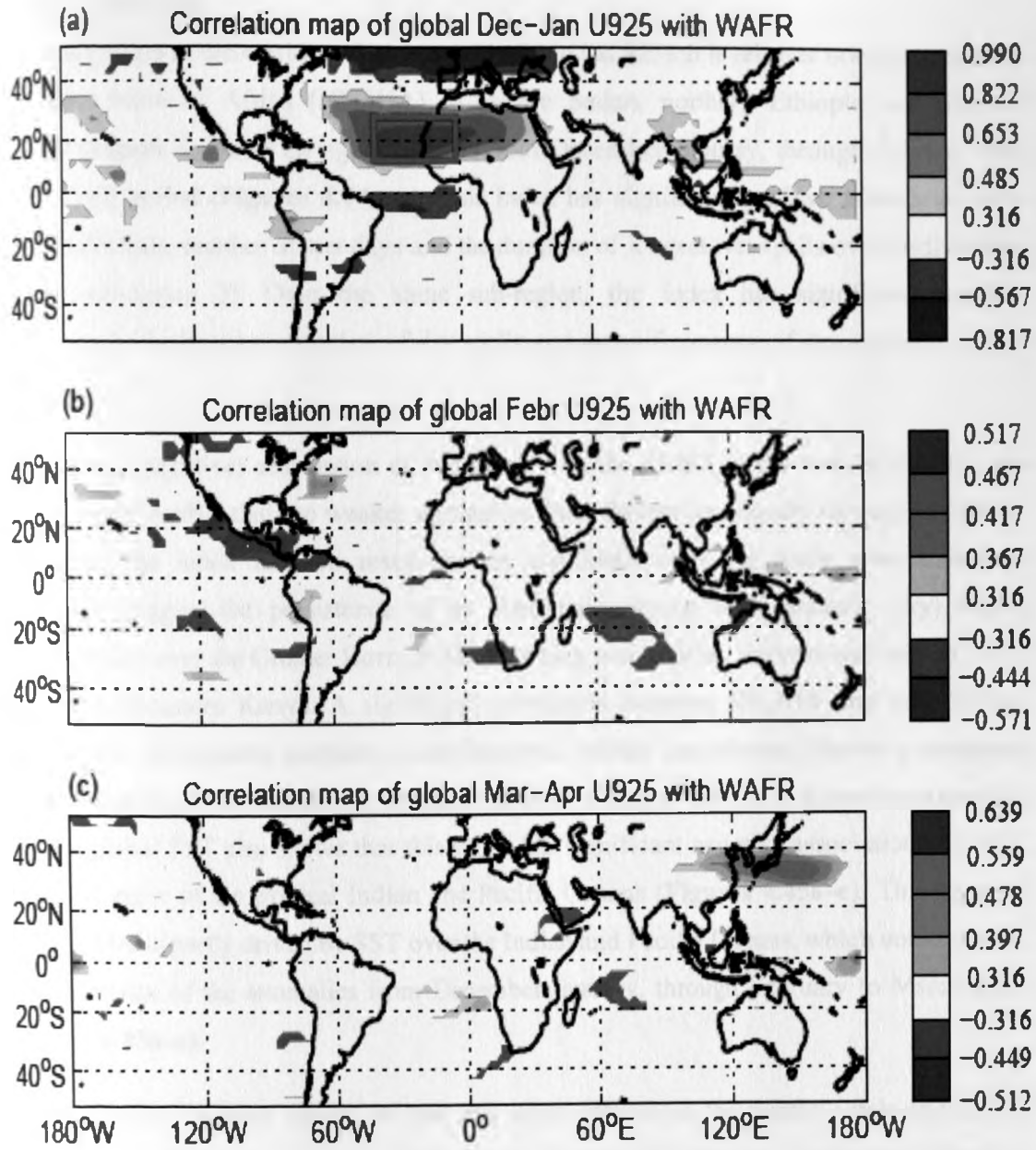


Figure 4.43: Map of significant correlation between western Africa (WAFR) zonal wind index and global U925 for (a) December-January, (b) February and (c) March-April. The green rectangle in (a) shows the approximate location of WAFR zonal wind index computed for December-January period from 1961 to 2000

(c) NEGHA

This index refers to the meridional component of wind at 925mb level over northeastern parts of Greater Horn of Africa (NEGHA) in eastern Sudan, northern Ethiopia and parts of Djibouti (**Figure 4.44a**). The signal persists from December-January, through February into March-April period (**Figures 4.44a–c**). This index has significant positive association with the rainfall totals, number of wet days and the duration of longest wet spells over northeastern Kenya (sub-region 3). Over the same sub-region, the index has significant negative association with the mean duration of dry spells and mean frequency of dry spells of 5 days or more.

The positive (negative) association of this index with the SRISS of the wet (dry) spells and rainfall totals implies that the weaker northerlies from December-January through to March-April over the index location result in wet conditions over the study area. Enhanced northerlies suggest the persistence of an abnormally strong north-easterly (dry) winter monsoon flow over the Greater Horn of Africa, which would delay the seasonal shift of ITCZ towards northeastern Kenya. A significant correlation between NEGHA and both 925mb zonal winds and specific humidity over equatorial Africa (not shown) denote a consistent pattern involving variations in the location / northern extent of the ITCZ. Correlation analysis with the global SST also shows that this index has significant negative association with SST over most parts of the tropical Indian and Pacific Oceans (**Figures 4.45a–c**). This suggests that NEGHA is partly driven by SST over the Indian and Pacific Oceans, which could explain the persistence of the anomalies from December-January, through February to March-April (**Figures 4.45a–c**).

The hypothesis provided earlier on how this index influences the rainfall totals and SRISS during the March-April period, coupled with the fact that this index is significantly related to the oceanic field (SST) and other atmospheric variables provides a strong case for its retention as an additional potential predictor.

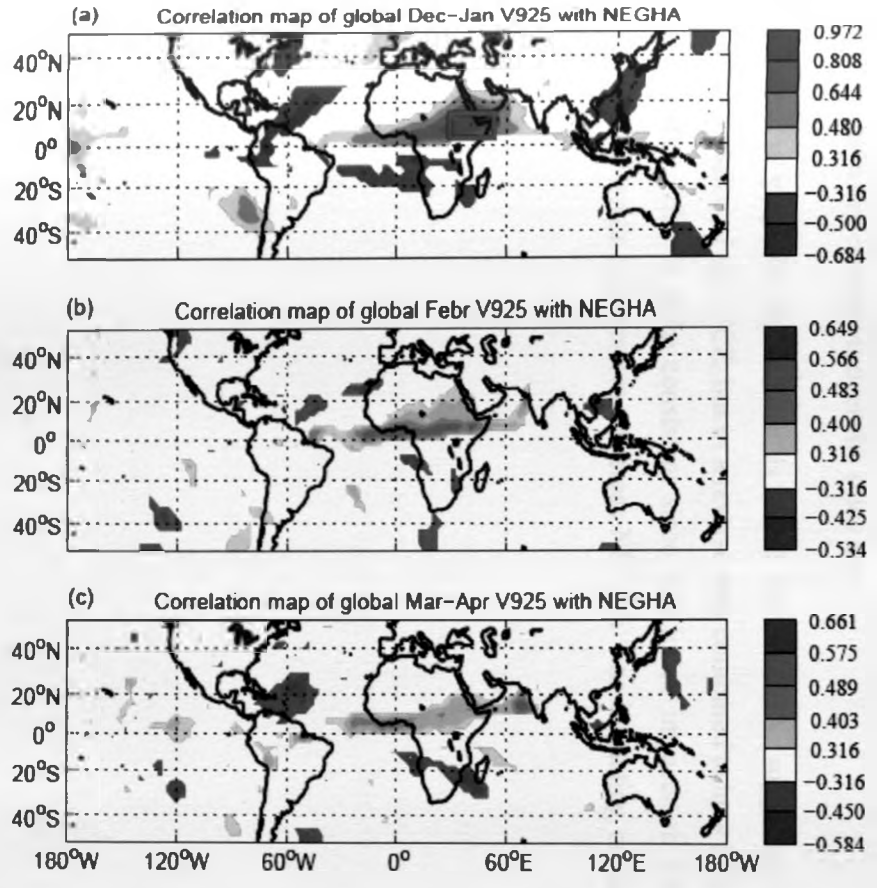


Figure 4.44: Map of significant correlation between northeastern parts of Greater Horn of Africa (NEGHA) meridional wind index and global V925 for (a) December-January, (b) February and (c) March-April. The green rectangle in (a) shows the approximate location of NEGHA meridional wind index computed for December-January period from 1961 to 2000

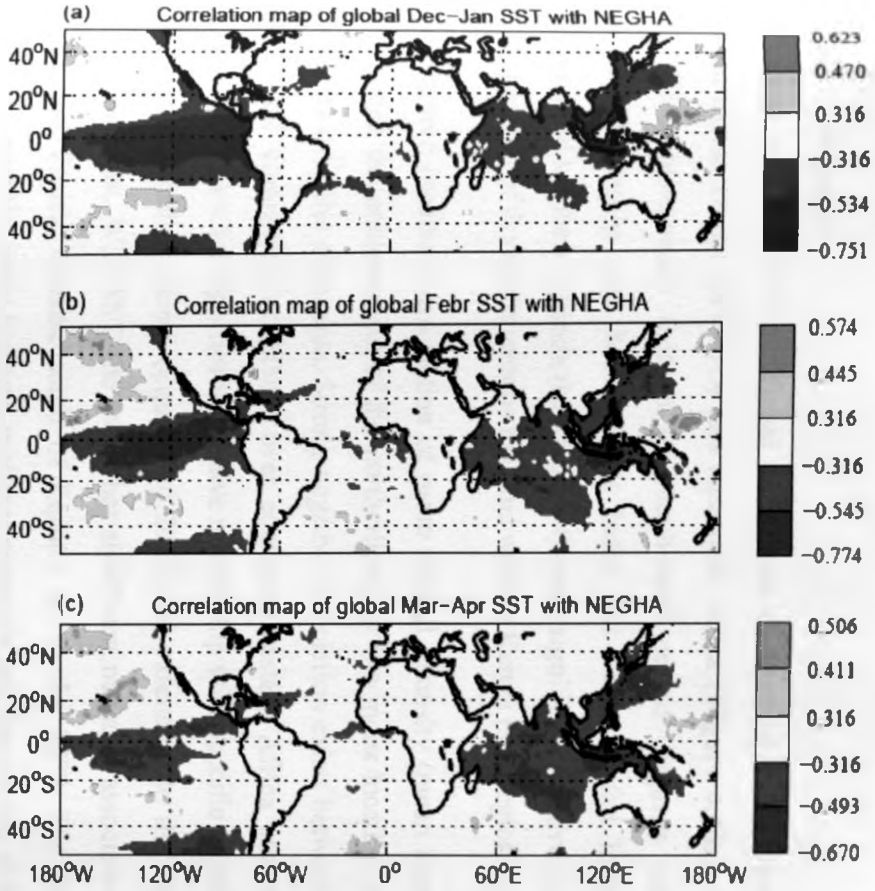


Figure 4.45: Map of significant correlation between northeastern parts of Greater Horn of Africa (NEGHA) meridional wind index and global SST for (a) December-January, (b) February and (c) March-April

(d) **WINDO**

This index refers to the meridional component of wind at 925mb level over equatorial western Indian Ocean (WINDO) and equatorial Africa (**Figure 4.46a**). This signal persists from December-January through to March-April period (**Figures 4.46a–c**). Significant inverse relationship exist between this index and mean rainfall intensity over western Tanzania and southern Uganda (sub-region 4) while significant positive association exist with mean duration of dry spells over central and western Kenya (sub-region 1).

The negative (positive) association of mean rainfall intensity (mean duration of dry spells) implies that the enhancement of the northerlies over the index location which results in dry conditions over the study area. Strong negative association exists between this index and the zonal wind component at 925mb over equatorial eastern Africa and the adjacent Indian Ocean. This index has significant negative relationship with specific humidity at 925mb level over southern Indian Ocean (Equator to 20°S) and extending to the adjacent parts of the Africa continent. With SST, the index has significant negative association over the tropical Indian and Pacific Oceans, south of Equator up to the 20°S. The enhancement of the northerlies over the index location, moisture reduction and the cooling of the SST over Indian Ocean results in dry conditions over the study area.

This index is very similar to what is found for NEGHA, suggesting variations in the latitudinal location of the ITCZ, but in a reverse way compared to how NEGHA impacts East Africa rainfall. This is quite consistent since this index impacts on sub-regions that are located further south and western, whereas NEGHA was impacting on northeastern Kenya rainfall.

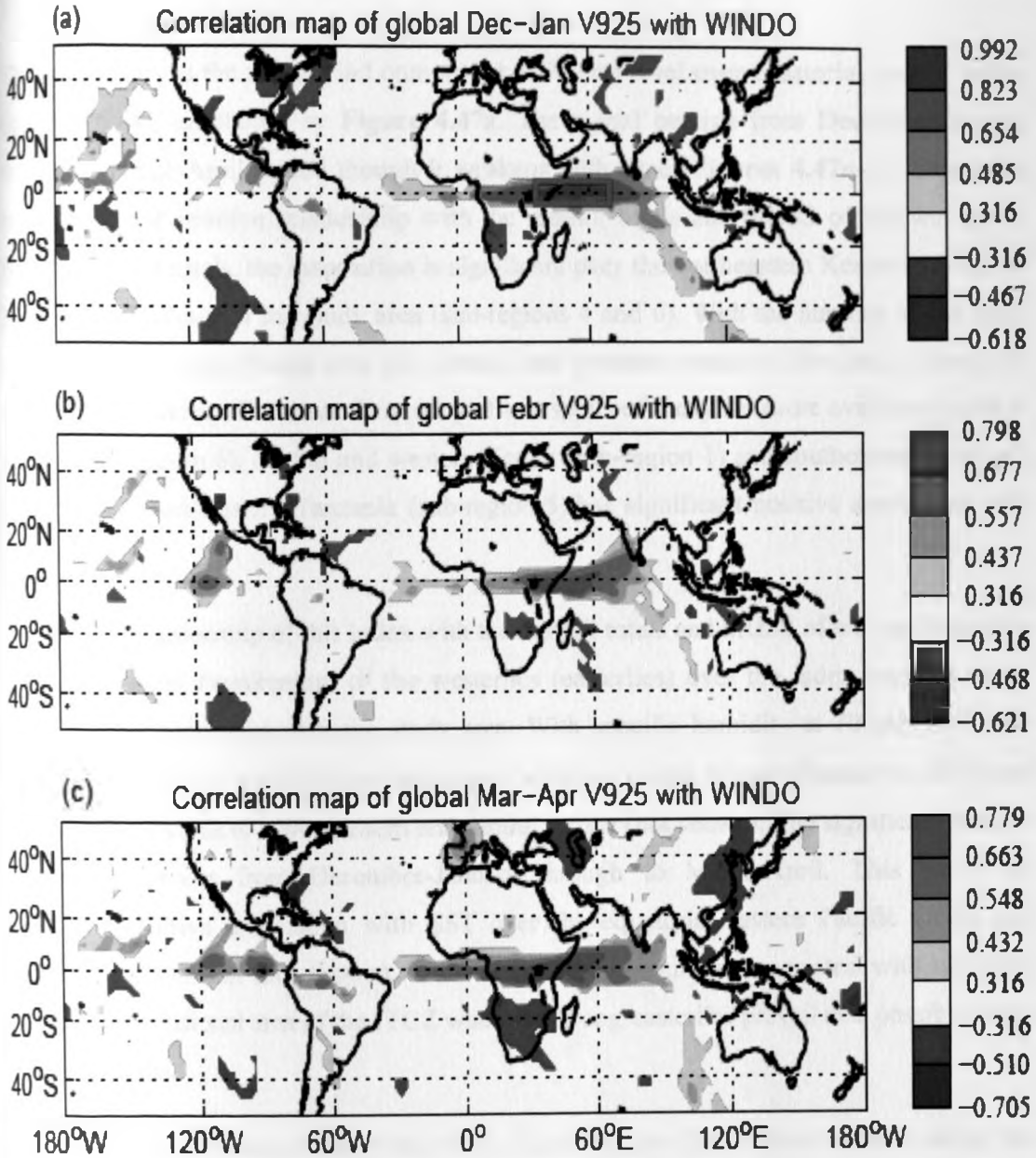


Figure 4.46: Map of significant correlation between equatorial western Indian Ocean (WINDO) meridional wind index and global V925 for (a) December-January, (b) February and (c) March-April. The green rectangle in (a) shows the approximate location of WINDO meridional wind index computed for December-January period from 1961 to 2000

(e) CINDO

This index refers to the zonal wind component at 700mb level over equatorial central Indian Ocean (CINDO) as shown by **Figure 4.47a**. The signal persists from December-January through to March-April period though it weakens with time (**Figures 4.47a–c**). This index has a significant positive relationship with the rainfall totals and SRISS of the wet spells. With the rainfall totals, the association is significant over the northeastern Kenya (sub-region 3) and western sector of the study area (sub-regions 4 and 6). With the number of wet days, the association is significant over the central and western blocks of the study area (sub-regions 1, 4, 5 and 6). The mean frequency of wet spells of 3 days or more over most parts of Uganda (sub-region 6), central and western Kenya (sub-region 1) and southeastern lowlands of Kenya and northeastern Tanzania (sub-region 5) has significant positive association with this index.

The positive relationship of this index with the rainfall totals and SRISS of wet spells implies that enhancement (weakening) of the westerlies (easterlies) over the index location which results in wet conditions over the study area. With specific humidity at 700mb level, this index has significant positive association over southern Indian Ocean (Equator to 20°S) and extending westwards to cover eastern and central Africa (not shown). The significant positive association persists from December-January through to March-April. This index has significant positive association with SST over the equatorial eastern Pacific Ocean and western Indian Ocean (not shown). The wet conditions could be associated with westward shift of the meridional arm of the ITCZ when the strong easterlies prevail and possibly, more stable easterlies.

Though the explanation on how this index influences the East Africa rainfall during the March-April period is still not straightforward, the index was retained as an additional potential predictor.

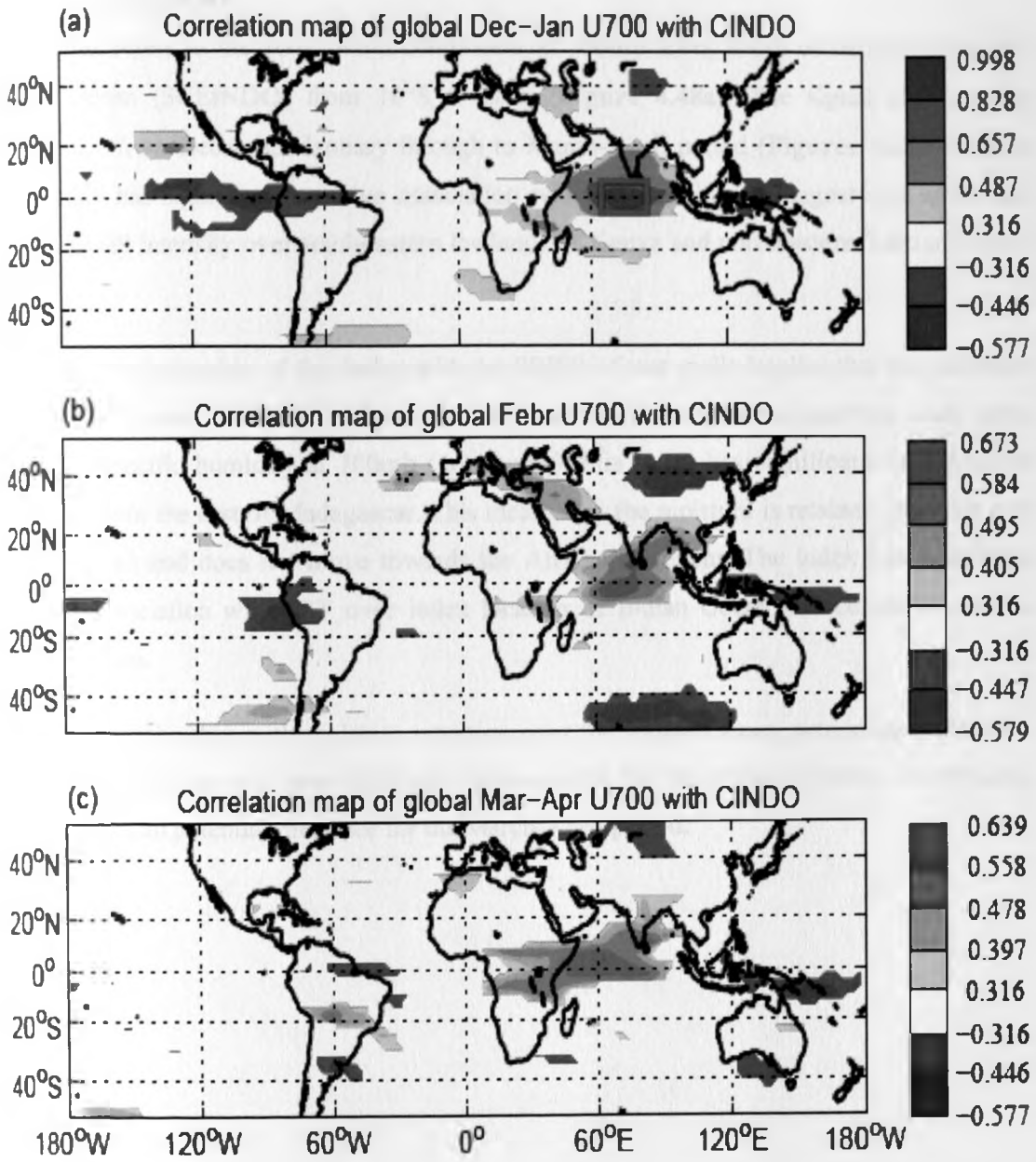


Figure 4.47: Map of significant correlation between equatorial central Indian Ocean (CINDO) zonal wind index and global U700 for (a) December-January, (b) February and (c) March-April. The green rectangle in (a) shows the approximate location of CINDO zonal wind index computed for December-January period from 1961 to 2000

(f) SCEINDO

This index refers to the zonal wind component at 700mb level, south of central equatorial Indian Ocean (SCEINDO) from 10°S to 20°S (**Figure 4.48a**). The signal shows weak persistence from December-January through to March-April period (**Figures 4.48a–c**). This index only has significant negative association with the duration of longest wet spells and mean rainfall intensity over southeastern lowlands of Kenya and northeastern Tanzania (sub-region 5).

The inverse relationship of this index with the SRISS of wet spells implies that the easterlies are weakened over the index location, which results in dry conditions over the study area. With the specific humidity at 700mb (not shown), this index has significant faint positive relationship to the east of Madagascar. This means that the moisture is retained over this part of the Ocean and does not move towards the African continent. The index has significant positive association with SST over index location in Indian Ocean and central equatorial Pacific Ocean.

Given its relationship with moisture retention over the Indian Ocean, providing a plausible physical explanation as to how this index influences the SRISS of the wet spells, it is retained as an additional potential predictor for the March-April period.

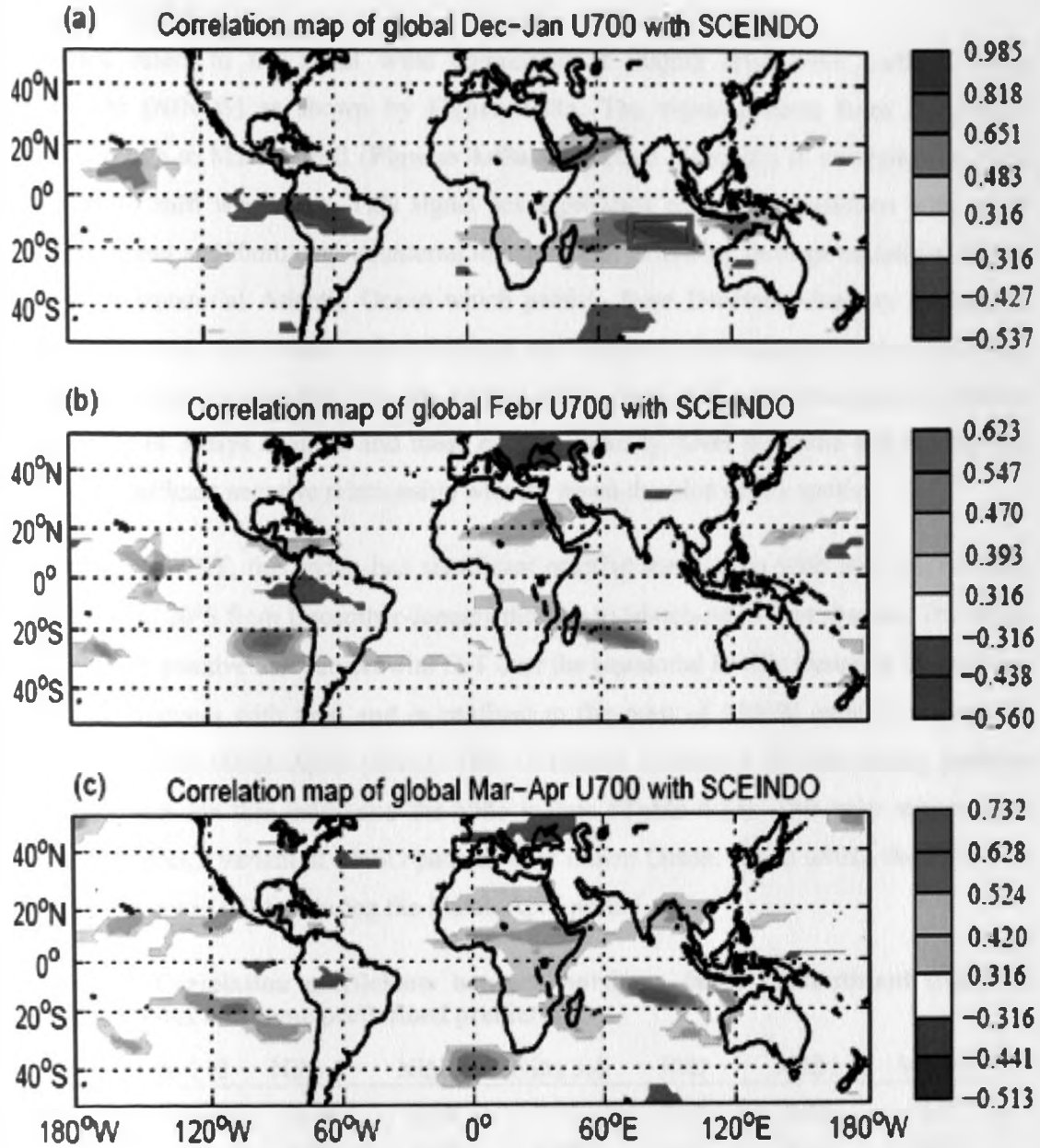


Figure 4.48: Map of significant correlation between south of central equatorial Indian Ocean (SCEINDO) zonal wind index and global U700 for (a) December-January, (b) February and (c) March-April. The green rectangle in (a) shows the approximate location of SCEINDO zonal wind index computed for December-January period from 1961 to 2000

(g) NINDS

This index refers to the zonal wind component at 200mb level over northern India subcontinent (NINDS) as shown by **Figure 4.49a**. The signal persists from December-January through to March-April (**Figures 4.49a–c**), though it reduces in strength with time and seems to shift westwards. This signal has significant negative association with zonal wind component at 200mb over equatorial Indian Ocean extending through equatorial Africa into eastern equatorial Atlantic Ocean which persists from December-January through to March-April. Over the coastal strip of Kenya and Tanzania (sub-region 2), the index has significant positive relationship with the rainfall totals, number of wet days, mean frequency of wet spells of 3 days or more and mean rainfall intensity. Over the same sub-region, this index has significant negative relationship with the mean duration of dry spells.

With the global SST, this index has significant positive association with SST over Indian Ocean north of 20°S from December-January through to March-April (not shown). The index has significant positive association with SST over the equatorial Pacific Ocean in December–January but reduces with time and is confined to the west of 120°W over the equatorial Pacific Ocean in March-April period. This is further confirmed by the strong positive association between this index and the Niño indices (**Table 4.23**). This may suggest that NINDS describes a variant of ENSO patterns over Indian Ocean, which unlike the ENSO, is a better predictor of rainfall during the March-April period.

Table 4.23: Correlation coefficients between northern India subcontinent (NINDS) zonal wind index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	MIB1	MAB3
Dec-Jan	**0.72**	**0.74**	**0.58**	**0.71**	0.26	**0.35**	-0.19
Mar-Apr	0.30	**0.40**	**0.40**	**0.50**	-0.14	0.13	0.03

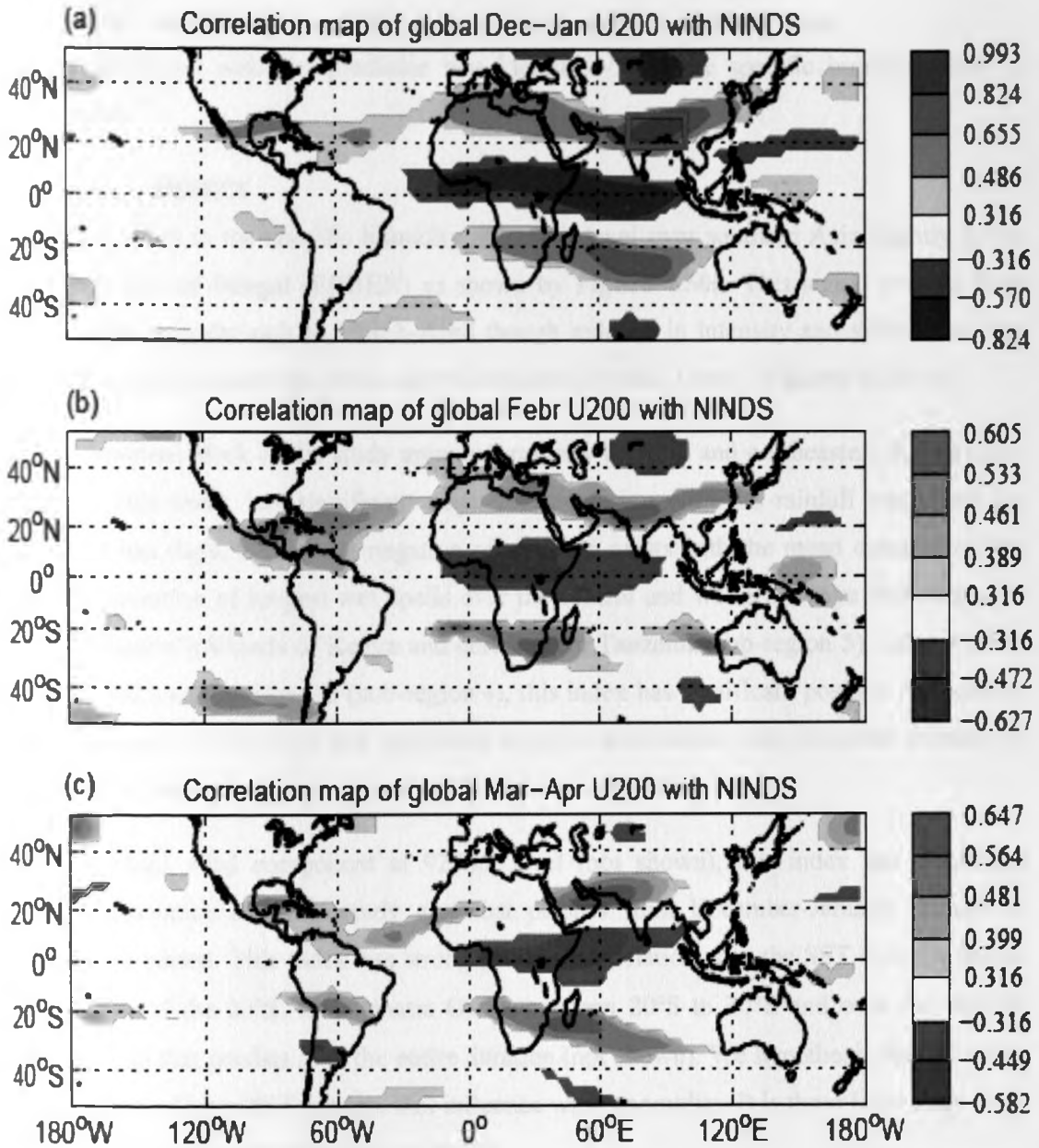


Figure 4.49: Map of significant correlation between northern India subcontinent (NINDS) zonal wind index and global U200 for (a) December-January, (b) February and (c) March-April. The green rectangle in (a) shows the approximate location of NINDS zonal wind index computed for December-January period from 1961 to 2000

4.5.2.1.2.2 Additional predictors from the specific humidity field

Another additional potential predictor was identified from the specific humidity field at 925mb level.

(a) EBBEN

This index refers to the specific humidity at 925mb level over southern Asia slightly to the east of the Bay of Bengal (EBBEN) as shown by **Figure 4.50a**. This signal persists from December-January through to March-April though reduced in intensity and shifts from Bay of Bengal towards equatorial Africa and the equatorial Indian Ocean (**Figures 4.50a–c**).

Over the western block of the study area (sub-regions 4 and 6) and northeastern Kenya (sub-region 3), this index has significant negative association with the rainfall totals and the number of wet days. Significant negative relationship exists with the mean duration of wet spells and duration of longest wet spells over the central and western Kenya (sub-region 1) and southeastern lowlands of Kenya and northeastern Tanzania (sub-region 5). Over western Tanzania and southern Uganda (sub-region 4), this index has significant positive relationship with the number of dry days and significant negative association with the mean duration of wet spells and mean frequency of wet spells of 3 days or more.

With the zonal wind component at 925mb level (not shown), this index has significant positive association over the study area that persists from December-January through to March-April period. This index has strong positive association with the SST over the Indian Ocean north of the 20°S, the Atlantic Ocean between 20°S to 30°S and over the tropical Pacific Ocean that persists over the entire duration (not shown). We hypothesis that the index is a reflection of large SST pattern, that influence wind anomalies. It is these large scale wind anomalies that influences East Africa rainfall.

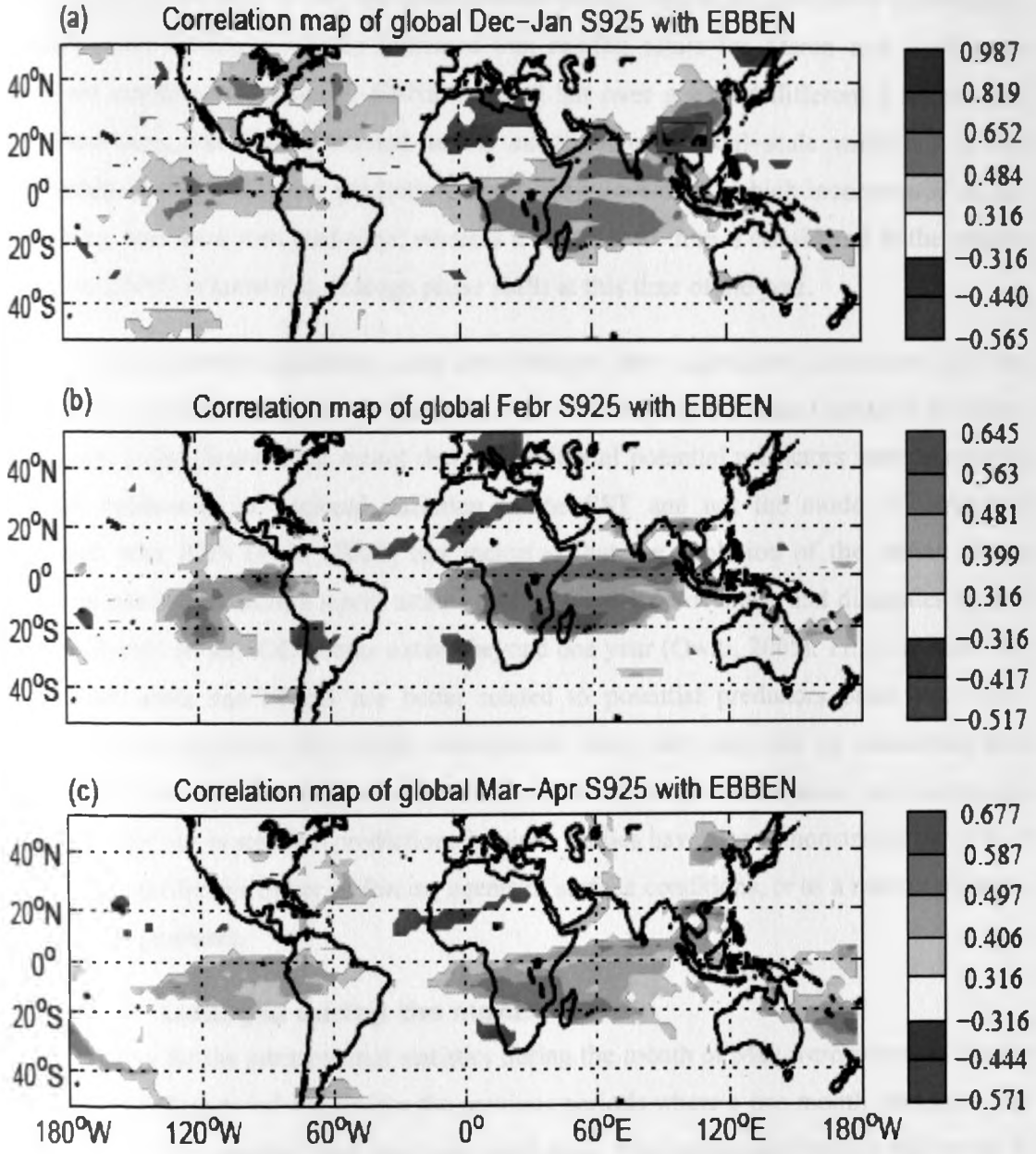


Figure 4.50: Map of significant correlation between east of the Bay of Bengal (EBBEN) specific humidity index and global S925 for (a) December-January, (b) February and (c) March-April. The green rectangle in (a) shows the approximate location of EBBEN specific humidity index computed for December-January period from 1961 to 2000

Most of the additional potential predictors identified earlier have significant association with the predefined indices especially the Niño indices and more specifically the Niño 1 + 2 and Niño 3 indices. This is despite the fact that the Niño indices did not show significant association with the rainfall totals and SRISS during the March-April period. This simply means that the additional potential predictors identified here carries with them part of the

ENSO signal, but that it is not this phenomenon which carries the predictive information. Camberlin and Philippon (2002) indicated that rainfall totals for March and April have significant negative responses to El-Niño events but over a period different from the one considered here, and they cautioned on the existence of decadal–scale variations in this relationship. Additionally, the prediction model they developed, which incorporated an El-Niño index, was for a zero lead-time, whereas a longer lead-time is considered in the present study, and ENSO is known to undergo phase shifts at this time of the year.

The additional potential predictors were also found to have significant association with the SST over the Indian Ocean and yet the association with the Indian Ocean Dipole (IOD) index was mostly insignificant. That meant that the additional potential predictors were associated with the basin-wide or regional variation of the SST and not the mode of variability associated with IOD. Owiti (2005) has indicated that the evolution of the Indian Ocean Dipole events begins around April, attains peak in October-November and dissipates around January. Rarely do the IOD events extend beyond one year (Owiti, 2005). This explains why the rainfall totals and SRISS are better related to potential predictors other than IOD. Although these predictors are mostly atmospheric ones, they may still be associated with SST, which is one of the most obvious features of the climate system, able to provide enough persistence for use in seasonal prediction. Previous studies have also demonstrated the skill of atmospheric predictors, either as forcing agents of surface conditions, or as a marker of large-scale energy gradients.

4.5.2.2 Linkages during the month of May

The predictors for the intraseasonal statistics during the month of May were averaged for the January and February values. Unlike the previous periods where a one month lead time was maintained, a two months' lead time was used here. The explanation behind this move is briefly discussed in **section 4.6.2**. An assessment of the association between the SRISS during the month of May and the predefined indices is discussed next.

4.5.2.2.1 Linkages with the predefined SST predictors

Figures 4.51a–j show graphical presentation of the total correlation coefficients between some of the predefined predictor indices on one hand and the SRISS and rainfall totals for the month of May on the other hand. Just like for the March-April period, the predefined predictors did not show many significant relationships with the SRISS and rainfall totals for the month of May. The Indian Ocean Dipole (IOD) index was the only predefined predictor

that has significant association with the rainfall totals and several of the SRISS of the wet and dry spells (**Figures 4.51a–c, f, h** and **i**). However, this was mainly over the western Tanzania and southern Uganda (sub-region 4). This is consistent with Zorita and Tilya (2002) who emphasized the zonal teleconnections across the Indian Ocean in May against meridional teleconnections in March-April.

4.5.2.2.2 Linkages with additional potential predictors

Partial correlation while controlling the effect of the IOD index identified ten (10) additional potential predictors from the oceanic and atmospheric fields. A brief description of these ten additional potential predictors is provided in **Table 4.24**. The association of these predictors with the rainfall totals and SRISS during the month of May is summarized by **Table 4.25**, **Figures 4.52a–e** and **4.53a–e**.

As earlier observed with the additional potential predictors of the March-April period of the long rainfall season, some of the predictors for the month of May had been identified during the short rainfall season or March-April period. These predictors are ECMAD-1 and WCAUS-2, all of which are from oceanic field (**Table 4.9**). These predictors are therefore not described in details in the subsequent sections. The rest of the potential predictors are from the atmospheric fields, with those from the wind field being discussed first in the next section.

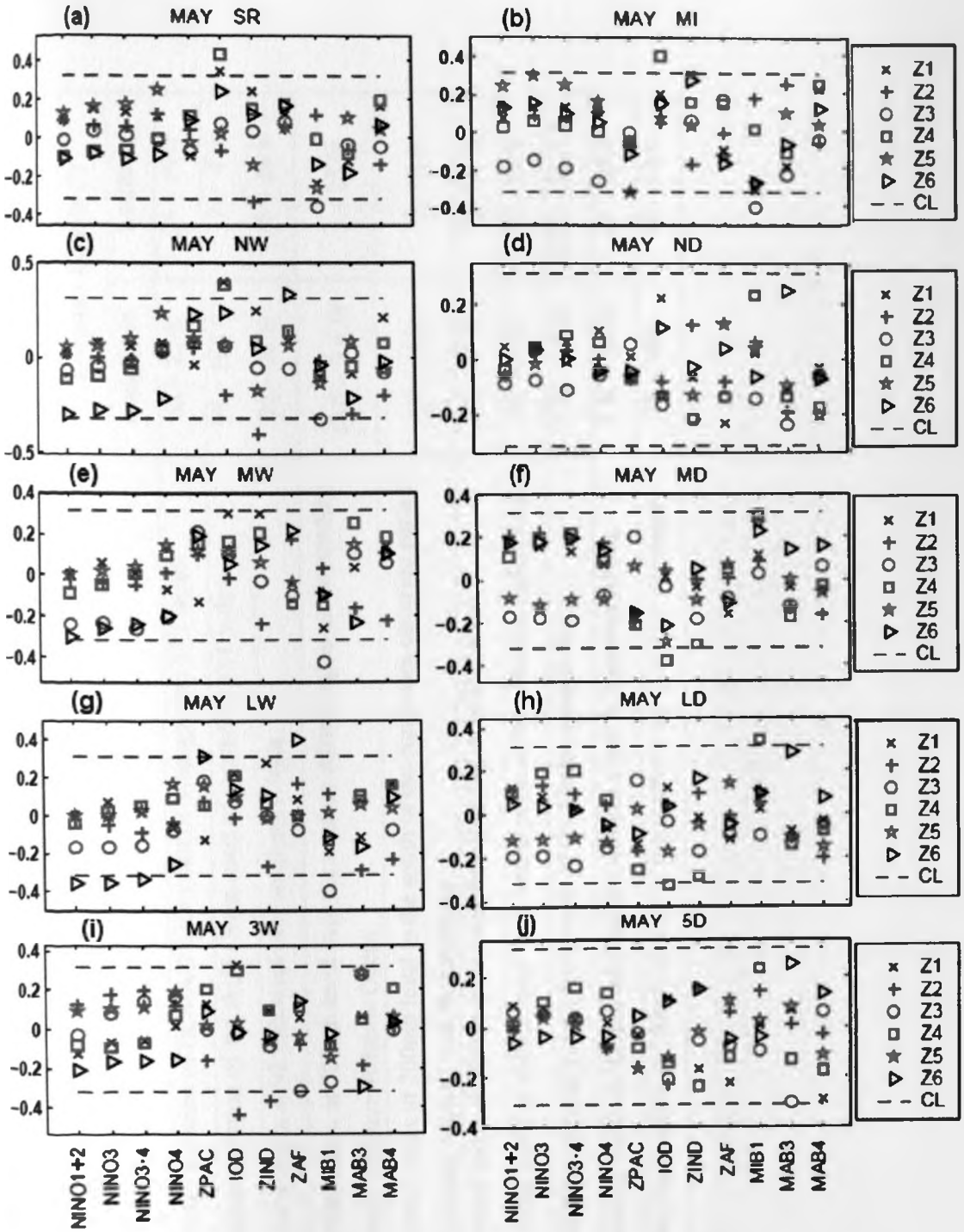


Figure 4.51: Correlation coefficient between predefined predictors averaged over January-February period (x-axis) and the areal-averaged (a) rainfall totals, (b) mean rainfall intensity, (c) number of wet days, (d) number of dry days, (e) mean length of wet spell, (f) mean length of dry spell, (g) longest wet spell, (h) longest dry spell, (i) frequency of 3 wet days or more, and (j) frequency of 5 dry days or more, for the month of May over the six rainfall sub-regions Z1 to Z6. CL shows the 95% confidence level threshold

Table 4.24: Brief description of the additional potential predictors for month of May during long rainfall season and their location details

Index Name	Description	Location Details (°)	
		Longitude	Latitude
ECMAD-1	A slight location variation of ECMAD (East coast of Madagascar) SST index	63 – 75 E	25 – 19 S
WCAUS-2	A slight location variation of WCAUS (West coast of Australia) SST index	90 – 100 E	12 – 4 S
SAFR	Meridional wind index at 925mb level over southern Africa	25 – 30 E	30 – 20 S
NEATO	Meridional wind index at 700mb level over the western Africa region and extending slightly over the northern Atlantic Ocean	12.5 – 7.5 W	5 – 25 N
SSA	Meridional wind index at 200mb level to the south of the study area covering parts of Southern Tanzania, Malawi and Mozambique	25 – 35 E	15 – 5 S
EQATO	Meridional wind index at 200mb level over the equatorial Atlantic Ocean	35 – 25 W	0 – 10 N
CSINDO	Meridional wind index at 200mb level over central parts of the southern Indian Ocean	80 – 90 E	30 – 20 S
SMESEA	Specific humidity index at 925mb level south of the Mediterranean Sea	5 W – 20 E	27.5 – 32.5 N
WCSOA	Specific humidity index at 925mb level on the western coast of southern Africa	5 W – 20 E	25 – 20 S
STAFR	Geopotential height index at 700mb level over the southern tip of Africa continent	20 – 30 E	40 – 35 S

Table 4.25: A summary of the association between the identified additional potential predictors and the sub-regional intraseasonal statistics of the wet and dry spells for the month of May of the long rainfall season and the most strongly correlated intraseasonal statistic and sub-region

Predictor	Atmospheric Level	Index Name	Number of SRISS associated with the predictor (out of 10)	Strongest total correlation		
				SRISS	Sub-region	Coefficient
SST	Surface	ECMAD-1	4	SR	1	0.44
		WCAUS-2	5	SR	4	*-0.43*
v-wind	925mb	SAFR	7	3W	6	*-0.52*
	700mb	NEATO	5	3W	3	*-0.41*
	200mb	SSA	4	LW/MI	4	*-0.41*
		EQATO	7	SR	5	0.45
		CSINDO	8	MW	2	*-0.52*
Specific humidity	925mb	SMESEA	5	NW	1	*-0.42*
		WCSOA	6	5D	5	*-0.49*
Geopotential heights	700mb	STAFR	5	3W	2	*-0.46*

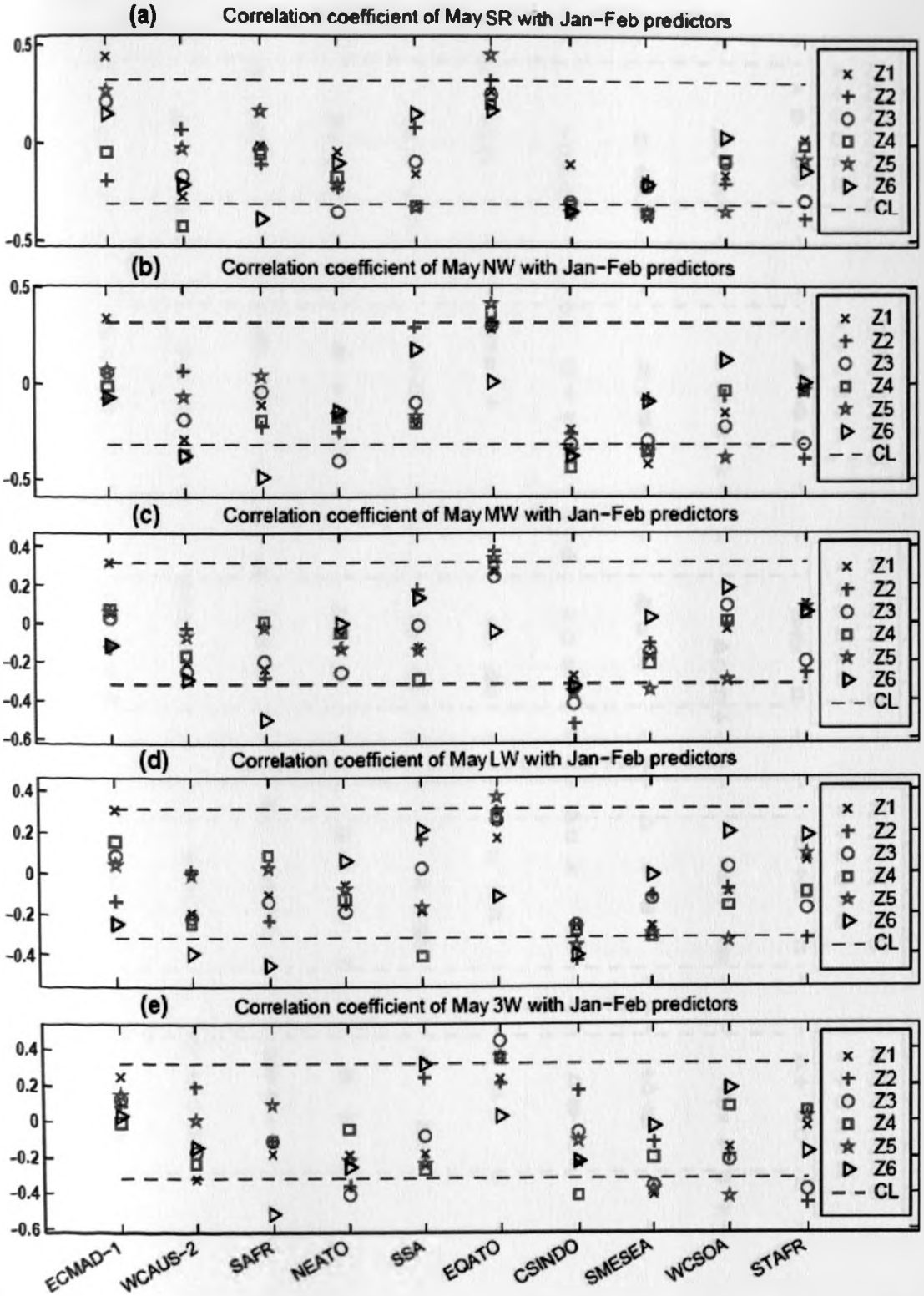


Figure 4.52: Correlation coefficient between the ten additional potential predictors identified averaged over January-February period and the areal-averaged (a) rainfall totals, (b) number of wet days, (c) mean length of wet spell, (d) longest wet spell, and (e) frequency of 3 wet days or more, for the month of May over the six rainfall sub-regions Z1 to Z6. CL shows the 95% confidence level threshold

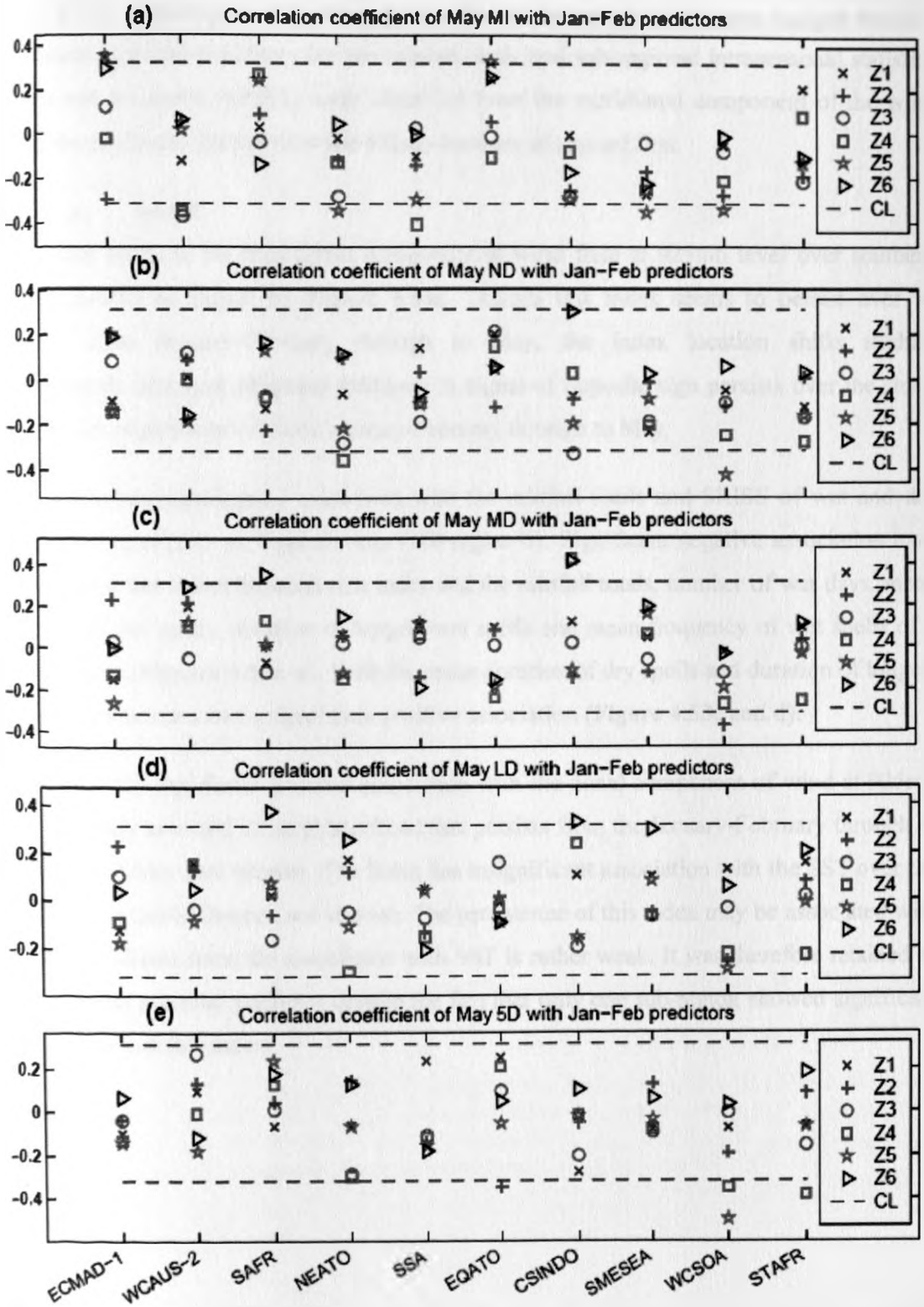


Figure 4.53: Correlation coefficient between the ten additional potential predictors identified averaged over January-February period and the areal-averaged (a) mean rainfall intensity, (b) number of dry days, (c) mean length of dry spell, (d) longest dry spell, and (e) frequency of 5 dry days or more, for the month of May over the six rainfall sub-regions Z1 to Z6. CL shows the 95% confidence level threshold

4.5.2.2.2.1 Additional predictors from the wind and geopotential height fields

Additional potential predictors for the rainfall totals and sub-regional intraseasonal statistics of wet and dry spells (SRISS) were identified from the meridional component of the wind field. The predictors identified at the 925mb level are discussed first.

(a) SAFR

This index refers to the meridional component of wind field at 925mb level over southern Africa (SAFR) as shown by **Figure 4.54a**. Though this index seems to persist over its location from January-February through to May, the index location shifts slightly equatorwards over time (**Figures 4.54a–c**). A signal of opposite sign persists over the study area and its neighbourhood from January-February through to May.

This index was significantly associated with the rainfall totals and SRISS of wet and dry spells over most parts of Uganda only (sub-region 6). Significant negative association over sub-region 6 was noted between this index and the rainfall totals, number of wet days, mean duration of wet spells, duration of longest wet spells and mean frequency of wet spells of 3 days or more (**Figures 4.52a–e**). With the mean duration of dry spells and duration of longest dry spells, this index had a significant positive association (**Figure 4.53c** and **d**).

This index has significant positive association with the zonal component of wind at 925mb over the study area and its neighbourhood that persists from the January-February through to the month of May (not shown). The index has insignificant association with the SST over the Indian and Atlantic Oceans (not shown). The persistence of this index may be associated with the land gradients since the association with SST is rather weak. It was therefore retained as an additional potential predictor despite the fact that only one sub-region showed significant association with this index.

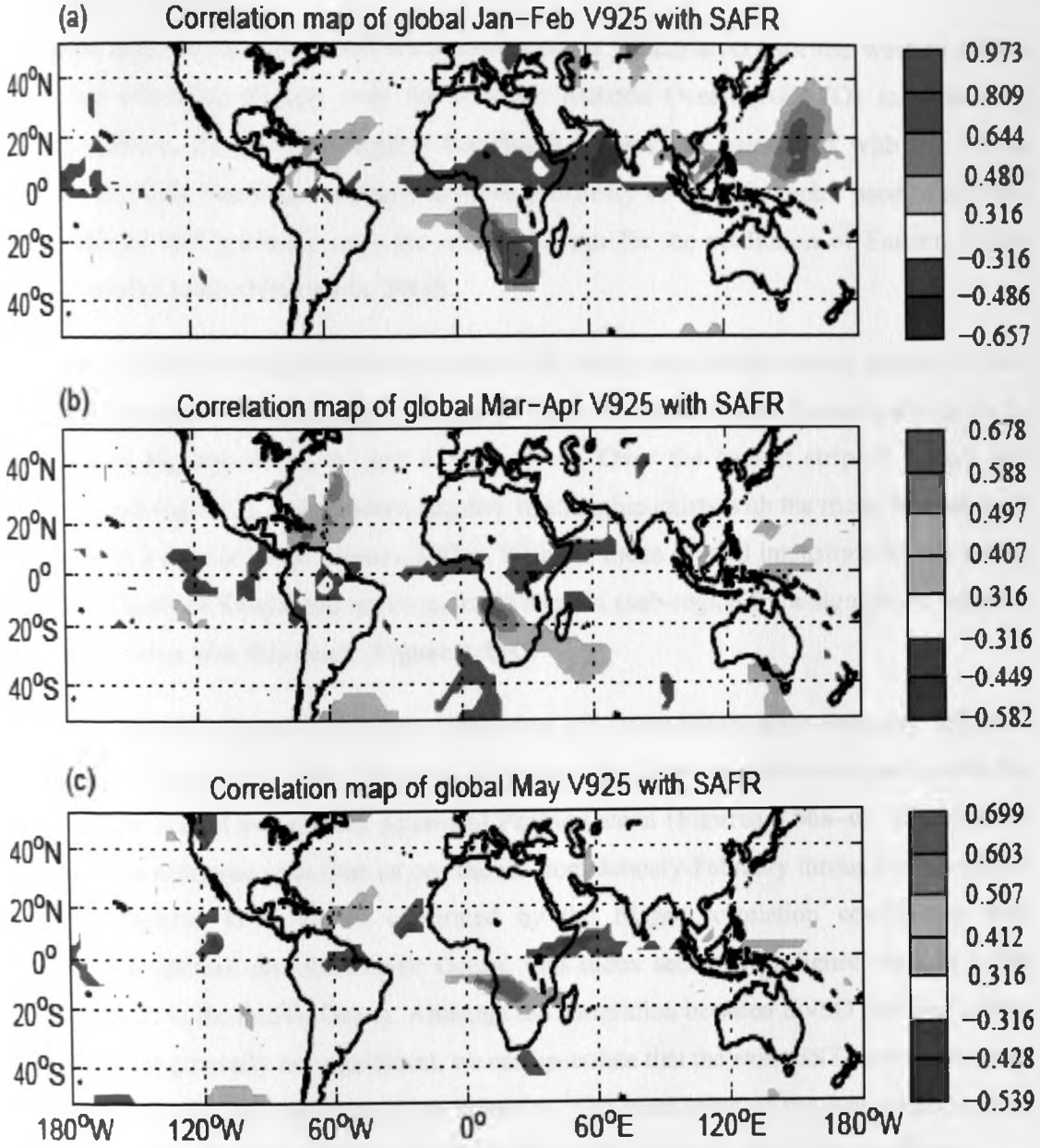


Figure 4.54: Map of significant correlation between southern Africa (SAFR) meridional wind index and global V925 for (a) January-February, (b) March-April and (c) May. The green rectangle in (a) shows the approximate location of SAFR meridional wind index computed for January-February period from 1962 to 2000

(b) NEATO

This index refers to the meridional wind component at 700mb level over the western Africa region and extending slightly over the northern Atlantic Ocean (NEATO) as shown by **Figures 4.55a–c**. Its location suggests that this index may be associated with the Africa Easterly Jet (AEJ). The index is also within the proximity of the pole centre used to develop the meridional SST gradients over the Atlantic Ocean for the prediction of Eastern Africa seasonal rainfall totals (Nyakwada, 2009).

Significant negative association exists between this index and rainfall totals, number of wet days, mean frequency of wet spells of 3 days or more over northeastern Kenya (sub-region 3) as shown in **Figures 4.52a, b,** and **e** respectively. Over the coastal strip of Kenya and Tanzania (sub-region 2), a significant negative relationship exists with the mean frequency of wet spells of 3 days or more (**Figure 4.52e**). With the mean rainfall intensity over the south-eastern lowlands of Kenya and north-eastern Tanzania (sub-region 5), a significant negative association exists with this index (**Figure 4.53a**).

The wind signal associated with this index dies off immediately after February (**Figures 4.55b & c**). With the SST, this index has persistent significant negative association with the SST over the central and eastern equatorial Pacific Ocean (**Figures 4.56a–c**). The negative association is enhanced with time as one moves from January-February through to the month of May (**Figures 4.56a–c**), as confirmed by the bigger correlation coefficients with predefined predictors over the Pacific Ocean. This index seems to prefigure changes in the ENSO context in the Pacific Ocean. Although the correlation between ENSO and East Africa May rainfall is generally not significant, we can speculate that the exact SST pattern shown to be associated with this index has more influence. The weakening of the westerlies coupled with the unavailable of the moisture supply in the month of May leads to dry conditions.

Despite its relationship with the East Africa rainfall not being straightforward, this index was retained as an additional potential predictor for the month of May.

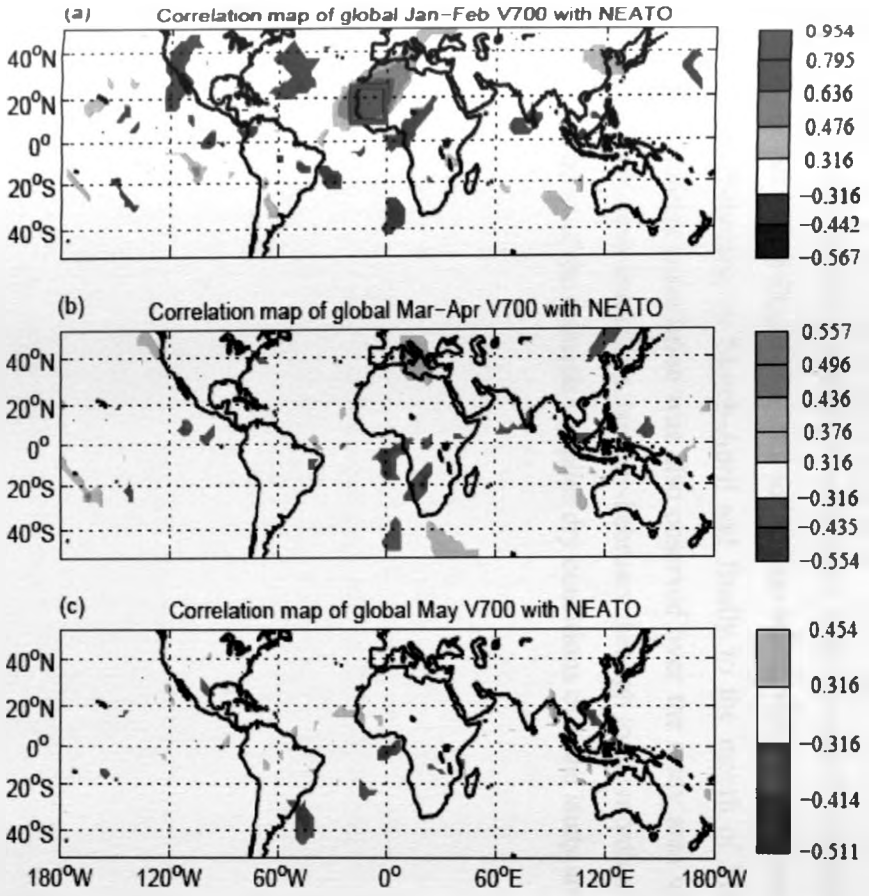


Figure 4.55: Map of significant correlation between northern Atlantic Ocean (NEATO) meridional wind index and global V700 for (a) January-February, (b) March-April and (c) May. The green rectangle in (a) shows the approximate location of NEATO meridional wind index computed for January-February period from 1962 to 2000

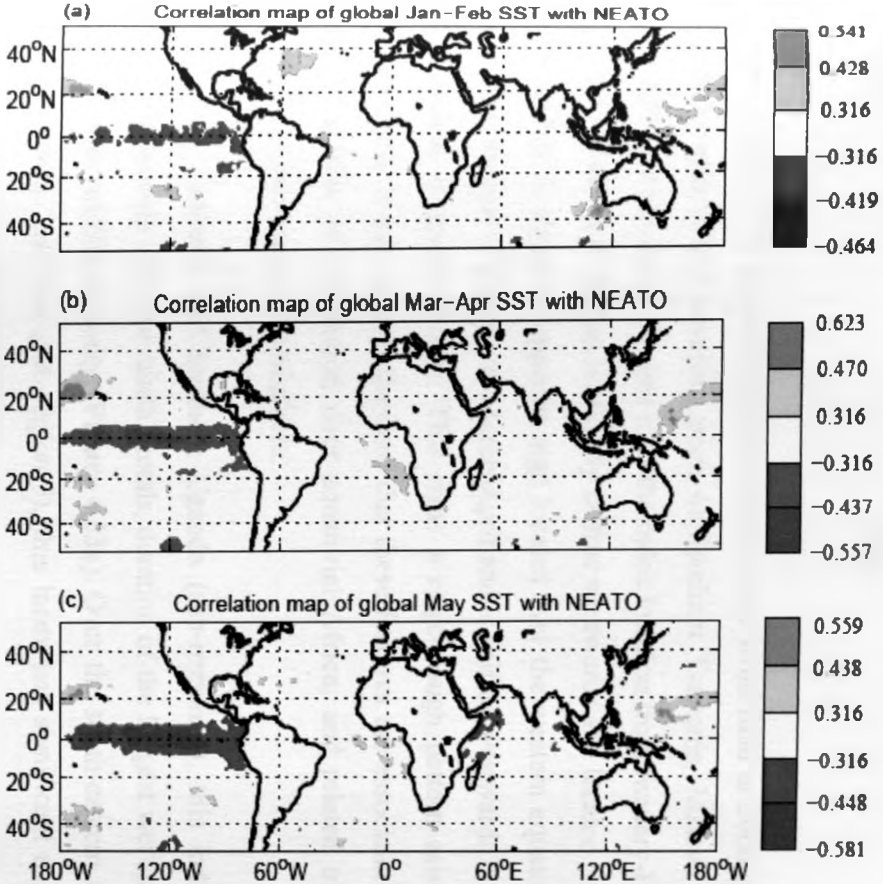


Figure 4.56: Map of significant correlation between northern Atlantic Ocean (NEATO) meridional wind index and global SST for (a) January-February, (b) March-April and (c) May

(c) SSA

This index refers to the meridional component of the wind field at 200mb level to the south of the study area (SSA) covering parts of Southern Tanzania, Malawi and Mozambique (**Figure 4.57a**). This signal persists over the index location from January-February to March-April (**Figures 4.57a & b**) and is slightly shifted westwards and located over Angola in May (**Figure 4.57c**). A signal of opposite sign located over the western equatorial Indian Ocean, persists from January-February to March-April and is shifted westwards in the month of May with reduction in spatial extent. This signs a ridge-trough pattern across equatorial and southern Africa. It can be hypothesized that these features are associated with shifts in the preferred location of convection over equatorial Africa, and related tropical-extratopical cloud bands in the southern hemisphere.

Over western Tanzania and southern Uganda (sub-region 4), this index has significant negative relationship with the rainfall totals, duration of the longest wet spells (**Figure 4.52a & d**) and mean rainfall intensity (**Figure 4.53a**). Over the south-eastern lowlands of Kenya and north-eastern Tanzania (sub-region 5), this index has significant negative relationship with the rainfall totals (**Figure 4.52a**).

The index has significant but inverse association with the specific humidity at 925mb level over the central Indian Ocean that tend to increase in intensity and spatial extend as one move from January-February, to March-April and finally to the month of May (not shown). A significant negative association was also observed over the study area with the zonal wind component at 925mb level from January-February through to the month of May (not shown). The enhancement of the easterlies implies dry conditions over the study area.

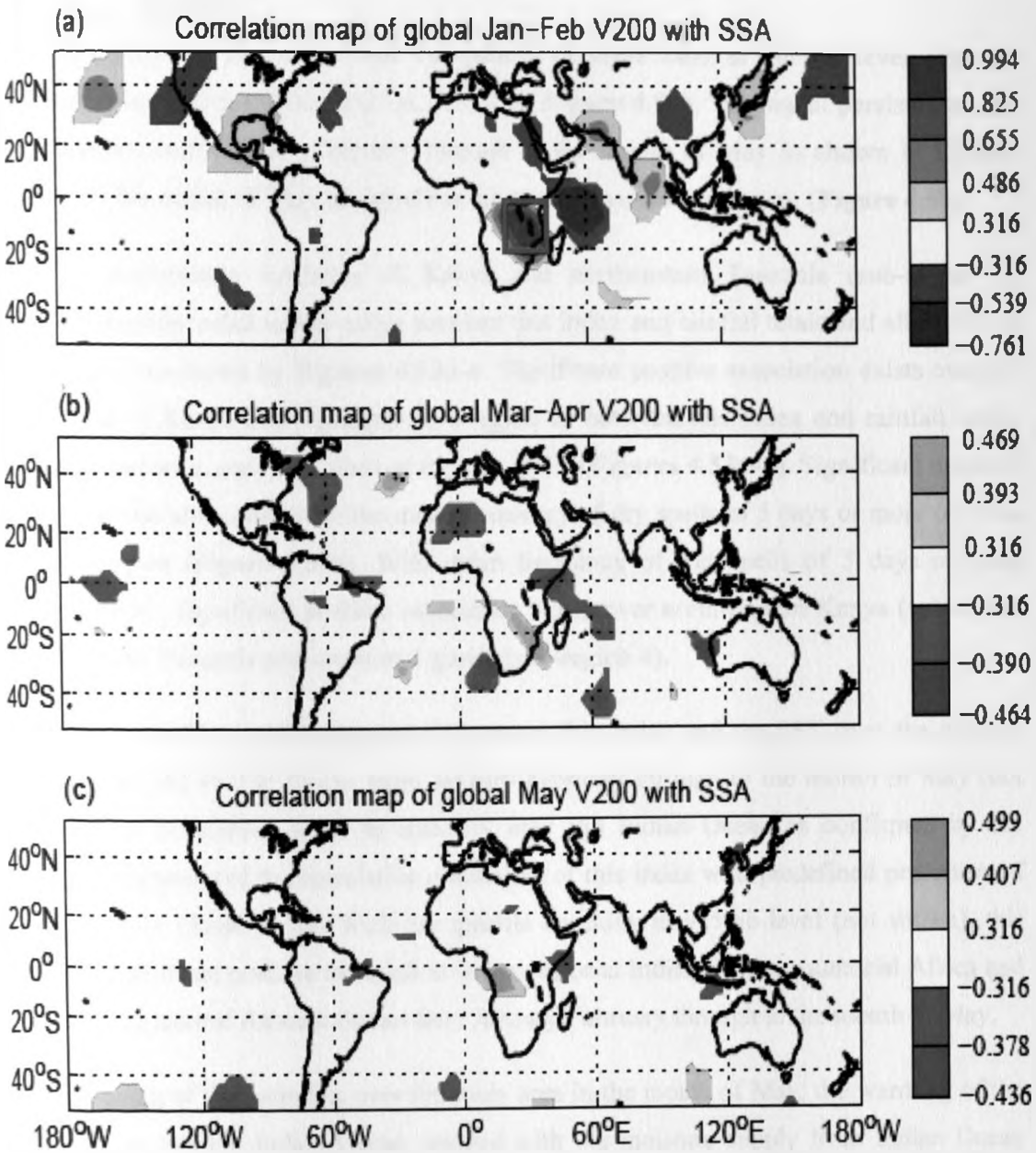


Figure 4.57: Map of significant correlation between south of the study area (SSA) meridional wind index and global V200 for (a) January-February, (b) March-April and (c) May. The green rectangle in (a) shows the approximate location of SSA meridional wind index computed for January-February period from 1962 to 2000

(d) EQATO

This index refers to the meridional component of wind field at 200mb level over the equatorial Atlantic Ocean (EQATO) as shown by **Figure 4.58a**. The signal persists over the index location from January-February through to the month of May as shown in **Figures 4.58a–c**. By the month of May, a signal has developed over the study area (**Figure 4.58c**).

Over the southeastern lowlands of Kenya and northeastern Tanzania (sub-region 5), significant positive relationship exists between this index and rainfall totals and all SRISS of the wet spells as shown by **Figures 4.53a–e**. Significant positive association exists over the coastal strip of Kenya and Tanzania (sub-region 2) between this index and rainfall totals, number of wet days, mean duration of the wet spells (**Figures 4.53a–c**). Significant negative relationship was also noted with the mean frequency of dry spells of 5 days or more over the same sub-region (**Figure 4.54e**). With mean frequency of wet spells of 3 days or more (**Figures 4.53e**), significant positive association exists over north-eastern Kenya (sub-region 3) and western Tanzania and southern Uganda (sub-region 4).

A significant positive relationship exists between this index and the SST over the western Indian Ocean and Pacific Ocean from January-February through to the month of May (not shown). This association tends to intensify over the Indian Ocean as confirmed by the increased magnitude of the correlation coefficient of this index with predefined predictors of IOD and MIB1 (**Table 4.26**). With the specific humidity at 925mb level (not shown), this index has significant positive association over equatorial Indian Ocean, equatorial Africa and parts of the equatorial Atlantic Ocean from January-February through to the month of May.

The weakening of the easterlies over the study area in the month of May, the warming of the SST over the western Indian Ocean coupled with the moisture supply from Indian Ocean leads to wet conditions over the study area. This index was therefore included as an additional potential predictor.

Table 4.26: Correlation coefficients between equatorial Atlantic Ocean (EQATO) meridional wind index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	MIB1	MAB3
Jan-Feb	**0.42**	**0.46**	**0.61**	**0.51**	0.05	-0.18	-0.03
May	0.29	**0.38**	**0.61**	**0.51**	**0.33**	-0.23	-0.14

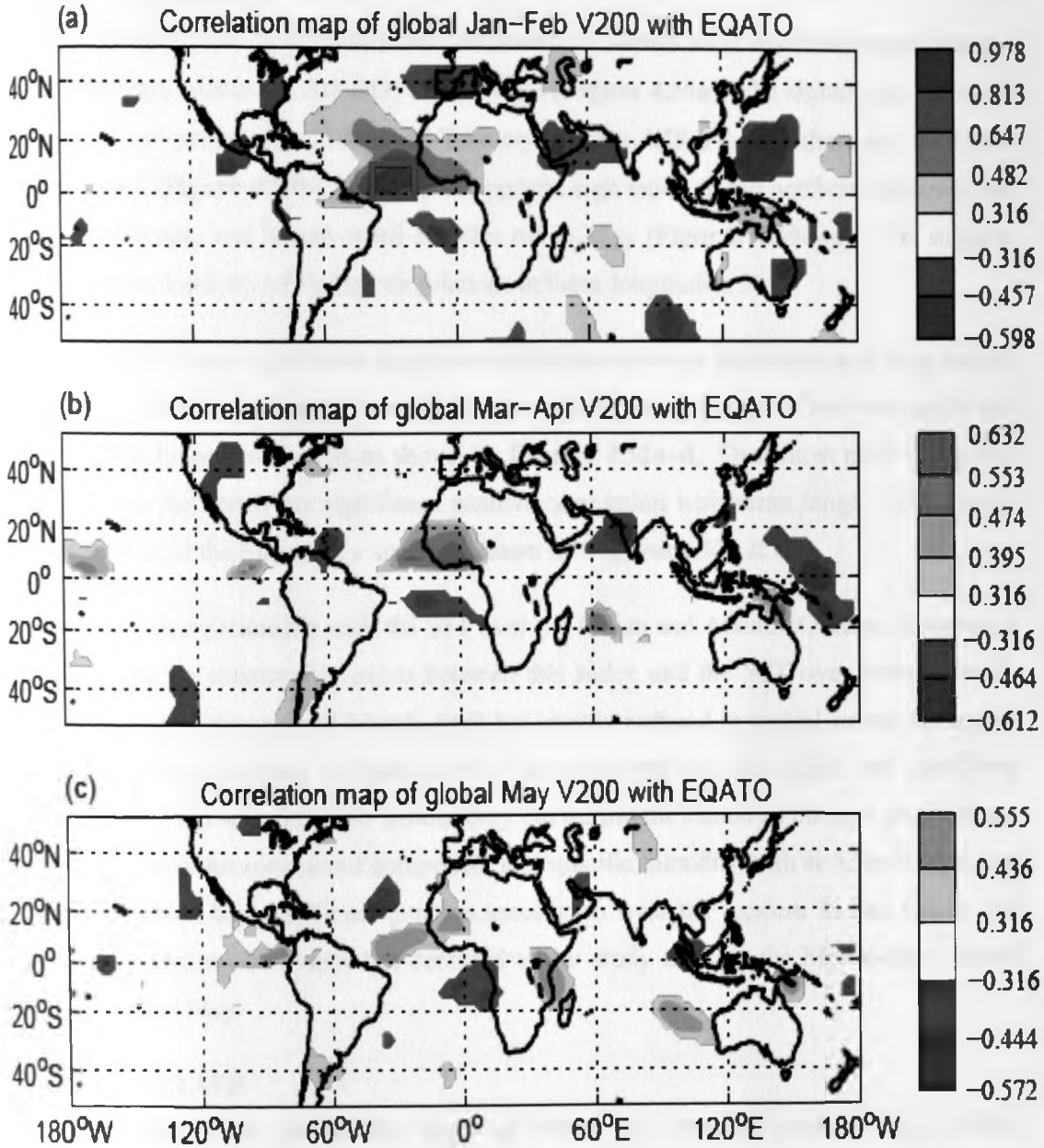


Figure 4.58: Map of significant correlation between equatorial Atlantic Ocean (EQATO) meridional wind index and global V200 for (a) January-February, (b) March-April and (c) May. The green rectangle in (a) shows the approximate location of EQATO meridional wind index computed for January-February period from 1962 to 2000

(e) CSINDO

This index refers to the meridional wind component at 200mb level over the central parts of the southern Indian Ocean (CSINDO) around 20°S (**Figure 4.59a**). The signal over the index location does not persist beyond January-February (**Figure 4.59a**) though there is a faint trace in March-April (**Figure 4.59b**). A signal of opposite sign exists in the northern Indian Ocean in January-February and March-April but dies off in May (**Figures 4.59a–c**). This suggests variations in the intensity of Hadley circulations at these longitudes.

Most sub-regions show significant negative correlations between this index and May rainfall and wet spells statistics especially number of wet days, mean length of the wet spells and duration of the longest wet spells as shown by **Figures 4.52a–d**. Over most parts of Uganda (sub-region 6), this index has significant positive correlation with mean length of dry spells and the duration of the longest dry spells as shown by **Figures 4.53c & d**.

This index has no relationship with the SST over the Indian and Atlantic Oceans. However, a significant positive relationship exists between this index and the SST over central Pacific Ocean in January-February and March-April but greatly reduced in spatial extent during the month of May (not shown). Consistent with these observations, the index had significant positive relationship with the Niño indices only but during the January-February period alone (not shown). With the zonal wind component and specific humidity both at 925mb level (not shown), this index has significant positive association over the western Indian Ocean and study area in January–February, but confined to the study area for the March-April period and the month of May.

(f) STAFR

This index refers to the geopotential height at 700mb level over the southern tip of Africa continent (STAFR) as indicated in **Figure 4.60a**. This signal does not persist beyond the January-February period (**Figures 4.60a–c**).

Over the coastal strip of Kenya and Tanzania (sub-region 2), significant negative relationship exists between this index and rainfall totals, number of wet days, duration of longest wet spells and mean frequency of wet spells of 3 days or more (**Figures 4.52a, b, d** and **e**). Significant negative relationship with the number of wet days and mean frequency of wet spells of 3 days or more over the northeastern Kenya (sub-region 3) was also noted (**Figures 4.52b** and **e**).

With the zonal wind component at 925mb and 700mb levels (not shown), this index has significant negative relationship over the study area and western parts of equatorial Indian Ocean from January-February through to the month of May. No significant association was noted with SST over the global Oceans. Its relationship with the East African rainfall is however not straightforward like most other additional potential predictors.

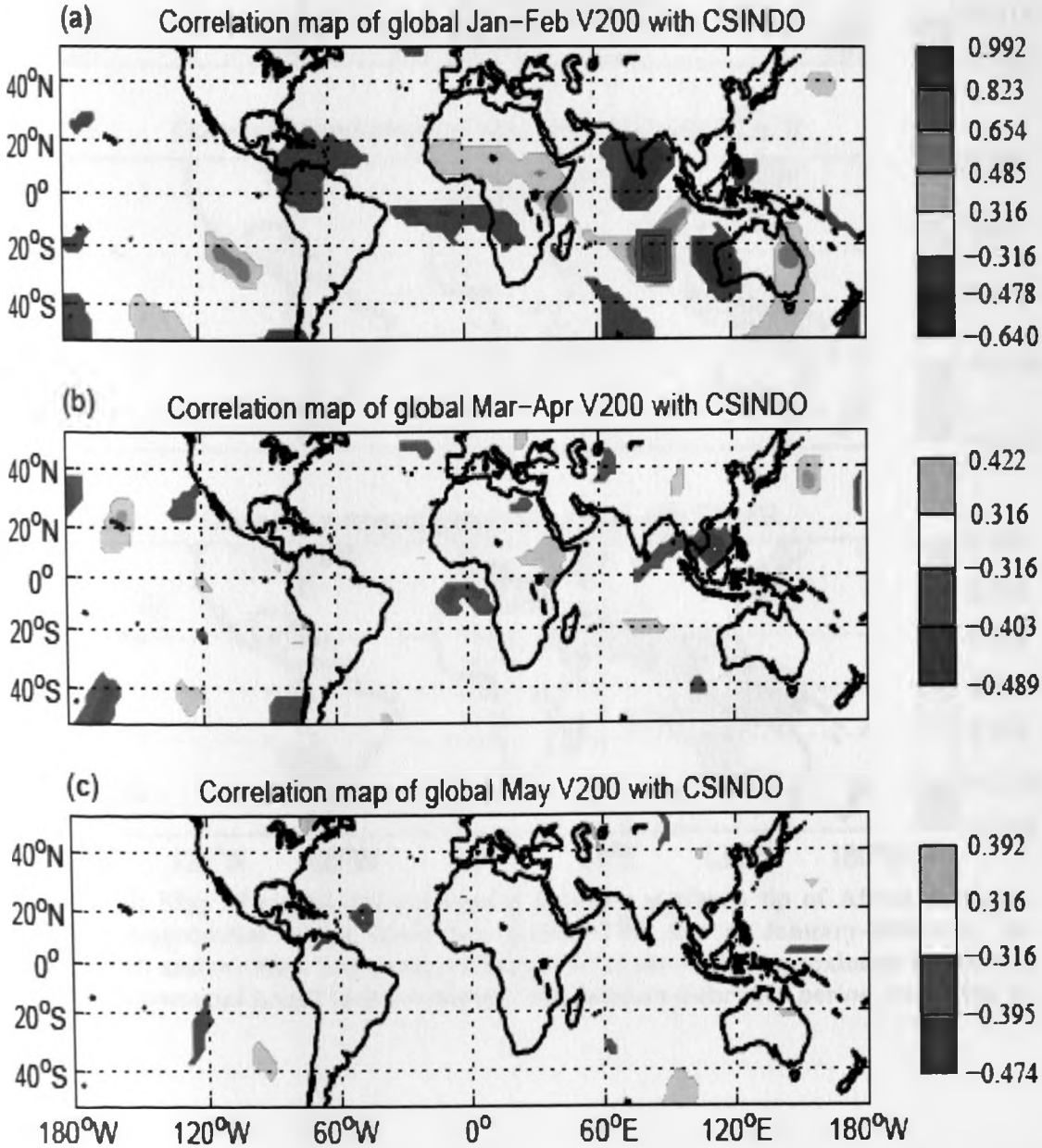


Figure 4.59: Map of significant correlation between central parts of the southern Indian Ocean (CSINDO) meridional wind index and global V200 for (a) January-February, (b) March-April and (c) May. The green rectangle in (a) shows the approximate location of CSINDO meridional wind index computed for January-February period from 1962 to 2000

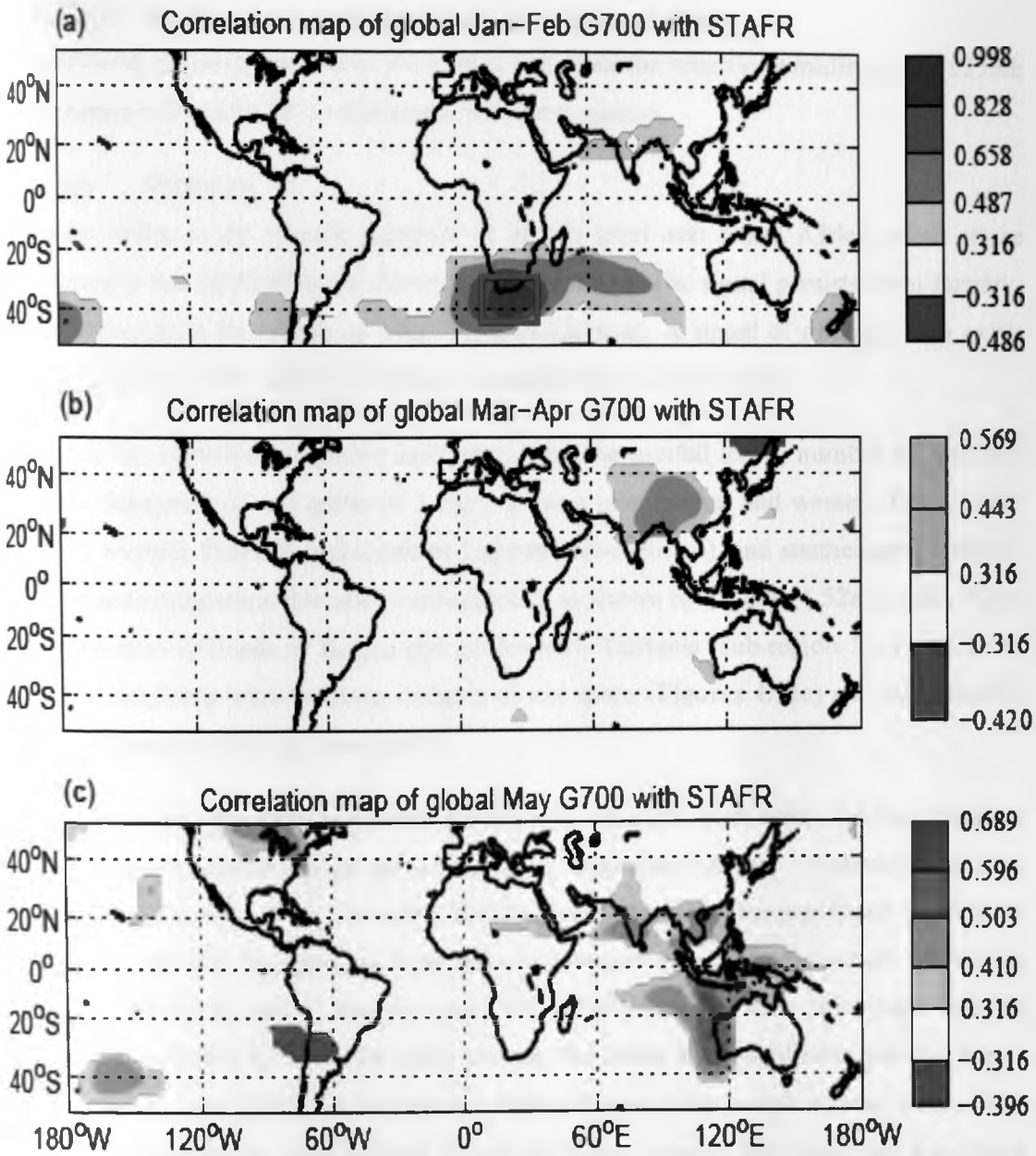


Figure 4.60: Map of significant correlation between southern tip of Africa continent (STAFR) geopotential height index and global G700 for (a) January-February, (b) March-April and (c) May. The green rectangle in (a) shows the approximate location of STAFR geopotential height index computed for January-February period from 1962 to 2000

4.5.2.2.2.2 Additional predictors from specific humidity

Two additional potential predictors were identified from the specific humidity at the 925mb levels (**Tables 4.24** and **4.25**) as discussed in the next section.

(a) SMESEA

This index refers to the specific humidity at 925mb level over North Africa, south of the Mediterranean Sea (SMESEA) as shown by **Figure 4.61a**. The signal persists from January-February through to the month of May (**Figures 4.61a–c**). A signal of opposite sign exists over the study area from January-February through to the month of May.

This index has significant negative association with the rainfall totals, number of wet days and mean frequency of wet spells of 3 days or more over central and western Kenya (sub-region 1), western Tanzania and southern Uganda (sub-region 4), and southeastern lowlands of Kenya and northeastern Tanzania (sub-region 5) as shown by **Figures 4.52a, b** and **e**. Over the southeastern lowlands of Kenya and northeastern Tanzania (sub-region 5), a significant negative relationship with the mean duration of wet spells (**Figures 4.52c**) and mean rainfall intensity (**Figures 4.53a**) was also noted.

Given that it is a continental signal and it has strong temporal persistence, we can speculate that this index is associated with the soil moisture anomalies (January – February is a rainy season in North Africa). Over the central Pacific Ocean, this index has significant but inverse association with SST that persists from January-February through to the month of May as shown by significant and increasing correlation coefficients between this index and the ENSO indices (**Table 4.27**). In the Indian Ocean, the index has a significant positive signal between Equator and 20°S over the western Indian Ocean in the month of May alone. With the zonal component of wind field at 925mb level (not shown), this index has significant inverse relationship over Gulf of Guinea in January–February, and shifting eastwards with time. By the month of May, the significant inverse association is slightly to the west of the study area and extending into Gulf of Guinea. The weakening of the westerlies over the study area, the cooling of SST over equatorial Pacific from January-February through to May results in dry conditions.

The robust physical explanation on how this index relates to the rainfall totals and SRISS coupled with the signals from the oceanic and atmospheric variables justify the inclusion of this index as an additional potential predictor.

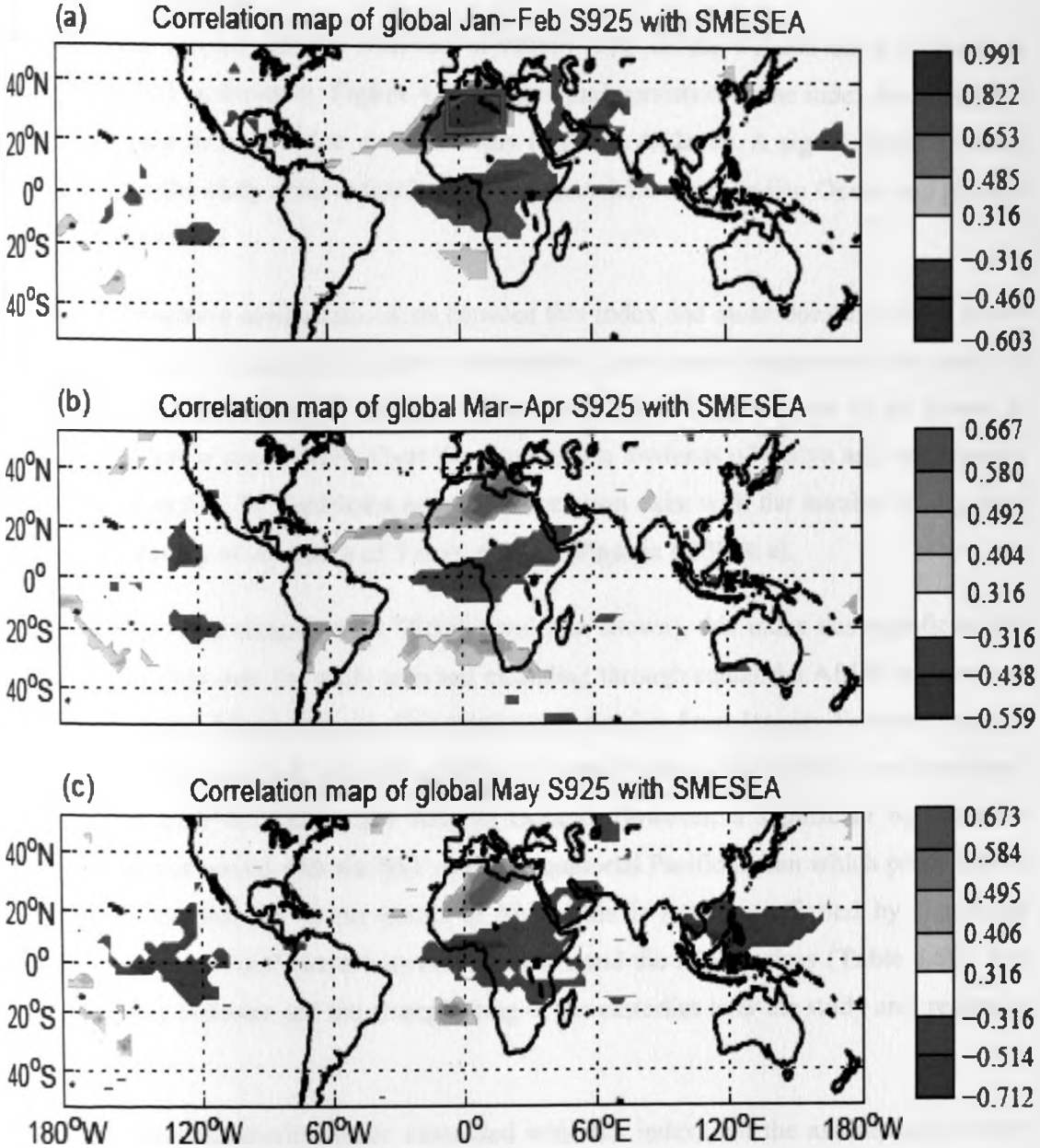


Figure 4.61: Map of significant correlation between south of the Mediterranean Sea (SMESEA) specific humidity index and global S925 for (a) January-February, (b) March-April and (c) May. The green rectangle in (a) shows the approximate location of SMESEA specific humidity index computed for January-February period from 1962 to 2000

Table 4.27: Correlation coefficients between south of the Mediterranean Sea (SMESEA) specific humidity index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	MIB1	MAB3
Jan-Feb	-0.31	**-0.37**	**-0.60**	**-0.45**	-0.09	-0.01	0.10
May	-0.22	**-0.36**	**-0.64**	**-0.54**	-0.18	-0.13	0.29

(b) WCSOA

This index refers to the specific humidity at 925mb level on the western coast of southern Africa (WCSOA) as shown by **Figure 4.62a**. The signal persists over the index location from January-February through to the month of May (**Figures 4.62a–c**). A signal of the opposite sign exists over the study area, extending to the equatorial western Indian Ocean and persists over the same period.

A significant negative association exists between this index and mean duration of dry spells over coastal strip of Kenya and Tanzania (sub-region 2) and mean frequency of dry spells of 5 days or more over western Tanzania and southern Uganda (sub-region 4) as shown in **Figures 4.53c** and **e** respectively. Over the southeastern lowlands of Kenya and northeastern Tanzania (sub-region 5), significant negative association exist with the number of dry days and mean frequency of dry spells of 5 days or more (**Figures 4.53b & e**).

With the zonal wind component at 925mb level (not shown), this index has significant but inverse relationship over the study area and extending through equatorial Africa into western parts of Equatorial Atlantic Ocean. This relationship persists from January-February through to the month of May though with the reduction in spatial extent. This index is not associated with the SST over the Indian and Atlantic Oceans. However, a significant but negative association was observed with the SST over the equatorial Pacific Ocean which persists from January-February through to the month of May. This is further confirmed by significant negative correlation coefficients between this index and the Niño indices (**Table 4.28**). The depletion of the moisture and the strengthening of the easterlies over the study area results in dry conditions.

The consistent atmospheric signals associated with this index, and the assumption on how they relates to the SRISS of dry spells provide a strong case for retention of this index as an additional potential predictor.

Table 4.28: Correlation coefficients between western coast of southern Africa (WCSOA) specific humidity index and some predefined predictors

	Niño 1+2	Niño 3	Niño 4	Niño 3.4	IOD	MIB1	MAB3
Jan-Feb	**-0.42**	**-0.48**	**-0.43**	**-0.51**	0.10	-0.16	0.02
May	**-0.41**	**-0.44**	**-0.46**	**-0.51**	0.11	-0.12	0.11

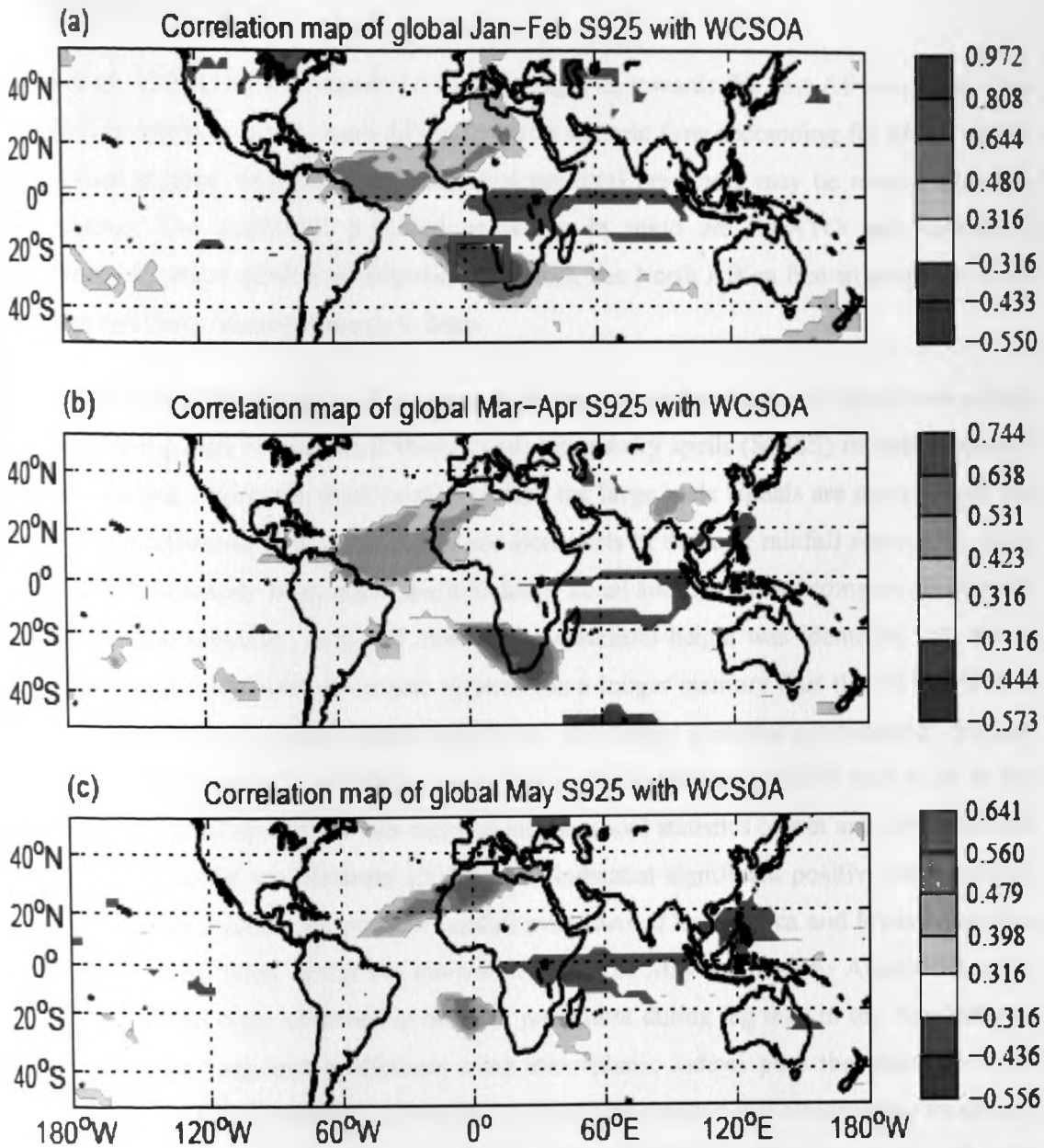


Figure 4.62: Map of significant correlation between western coast of southern Africa (WCSOA) specific humidity index and global S925 for (a) January-February, (b) March-April and (c) May. The green rectangle in (a) shows the approximate location of WCSOA specific humidity index computed for January-February period from 1962 to 2000

As earlier observed, most of the potential predictors identified have significant association with the predefined predictor indices. This is despite the fact that most of the predefined SST predictors did not show significant association with the rainfall totals and SRISS for the month of May. This simply means that the additional potential predictors identified here carries with them part of the ENSO signal, but that it is not this phenomenon which carries

the predictive information.

Henne *et al.*, (2008) have documented six flow regimes towards the east Africa region. One of these flow regimes is the North Africa free tropospheric flow accounting for 6% of all the observations studied. Several of the additional potential predictors may be related with this flow regime. The additional potential predictors in mind are NEATO and SMESEA. Consistent with these additional potential predictors, the North Africa free tropospheric flow was observed from January through to May.

In conclusion, the identification of the large scale oceanic and atmospheric signals associated with the sub-regional intraseasonal statistics of wet and dry spells (SRISS) including rainfall totals shown that during the short rainfall season, the large scale signals are mainly from the oceanic field. However during the earlier and later parts of the long rainfall season, the large scale signals are mainly from atmospheric fields of zonal and meridional components of wind and the specific humidity. A signal from the geopotential height was identified only once. The ocean (as a component of climate system) has a longer memory thus the SST field as a climatic variable has a greater persistence hence the higher potential predictability already observed. The atmospheric variables have a rather lower persistence which may point to the lower potential predictability of sub-regional intraseasonal statistics of wet and dry spells and rainfall totals. Indeje and Semazzi (2000) have indicated significant positive simultaneous and non-zero lag correlations between rainfall over parts of East Africa and lower equatorial stratospheric zonal wind during the months of March to May and June to August did exist. These associations were observed to be more prominent during lag than in the simultaneous correlations. The long lead predictions using atmospheric indices pose the question of the physical basis of the relationships. However, it should be recalled that atmospheric variability may reflect land and/or ocean surfaces, both having a relatively longer 'memory'. In such case, the atmospheric predictor can be viewed as a proxy of climate memory associated with these surface conditions. Surface conditions (esp. land) cannot be always directly captured by available data sets.

The additional potential predictors (both oceanic and atmospheric) for the earlier and later parts of the long rainfall seasons were all from within the African continent and the two adjacent oceans. The oceanic indices associated with the SRISS and rainfall totals of earlier and later parts of the long rainfall season happened to be some of the oceanic indices already identified during the short rainfall season though with slight displacement in location.

Having identified the additional potential predictors for rainfall totals and SRISS for the two rainfall seasons, the multivariate linear regression (MLR) models were developed and their performance evaluated. These MLR models are discussed in the next section.

4.6 Regression models for sub-regional intraseasonal statistics of wet and dry spells

The methodology used to develop the multivariate linear regression (MLR) models at sub-regional level and assess the performance of these models was discussed in **section 3.2.6**. The results of the MLR models developed for rainfall totals and SRISS and their performance assessment are shown and discussed in the subsequent sections starting with those of short rainfall season.

4.6.1 Regression models during the short rainfall season

The seasonal rainfall totals and sub-regional intraseasonal statistics of the wet and dry spells (SRISS) during the short rainfall season were found to be spatially more coherent, suggesting higher potential predictability as compared to those of the long rainfall season (**Figures 4.19** and **4.20**). This is consistent with previous studies that have found significant concurrent and lagged association with the Niño, IOD and SST gradient indices (Ogallo, 1988; Mutemi, 2003; Black *et al.*, 2003; Black, 2005; Owiti, 2006; Nyakwada, 2009).

The list of the predictors from which the regression models were developed was shown in **Tables 4.9** and **4.10.** Two predictors from the predefined indices (ZIND and Niño 3.4) and nine additional predictors from the oceanic and atmospheric fields are used. The total correlation of each of this additional potential predictor with the rainfall totals and intraseasonal statistics was shown in **Figures 4.23a–e** and **4.24a–e.** As indicated earlier, the predictor indices were averaged for the months of July-August and used to develop the OND MLR models for seasonal rainfall totals and SRISS. The predicted values directly obtained from the MLR model developed and the MLR cross-validated model are shown as graphs while the predictors that are picked and the assessment of performance are tabulated. For the cross-validated MLR models, three observations were left out.

4.6.1.1 Seasonal rainfall totals

Table **4.29** illustrates on how the final number of predictors to be retained was determined for the seasonal rainfall totals during the OND season over sub-region 1 (central Kenya and southeastern lowlands of Kenya) based on the R-adjusted consideration with the addition of one extra predictor at a time. The predictor that was most strongly associated with seasonal rainfall totals over sub-region 1 was SINDS with a correlation coefficient of 0.543 and adjusted correlation coefficient of 0.276. In the cross-validated mode, this predictor had a correlation coefficient of 0.468 and adjusted correlation coefficient of 0.198. In the second step, predictor BoBEN was picked. The two predictors had a multiple correlation coefficient of 0.742 with seasonal rainfall totals while the adjusted correlation coefficient was 0.526. In the cross-validated mode, the two predictors had a multiple correlation coefficient of 0.686 with seasonal rainfall totals and the adjusted correlation coefficient of 0.441. In the third step SWAFRC was picked, the fourth step gave SWHAW and so on.

A close look at this table shows that the multiple correlation coefficient for the developed MLR model and its adjusted correlation coefficient as well as multiple correlation coefficient for the cross-validated model has been increasing at each step. However, the adjusted correlation coefficient for the cross-validated model starts to decrease after step 4. This means that the additional predictor, (SWHAW), makes little marginal changes in the unexplained variance and hence should therefore be dropped. The first four predictors can thus be used to develop the multivariate linear regression (MLR) model for seasonal rainfall totals over this particular sub-region. Multi-collinearity assessment further shows that SINDS and SWHAW are significantly inverse correlated (r=-0.406) at 95% confidence level. Since the Variance Inflation Factor (VIF) was not calculated, only one of these two predictors should be used to avoid the inflation of the variance and loss of degrees of freedom (Krishna Kumar *et al.*, 1995). Thus the regression model developed for the seasonal rainfall totals over sub-region 1 was based on the first three predictors shown in **Table 4.29.** The predictors to be retained for other sub-regions and the SRISS were similarly obtained.

Table 4.29: Forward stepwise fitting of the multivariate regression model for OND areal-averaged seasonal rainfall totals over sub-region 1

Step	Predictor included	Multiple Correlation Coefficient			
		R	Adjusted R	R cv	Adjusted R cv
1	SINDS	0.543	0.276	0.468	0.198
2	BoBEN	0.742	0.526	0.686	0.441
3	SWAFRC	0.789	0.591	0.732	0.497
4	SWHAW	0.808	0.611	0.746	0.504
5	WCAUS	0.818	0.618	0.748	0.493
6	EQAFR	0.818	0.607	0.740	0.463

Figures 4.63a–f show the time series plots from the developed and cross-validated multivariate linear regression (MLR) models as well as the actual observations for the seasonal rainfall totals (SR) while **Table 4.30** summarizes the predictors used and the skill score of the models. The figures show that the developed models capture the direction of the observation quite well though at times the magnitudes are not attained. From a list of four predictors, two sets of combinations of these predictors were adequate to describe the interannual variability of the seasonal rainfall totals over the six sub-regions during the short rainfall season. The atmospheric predictor SINDS (a July-August U-wind index at 925mb over southern tip of India sub-continent) and oceanic predictor BoBEN (a July-August SST index over Bay of Bengal) were common to all the MLR models. This was closely followed by SWAFRC (a July-August specific humidity index at 700mb over the southwestern Africa) which was picked in five models. It should be observed from **Table 4.30** that none of the models picked the Nino 3.4 index as a predictor while ZIND index was only picked once. This does not mean that Niño 3.4 index (a representative of the ENSO indices) is not related to Equatorial Eastern Africa seasonal rainfall totals, but rather the predictive signal in ENSO is contained in the other predictors from the Indian Ocean region.

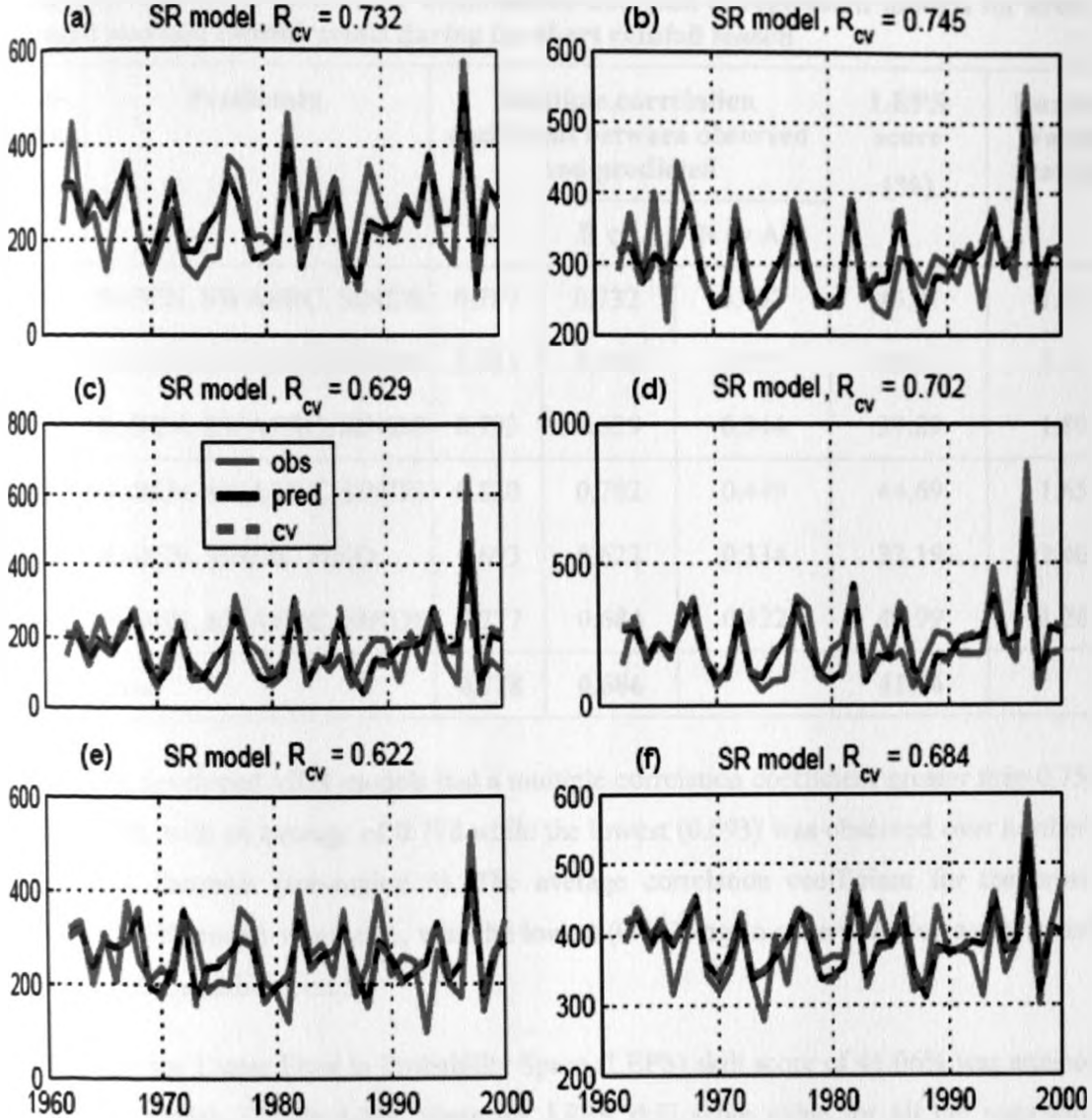


Figure 4.63: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for October-November-December areal-averaged rainfall totals (SR) over (a) Central highlands and southeastern lowlands of Kenya, (b) Western Kenya and most parts of Uganda, (c) Northeastern Kenya, (d) Coastal strip of Kenya and Tanzania, (e) Central and northern Tanzania, and (f) Western of Lake Victoria and western Tanzania. R_{cv} shows the multiple correlation coefficient for cross-validated model

Table 4.30: The list of predictors' combination and skill of regression models for areal-averaged seasonal rainfall totals during the short rainfall season

Sub-region	Predictors	Multiple correlation coefficient between observed and predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	BoBEN, SWAFRC, SINDS	0.789	0.732	0.497	43.31	1.40
2	BoBEN, SWAFRC, SINDS	0.795	0.745	0.517	42.27	2.11
3	BoBEN, SWAFRC, SINDS	0.793	0.629	0.344	39.89	1.89
4	BoBEN, SWAFRC, SINDS	0.820	0.702	0.449	44.69	1.65
5	BoBEN, SINDS, ZIND	0.693	0.622	0.334	32.19	2.40
6	BoBEN, SWAFRC, SINDS	0.777	0.684	0.422	43.99	1.26
Mean Value		**0.778**	**0.686**		**41.06**	

Most of the developed MLR models had a multiple correlation coefficient greater than 0.750 (**Table 4.30**), with an average of 0.778 while the lowest (0.693) was observed over northern and central Tanzania (sub-region 5). The average correlation coefficient for the cross-validated MLR models was 0.686, with the lowest (0.622) again observed over northern and central Tanzania (sub-region 5).

On average, the Linear Error in Probability Space (LEPS) skill score of 41.06% was attained for the six models (**Table 4.30**). Since the LEPS skill score value for all the regression models are positive, it means that the models output (forecast) are much better than climatology. Results of the one sample Kolmogorov-Smirnov test analysis shows that the residuals from the cross-validated MLR models are normally distributed. The computed Durbin-Watson statistic indicates that the residuals from the cross-validated MLR models over western Kenya and most parts of Uganda (sub-region 2) and northern and central Tanzania (sub-region 5) had negative autocorrelation (the value is greater than 2) while the rest of the study area had positive autocorrelation (the value is less than 2). Comparison with the tabulated critical values by Farebrother (1980) shown that the residuals over the northern sector of the study were not significantly autocorrelated. Over the rest of Kenya, northern, eastern and central Tanzania (sub-regions 1, 2 and 3), the significant test was inconclusive while the residuals over the southern Uganda and western Tanzania had positive first-order

autocorrelation. This means that the regression models could be improved further by adding an autoregressive term. The assessment of the cross-validated MLR models and the residual analysis show that these models are robust and can be incorporated for operational uses.

In the subsequent sub-sections, only the skill scores of the cross-validated MLR model will be discussed since it is these models that should be used for operational forecasting work. Also only those sub-regions for which the multiple correlation coefficient between the time series of observed and cross-validated MLR model output is equal or greater than 0.5 will be discussed since only such models can be used for operational forecasting (Philippon *et al.*, 2009).

4.6.1.2 Number of wet days in a season

Figures 4.64a–f show the time series plots from the developed and cross-validated MLR models and the actual observations for the number of wet days in a season (NW) while **Table 4.31** shows the combination of predictors used and the skill of the models. The figures show that the models developed capture the peaks quite well but not so well for the lows. From a list of five predictors, three combinations were adequate to describe the interannual variability of the number of wet days over the six sub-regions (**Table 4.31**). BoBEN, SWAFRC and SINDS were each picked in five out of the six MLR models. Two predictors, BoBEN and SINDS were adequate to describe the interannual variability of number of wet days over western sector of the study area (sub-regions 2 and 6).

The multiple correlation coefficient between the cross-validated MLR model outputs and the actual observations of the number of wet days for the six sub-regions range from 0.60 to 0.70, with an average of 0.65 (**Table 4.31**). According to the LEPS skill score, an average value of 37.2% was obtained for the six cross-validated MLR models. It should be observed that the skill of the multiple correlation coefficient and the LEPS skill score for the cross-validated MLR models for the number of wet days are comparable to those obtained for the seasonal rainfall totals (**Table 4.30**) though slightly lower. This is consistent with the spatial coherence results which showed that the two are almost equally potentially predictable (**Figures 4.19b** and **4.20**).

The residuals from the six cross-validated MLR models are normally distributed according to one sample Kolmogorov-Smirnov test. The Durbin–Watson statistic over southern Uganda and western Tanzania (sub-region 6) indicates that the residuals from the cross-validated

MLR models have negative autocorrelation while the rest of the study area had positive autocorrelation. Compared to the tabulated critical values, the residuals from the cross-validated MLR models over western block of the study area (sub-regions 2, 5 and 6) were not significantly autocorrelated. For the eastern block of the study area (sub-regions 1, 3 and 4), the test for significant autocorrelation was inconclusive.

Table **4.31: The list of predictors' combination and skill of regression models for areal-averaged number of wet days during the short rainfall season**

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	BoBEN, SWAFRC, SINDS	0.770	0.697	0.442	41.84	1.43
2	BoBEN, SINDS	0.695	0.627	0.359	33.17	1.86
3	BoBEN, SWAFRC, SINDS	0.746	0.649	0.372	34.15	1.50
4	BoBEN, SWAFRC, SINDS	0.788	0.698	0.443	43.26	1.55
5	BoBEN, SWAFRC, SINDS	0.689	0.625	0.338	35.11	1.84
6	BoBEN, SINDS	0.689	0.598	0.322	35.49	2.04
Mean Value		**0.729**	**0.649**		**37.17**	

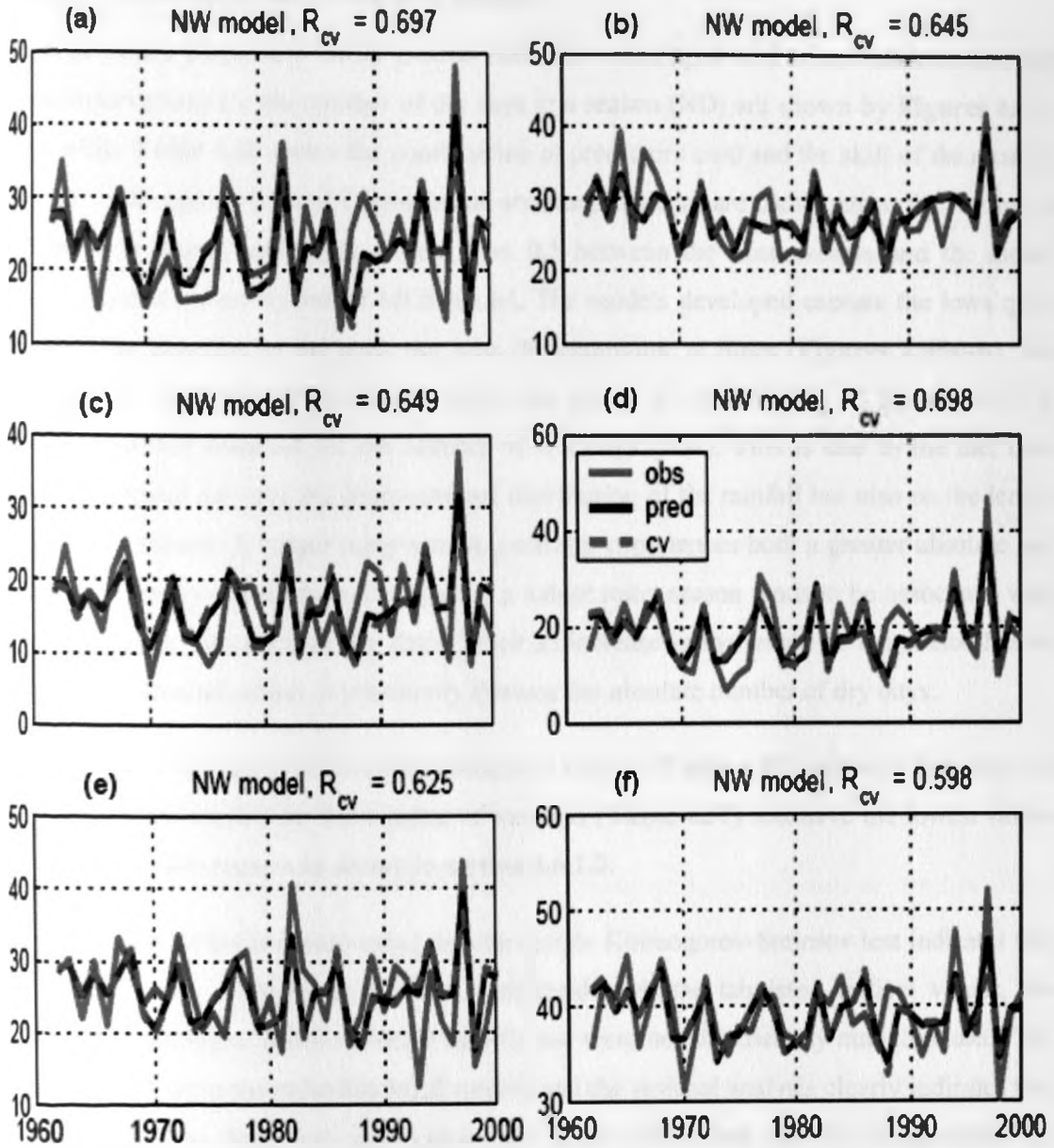


Figure 4.64: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for October-November-December areal-averaged number of wet days (NW) over (a) Central highlands and southeastern lowlands of Kenya, (b) Western Kenya and most parts of Uganda, (c) Northeastern Kenya, (d) Coastal strip of Kenya and Tanzania, (e) Central and northern Tanzania, and (f) Western of Lake Victoria and western Tanzania. R_{cv} shows the multiple correlation coefficient for cross-validated model

4.6.1.3 Number of dry days in a season

The time series plots from MLR models that were developed and cross-validated and the actual observations for the number of dry days in a season (ND) are shown by **Figures 4.65a** and **b** while **Table 4.32** shows the combination of predictors used and the skill of the models. Only two sub-regions (inland Tanzania and southern Uganda) are shown since the rest had a multiple correlation coefficient of less than 0.5 between the observations and the model outputs from the cross-validated MLR model. The models developed capture the lows quiet well and the direction of the peak but miss the magnitude at times (**Figures 4.65a–b**). We observe that much lower prediction skills was found for the number of dry days (ND), compared to that obtained for the number of wet days (NW). This is due to the fact both variables depend not only the intra-seasonal distribution of the rainfall but also on the length of the rainy season. A longer rainy season generally experiences both a greater absolute and relative frequency of rain days. By contrast a longer rainy season tends to be associated with a lesser relative frequency of dry days; hence an increase in the length of the season has an inverse (mechanical) effect to potentially increase the absolute number of dry days.

The good skill of these models over sub-regions 5 and 6 (**Table 4.32**) indicates that they can complement the models for the number of wet days (**Table 4.31**) that have the lowest values over the same sub-regions as shown in **section 4.6.1.2**.

An assessment of the residuals using the one sample Kolmogorov-Smirnov test indicates that the residuals are normally distributed. Compared with the tabulated critical values, the residuals according to Durbin-Watson statistic test were not significantly autocorrelated. The assessment of the cross-validated MLR models and the residual analysis clearly indicates that the models over these two sub-regions are quite robust and can be incorporated for operational uses.

Table 4.32: The list of predictors' combination and skill of regression models for areal-averaged number of dry days during the short rainfall season

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
5	SWAFRC, SINDS	0.739	0.691	0.448	30.48	1.79
6	BoBEN, SWHAW, NINO3.4	0.748	0.682	0.420	40.11	2.17

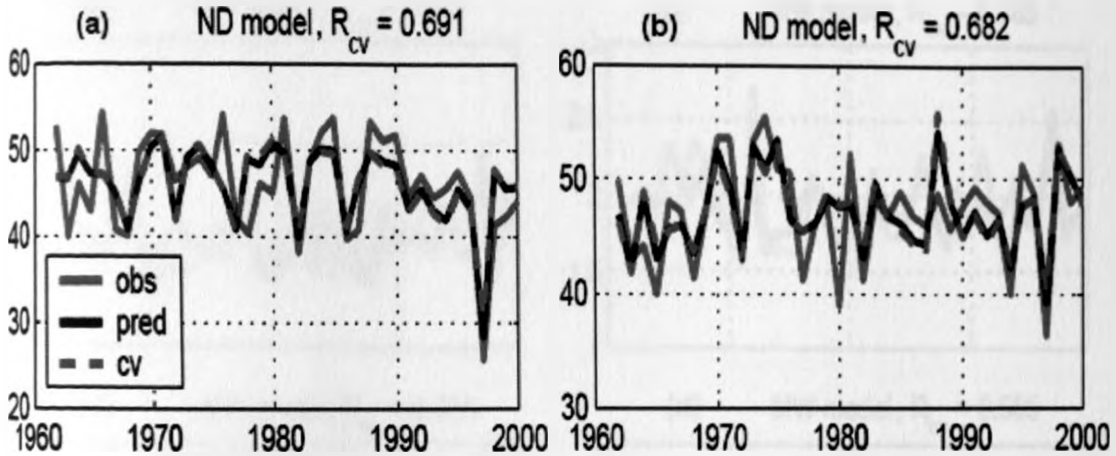


Figure 4.65: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for October-November-December areal-averaged number of dry days (ND) over (a) Central and northern Tanzania, and (b) Western of Lake Victoria and western Tanzania. R_{cv} shows the multiple correlation coefficient for cross-validated model

4.6.1.4 Mean length of wet spells

The time series plots from the developed and cross-validated MLR models and the actual observations for the mean duration of wet spells (MW) are shown by **Figures 4.66a–f** while **Table 4.33** shows the combination of predictors used and the skill of the models.

The multiple correlation coefficient for the cross-validated models are high (**Table 4.33**), though lower than for the seasonal rainfall totals (**Table 4.30**), ranging between 0.57 and 0.71, with an average of 0.63. Most of the additional potential predictors (**Table 4.33**) picked by these models are similar to the ones retained for seasonal rainfall totals (**Table 4.30**) and number of wet days in a season (**Table 4.31**). According to the LEPS skill score, an average value of 37.3% was obtained for the six MLR models. The lowest multiple correlation coefficients and the lowest LEPS skill scores were obtained over the western sector of the study area (sub-regions 2 and 6), which was closely followed by the coastal strip of Kenya and Tanzania (sub-region 4).

One sample Kolmogorov-Smirnov test indicates that the cross-validated MLR model residuals are normally distributed for the six sub-regions. The significant test for autocorrelation of residuals from the cross-validated MLR models over coastal strip of Kenya and Tanzania (sub-region 4) and central Kenya and southeastern lowlands (sub-region 1) was inconclusive. The residuals from cross-validated MLR models over the rest of the study area were not significantly autocorrelated.

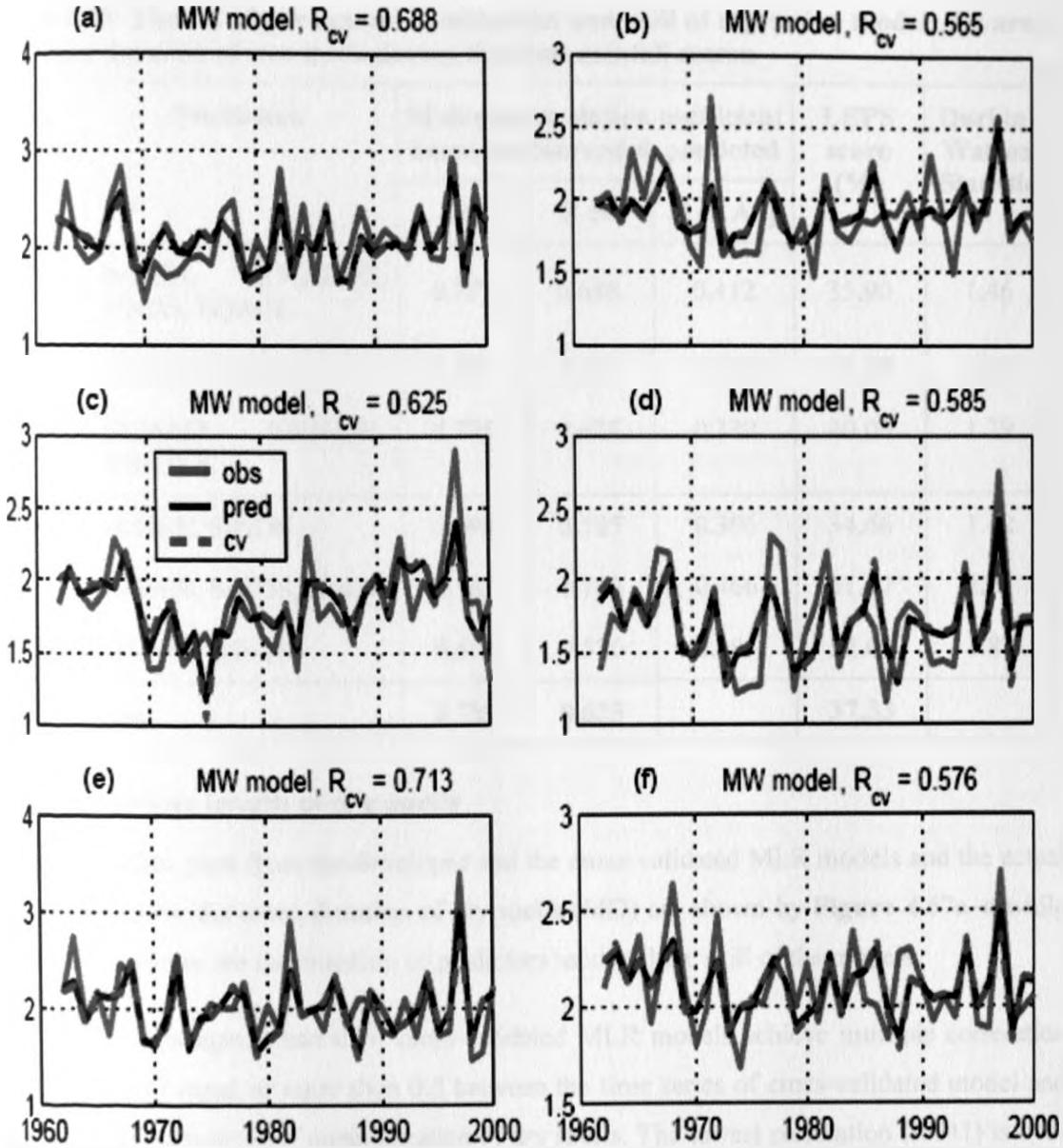


Figure 4.66: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for October-November-December areal-averaged duration of wet spells (MW) over (a) Central highlands and southeastern lowlands of Kenya, (b) Western Kenya and most parts of Uganda, (c) Northeastern Kenya, (d) Coastal strip of Kenya and Tanzania, (e) Central and northern Tanzania, and (f) Western of Lake Victoria and western Tanzania. R_{cv} shows the multiple correlation coefficient for cross-validated model

Table 4.33: The list of predictors' combination and skill of regression models for areal-averaged duration of wet spells during the short rainfall season

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	BoBEN, SWAFRC, SINDS, EQAFR	0.757	0.688	0.412	35.90	1.46
2	BoBEN, SINDS	0.633	0.565	0.282	30.19	1.93
3	ECMAD, SWHAW, NINO3.4	0.725	0.626	0.339	40.06	1.79
4	BoBEN, SINDS	0.698	0.585	0.306	34.66	1.42
5	BoBEN, SINDS, ZIND	0.783	0.713	0.466	51.17	2.25
6	BoBEN, SINDS	0.660	0.576	0.295	32.01	1.87
Mean Value		**0.709**	**0.625**		**37.33**	

4.6.1.5 Mean length of dry spells

The time series plots from the developed and the cross-validated MLR models and the actual observations for the mean duration of dry spells (MD) are shown by **Figure 4.67a–c** while **Table 4.34** shows the combination of predictors used and the skill of the models.

Only three sub-regions had their cross-validated MLR models achieve multiple correlation coefficients of equal or more than 0.5 between the time series of cross-validated model and the actual observations of mean duration of dry spells. The lowest correlation (0.541) is over southern Uganda and western Tanzania (sub-region 6) as shown by **Figure 4.67c**, followed by central Kenya and southeastern lowlands (sub-region 1) with 0.604 as shown by **Figure 4.67a**, and the highest (0.672) over the coastal strip of Kenya and Tanzania (sub-region 4) as shown by **Figure 4.67b**. The LEPS skill scores are relatively good for the three cross-validated MLR models (**Table 4.34**). The lowest LEPS skill score was again obtained over southern Uganda and western Tanzania. It is interesting to know that the average duration of dry spells over southern Uganda and western Tanzania is influenced by NIÑO3.4 index alone.

An assessment using the one sample Kolmogorov-Smirnov test indicates that the residuals are normally distributed. The Durbin-Watson statistic shows that the residuals autocorrelation were not significant over the three sub-regions, which means that the residuals were independent of each other.

The skill assessment of the cross-validated MLR models and analysis of the cross-validated MLR model residuals show that these models perform much better than climatology, are very robust and can therefore be incorporated for operational uses.

Table 4.34: The list of predictors' combination and skill of regression models for areal-averaged duration of dry spells during the short rainfall season

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	BoBEN, SINDS	0.669	0.604	0.330	35.95	1.73
4	BoBEN, SWHAW, SWAFRC	0.738	0.672	0.405	40.26	2.10
6	NIÑO3.4	0.613	0.541	0.274	30.42	2.03

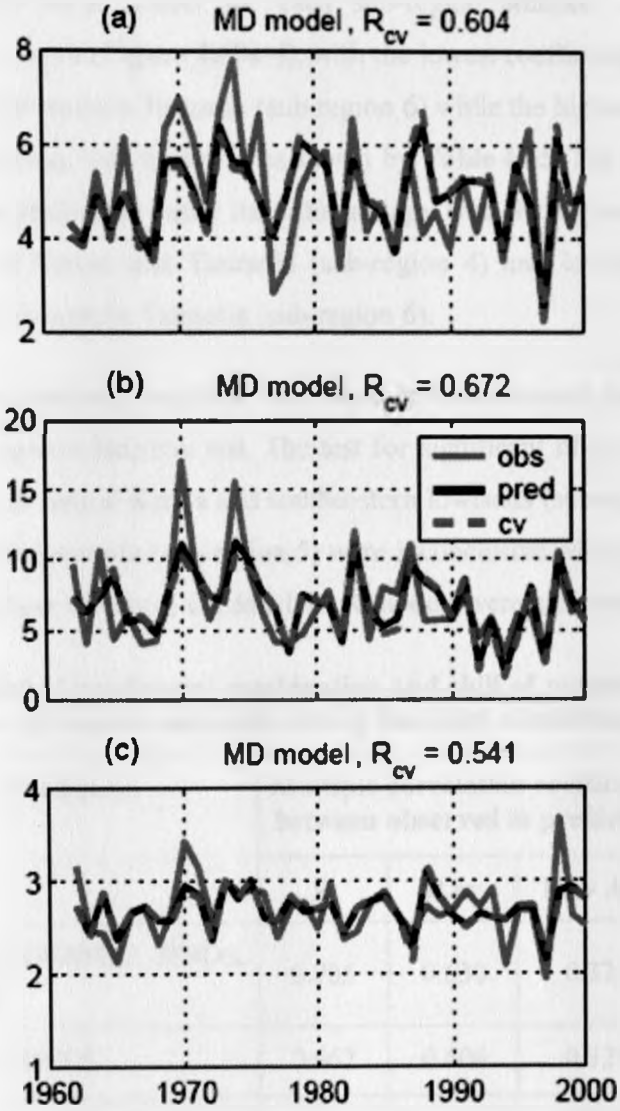


Figure 4.67: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for October-November-December areal-averaged duration of dry spells (MD) over (a) Central highlands and southeastern lowlands of Kenya, (b) Coastal strip of Kenya and Tanzania and (c) Western of Lake Victoria and western Tanzania. R_{cv} shows the multiple correlation coefficient for cross-validated model

4.6.1.6 Duration of longest wet spell

Table 4.35 shows the combination of predictors used and the skill of the models while **Figures 4.68a–f** shows the time series plots from developed and cross-validated MLR models and the actual observations for duration of longest wet spells (LW) during the short rainfall season.

The cross-validated MLR model for each sub-region attained a multiple correlation coefficient of 0.5 or more (**Figure 4.68a–f**), with the lowest coefficient (0.526) obtained over southern Uganda and western Tanzania (sub-region 6) while the highest (0.633) was obtained over northeastern Kenya (sub-region 3) as shown by **Table 4.35**. The LEPS skill score show that all the models performed better than climatology with the highest score (36.89%) over the coastal strip of Kenya and Tanzania (sub-region 4) and lowest skill (26.62%) over southern Uganda and western Tanzania (sub-region 6).

The residuals from the cross-validated MLR models had a normal distribution according to one sample Kolmogorov-Smirnov test. The test for significant of the autocorrelation of the model residuals over central Kenya and southeastern lowlands (sub-region 1) as well as over northern and central Tanzania (sub-region 5) were inconclusive while in the rest of the study area, the residuals from the cross-validated MLR models were independent.

Table 4.35: The list of predictors' combination and skill of regression models for areal-averaged duration of longest wet spells during the short rainfall season

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	BoBEN, SWAFRC, SINDS, EQAFR	0.705	0.630	0.326	34.33	1.60
2	BoBEN, SINDS	0.662	0.604	0.329	31.41	1.93
3	ECMAD, SINDS, MARCON	0.717	0.633	0.350	31.49	2.10
4	BoBEN, SWAFRC, SINDS	0.711	0.623	0.336	36.89	1.66
5	BoBEN, SINDS	0.663	0.586	0.307	27.62	1.90
6	BoBEN, SINDS	0.602	0.526	0.236	26.62	1.99
Mean Value		**0.677**	**0.600**		**31.39**	

4.6.1.7 Duration of longest dry spell

The time series plots from the developed and cross-validated MLR models as well as the actual observations for duration of longest dry spells (LD) are shown by **Figures 4.69a–c** while the combination of the predictors used and the skill of the cross-validated MLR models are shown by **Table 4.36**.

Only three sub-regions had the multiple correlation coefficient between the actual observations and cross-validated MLR time series of 0.5 or greater (**Figures 4. 69a–c**). These sub-regions were central highlands and southeastern lowlands of Kenya (sub-region 1), western Kenya and most parts of Uganda (sub-region 2) and coastal strip of Kenya and Tanzania (sub-region 4). The Linear Error in Probability Space (LEPS) skill score indicates that these cross-validated MLR models were better off than climatology (**Table 4.36**). However, the skills are lower than for most other variables, which are confirmed by some disagreements between the observed and predicted values (**Figures 4.69a–c**).

Residuals analysis indicates that the residuals for each cross-validated MLR model were normally distributed according to one sample Kolmogorov-Smirnov test. The Durbin-Watson statistic indicate that the residuals from the cross-validated MLR models over central Kenya and southeastern lowlands of Kenya (sub-region 1) and western Kenya and most parts of Uganda (sub-region 2) had positive autocorrelation while those over the coastal strip of Kenya and Tanzania (sub-region 4) had negative autocorrelation. The test of significance of the autocorrelation shows that the residuals over eastern block of the study area, south of Equator (sub-regions 1 & 4) were independent. However, the test was inconclusive over western Kenya and most parts of Uganda (sub-region 2).

The skill assessment of the cross-validated MLR models and the residual analysis from these models clearly indicate that the three models are robust and better off than climatology, and the residuals are independent of each other. Hence these cross-validated MLR models can therefore be incorporated for operational uses.

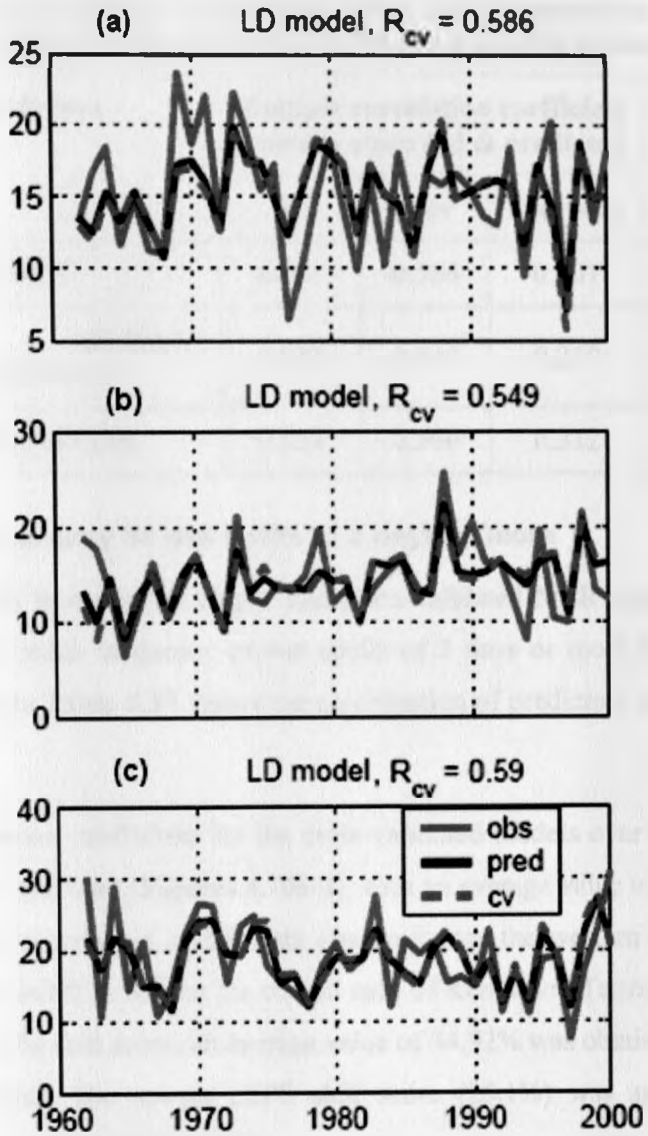


Figure 4.69: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for October-November-December areal-averaged duration of longest dry spell (LD) over (a) Central highlands and southeastern lowlands of Kenya, (b) Western Kenya and most parts of Uganda and (c) Coastal strip of Kenya and Tanzania. R_{cv} shows the multiple correlation coefficient for cross-validated model

Table 4.36: **The list of predictors' combination and skill of regression models for areal-averaged duration of longest dry spells during the short rainfall season**

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	BoBEN, SINDS	0.654	0.586	0.307	29.59	1.62
2	BoBEN, SWHAW, WCAUS, NIÑO3.4	0.687	0.549	0.219	31.07	1.42
4	SWHAW, MARCON	0.654	0.590	0.312	31.75	2.31

4.6.1.8 Mean frequency of wet spells of 3 days or more

The time series plots from the developed and cross-validated MLR models and the actual observations for the mean frequency of wet spells of 3 days or more (3W) are shown by **Figures 4.70a–f** while **Table 4.37** shows the combination of predictors used and the skill of the models.

The multiple correlation coefficient for the cross-validated models over the six sub-regions ranges between 0.57 and 0.69 (**Figures 4.70a–f**), with an average value of 0.62 (**Table 4.37**). The lowest multiple correlation coefficients observed over the western sector of the study area (**Figures 4.70b** and **f**) as well as the coastal strip of Kenya and Tanzania (**Figure 4.70d**). According to the LEPS skill score, an average value of 34.92% was obtained for the six MLR cross-validated models. The lowest LEPS skill score (25.1%) was again observed over western Kenya and most parts of Uganda.

According to one sample Kolmogorov-Smirnov test, the residuals from these cross-validated MLR models are normally distributed. The Durbin-Watson statistic shows that the residuals from the cross-validated MLR models over central Kenya and southeastern lowlands of Kenya (sub-region 1) and coastal strip of Kenya and Tanzania (sub-region 4) had negative autocorrelation while those from the rest of the study area had positive autocorrelation. The test for significance of the residuals autocorrelation was inconclusive over the northeastern Kenya (sub-region 3) while over the rest of the study area, the residuals from the cross-validated MLR models were independent of each other.

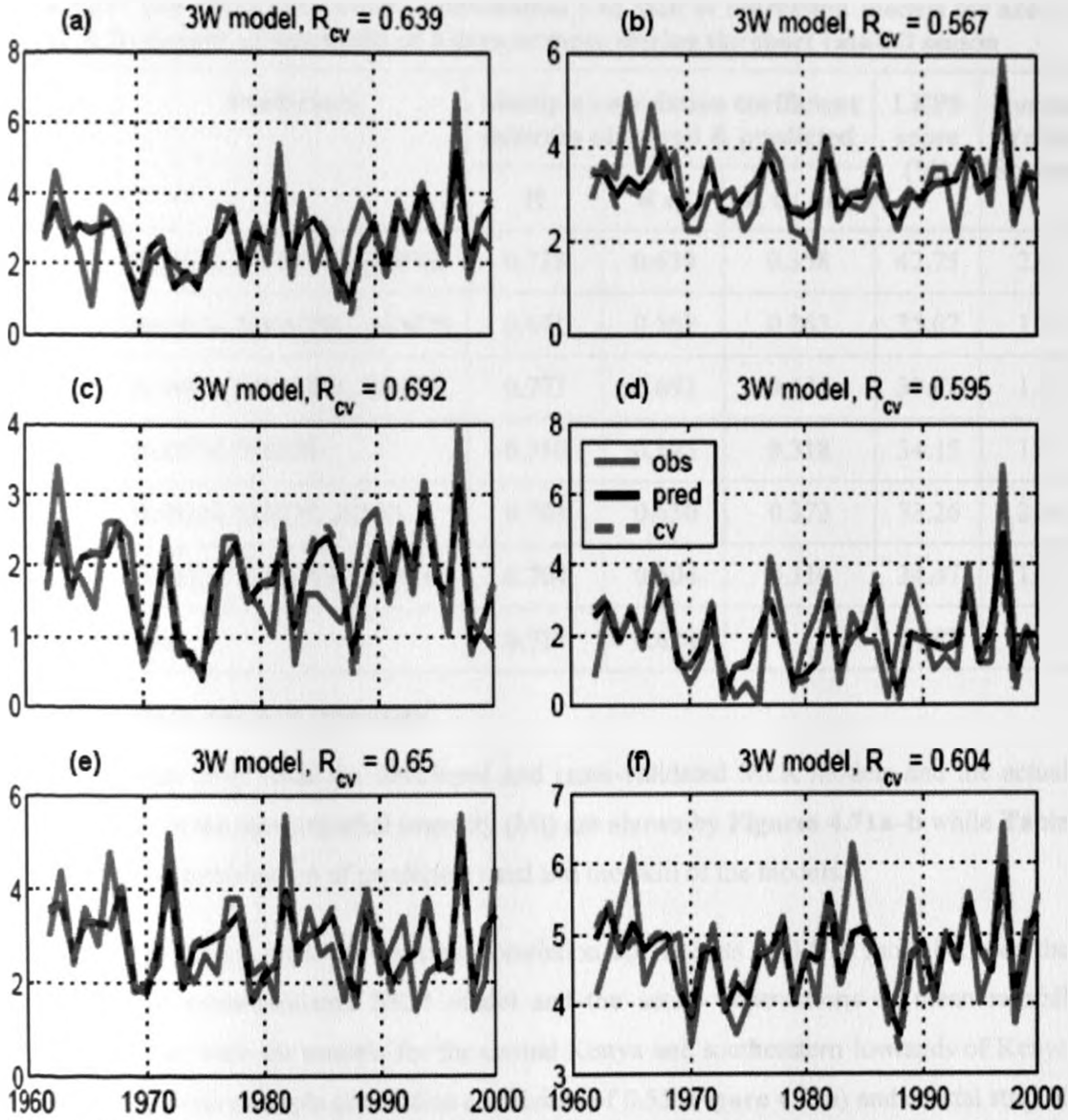


Figure 4.70: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for October-November-December areal-averaged frequency of wet spells of 3 days or more (3W) over (a) Central highlands and southeastern lowlands of Kenya, (b) Western Kenya and most parts of Uganda, (c) Northeastern Kenya, (d) Coastal strip of Kenya and Tanzania, (e) Central and northern Tanzania, and (f) Western of Lake Victoria and western Tanzania. R_{cv} shows the multiple correlation coefficient for cross-validated model

Table 4.37: The list of predictors' combination and skill of regression models for areal-averaged frequency of wet spells of 3 days or more during the short rainfall season

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	BoBEN, SWHAW, EQIND	0.738	0.639	0.358	42.75	2.06
2	BoBEN, SWAFRC, SINDS	0.671	0.567	0.263	25.07	1.75
3	BoBEN, SWHAW, SINDS	0.773	0.692	0.435	39.81	1.45
4	BoBEN, SINDS	0.710	0.595	0.318	34.15	1.67
5	BoBEN, SINDS, ZIND	0.703	0.650	0.372	33.26	2.34
6	BoBEN, SWHAW, EQIND	0.704	0.604	0.310	34.47	1.81
Mean Value		**0.717**	**0.624**		**34.92**	

4.6.1.9 Mean rainfall intensity

The time series plots from the developed and cross-validated MLR models and the actual observations for the mean rainfall intensity (MI) are shown by **Figures 4.71a–b** while **Table 4.38** shows the combination of predictors used and the skill of the models.

Only two sub-regions achieved multiple correlation coefficients of 0.5 or more between the time series of cross-validated MLR model and the actual observations of mean rainfall intensity. These were the models for the central Kenya and southeastern lowlands of Kenya (sub-region 1) with multiple correlation coefficient of 0.55 (**Figure 4.71a**) and coastal strip of Kenya and Tanzania (sub-region 4) with multiple correlation coefficient of 0.682 (**Figure 4.71b**). The LEPS skill score was 28.08% for the model of sub-region 1 and 36.01% for sub-region 4 (**Table 4.38**). These sub-regions are close to the Indian Ocean. The hypothesis is that SST off East Africa exerts a direct control on the intensity of local convective activity, whereas inland the intensity of the rains is more random. The skill for the coast is particularly high, which was not necessarily expected for a variable which has low spatial coherence and is considered to be less predictable than the occurrence of the rains.

Results indicate that the residuals for each model were normally distributed according to one sample Kolmogorov-Smirnov test. The Durbin-Watson statistics indicate that the residuals from the cross-validated MLR model over the central Kenya and southeastern lowlands of

(coastal strip of Kenya and Tanzania) had negative (positive) autocorrelation. The test of significance of the autocorrelation shows that the residuals over central Kenya and southeastern lowlands of Kenya were independent of each other while the test was inconclusive for the residuals over the coastal strip of Kenya and Tanzania.

Table 4.38: The list of predictors' combination and skill of regression models for areal-averaged rainfall intensity during the short rainfall season

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	BoBEN, SINDS	0.591	0.553	0.267	28.08	2.14
4	BoBEN, SWHAW, EQIND	0.760	0.682	0.419	36.01	1.51

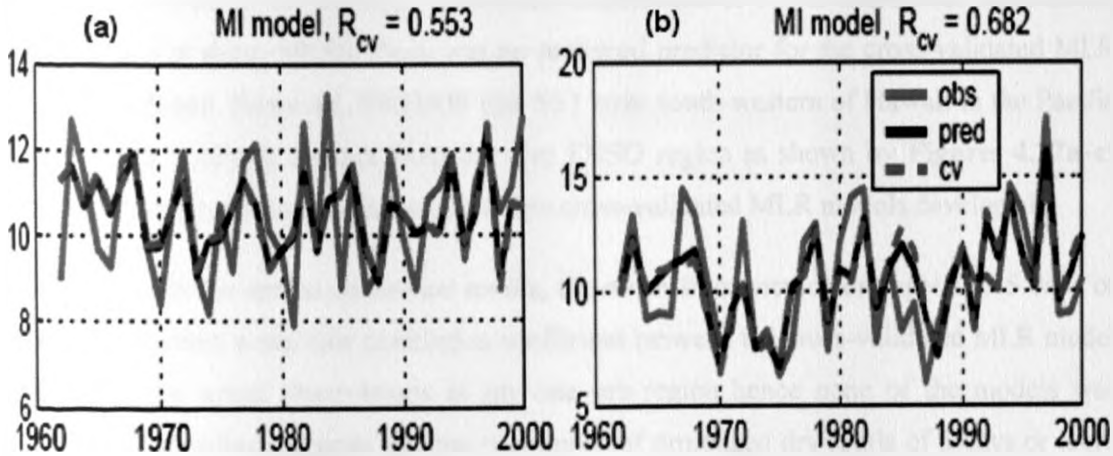


Figure 4.71: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for October-November-December areal-averaged rainfall intensity over (a) Central highlands and southeastern lowlands of Kenya, and (b) Coastal strip of Kenya and Tanzania. R_{cv} shows the multiple correlation coefficient for cross-validated model

As a conclusion to the development of prediction models for the OND season, it was found that for most variables and sub-regions, it was possible to produce skill models (the multiple correlation coefficients between the areal-averaged observations and the cross-validated MLR model output equal or greater than 0.5). The residuals from the cross-validated MLR models were normally distributed according to one sample Kolmogorov-Smirnov test. The significance test of the calculated Durbin-Watson statistics against those tabulated further shown that for most models, the residuals were independent from each other. Occasionally,

the test was inconclusive. This occurs when the calculated value lies between the lower and upper boundary of the critical values.

The cross-validated MLR models for sub-regional intraseasonal statistics (SRISS) of wet spells and seasonal rainfall totals picked the predictors around Bay of Bengal mostly. These predictors are BoBEN (the SST over the Bay of Bengal extending to west coast of Malaysia and Indonesia as shown by **Figures 4.26a–c**) and SINDS (the zonal wind component at 925mb to the south of the Bay of Bengal and near the southern tip of India as shown by **Figures 4.30a–c**). This shows that oceanic and atmospheric conditions during the July-August period around the Bay of Bengal and largely associated Asian monsoon dynamics provide a lot of predictive information for the SRISS of wet spells and seasonal rainfall totals during the short rainfall season.

In the case of the SRISS of dry spells and mean rainfall intensity, the cross-validated MLR models did not attain the multiple correlation coefficients of 0.5 or more over most of the sub-regions. For these SRISS, there was no preferred predictor for the cross-validated MLR models developed. However, SWHAW (the SST over south-western of Hawaii in the Pacific Ocean which is clearly distinct from the core ENSO region as shown by **Figures 4.27a–c**) was the frequently picked predictor by the few cross-validated MLR models developed.

Consistent with the spatial coherence results, the mean frequency of dry spells of 5 days or more did not attain a multiple correlation coefficient between the cross-validated MLR model output and the actual observations at any one sub-region hence none of the models was shown. This therefore suggests that the occurrence of prolonged dry spells of 5 days or more could be mainly influenced by local factors. The large scale climate fields are modified by the local factors such that they loose most of their properties thus cannot be used to predict this statistic with a sufficiently good skill for a lead-time of one month.

4.6.2 Regression models during the long rainfall season

As already mentioned, the time series for the long rainfall season was split into two parts due to the low temporal homogeneity of this season, which results into a low spatial coherence for the SRISS obtained during the long rainfall season (**Figures 4.19a** and **4.20**) and to insignificant lagged correlations with predefined SST predictors (**Figures 4.39a–j** and **4.51a–j**). The first part constitute the earlier months of March and April, while the second part was for the month of May.

For the purpose of developing multivariate linear regression (MLR) models, the predefined and additional potential predictors considered for the earlier part of the season were averaged for the months of December and January. For the later part, the months of January and February were averaged. The rationale behind the use of the February predictors to update the long rainfall season forecast was that previous studies have shown that the mean rainfall onset date for Kenya and northeastern Tanzania is around 12 – 16 March (Alusa and Mushi, 1974), around 22 – 26 March (Asnani, 1993), 25 March (Camberlin and Okoola, 2003). The small discrepancies between the mean onset dates by Alusa and Mushi (1974) and those of Asnani (1993) and Camberlin and Okoola (2003) were attributed to a small trend in recent years, toward a delayed onset of the rains (Camberlin and Okoola, 2003). The atmospheric and oceanic fields used in this study at the monthly timescale are released towards the second week of the following month. The update will thus be issued on time before the actual onset of the rainfall occurs. Alternatively most of the predictors from Hadley Centre and ECMWF can be forecasted with a one month lead in which case, they would be available and used in the regression models before the start of the season.

4.6.2.1 Regression models during the March-April period

The list of the additional potential predictors from which the multivariate linear regression (MLR) models were developed was shown in **Table 4.21**. Two other predictors from the predefined indices (MIB1 and MAB3) were also used. As already mentioned in **section 4.6.1.1**, only those MLR models for which the cross-validated multiple correlation coefficient is equal to 0.5 or more are discussed. Most of the SRISS did not attain this value. None of the SRISS of dry spells actually attained this value. Only the rainfall totals, number of wet days and mean frequency of wet spells of 3 days or more attained this value in two or more sub-regions out of the possible six, while the rest had this value in one sub-region or none. The cross-validated MLR models for the rainfall totals are discussed in the next section.

4.6.2.1.1 Rainfall totals

The time series plots from the developed and cross-validated MLR models as well as the actual observations for rainfall totals (SR) during the March-April period of the long rainfall season are shown by **Figures 4.72a–d** while the combination of the predictors used and the skill of the cross-validated MLR models are shown by **Table 4.39**.

Only four sub-regions had the multiple correlation coefficient of 0.5 or more between the actual observations and the cross-validated MLR model time series. These sub-regions were

central and western Kenya (**Figure 4.72a**), northeastern Kenya (**Figure 4.72b**), southeastern lowlands of Kenya and northeastern Tanzania (**Figure 4.72c**) and most parts of Uganda (**Figure 4.72d**). The Linear Error in Probability Space (LEPS) skill score indicates that these cross-validated MLR models were better off than climatology (**Table 4.39**).

Results of residuals analysis indicate that the residuals for each model were normally distributed according to one sample Kolmogorov-Smirnov test. The test of significance of the first-order autocorrelation according to Durbin-Watson statistic shows that the residuals from each model in the four sub-regions were independent.

The skill assessment of the cross-validated MLR models indicate that the four models are better off than climatology and robust while residual analysis results show that the residuals are normally distributed and are independent of each other. These models can therefore be incorporated for operational uses. However, the fact that not all the sub-regions had regression models developed and also the fact that the predictors picked often differ between the sub-regions suggests that care should be taken when using these models. A few anomalous years (for instance the 1993 drought in sub-regions 1 and 5 in **Figures 4.72a** and **c** respectively) are not detected by the models.

Table 4.39: The list of predictors' combination and skill of regression models for areal-averaged rainfall totals during the March-April period of the long rainfall season

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	BoBEN-1, NEGHA, EQAFR-1, MAB3	0.673	0.597	0.280	30.23	2.16
3	WCAUS-1, WINDO, EBBEN, SCEINDO	0.727	0.636	0.335	43.43	2.19
5	BoBEN-1, SCEINDO, EQAFR-1	0.688	0.594	0.298	38.20	2.11
6	CINDO, NINDS, MAB3	0.731	0.647	0.350	37.76	1.80

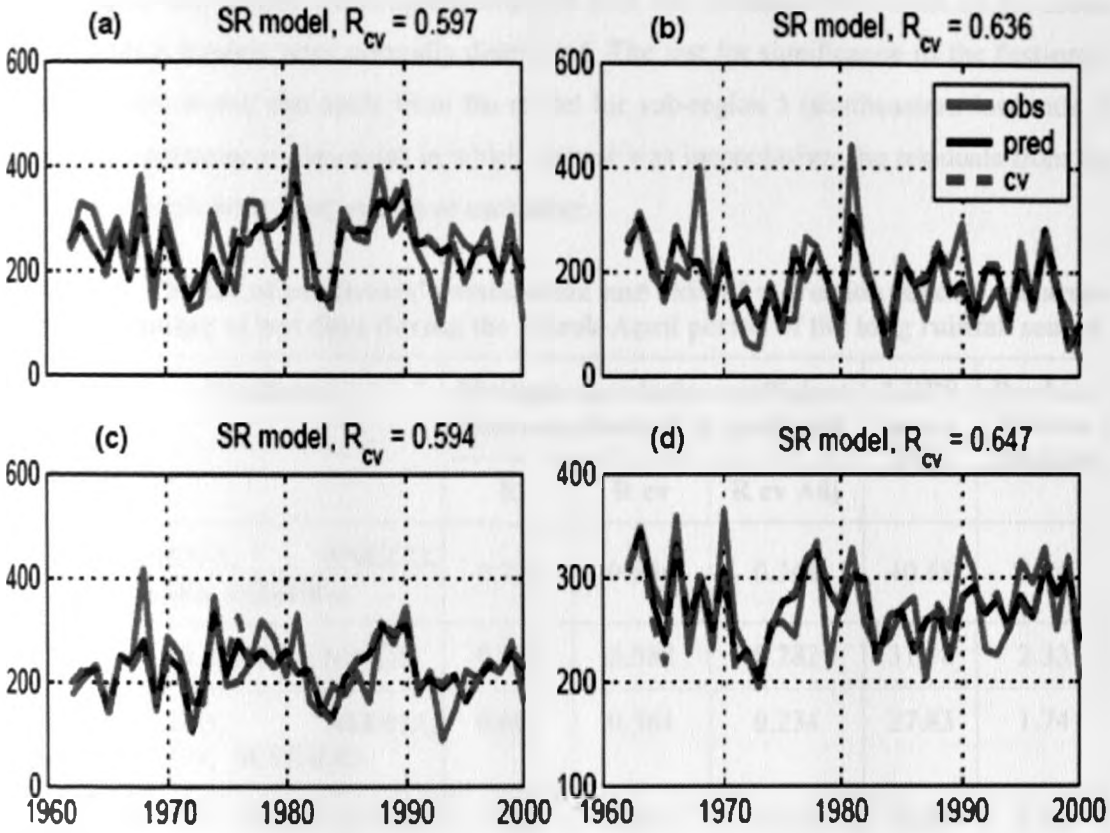


Figure 4.72: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for March-April areal-averaged rainfall totals (SR) over (a) central and western Kenya, (b) northeastern Kenya, (c) southeastern lowlands of Kenya and northeastern Tanzania, and (d) most parts of Uganda. R_{cv} shows the multiple correlation coefficient for cross-validated model

4.6.2.1.2 Number of wet days

Figures 4.73a–j show time series plots from the developed and cross-validated MLR models as well as the actual observations for the number of wet days (NW) during the March-April period while **Table 4.40** shows the combination of predictors used and the skill of the models.

Unlike the case of rainfall totals, the number of wet days had a multiple correlation coefficient value of 0.5 or greater over the six sub-regions (**Figures 4.73a–j**), with an average value of 0.58 (**Table 4.40**). The six cross-validated MLR models had an average LEPS skill score of 31.5%, which means they were better off than climatology. The predictors picked differ from one sub-region to the other indicating that they are unstable. However, CINDO appears in all models except for sub-region 3 (northeastern Kenya) as shown in **Table 4.40**.

According to one sample Kolmogorov-Smirnov test, the residuals from each of the cross-validated MLR models were normally distributed. The test for significance of the first-order autocorrelation shows that apart from the model for sub-region 5 (southeastern lowlands of Kenya and northeastern Tanzania) in which the test was inconclusive, the residuals from the rest of the models were independent of each other.

Table 4.40: The list of predictors' combination and skill of regression models for areal-averaged number of wet days during the March-April period of the long rainfall season

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	BoBEN-1, ANGCO, CINDO, EQAFR-1	0.724	0.656	0.364	40.68	2.26
2	WAFR, CINDO, NINDS	0.645	0.582	0.282	31.38	2.33
3	ANGCO, NEGHA, EBBEN, SCEINDO	0.691	0.561	0.234	27.83	1.74
4	EBBEN, CINDO, NINDS	0.641	0.539	0.229	26.86	2.01
5	BoBEN-1, SINDS-1, CINDO, EQAFR-1	0.725	0.560	0.233	32.71	1.70
6	WAFR, SINDS-2, CINDO	0.649	0.566	0.262	29.61	2.32
Mean Value		**0.679**	**0.577**		**31.51**	

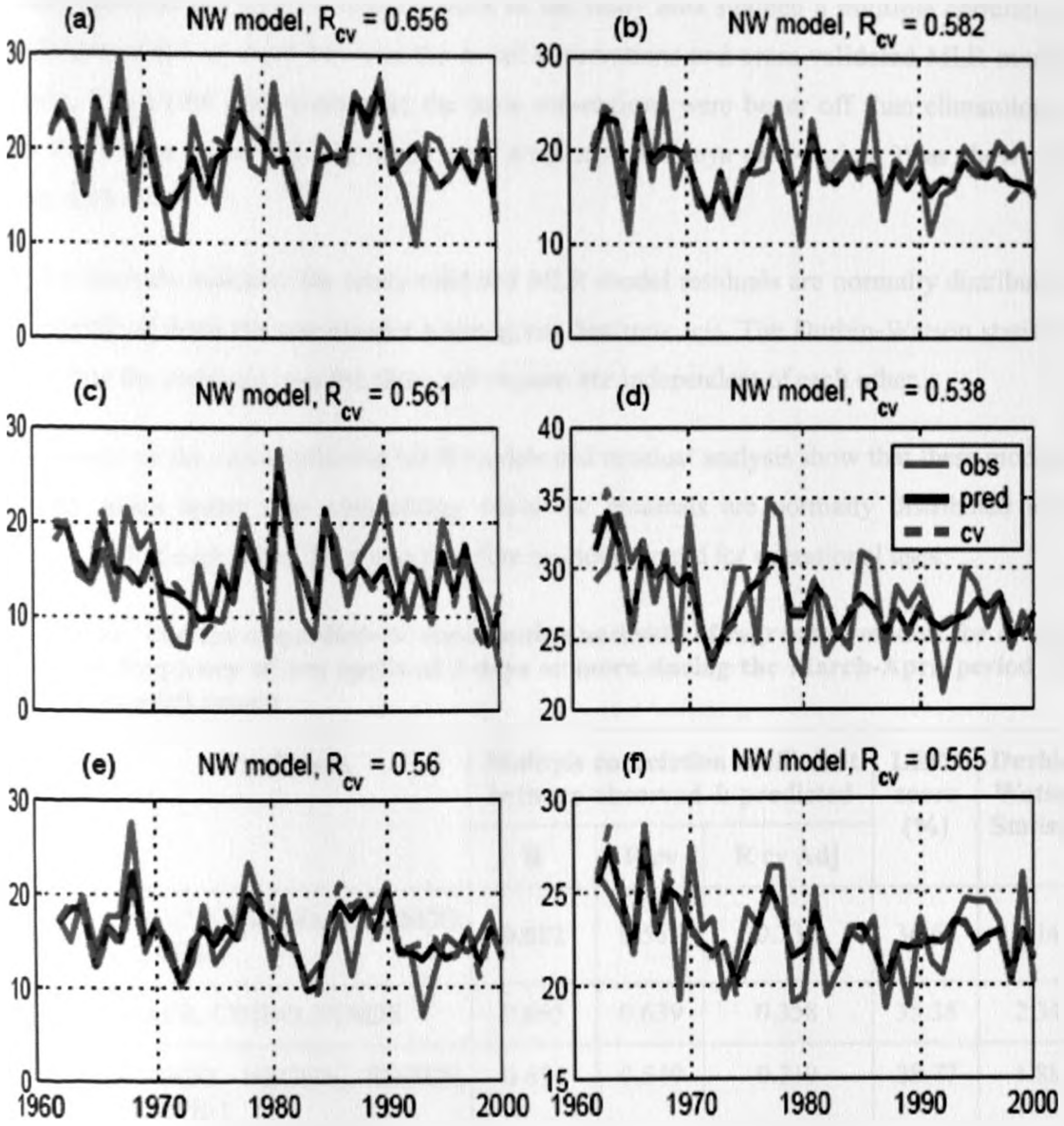


Figure 4.73: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for March-April areal-averaged number of wet days (NW) over (a) central and western Kenya, (b) coastal strip of Kenya and Tanzania, (c) northeastern Kenya, (d) western Tanzania and southern Uganda, (e) southeastern lowlands of Kenya and northeastern Tanzania, and (f) most parts of Uganda. R_{cv} shows the multiple correlation coefficient for cross-validated model

4.6.2.1.3 Mean frequency of wet spells of 3 days or more

Table 4.41 shows the combination of predictors used and the skill of the MLR models while Figures 4.74a–c show the time series plots from the developed and cross-validated MLR models, together with the actual observations for mean frequency of wet spells of 3 days or more (3W) during the March-April period of the long rainfall season. Only three sub-regions

which constitute the most of eastern block of the study area attained a multiple correlation coefficient of 0.5 or more between the actual observations and cross-validated MLR model outputs. The LEPS skill score over the three sub-regions were better off than climatology, with the lowest value (28.8%) noted over northeastern Kenya (sub-region 3) as shown in **Table 4.41.**

Further analysis indicates the cross-validated MLR model residuals are normally distributed as determined from the one sample Kolmogorov-Smirnov test. The Durbin-Watson statistic shows that the residuals over the three sub-regions are independent of each other.

Assessment of the cross-validated MLR models and residual analysis show that these models perform much better than climatology while the residuals are normally distributed and independent of each other. They can therefore be incorporated for operational uses.

Table 4.41: The list of predictors' combination and skill of regression models for areal-averaged frequency of wet spells of 3 days or more during the March-April period of the long rainfall season

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	ANGCO, CINDO, SCEINDO, EQAFR-1	0.682	0.565	0.239	36.60	2.14
2	WAFR, CINDO, NINDS	0.695	0.639	0.358	35.38	2.34
3	ANGCO, NEGHA, EBBEN, EQAFR-1	0.674	0.549	0.219	28.77	1.88

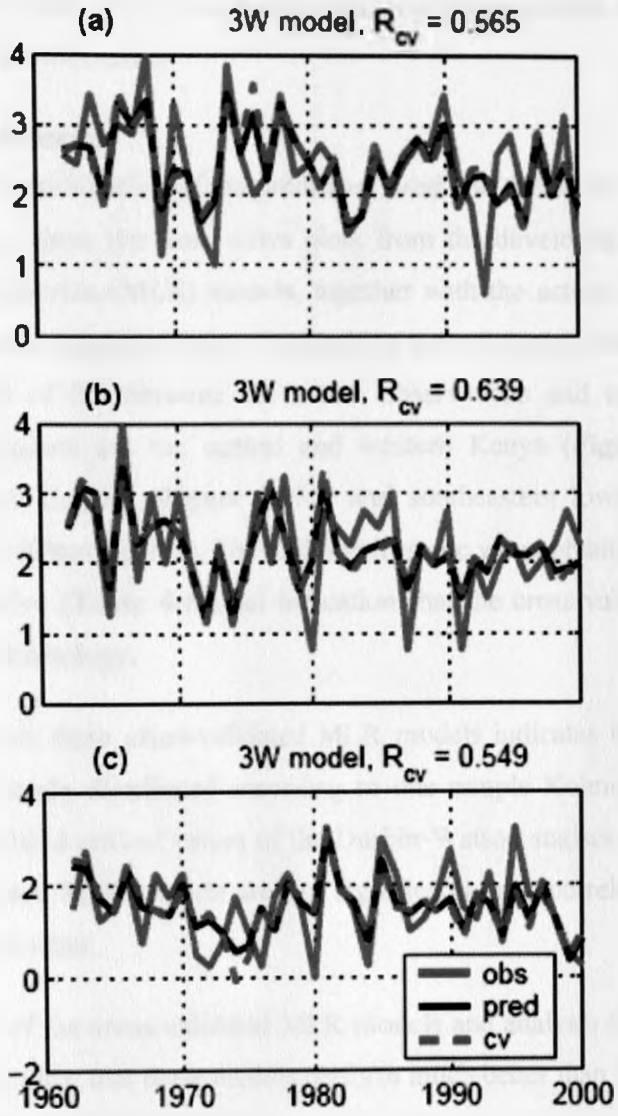


Figure 4.74: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for March-April areal-averaged frequency of wet spells of 3 days or more over (a) central and western Kenya, (b) coastal strip of Kenya and Tanzania, and (c) northeastern Kenya. R_{cv} shows the multiple correlation coefficient for cross-validated model

4.6.2.2 Regression models for the month of May

The IOD index and list of the additional potential predictors shown in **Table 4.24** were used to develop multivariate linear regression (MLR) models for the rainfall totals and SRISS for the month of May. The cross-validated MLR models for rainfall totals, number of wet days and mean frequency of wet spells of 3 days or more attained the threshold of multiple correlation coefficient of 0.5 or more with the actual observations in two or more sub-regions and hence are discussed in this section. None of the SRISS of dry spells attained the 0.5

correlation coefficient value. In the next section, the regression models for the rainfall totals for the month of May are discussed.

4.6.2.2.1 Rainfall totals

Table 4.42 shows the combination of the predictors used and the skills of the MLR models while **Figures 4.75a–c** show the time series plots from the developed and cross-validated multivariate linear regression (MLR) models, together with the actual observations for the rainfall totals (SR) for the month of May. Only half of the sub-regions did attain the multiple correlation coefficient of 0.5 between the actual observations and cross-validated MLR models. These sub-regions are the central and western Kenya (**Figure 4.75a**), western Tanzania and southern Uganda (**Figure 4.75b**) and southeastern lowlands of Kenya and northeastern Tanzania (**Figure 4.75c**). The LEPS skill score value obtained for the three sub-regions were all positive (**Table 4.42**), an indication that the cross-validated MLR models were better off than climatology.

Residuals analysis from these cross-validated MLR models indicates that the residuals for each model were normally distributed according to one sample Kolmogorov-Smirnov test. Compared to the tabulated critical values of the Durbin-Watson statistics, the residuals from the three cross-validated MLR models are not significantly autocorrelated, hence they are independent from each other.

The skill assessment of the cross-validated MLR models and analysis of the cross-validated MLR model residuals show that these models perform much better than climatology, are very robust and can therefore be incorporated for operational uses.

Table 4.42: The list of predictors' combination and skill of regression models for areal-averaged rainfall totals for the month of May

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
1	ECMAD-1, WCAUS-2, SMESEA	0.676	0.565	0.261	35.61	2.27
4	WCAUS-2, SSA, SMESEA	0.689	0.624	0.337	36.17	2.10
5	ECMAD-1, SSA, WCSOA, CSINDO	0.737	0.645	0.347	37.56	1.58

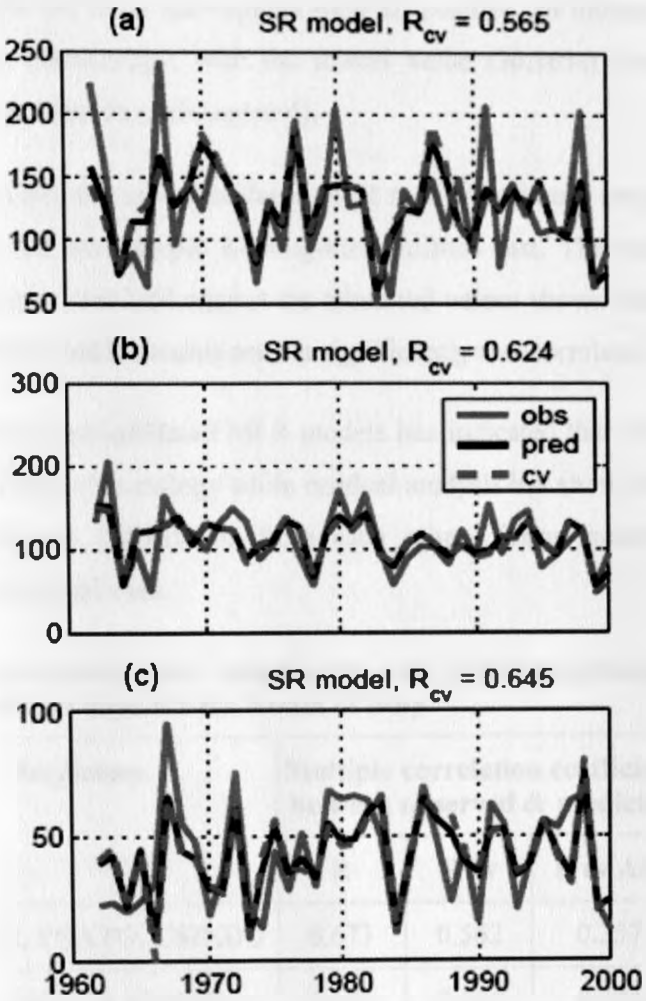


Figure 4.75: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for areal-averaged rainfall totals (SR) for the month of May over (a) central and western Kenya, (b) western Tanzania and southern Uganda, and (c) southeastern lowlands of Kenya and northeastern Tanzania. R_{cv} shows the multiple correlation coefficient for cross-validated model

4.6.2.2.2 Number of wet days

The time series plots from the developed and cross-validated MLR models as well as the actual observations for the number of wet days (NW) for the month of May are shown by **Figures 4.76a–c** while the combination of the predictors used and the skill of the cross-validated MLR models are shown by **Table 4.43**.

Only three sub-regions had the multiple correlation coefficient of 0.5 or more between the actual observations and the cross-validated MLR models time series. These sub-regions were western Tanzania and southern Uganda (**Figure 4.76a**) southeastern lowlands of Kenya and northeastern Tanzania (**Figure 4.76b**), and most parts of Uganda (**Figure4.76c**). The LEPS

skill score attained for the three sub-regions were all positive, an indication that the models were better off than climatology, with the lowest value (30.16%) obtained over western Tanzania and southern Uganda (sub-region 4).

Further analysis indicates the cross-validated MLR models residuals are normally distributed as determined from the one sample Kolmogorov-Smirnov test. The test of significance of Durbin-Watson statistics obtained against the tabulated values shown that the residuals from the three cross-validated MLR models are not significantly autocorrelated.

An assessment of the cross-validated MLR models has indicated that the three models were robust and better off than climatology while residual analysis has shown that the residuals are normally distributed and independent from each other. These models can therefore be incorporated for operational uses.

Table 4.43: The list of predictors' combination and skill of regression models for areal-averaged number of wet days for the month of May

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
4	WCAUS-2, EQATO, CSINDO	0.673	0.562	0.257	30.16	2.13
5	EQATO, WCSOA, CSINDO	0.689	0.614	0.324	37.12	1.90
6	IOD, SAFR, NEATO	0.643	0.579	0.278	34.36	2.19

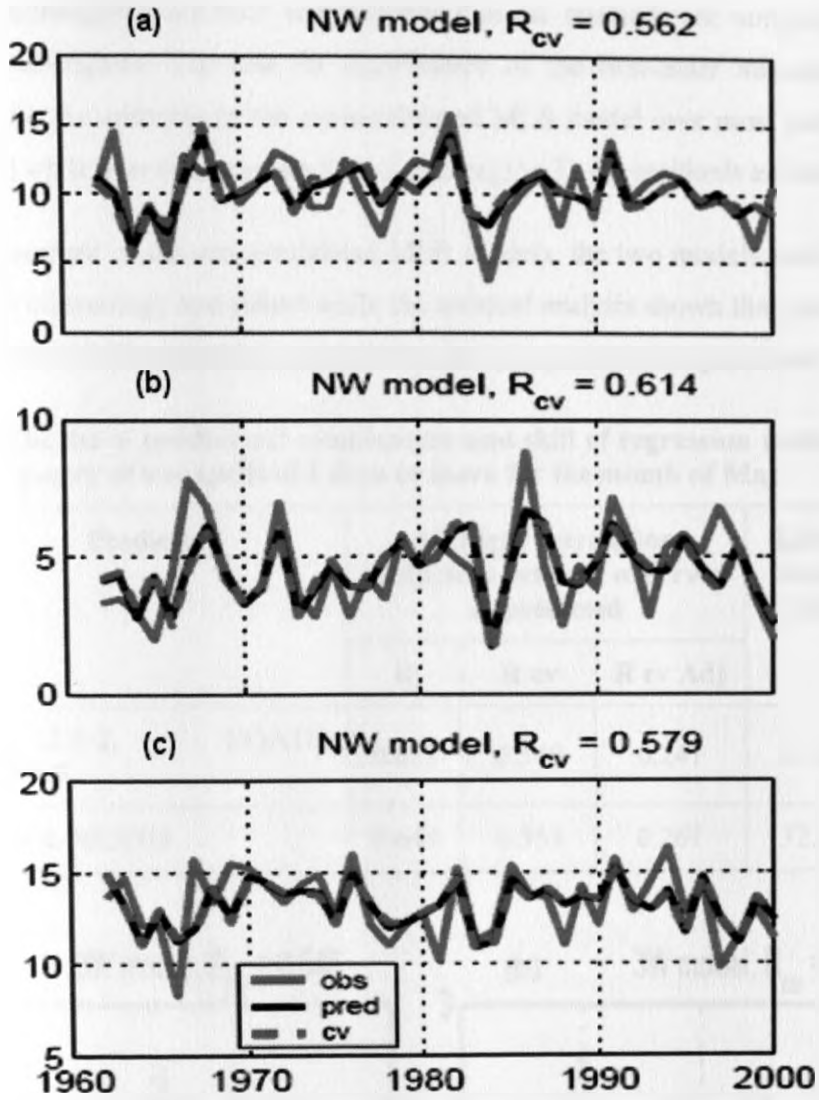


Figure 4.76: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for areal-averaged number of wet days (NW) for the month of May over (a) western Tanzania and southern Uganda, (b) southeastern lowlands of Kenya and northeastern Tanzania, and (c) most parts of Uganda. R_{cv} shows the multiple correlation coefficient for cross-validated model

4.6.2.2.3 Mean frequency of wet spells of 3 days or more

Only two sub-regions had the multiple correlation coefficient equal or greater than 0.5 (**Table 4.44** and **Figures 4.77a –b**). These were northeastern Kenya (sub-region 3) and most parts of Uganda (sub-region 6) as shown by **Figures 4.77a** and **b** respectively. Over these two sub-regions, the cross-validated MLR models were better off than climatology as indicated by the LEPS skill score (**Table 4.44**).

One sample Kolmogorov-Smirnov test indicates that the residuals are normally distributed for the two sub-regions. The test for significance of the first-order autocorrelation was inconclusive for the residuals of the cross-validated MLR model over most parts of Uganda (sub-region 6) while over northeastern Kenya (sub-region 3), the residuals are independent.

From the assessment of the cross-validated MLR models, the two models were found to be better off than climatology and robust while the residual analysis shown that the residuals are normally distributed. These models can therefore be incorporated for operational uses.

Table 4.44: The list of predictors' combination and skill of regression models for areal-averaged frequency of wet spells of 3 days or more for the month of May

Sub-region	Predictors	Multiple correlation coefficient between observed & predicted			LEPS score (%)	Durbin-Watson Statistic
		R	R cv	R cv Adj		
3	WCAUS-2, EQATO, STAFR	0.622	0.548	0.241	25.68	2.17
6	SAFR, NEATO	0.644	0.553	0.267	32.59	1.53

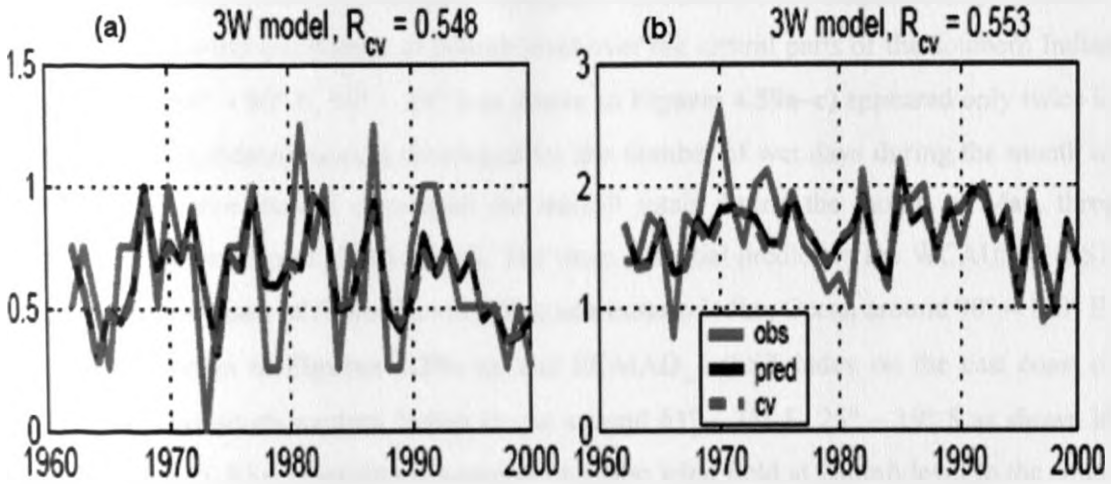


Figure 4.77: Time series plot of the observed (obs), regression model (pred) and cross-validated model (cv) estimates for areal-averaged frequency of wet spells of 3 days or more (3W) for the month of May over (a) northeastern Kenya, and (b) most parts of Uganda. R_{cv} shows the multiple correlation coefficient for cross-validated model

The poor pattern especially for the peaks and lows and the low multiple correlation coefficient between the cross-validated MLR models and actual observations for the month of

May was associated with the two months' lead-time in the predictors used to develop these models among the other factors. This lead time is considered to be quite long especially given that the predictors for the sub-regional intraseasonal statistics of wet spells for long rainfall season are mainly from the atmospheric fields that have lower persistence. The atmospheric systems thus evolve quite fast.

Consistent with the spatial coherence analyses, most of the variables during the long rainfall could not have skillful models. None of the dry spells variables could produce skill models. For the wet spells variables, only the rainfall totals, number of wet days and mean frequency of wet spells of 3 days or more had slightly skillful models for the two split parts of the long rainfall season. The cross-validated MLR models for March-April period of the long rainfall did not have any preferred predictors. However, for the number of wet days, CINDO (zonal wind component at 700mb level over equatorial central Indian Ocean around 70° – 80° E, 2.5° S – 2.5° N as shown in **Figures 4.47a–c**) appears in five out of the six cross-validated models developed. At a distant second was the EQAFR_1 (Zonal wind component index at 200mb level extending from Equatorial Eastern Africa into Equatorial Atlantic Ocean around 10° – 20° E, 10° – 5° S as shown in **Figures 4.32a–c**) which appeared in two models for each of the three variables for which the models were developed. None of the additional predictors appeared across the models for the three wet spells variables for the month of May. CSINDO (the meridional wind component at 200mb level over the central parts of the southern Indian Ocean around 80° – 90° E, 30° – 20° S as shown in **Figures 4.59a–c**) appeared only twice in the three cross-validated models developed for the number of wet days during the month of May. For the three models developed for rainfall totals during the month of May, three potential predictors appeared twice each. The three potential predictors are WCAUS_2 (SST index on western coast of Australia over the south-eastern Indian Ocean around 90° – 100° E, 12° – 4° S shown in **Figures 4.29a–c**), and ECMAD_1 (SST index on the east coast of Madagascar over south-western Indian Ocean around 63° – 75° E, 25° – 19° S as shown in **Figures 4.25a–c**), SSA (meridional component of the wind field at 200mb level to the south of the study area around 25 – 35 E, 15 – 5 S as shown in **Figures 4.57a–c**). The fact that there is no preferred predictor for the long rainfall season may suggest that these predictors could be unstable and highly variable. This could be attributed to the fact that the long rainfall marks the transition of the phase shift for ENSO, that more of the predictors are from atmospheric fields which have lower persistence and also the long lead time used for later period of the rainfall season.

In conclusion, the use of the adjusted correlation coefficient ensured that only those predictors that significantly contributed to explain the variance are included in the MLR models for the rainfall totals and SRISS of both rainfall seasons. Most of the cross-validated MLR models that are shown were developed using two or three predictors, and occasionally one or four predictors. This is consistent with Delsole and Shukla (2002) and Nyakwada (2009) who have indicated that fewer predictors tend to produce better models than those developed using large numbers of predictors. It was also observed that the correlation coefficient of the developed and cross-validated MLR models were slightly different. This was mainly attributed to the fact that in the cross-validated model, three observation values at a go were left out each time and regression models developed with the remaining observations. The LEPS skill score was positive for all the cross-validated MLR models, an indication that the models performance much better than the climatology. According to one sample Kolmogorov-Smirnov test, the residuals from the MLR models were normally distributed. Comparison of the calculated with the tabulated critical values of Durbin-Watson statistics indicated that the residuals for most of the cross-validated MLR models were independent from each other.

While the cross-validated MLR models were developed for rainfall totals, all SRISS of wet spells and most of SRISS of dry spells during the short rainfall season, the two parts of the long rainfall season has MLR regression models for the rainfall totals, number of wet days and mean frequency of wet spells of 3 days or more only. For the long rainfall season, the skills for the other statistics were not high enough to justify their discussion and future adopted for operational uses.

For the first time, this study has produced cross-validated MLR models for the number of wet days and mean frequency of wet spells of 3 days or more in additional to the routine seasonal rainfall MLR models developed by the IGAD Climate Prediction and Applications Centre (ICPAC) and National Meteorological and Hydrological Services (NMHS) for their operational use. The skills for number of wet days are similar or slightly lower than those of rainfall totals, but the cross-validated regression models were developed for a larger number of sub-regions, suggesting that this variable is spatially more robust / consistent than the seasonal rainfall totals (Moron *et al.*, 2006; 2007; Robertson *et al.*, 2009).

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary and Conclusions

This thesis is organised into five chapters, which are summarised independently in this section. Chapter one is subdivided into six major sections. A general introduction and the problem statement are presented in section one and two respectively. Many past studies have focused on understanding the rainfall variability at monthly, seasonal, and interannual time scales. Recent studies over the East Africa region have concentrated on the understanding of atmospheric processes and prediction of rainfall at different timescales, especially at monthly and seasonal timescales. Few studies have considered the intraseasonal models of rainfall variability over the region. However, it is still unclear how well do intraseasonal wet and dry spells which depicting the distribution of the rains relate to seasonal rainfall anomalies, their variability over time and space, and how predictable they are. This was the focus of the study.

In the third section of chapter one, the overall and specific objectives pursued in this study are highlighted. The overall objective of the study was to investigate the structure of the rainfall season in terms of distribution of the wet and dry spells and its variation in space and time over Equatorial Eastern Africa. Three specific objectives were therefore; to delineate and diagnose some aspects of the distribution of the wet and dry spells at interannual scale; investigate the linkage between these aspects and the dominant large scale climate fields that drive the global climate during specified seasons; and assess the predictability of the aspects of wet and dry spells for the improvement of early warning systems in the region.

Section four provides a justification for carrying out the study. Advance information of forthcoming distribution of wet/dry spells could be used to strategize on agricultural and water management policies as well as mitigating the adverse effects of recurring extreme climate events while fully capitalizing when more abundant and evenly spread rainfall occurs. Previous studies have revealed significant associations between rainfall season onsets, cessations and occurrence of wet/dry spells on one hand and end-of-season agricultural yields on the other hand.

The domain of the study was discussed in section five of chapter one. Three countries of the eastern Africa region namely Kenya, Uganda and Tanzania which are located within the latitudes 5° N and 12° S, and longitudes 29° E and 42° E, constitute the study domain. The

physical features and rainfall climatology are elaborated in this section. The study domain has large diversity of topographic features that includes the eastern and western highlands with snow-capped mountains; and the Great Rift Valley with large inland water bodies in form of deep fault lakes. Most parts of this study region have two wet (rainfall) and two dry seasons within the year. The rainfall seasons are locally referred to as long (observed from the months of March to May) and short (concentrated within the months of October to December) rainfall seasons. In the last section, an overview of this thesis is finally provided.

Previous literature on the intraseasonal aspects of the rainfall distribution over the study region and elsewhere are reviewed in details in chapter two. The first two sections of this chapter reviewed studies aimed at understanding the processes and systems associated with the spatio-temporal rainfall variability; and those studies aimed at assessing the predictability and development of forecast models. Previous studies have related the occurrence of wet and dry spells with various circulation regimes over the study area and the surrounding area. Others have showed that the first-order Markov chain models describe the occurrence of wet and dry spells over the Eastern Africa region quite well. However, those that attempted to develop prediction models for the occurrence of wet and dry spells did not have an adequate lead-time for operational applications and practice. An increase in the lead-time between the prediction and observation time is one of the aim that this study wish to achieve.

The third section of chapter two reviewed the systems that influence rainfall occurrence over the study domain. Such systems are the inter-tropical convergence zone, monsoons, tropical cyclones, subtropical anticyclones, jet streams, global and regional teleconnections such as intraseasonal oscillations, quasi-biennial oscillations, El Niño/southern oscillation and Indian Ocean dipole, and the mesoscale systems.

The third chapter addressed the datasets and the methods of analysis used in order to fulfill the overall and specific objectives of the study. The first section covered the secondary datasets used. The secondary data sets used included observed daily rainfall, reanalysis data, sea surface temperature and radiosonde data. The observed daily gauge rainfall observations over 36 locations distributed evenly over the study domain and spanning 39 years (1962 – 2000) were used. The rainfall dataset was provided by the National Meteorological and Hydrological Services of the respective countries and the IGAD Climate Prediction and Applications Centre. The radiosonde data over Bangui (Central Africa Republic) and Nairobi (Kenya) was used to assess the reliability in the use of the re-analysis data from the National

Centre for Environmental Prediction / National Centre for Atmospheric Research and the European Centre for Medium range Weather Forecasting, both of which are gridded at a horizontal resolution of 2.5° latitudes by 2.5° longitudes. The Bangui upper-air station was used simply because it was outside the study area but within the equatorial region and the observed data has a common time overlay with the Nairobi upper-air data. The atmospheric variables of the re-analysis dataset considered were the zonal and meridional components of the wind vector, the geopotential heights and the specific humidity at three standard atmospheric levels of 925mb, 700mb and 200mb representing the lower, middle and upper air levels. The specific humidity at 200mb level was however not considered since it is negligible at this level. The four constituted the atmospheric variables from which a list of additional potential predictors could be derived. The Hadley Centre Sea Surface Temperature (SST) gridded at a horizontal resolution of 1° latitude by 1° longitude provided the oceanic variable from which a list of additional potential predictors could be obtained. Other climatic indices considered were those previously published depicting El-Niño, the Indian Ocean Dipole (IOD) and SST gradient indices.

The second section of chapter three covered the methodology of analysis. This study adopted the statistical research design. Initially the missing daily rainfall observations were estimated and quality of the datasets was assessed using the graphical and statistical techniques. The missing rainfall data was estimated using the correlation and regression analyses. Less than seven percent of the daily rainfall was estimated. The double mass analysis was used to assess the quality of rainfall dataset. The other method that was used to test the quality of the data sets used was the computing of simple correlation analysis between zonal and meridional components of the radiosonde wind data at both Nairobi and Bangui with re-analysis dataset at the closed grid point to the radiosonde station.

S-mode Rotated Principal Component Analysis (RPCA) was then used to delineate areas with similar daily rainfall characteristics. Frequency distribution of the wet/dry spells based on 1mm threshold was determined; intraseasonal wet and dry spells at local and sub-regional levels were then derived. The Pearson correlation analysis was computed between the seasonal rainfall totals and various aspects of intraseasonal wet and dry spells at local (station) and sub-regional (near-homogeneous zone) levels. Using the non-parametric Spearman rank correlation analysis, the trend of the seasonal rainfall totals and various aspects of intraseasonal wet and dry spells at local and sub-regional levels was finally

determined.

The inter-station correlation analyses of the various aspects of intraseasonal wet and dry spells over a given sub-region were computed in order to assess their spatial coherence, an indirect measure of potential predictability. Low spatial coherence indicate that the signal is localized and thus the predictability potential is reduced, since any large scale forcing may be masked by stronger local effects. Total and partial simple correlation analyses were computed to quantify the linkages between the various aspects of intraseasonal wet and dry spells including rainfall totals at sub-regional level and large scale climate fields that drive global climate. The locations of additional potential predictors were noted and indices extracted over these locations. Attempts were made to provide plausible physical/dynamical explanation on how the various aspects relates to additional potential predictors and comparative location assessment of the additional potential predictor indices helped in reducing the number of the robust additional potential predictor indices.

Stepwise multivariate linear regression (MLR) technique was used to develop empirical statistical prediction models with sufficient lead time for improving the existing early warning systems. The concept of the adjusted correlation coefficient was used to determine the optimum number of predictors retained in the models. The cross-validated (leaving out three observations each time) method and calculation of the Linear Error in Probability Space skill score were used to assess the skill of the developed MLR models. The residuals from these MLR models were finally evaluated for independence using the Durbin-Watson statistics and Kolmogorov-Smirnov test to ascertain that these residuals had a normal distribution. The final section of this chapter highlighted the major limitations that were encountered and assumptions that were made in order for the research to move forward. The results obtained and conclusions drawn were based on the assumptions, despite the limitations.

The results obtained from the various methods of analysis are discussed in chapter four. Results from the double mass curve analysis of the gap-filled daily rainfall data indicated that a single straight line could be fitted to the cumulative seasonal rainfall totals for any two chosen stations. The gap-filled and quality controlled daily rainfall observations were of good quality hence suitable for further analyses in order to achieve the overall and specific objectives of the study. For the re-analysis dataset, it was quite clear that the correlation coefficients between radiosonde observations and ERA40 and NCEP/NCAR re-analysis at

most standard pressure levels are of high skills with Nairobi data, but relatively low for Bangui. The ERA40 accounts for slightly higher variance of the radiosonde data observations for both Nairobi and Bangui at all standard pressure levels considered compared to NCEP/NCAR re-analysis. The ERA40 was therefore used in this study. These datasets formed the foundation of the analysis for the current study.

For the first time, application of the rotated principal component analysis and simple correlation analysis on the square-root transformed daily rainfall observations showed that the occurrence and amounts of daily rainfall over the equatorial eastern Africa can be broadly classified into six near-homogeneous rainfall regimes during both the March-May and October-December rainfall seasons. The spatial patterns for the six near-homogeneous rainfall regimes have slight variations between the two rainfall seasons which may point to the different atmospheric and oceanic dynamics responsible for the behavior of climate during the various seasons of the year over the study area. In conclusion, though the total variance explained remain low, the six near-homogeneous sub-regions obtained are climatologically reasonable, related to specific topographic contexts and they were linked to known local and regional climate processes. The low percentage of total variance obtained in this study was attributed to the fact that for daily rainfall observations, both the intraseasonal and interannual variability are in play while at higher timescales such as month and seasonal, only the interannual variability is considered.

In the general terms of wet and dry spells, the long rainfall season has longer (shorter) mean durations of the wet (dry) spells and records the longest wet spell. Higher (lower) frequency of wet (dry) spells of 3 (5) days or more were also obtained during the long rainfall season at both local and sub-region levels. There are more wet days during long rainfall season as compared to the short rainfall season which has more dry days. The sub-regional intraseasonal statistics of the wet and dry spells (SRISS) including seasonal rainfall totals obtained from the averaging the local intraseasonal statistics of the wet and dry spells (LISS) at sub-regional level and those from PC scores are quite comparable. The SRISS obtained from averaging the daily rainfall amounts from the individual stations at a specific sub-region are the most unrealistic and thus should not be used. This approach tends to overestimate the components of the wet statistics while underestimating the components of the dry statistics. In conclusion, only the SRISS obtained from the PCA scores and those obtained from averaging the LISS were subjected to further analysis.

For the first time, the results at both local and sub-regional levels shown that the seasonal rainfall totals has positive (negative) linear associations with the intraseasonal statistics of the wet (dry) spells for both seasons. While the relationships with the intraseasonal statistics of the wet spells are mainly statistically significant (at 95% confidence levels) over most locations, those of the dry spells mostly remain statistically insignificant. The mean frequency of dry spells of 5 days or more (the number of wet days) has the least (strongest) association with the seasonal rainfall totals. The associations are stronger in the short rainfall season than the long rainfall season. However, during the short rainfall season and over the arid and semi-arid lands, the seasonal rainfall totals had significant positive association with the number of dry days and mean frequency of dry spells of 5 days or more. This meant that as seasonal rainfall total reduced, the number of dry days and the mean frequency of the dry spells of 5 days or more also reduced since the rest of the October-December period constitutes the dry season that was not analyzed.

The results of the trend analysis showed that during the long rainfall season, several locations had significant (at 95% confidence level) decreasing trend between 1962 and 2000 in the mean duration of wet spells, followed by the number of wet days and the mean rainfall intensity during the wet spells. However, these locations did not have an organized pattern. At least one in every six stations showed significant increasing trend in the mean frequency of wet spells of 3 days or more and duration of the longest wet spells during the short rainfall season. For both rainfall seasons, one in every three stations has significant increasing trend in the mean frequency of dry spells of 5 days or more. The stations with significant increasing trend in occurrence of prolonged dry spells of 5 days or more showed an organized pattern. In conclusion, though the seasonal rainfall totals seem neither to have significantly increased nor decreased, the significant increase in the occurrence of prolonged dry spells within the rainfall season may help to explain the recent poor agricultural performance and lower yields. Climate change is becoming a major development concern not only over the Equatorial Eastern Africa region but the world over. Further studies are therefore required to examine whether the trends observed in the daily rainfall spells in this study reflect any regional climate change signals.

Previous studies have relied mainly on the use of a representative station for a given near-homogeneous sub-region based on communality analysis. This study has clearly demonstrated the discrepancies associated with the use of a representative station especially

in terms of the intraseasonal distribution of the daily rainfall events and the associated intraseasonal statistics of wet and spells. This study has shown that the different methods used to aggregate rainfall stations at sub-regional scales did not have the same efficiency/effectiveness. The study has shown that for highly variable climate element such as rainfall, the sub-regional indices provides superior results since they minimize the localized effects such as the orography, the errors associated with the individual measurements and provides results that are more representative of the general synoptic scale features.

Consistent with previous studies, the spatial coherence and potential predictability analysis results indicated that the number of wet days was the spatially most coherent SRISS and closely followed by the seasonal rainfall totals, while the mean frequency of dry spells of 5 days or more was spatially least coherent SRISS and hence the least predictable. The results further shown that the sub-regional intraseasonal statistics derived from the scores of PCA are less spatially coherent compared to those derived from the areal-average of the local intraseasonal statistics. This study has for the first time shown that the percentage of the local variance explained for the whole study region during the two rainfall seasons was more than 30% for the seasonal rainfall totals and number of wet days for both the PCA-based and arithmetic areal-average based SRISS. Consistent with earlier studies on the seasonal rainfall totals, the intraseasonal statistics of wet and dry spells during the short rainfall season are more coherent and potentially more predictable, compared to those of the long rainfall season. The PCA-based SRISS has the least spatial coherence with the percentage of the local variance explained by SRISS remaining below 20% (10%) for all the SRISS apart from the seasonal rainfall totals and number of wet days during the short (long) rainfall season. The study concluded that the arithmetic areal-averaged based SRISS explained slightly higher percentage of local variance of SRISS for any statistic considered and was therefore subjected to further analysis.

Results of simple lagged total correlation analysis showed that a two-month average of the predefined indices (Niño, IOD and SST gradient) with a one month lead time gives the optimum significant stable correlation coefficient with the sub-regional intraseasonal of wet and dry spells (SRISS). This lead time is adequate for prediction purposes since it provides an adequate time to update the indices before the start of the rainfall season. The search for the additional potential predictors using the simple lagged partial correlation analysis was

therefore based on a lead time of one month and averaging the predictor fields for two months.

Results from total and partial correlation analyses identified several large scale oceanic and atmospheric signals with robust physical/dynamical linkages with arithmetic areal-averaged based SRISS including rainfall totals. For the short rainfall season, the simple lagged partial correlation analysis identified nine (9) additional potential predictor indices, four from the oceanic field and five from the atmospheric fields of zonal wind component and specific humidity and span across the whole globe within the latitudes 45° N and 45° S. The long rainfall season was split into two parts. For the first part (March-April period), thirteen (13) additional potential predictor indices were identified, two from the oceanic field and the rest from the atmospheric fields of zonal and meridional wind components as well as specific humidity. Of these thirteen (13) indices, five had slight variation in locations to those already identified during the short rainfall season. The slight variation was mainly attributed to the evolution of the global climate system with time. Several of the additional potential predictors had significant association with the predefined indices especially the Niño indices despite that fact that these did not show significant association with the rainfall totals and SRISS during the March-April period. The additional potential predictor indices for this part of the long rainfall season are partly related to basin-wide variation of the SST and not the mode of variability associated with IOD. An earlier study had indicated that the evolution of the IOD events begins around April, attains peak in October-November and dissipates around January and rarely do these events extend beyond one year.

In the later part of the long rainfall season consisting the month of May, 10 additional potential predictor indices were identified. It should however be remembered that for the later part of the long rainfall season. a two-month lead time was used. The only two indices from the oceanic field had already been identified in the early part of the long rainfall season or the short rainfall season. Some of the atmospheric indices identified were often associated with the North Africa free tropospheric flow regime towards east Africa region. This flow regime accounted for 6% of all the observations in an earlier study. Consistent with this study, the flow regime was observed from January through to May. Each of these additional potential predictor indices had a plausible and robust physical/dynamical association with the SRISS. Unlike in the short rainfall season, the large scale potential predictor indices for the earlier and later parts of the long rainfall season were all from within African continent and the two

adjacent oceans. By identifying stronger linkages between SRISS of wet spells for long (short) rainfall season and the atmospheric (oceanic) variables, the study has for the first time provided some insights to the prediction challenges for the specific seasons. Although SST forms the basis of most prediction schemes, the inclusion of atmospheric predictors was found to improve the skill of the predictions. The atmosphere has less climatic memory when compared with the oceans. However, it is hypothesized that such atmospheric predictors are proxies of large-scale land and/or oceanic energy gradients, with an inherent memory. The study therefore concludes that future predictability efforts for the long rainfall season should encourage the inclusion of atmospheric variables in the prediction models.

Skillful cross-validated multivariate linear regression (MLR) models were developed over all the six sub-regions for the sub-regional intraseasonal statistics (SRISS) of wet spells including rainfall totals during the short rainfall season. Most of these models picked additional potential predictor indices around the Bay of Bengal. The predictor indices were the BoBEN (the SST over the Bay of Bengal extending to west coast of Malaysia and Indonesia) and SINDS (the zonal wind component at 925mb to the south of the Bay of Bengal and near the southern tip of India). This clearly indicated that the oceanic and atmospheric conditions during the July-August period around the Bay of Bengal provide a lot of predictive information for the SRISS of wet spells and rainfall totals during the short rainfall season. For the SRISS of dry spells and mean rainfall intensity, most of the sub-regions didn't have useful cross-validated MLR models. There was no preferred additional potential predictor index for the cross-validated MLR models developed. However, SWHAW (the SST over south-western of Hawaii in the Pacific Ocean) was the frequently picked additional potential predictor index. Consistent with the spatial coherence results obtained earlier, none of the six sub-regions have a skillful cross-validated MLR model for the mean frequency of the dry spells of 5 days or more. This suggests therefore that the occurrence of prolonged dry spells could be mainly influenced by local factors.

For the earlier and later parts of the long rainfall season, skillful cross-validated MLR models could only be developed for two SRISS of the wet spells and the rainfall totals. These statistics are the number of wet days and the mean frequency of wet spells of 3 days or more. For these statistics, they are no preferred additional potential predictor indices. Consistent with earlier studies, the skill of these models were low compared to those of the short rainfall season and also the models for the number of wet days sometimes had higher skill compared

to those of rainfall totals. Generally the skills of the model for both parts of the long rainfall season were lower. There was no any preferred additional predictor which appeared in both parts of the season and for the different SRISS. However, the long rainfall season is a difficult season, during which the climatic system undergoes phase shift and also it displays a lesser ocean-atmosphere coupling. The models presented still complement the prediction schemes currently available

The principle of the adjusted correlation coefficient has clearly provided a criterion for determining the optimum number of predictor indices that should be included in the MLR models. With the adjusted correlation coefficient criteria, only those predictor indices that significantly contributed to explain the variance are included in the MLR models. This gave two or three predictor indices which is consistent with previous studies which indicated that fewer predictors produce better models than those developed using large number of predictor indices.

In conclusion, this study has therefore for the first time produced skillful cross-validated multivariate linear regression (MLR) models for predicting some intraseasonal characteristics of wet spells that can be used to support the current generation of seasonal rainfall totals prediction models being used by the IGAD Climate Prediction and Applications Centre (ICPAC) and National Meteorological and Hydrological Services (NHMS). The residuals from these models are normally distributed and independent from each others. These models together with the information on the likely dates of onset should provide a more clear picture of the likely performance of the rainfall activities within the season.

The results obtained from the current study can be applied in a number of ways to achieve sustainable development in the eastern Africa region. These results have showed significant increasing trend in the occurrence of prolonged dry spells within the rainfall season over the region. There is therefore the need for proper planning, development and management of water resource uses to match the water availability as supplied by the wet spells. A study done over Machakos in southeastern Kenya has demonstrated that mitigating dry spells through irrigation and supplementing soil nutrients by application of fertilizers can increase the current food production by three to five times. This ensures that the region attains food security and further improves the economic status of the farmers. The results can also be incorporated in the early warning systems aimed at reducing the climate risks that have been associated with heavy losses in hydropower generation, agricultural production and other

rain-dependent socio-economic sectors.

5.2 Recommendations

The recommendations from this study target the climate scientists and related research institutions, the policy makers, the users of the climate information and prediction products and other stakeholders in the field of climate and weather. The recommendations address issues related to the generation of wet/dry spells distribution within the season, in addition to the seasonal rainfall outlook provided by the national and regional meteorological services.

5.3.1 Recommendations to climate scientists and research institutions

This study has not determined the onset and cessation dates for the two rainfall seasons considered. Instead the study considered the first and last wet day as representative of the onset and cessation days respectively. Such definition of onset and cessation dates are greatly affected by false starts which in turn affects the distribution of wet and dry spells. The study therefore suggests that the actual onset and cessation dates based on daily rainfall be determined before the frequency distribution of wet and dry spells are derived.

Another suggestion would be to use a sub-regional definition of wet / dry spells as opposed to the local definition used in this study. This may possibly reduce the signal-to-noise ratio associated with the local definition of wet / dry spells and consequently increase spatial coherence of the various sub-regional intraseasonal statistics of wet and dry spells.

The low spatial coherence obtained for most of the intraseasonal statistics was attributed to the influence of the stronger local factors that may mask the large scale atmospheric and oceanic influence. The study suggests an alternative approach in which the intraseasonal statistics are first derived and individual statistic used to regionalize the study region. From this, one expects to yield slightly higher spatial coherence of each intraseasonal statistics and thus improved/increased predictability.

This study did not analyze the individual seasons with anomalously high or low occurrence of the wet and dry spells. The study therefore recommends that these seasons should be identified and the atmospheric and oceanic conditions associated with them clearly studied. This will provide more insights on the precursor atmospheric and oceanic conditions before the occurrence of the anomalously high or low occurrence of the wet and dry spells.

The oceanic and atmospheric predictors identified by the current study could be used to force dynamical models or existing (forced) numerical simulations. The results would then be compared with the statistical ones obtained in this study. This will confirm the reliability in the use of the identified additional predictors to explain the interannual variability of SRISS and rainfall totals over Equatorial Eastern Africa and possibly beyond.

This study recommends further exploration on the accuracy of the forecasts of atmospheric and oceanic variables used here. The ECMWF issues monthly forecasts of the zonal and meridional winds, geopotential heights and specific humidity at the standard atmospheric levels. From these numerical forecasts, one can extract the key indices depicting atmospheric features known to be related to East African rainfall, then statistically relate these predicted atmospheric indices to the rainfall. This approach (dynamical-statistical prediction) has over some regions been shown to perform better than direct forecasts of seasonal rainfall. A month by month development of empirical models for the different intraseasonal statistics especially during the long rainfall season could also be assessed.

5.3.2 Recommendations to policy makers

This study has been limited by sparse distribution of the daily rainfall data over the study region. It has also been observed that most of the daily data acquired in the recent years has not been computerized further limiting the length of the time series used for the study. Most of the pre-colonial rainfall stations has also been closed thus further limiting the number of stations to be studied. The study therefore recommends the provision of resources which will enable the computerization of the data such that it can be available in electronic form and also revive the stations already closed.

The study has further indicated that the ERA40 dataset is slightly superior over the NCEP/NCAR re-analysis over the study area and its neighbourhood. However, the recent ERA40 data is out of public domain thereby inhibiting its accessibility and usage for research purposes to climate scientists in particular and to the general public at large. This study therefore recommends that all the ERA40 dataset should be availed to the public domain which will enable the comparison of the two major re-analysis datasets in the course of the research work.

5.3.3 Recommendations to users of climate information and prediction products and other stakeholders

The results of the current study can be incorporated by researchers in the agricultural and food production sectors to ensure that they breed crops (seedlings) that can withstand the increasing prolonged dry spells within the rainfall season. The plant-breeders should ensure that the phenology of such crops matches the distribution of the wet and dry spells within the growing season. In so doing, loss of lives and livestock and famine emanating from the crop failure can be avoided.

With the knowledge of the crop water requirements, the farmers and agricultural officers can utilize the models developed in this study as a guide in planning of the agricultural activities such as weeding, spraying and harvesting; and other socio-economic activities such as transportation to the market.

Before the advent of modern scientific methods, rural communities had realized that changes in behaviour by some animals, birds, insects and plants had the capacity to detect and respond to changes in the atmospheric conditions. Over Chitora in Zimbabwe for example, the emergence of black and brown ants from their holes to collect food in the houses in large numbers is associated with an impending long wet spell while the appearance of the same bringing out the dead and damp food would imply an impending dry spell. Also the redness of the sky at sunrise and sunset which depends on the amount of dust particles in the air is regarded as precursor of long dry spell. This clearly indicates that the modern scientific studies can greatly benefit from indigenous knowledge. It is imperative therefore to integrate the traditional knowledge systems with the modern science to further our understanding and thereby ensure effective agricultural and disaster management practices. Documentation of such traditional knowledge systems is thus recommended through collaborative research between climate scientists and stakeholders from other relevant fields.

REFERENCES

Alusa, A. and M. Mushi, 1974: A study of the onset, duration and cessation of the rains in East Africa. *Preprints, International Tropical Meteorology Meeting, Nairobi, 21 Jan – 7 Feb 1974.* AMS, Boston, 133–140

Alusa, A. L. and P. M. Gwage, 1978: The occurrence of dry spells during the East African long rains. *Kenya Met. Dept. Res. Report*, **2/78**, 1 – 19

Ambaum, M. H.P., B. J. Hoskins and D. B. Stephenson, 2001: Arctic oscillation or North Atlantic Oscillation? *J. Climate,* **14**, 3495 – 3507

Ambaum, M. H.P., B. J. Hoskins and D. B. Stephenson, 2002: Corrigendum: arctic oscillation or North Atlantic Oscillation? *J. Climate,* **15**, 553

Ambenje, P. G., P. Y. Groiman, L. Ogallo, and T. R. Karl, 2001: Trends in seasonal precipitation and frequency over eastern and southern Africa. *Climate Network Africa, CNA*

Anderson, T. W. 1963: Asymptotic theory for principle component analysis. *Annals of Mathematical Statistics,* **34**, 122 – 148

Annamalai, H., J. M. Slingo, K. R. Sperber and K. Hodges, 1999: The mean evolution and variability of the Asian summer monsoon: Comparison of ECMWF and NCEP–NCAR Re-analyses. *Mon. Wea. Rev.,* **127**, 1157 – 1186

Anyamba, E. K. 1990: A diagnostic study of low frequency oscillations in tropical outgoing long wave radiation. *PhD thesis, University of California, Davis*

Asnani, G. C. and J. H. Kinuthia, 1979: Diurnal variation of precipitation in East Africa. *Tech. Res. Rep.* **8/79**, Kenya Meteolo. Dept., 1 – 58

Asnani, G. C., 1993: Tropical meteorology. Indian Institute of Tropical Meteorology: Poona, **2**, 1202 pp

Asnani, G.C., 2005: Tropical Meteorology, revised edition. Indian Institute of Tropical Meteorology, Pune, India **2**, 753pp

Bamanya, D., 2007: Intraseasonal characteristics of daily rainfall over Uganda during wet seasons. *MSc thesis, Department of Meteorology, University of Nairobi, Kenya*

Barnston, A. G., W. Thiao, and V. Kumar, 1996: Long-lead forecasting of seasonal

precipitation in Africa using CCA. *Weather and Forecasting*, **11**, 506 – 520

Bärring, L., 1988: Regionalization of daily rainfall in Kenya by means of common factor analysis. *Int J Climatol.*, **8**, 371 – 389

Bärring, L., T. Holt, M. L. Linderson, M. Radziejewski, M. Moriondo and JP Palutikof 2006: Defining dry/wet spells for point observations, observed area averages and regional climate model gridboxes in Europe. *Clim. Res.*, **31**, 35 – 49

Barr-Kumarakulasinghe, S. A. and K. M. M. Lwiza, 1998: Deep convective cloud scales and the direct adjustment of upper troposphere moisture in the tropical western Pacific Environment. *Meteor. and Atmos. Phys.*, **66**, 35 – 50

Barrow, J., 2004: Dry spell mitigation to upgrade semi-arid agriculture: Water harvesting and soil nutrient management for smallholder maize cultivation in Machakos, Kenya. *PhD thesis in Natural Resources management, Department of Systems Ecology, Stockholm University, Sweden*

Barrow, J., J. Rockstrom, F. Gichuki and N. Hatibu, 2003: Dry spell analysis and maize yields for two semi-arid locations in East Africa. *Agricultural and Forest Meteorology*, **117**, 23 – 27

Basalirwa, C., 1991: Raingauge network designs for Uganda. *PhD thesis, Department of Meteorology, University of Nairobi, Kenya*

Bazira, E. and Ogallo, L. J., 1985: Daily Rainfall Characteristics over Uganda. *Proceedings of the first Technical Conference on Meteorological Research in Eastern and Southern Africa, Nairobi, Kenya.* 6-9 January 1985; 152 – 155

Behera, S. K., J. Luo, S. Masson, P. Delecluse, S. Gualdi, A. Navarra and T. Yamagata, 2005: Paramount Impact of the Indian Ocean Dipole on the East African short rains: A CGCM study. *J. Climate.*, **18**, 4514 – 4530

Biasutti, M., D. S. Battisti, and E. S. Sarachik, 2004: Mechanisms controlling the annual cycle of precipitation in the Tropical Atlantic Sector in an Atmospheric GCM., *J. Climate*, *17*, 4708 – 4723

Black, E., 2005: The relationship between Indian Ocean sea-surface temperature and East African rainfall. *Phil. Trans. R. Soc. A*, **363**, 43 – 47

Black, E., J. M. Slingo and K. R. Sperber, 2003: An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean

SST. *Mon. Wea. Rev.*, **103**, 74 – 94

Bryan, R. B. and R.A. Southerland, 1989: Erosion and soil conservation in semi-arid tropical region: Rift Valley Province, Kenya. *Sixth International Soil Conservation Conference*, Ethiopia, 6–18 November 1989.

Cadet, D. L., 1985: The Southern Oscillation over the Indian Ocean. *J. Climatol.*, **5**, 189 – 212

Camberlin, P. and J. Wairoto, 1997: Intraseasonal wind anomalies related to wet and dry spells during the "long" and "short" rainy seasons in Kenya. *Theor. Appl. Climatol.* **58**, 57 – 69

Camberlin, P. and N. Philippon, 2002: The East African March–May rainy season: Associated atmospheric dynamics and predictability over the 1968–1997 period. *J Climate,* **15**, 1002 – 1019

Camberlin, P. and R. E. Okoola, 2003: The onset and cessation of the "long rains" in eastern Africa and their interannual variability. *Theor. Appl. Climatol.* **75**, 43 – 54

Camberlin, P., S. Janicot, and I. Poccard, 2001: Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea-surface temperature: Atlantic vs. ENSO. *Inter. J. Climatol.*, **21**, 973 – 1005

Castell, R. B., 1966: The scree-test for the number of factors. *Multvar. Behav. Res.,* **1**, 245 – 276

Ceballos, A., J. Martinez-Fernandez, F. Santos, and F. Alonso, 2002: Soil water behaviour of sandy soils under semiarid conditions in the Duero Basin (Spain). *J. Arid Environ.*, **51**, 501 – 519

Ceballos, A., J. Martínez-Fernández, and M. Á. Luengo-Ugidos, 2004: Analysis of rainfall trends and dry periods on a pluviometric gradient representative of Mediterranean climate in the Duero Basin, Spain. *J. Arid Environ.*, **58**, 215 – 233

Chapman, T. G., 1998: Stochastic modeling of daily rainfall: the impact of adjoining wet days on the distribution of rainfall amounts. *Environmental Modeling and Software*, **13**, 317 – 324

Clark, C. O., P. J. Webster and J. E. Cole, 2003: Interdecadal variability of the relationship between the Indian Ocean Zonal Mode and East African coastal rainfall

anomalies. *J. Climate*, **16**, 548 – 554

Cornelius. J. and J. Reynolds, 1991: On determining the statistical significance of discontinuities within ecological data. *Ecology*, **72**, 2057 – 2070

Cressie, N. A. C., 1990: The Origins of Kriging, *Mathematical Geology*, **22**, 239 – 252

Cressie, N. A. C., 1991: Statistics for Spatial Data, *John Wiley and Sons, Inc., New York*, 900 pp

Delsole, T. and J. Shukla, 2002: Linear Prediction of Indian Monsoon Rainfall. *J. Climate*, **15**, 3645 – 3658

Dobi-Wantuch, I., J. Mika and L Szeidi, 2000: Modeling wet and dry spells with mixture distributions. *Meteo. Atmos. Phys.* **73**, 245 -256

Dommenget, D. and M. Latif, 2002: A cautionary note on the interpretation of EOFs. *J. Climate*, **15**, 216 – 225

Douguedroit, A., 1987: The variation of dry spells in Marseilles from 1865 to 1984. *J. Climatol.*, **7**, 541 – 551

Duree, I., R. S. Vose and D. B. Wuertz, 2006: Overview of the Integrated Global Radiosonde Archive. *J. Climate*, **19**, 53 – 68

Engelen R. J., I. L. Wittmeyer, and G. L. Stephens, 1998: Assessment of reanalysis hydrology and energy budgets: Water vapour and radiative fluxes. *Proc. First Int. WCRP Conf. on Reanalysis,* Silver Spring, MD, World Climate Research Programme, WCRP-104 (WMO/TD-876), 175 – 178

Farebrother, R. W., 1980: The Durbin-Watson test for serial correlation when there is no intercept in the Regression. *Econometrica*, **48(6)**, 1553 – 1563

Farmer, G. 1988: Seasonal forecasting of the Kenya coast short rains, 1901–1984. *J. Climatol.*, **8**, 489 – 497

Findlater, J., 1966: Cross-equatorial jet streams at low-levels over Kenya. *Meteor. Mag.* **95**, 353 – 364

Findlater, J., 1977: Observational aspects of the low level cross-equatorial jet stream of the western Indian Ocean. *Pure Appl. Geophys.* **115**, 1251 – 1262

Fox, J., 1991. Regression diagnostics: An Introduction. *Sage University Paper series on Quantitative Applications in the Social Sciences, Series No. 07-079*, Sage:

Newbury Park, CA

Frei, C., Christensen, J H. Deque M, Jacob D, Jones RG, Vidale PL, 2003: Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps. *J Geophys Res.*, **108**, 4124.

Gatebe, C. K., P. D. Tyson, H. Annegarn, S. Piketh, and G. Helas, 1999: A seasonal air transport climatology for Kenya, *J. Geophys. Res.*, **104(D12)**, 14237–14244

Gitau, W. 2005: Characteristics of the wet and dry spells during the wet seasons over Kenya. *MSc thesis, Department of Meteorology, University of Nairobi, Kenya*

Gitau, W., L. Ogallo and J. N. Mutemi, 2008: Intraseasonal characteristics of wet and dry spells over Kenya. *JKMS*, **2(1)**, 19 – 30

Goddard, L. and N. E. Graham, 1999: The importance of the Indian Ocean for simulating precipitation anomalies over eastern and southern Africa. *J. Geophys. Res.*, 104, 19099 – 19116

Godínez-Domínguez, E., J. Rojo-Vázquez, V. Galván-Piña and B. Aguilar-Palomino, 2000: Changes in the structure of a coastal fish assemblage exploited by a small scale gillnet fishery during an El Niño–La Niña Event. *Estuarine, Coastal and Shelf Science*, **51(6)**, 773 – 787

Gong, X., A. G. Barnston, and M. N. Ward, 2003: The effect of spatial aggregation on the skill of seasonal precipitation forecasts. *J. Climate*, **16**, 3059 – 3071

Goswani, B. N., R. S. Ajayamohan, P. K. Xavier and D. Sengupta, 2003: Clustering of synoptic activity by India Summer Monsoon intraseasonal oscillations. *Geophy. Resea. Let.*, **30**, 1431 – 1434

Griffiths, J.F. and K. H. Solimani, 1972: Climates of Africa. *World Survey of Climatology, Elsevier,* **10**, 604

Hannachi, A., I. T. Jolliffe and D. B. Stephenson, 2007: Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. Climatol.*, **27**, 1119 – 1152

Hastenrath, S., 2007: Circulation mechanisms of climate anomalies in East Africa and equatorial Indian Ocean. *Dynamics of Atmospheres and Oceans*, **43**, 25 – 35

Hastenrath, S., A. Nicklis, and L. Greischar, 1993: Atmospheric–hydrospheric mechanisms of climate anomalies in the western equatorial Indian Ocean. *J. Geophys. Res.*

Oceans, **98**(C11), 20 219–20 235

Haylock, M. and J. McBride, 2001: Spatial coherence and predictability of Indonesian wet season rainfall. *J. Climate*, **14**, 3382 – 3887

Helsel, D. R. and R. M. Hirsch, 1992: Statistical Methods in Water Resources. Book 4, Hydrologic Analysis and Interpretation

Henne, S., J. Klausen, W. Junkermann, J. M. Kariuki, J. O. Aseyo, and B. Buchmann, 2008: Representativeness and climatology of carbon monoxide and ozone at the global GAW station Mt. Kenya in equatorial Africa. *Atmos. Chem. Phys.*, **8**, 3119 – 3139

Horel, J. D. 1981: A rotated principal component analysis of the interannual variability of the Northern Hemisphere 500 mb height field. *Mon. Wea. Rev.*, **109**, 2080 – 2092

Huffman, G. J., R. F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, and U. Schneider, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *BAMS*, **78(1)**, 5 – 20

ICPAC, 2006: Factoring Climate Products for Improved Agriculture Production and Food Security for East Africa Region: *Proceedings of Regional Capacity Building and Training Workshop, 21 – 28 August 2006*, Nairobi, Kenya

Indeje, M., 2000: Prediction and numerical simulation of the regional climate of equatorial eastern Africa. *PhD. dissertation, Marine, Earth and Atmospheric sciences, North Carolina State University, Raleigh, USA*, 327pp

Indeje, M. and F. H. M. Semazzi, 2000: Relationship between QBO in the lower equatorial stratospheric zonal winds and East African Seasonal Rainfall. *Meteorol. Atmos. Phys.*, **73**, 227 – 244

Indeje, M., F. H. M. Semazzi, and L. J. Ogallo, 2000: ENSO signals in East African Rainfall and their prediction potentials. *Int. J. Climatol.* **20**, 19 – 46

Indeje, M., F. H. M. Semazzi, L. Xie and L. J. Ogallo, 2001: Mechanistic model simulations of the East African Climate using NCAR Regional Climate Model: Influence of large scale orography on the Turkana Low-Level Jet. *J. Climate*, **14**, 2710 – 2724

Ininda, J.M., 1995: Numerical simulation of the influence of sea surface temperature

anomalies on the East African seasonal rainfall. *PhD thesis, Department of Meteorology, University of Nairobi, Kenya*

Isaaks, E. H. and R. M. Srivastaka, 1989: Applied Geostatistics. *Oxford University Press, New York*

Jagadheesha, D., M. Tadross, and C. J. C. Reason, 2003: Characteristics of intraseasonal OLR Variability over Southern Africa during ENSO. *Clivar. Exchanges*, **27**

Janowiak, J. E., 1988: An investigation of interannual rainfall variability in Africa. *J. Climate*, **1**, 165 – 179

Jennrich, R. I., 2001: A simple general procedure for orthogonal rotation. *Psychometrika*, **66**, 289 – 306

Jennrich, R. I., 2002: A simple general procedure for oblique rotation. *Psychometrika* **67**, 7 – 19

Jolliffe, I. and D. B. Stephenson, 2003: Forecast Verification: A Practitioner's Guide in Atmospheric Science. *John Wiley and sons*

Jolliffe, I. T., 2002: *Principal Component Analysis*, 2nd edn. Springer: New York.

Jones, C., D. F. Waliser, J. K. E. Schemm and W. K. M. Lau, 2000: Prediction skill of the Madden and Julian Oscillation in dynamical extended range forecasts. *Clim. Dyn.*, **16**, 273 – 289

Jury, M. R. and B. Pathack, 1993: Composite climatic patterns associated with extreme modes of summer rainfall over southern Africa: 1975 – 1984. *Theor. Appl. Climatol.*, **47**, 137 – 145

Jury, M. R., B. Pathack and D. Waliser, 1993: Satellite OLR and microwave data as a proxy for summer rainfall over subequatorial Africa and adjacent oceans. *Int. J. Climatol.*, **13**, 257 – 269

Jury, M. R., E. Matari and M. Matitu, 2009: Equatorial African climate teleconnections. *Theor. Appl. Climatol.*, **95**, 407 – 416

Kaiser, H. F., 1959: The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187 – 200

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, L. Iredel, H. Saha, G. White, J. Woolle, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J.

Jonawiak, C. K. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-Year Re-analysis Project. *Bull. Amer. Meteo. Soc.*, 77, 437 – 471

Kanamitsu, M., W. Ebisuzaki, J. Wollen, S. K Yang, J. J. Hnilo, M. Fiorino and G. L. Potter, 2002: NCEP–DOE AMIP–II reanalysis, *BAMS*, **83**, No. 11, 1631 – 1643

Kazutoshi, O., H. Koide, M. Sakamoto, S. Kobayashi, J. Tsutsui, H. Hatsushika, T. Matsumoto, N. Yamazaki, H. Kamahori, K. Takahashi, K. Kato, R. Oyama, T. Ose, S. Kadokura and K Wada, 2005: JRA-25: Japanese 25-year re-analysis project – progress and status. *Q. J. R. Meteorol. Soc.* **131**, 3259 – 3268

Kemp, P., J. Cornelius, and J. Reynolds, 1994: Temporal discontinuities in precipitation in the central North American Prairie. *Int. J. Climatol.*, **14**, 539 – 557

Kendall, M. G., 1976: Time Series. *Charles Griffin, London*

Kendall, M. G. and A. Stuart, 1961: Advanced theory of Statistics. *Charles Griffins, London*

Kiangi, P. M. R., M. M. Kavishe and J. K. Patnaik, 1981: Some aspects of mean tropospheric motion in East Africa during the long rain season. *Kenya J. Sci. Tech. Ser. A,* **2**, 91 – 103

Kinuthia, J. H., 1992: Horizontal and vertical structure of the Lake Turkana jet. *J. Appl. Meteor.*, **31**, 1248 – 1274

Kinuthia, J. H. and G. C. Asnani, 1982: A newly found jet in North Kenya (Turkana Channel). *Mon. Wea. Rev.*, **110**, 1722 – 1728

Komutunga, T. E., 2006: Optimum cropping calendars derived for rain-fed agriculture of Uganda from rainfall data. *PhD thesis, Department of Meteorology, University of Nairobi, Kenya*

Kongoti, J. G., 1989: Estimating daily rainfall from meteosat satellite imagery over Kenya. *MSc thesis, University of Michigan, USA*

Korecha, D. and A. N. Barnston, 2007: Predictability of June-September rainfall in Ethiopia. *Mon. Wea. Rev.* **135**, 628 – 650

Kripalani, R. H., A. Kulkarni, S. S. Sabade, J.V. Revadekar, S. K. Patwardhan, and J. R. Kulkarni, 2004: Intraseasonal oscillations during monsoon 2002 and 2003. *Current Science*, **87**, 325 – 331

Krishna Kumar, K., M. K. Soman, and K. Rupa Kumar, 1995: Seasonal forecasting of Indian summer monsoon rainfall. *Weather*, **50**, 449 – 467

Krishnamurti, T. N. and P. Ardunay, 1980: The 10 to 20 day westward propagating mode and breaks in the monsoon. *Tellus*, **32**, 15 – 20

Krishnamurti, T. N., J. Molinari, and H. L. Pan, 1976: Numerical simulation of the Somali jet. *J. Atmos. Sci.* **33**, 2350 – 2362

Latif, M., D. Dommenget, M. Dima, and A. Grotzner, 1999: The role of Indian Ocean sea surface temperature in forcing East African rainfall anomalies during December–January 1997/98. *J. Climate*, **12**, 3497 – 3504

Lazaro, R., F. S. Rodrigo, L. Gutierrez, F. Domingo, and J. Puigdefabregas, 2001: Analysis of a 30-year rainfall record (1967–1997) in semi–arid SE Spain for implications on vegetation. *J. Arid Environ.*, **48**, 373 – 395

Li, S., W. A. Robinson, and S. Peng, 2003: Influence of the North Atlantic SST tripole on northwest African rainfall. *J. Geophys. Res.*, **108 (D19)**, doi: 10.1029/2002JD003130

Liao, J. G. and Dan McGee, 2003: Adjusted Coefficients of Determination for Logistic Regression, *the American Statistician*, **57**, 161-165

Lilliefors, H. W., 1967: On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *J. American Statistical Association*. **62**, 399 – 402

Lucio, P. S., F. C. Conde and A. M. Ramos, 2006: A Bayesian approach for recovering and homogenising meteorological time series. *Proceedings of 8 ICSHMO, Foz do Iguacu, Brazil, April 24-28, INPE*, 29-36

Lynch, S. D. and R. E. Schulze, 1995: Techniques for estimating areal daily rainfall. *ESRI User conference proceedings, California, USA*

Makridakis, S., S. C. Wheelwright, and R. B. Hyndman, 1998: Forecasting: Methods and Applications. *John Wiley and Sons: New York*, 241 – 310

Martin-Vide, J. and L. Gomez, 1999: Regionalisation of Peninsular Spain based on the length of dry spells, *Internat. J. Climatol.*, **19**, 537 – 555

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

Matari, E. E., 2002: Impacts of Congo convection on tropical Africa's circulation, rainfall and resources. *MSc thesis, Univ. Zululand,* 172 pp

Mc Cuen, R. H., 1985: Statistical Methods for Engineers. *Prentice Hall: New Jersey,* 439

Meyers, G., 1996: Variation of Indonesian Throughflow and El Niño-Southern Oscillation. *J. Geophys. Res.*, **101**, 12255 – 12263

Misra, V., 2003: The influence of Pacific SST on the precipitation over southern Africa diagnosed from an AGCM. *J. Climate*, **16**, 2408 – 2418

Moon, S. E., S. B. Ryoo and J. G. Kwon, 1994: A Markov chain model for daily precipitation occurrence in South Korea. *Int. J. Climatol.*, **14**, 1009 – 1016

Moron, V., A. W. Robertson, and M. N. Ward, 2006: Seasonal predictability and spatial coherence of rainfall characteristics in the tropical setting of Senegal. *Mon. Wea. Rev.*, **134**, 3248 – 3262

Moron, V., A. W. Robertson, M. N. Ward and P. Camberlin, 2007: Spatial coherence of tropical rainfall at the regional scale. *J. Climate,* **20**, 5244 – 5263

Moron, V., N. Philippon, and B. Fontaine, 2003: Skill of Sahel rainfall variability in four atmospheric GCMs forced by prescribed SST. *Geophys. Res. Lett.*, 30(23), 2221, doi: 10.1029/2003GL018006.

Mpeta, E. J. and M. R. Jury, 2001: Intraseasonal convective structure and evolution over tropical East Africa. *Clim. Res.*, **17**, 83 – 92

Mukabana, J. R., 1992: Numerical simulation of the influence of the large scale monsoon flow on the weather patterns over Kenya using a three-dimensional limited area model. *PhD thesis, Department of Meteorology, University of Nairobi, Kenya*

Mukabana, J. R. and R. A. Pielke, 1996: Investigating the influence of synoptic-scale monsoonal winds and mesoscale circulations on diurnal weather patterns over Kenya using a mesoscale numerical model. *Mon. Wea. Rev.*, **124**, 224 – 243

Mungai, D. N., 1984: Analysis of some seasonal rainfall characteristics in the Lake Victoria region of Kenya. *MA thesis, Department of Geography, University of Nairobi, Kenya*

Mutai, C. C., 2000: Diagnosis and predictability of East African rainfall on intraseasonal to

interannual timescales. *PhD thesis, Department of Meteorology, University of Nairobi, Kenya*

Mutai, C. C. and M. N. Ward, 2000: East African rainfall and the tropical circulation/convection on intraseasonal to interannual timescales. *J Climate*, **13**, 3915 – 3939

Mutai, C. C., M. N. Ward and A. W. Colman, 1998: Towards the prediction of the East Africa short rains based on sea-surface temperature coupling. *Int. J. Climatol.*, **18**, 975 – 997

Mutemi, J. N., 2003: Climate anomalies over Eastern Africa associated with various ENSO evolution phases. *PhD thesis, Department of Meteorology, University of Nairobi, Kenya*

Nayagam, L. R., R. Janardanan and H. S. Ram Mohan, 2008: An empirical model for the seasonal prediction of southwest monsoon rainfall over Kelala, a meteorological subdivision of India. *Inter. J. Climatol.*, **28**, 823 – 831

Newman, M., P. D. Sardeshmukh and J. W. Bergman, 2000: An assessment of the NCEP, NASA and ECMWF reanalyses over tropical west Pacific Warm Pool. *BAMS*, **81**, 41 – 48

Ngigi, S. N., H. H. G. Savenije, J. N. Thome, J. Rockstrom, and F.W.T. Penning de Vries, 2005: Agro-hydrological evaluation of on-farm rainwater storage systems for supplemental irrigation in Laikipia district, Kenya. *Agriculture Water Management*, **73**, 21 – 41

Nicholson S. and J. Kim, 1997: The relationship of the El Niño–Southern Oscillation to African rainfall. *Int. J. Climatol.*, **17**, 117 – 135

Nicholson, S. E. 1996. A review of climate dynamics and climate variability in eastern Africa, in Johnson, T. C., and Odada, E. (eds), *The Limnology, Climatology and Paleoclimatology of the East African Lakes*, Gordon and Breach, Melbourne, Australia, 25 - 56

Njau, L. N., 2006: Diagnostics and predictability of East African rainfall with tropospheric circulation parameters. *PhD thesis, Department of Meteorology, University of Nairobi, Kenya*

Nogues-Paegle, J. and K.C. Mo, 1997: Alternative wet and dry conditions over South America during summer. *Mon. Wea. Rev.*, **125**, 279 - 291

North, G. R., T. L. Bell, R. F. Cahalan and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.*, **110**, 699 – 706

Nyakwada, W., 1991: Relationship between satellite derived outgoing longwave radiation (OLR) and some important meteorological parameters. *MSc thesis, Department of Meteorology, University of Nairobi, Kenya*

Nyakwada, W., 2009: Predictability of East African seasonal rainfall with sea surface temperature gradient modes. *PhD thesis, Department of Meteorology, University of Nairobi, Kenya*

Ochola, W. O. and P. Kerkides, P. 2003: A Markov chain simulation model for predicting critical wet and dry spells in Kenya: Analysing rainfall events in the Kano plains. *Irrigation and Drainage*, **52**, 327 – 342

Oettli, P. and P. Camberlin, 2005: Influence of topography on monthly rainfall distribution over East Africa. *Climate Research*, **28**, 199 - 212

Ogallo, L. J. and A. W. Chillambo, 1982: The characteristics of wet spells in Tanzania. *E. Afr. Agric. For. J.*, **47 (4)**, 87 – 95

Ogallo, L. J., 1980: Regional classification of the East African Rainfall stations into homogeneous groups using the method of Principal Component Analysis. *Stat. Climatol. Devel. Atmos. Sci.*, **13**, 255 – 266

Ogallo, L. J., 1982: Quasi-periodic patterns in the East African rainfall records. *Kenya J. Science and Techn.*, **A3**, 43 – 54

Ogallo, L. J., 1988: Relationship between seasonal rainfall in East Africa and Southern Oscillation. *Int. J. Climatology*, **8**, 31 – 43

Ogallo, L. J., 1989: The spatial and temporal patterns of the East African seasonal rainfall derived from Principle Component Analysis. *Inter. J. Climatol.*, **9**, 145 – 167

Ogallo, L. J., 1993: Dynamics of the East African climate. *Proc. Indian Acad. Sci. (Earth Planet. Sci.)*, **102**, 203 – 217

Ogallo, L. J., J. E. Janowiak, and M. S. Halpert, 1988: Teleconnection between seasonal

rainfall over Eastern Africa and Global Sea surface temperature anomalies. *J. Meteor. Soc., Japan*, **66**, 807 – 822

Ogallo, L. J., M. S. Boulahya, and T. Keane, 2000: Applications of seasonal to interannual climate prediction in agricultural planning and operations. *Agric. Fore. Meteoro.*, **103**, 159 – 166

Ogallo, L. J., R. E. Okoola and D. N. Wanjohi, 1994: Characteristics of quasi-biennial oscillation over Kenya and their predictability potential for the seasonal rainfall. *Mausam*, **45**, 57 – 62

Okeyo, A. E., 1987: Towards the development of a forecasting numerical model for Kenya. *PhD thesis, Department of Meteorology, University of Nairobi, Kenya*

Okoola, R. E., 1982: Monsoon system over Southwest India Ocean during northern summer of 1979. *MSc thesis, Department of Meteorology, University of Nairobi, Kenya*

Okoola, R. E., 1989: Westwards moving disturbances in the southwest India Ocean. *Met. Atmos. Phys.*, **41**, 33 – 44

Okoola, R. E., 1996: Space-time Characteristics of the ITCZ over equatorial Eastern Africa during anomalous rainfall years. *PhD thesis Department of Meteorology, University of Nairobi, Kenya*

Okoola, R. E., 1998: Spatial evolutions of the active convective patterns across the equatorial eastern Africa during Northern Hemisphere spring season using outgoing longwave radiation records. *Meteorol. Atmos. Phys.*, **66**, 51 – 63

Okoola, R. E., 1999a: Synoptic systems affecting Eastern Africa. *Lecture notes of the first Drought Monitoring Centre (DMC), Nairobi climate prediction capacity building training workshop for the Greater Horn of Africa, Nairobi, Kenya. Drought Monitoring Centre – Nairobi*, 51 – 62

Okoola, R. E., 1999b: A diagnostic study of the eastern Africa monsoon circulation during the northern hemisphere spring season. *Int. J. Climatol.*, **19**, 143 – 168

Okoola, R. E., 1999c: Midtropospheric circulation patterns associated with extreme dry and wet episodes over equatorial Eastern Africa during the Northern Hemisphere spring. *J. Appl. Meteorol.* **38**, 1161 – 1169

Okoola, R. E., J. M. Ininda and P. Camberlin, 2008: Wet periods along East Africa coast and

extreme wet spell event of October 1997. *JKMS*, **2(1)**, 67 – 83

Oladipo, E. O. and J. O. Kyari, 1993: Fluctuations in the onset, termination and length of the growing season in Northern Nigeria. *Theor. Appl. Climatol.* **47**, 241 – 250

Omeny, P. A., 2006: East African Rainfall variability associated with Madden-Julian Oscillation. *MSc. Dissertation, Department of Meteorology, University of Nairobi, Kenya*

Omeny, P. A., L. Ogallo, R. E. Okoola, H. Hendon and M. Wheeler, 2008: East African Rainfall variability associated with Madden-Julian Oscillation. *JKMS*, **2(2)**, 109 – 118

Oteng'i, S. B. B. and L. J. Ogallo, 1984: Persistence of daily rainfall over some parts of Kenya. *WMO Regional Scientific conf. on GATE, EAMEX and tropical Met. in Africa, 10th to 14th Dec. 1984, Dakar, Senegal*

Ouma G. O., 2000: Use of satellite data in the monitoring and prediction of rainfall over Kenya. *PhD thesis, Department of Meteorology, University of Nairobi, Kenya*

Overland, J. E. and R. W. Preisendorfer, 1982: A significance test for principal components applied to a cyclone climatology. *Mon. Wea. Rev.*, **110**, 1 – 4

Owiti, Z. O., 2005: Use of the Indian Ocean dipole indices as a predictor of East African rainfall anomalies. *MSc thesis, Department of Meteorology, University of Nairobi, Kenya*

Owiti, Z. O., L. A. Ogallo and J. Mutemi, 2008: Linkages between the Indian Ocean Dipole and East African Seasonal Rainfall Anomalies. *JKMS*, **2(1)**, 3 – 17

Paeth, H. and Friederichs, P., 2004: Seasonality and time scales in the relationship between global SST and African rainfall. *Clim. Dynamics*, **23**, 815 – 837

Peña, M. and M. W. Douglas, 2002: Characteristics of wet and dry spells over the Pacific side of Central America during the rainy season. *Mon. Wea. Rev.*, **130**, 3054 – 3073

Perzyna, G., 1994: Spatial and temporal characteristics of maximum dry spells in southern Norway. *Int. J. Climatol.*, **14**, 895 – 909

Philippon, N., 2002: Une nouvelle approche pour la prevision statistique des precipitations saisonnieres en afrique de l'Ouest et de l'Est: methods, diagnostics (1968 – 1998) et applications (2000-2001). *Docteur These, Centre de Recherches de*

Climatologie, Universite de Bourgogne, France

Philippon, N., P. Camberlin and N. Fauchereau, 2002: Empirical predictability study of October-December East African rainfall. *Q. J. R. Meteorol. Soc.*, **128**, 2239 – 2256

Philippon, N., N. Martiny and P. Camberlin, 2009: Forecasting the vegetation photosynthetic activity over the Sahel: a Model Output Statistics approach, *Int. J. Climatol.*, **29**, 1463 – 1477

Phillips, I. D. and H. Denning, 2007: Winter daily precipitation variability over the South West Peninsula of England. *Theor. Appl. Climatol.*, **87**, 103 – 122

Phillips, J. and B. McIntyre, 2000: ENSO and interannual variability in Uganda: Implications for agricultural management. *Inter. J. Climatol.*, **20**, 171 – 182

Pohl, B. and P. Camberlin, 2006: Influence of the Madden-Julian Oscillation on East African rainfall. 1: Intraseasonal variability and regional dependency. *Q. J. R. Meteorol. Soc.*, **132**, 2521 – 2539

Pohl, B., P. Camberlin and P. Roucou, 2005: Typology of pentad circulation anomalies over the Eastern Africa–Western Indian Ocean region, and their relationship with rainfall. *Clim. Res.*, **29**, 111–127

Potts, J. M., C. K. Folland, I. T. Jolliffe, and D. Sexton, 1996: Revised "LEPS" scores for assessing climate model simulations and long-range forecasts. *J. Climate*, **9**, 34 – 53

Rayner, N. A., D. E. Parker, E. B Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice and night marine air temperature since the late nineteenth century. *J. Geophys. Res.* **108**, D14, 4407 10.1029/2002JD002670

Reason, C. J. C., 2001: Subtropical Indian Ocean SST dipole events and southern African rainfall. *Geo. Res. Let.*, **28**, 2225 – 2227

Reason, C. J. C., 2002: Sensitivity of the southern African circulation to dipole SST patterns in the south Indian Ocean. *Int. J. Climatol.*, **22**, 377 – 393

Reason, C. J. C. and J. R. E. Lutjeharms, 1998: Variability of the south Indian Ocean and implications for southern African rainfall, *South African J. of Sci.* **94**, 115 – 123

Richman, M. B., 1986: Rotation of principal components. *J. Climatology*, **6**, 293 – 335

Robertson, A. W., V. Moron and Y. Swarinoto, 2009: Seasonal predictability of daily rainfall statistics over Indramayu district, Indonesia. *Int. J. Climatol.*, **29**, 1449 – 1462

Ropelweski, C. F. and M. S. Halpert, 1987: Precipitation patterns associated with El Niño/Southern Oscillation. *Mon. Wea. Rev.*, **115**, 1606 – 1626

Rowell, D. P., J. M. Ininda, and M. N. Ward, 1994: The impacts of global sea surface temperature patterns on seasonal rainfall in East Africa. *Proceedings International Conference on Monsoon Variability and Prediction*, Trieste, Italy, 9–13 May 1994, WMO, 666 - 672

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. *Nature*, **401**, 360 – 363

Schafer, N. W., 1991: Modelling the areal distribution of daily rainfall. *MSc Eng. thesis, Department of Agricultural Engineering, University of Natal, Pietermaritzburg, South Africa*

Semazzi, H. F. M., B. Burns, N. H. Lin and J. E. Schemm, 1996: A GCM study of the teleconnections between the continental climate of Africa and global sea-surface temperature anomalies. *J. Climate*, **9**, 2480 – 2497

Sharma, T. C., 1996: Simulation of the Kenyan longest dry and wet spells and the largest rainsums using a Markov model. *J. Hydrol.*, **178**, 55 – 67

Shinoda, M. and R. Kawamura, 1996: Relationships between rainfall over semi-arid Southern Africa, geopotential heights and sea surface Temperatures. *Journal of the Meteorological Society of Japan*, **74(1)**, 21 – 36

Sikka, D. R. and S. Gadgil, 1980: On the maximum cloud zone and the ITCZ over India longitudes during the southwest monsoon. *Mon. Wea. Rev.*, **108**, 1840 – 1853

Sivakumar, M. V. K., 1992: Empirical analysis of dry spells for agricultural application in West Africa. *J. Climate*, **5**, 532 – 539

Soden, B. J. and R. Fu, 1995: A satellite analysis of deep convection upper tropospheric humidity, and the greenhouse effect. *J. Climate*, **8**, 2333 – 2351

Stendel, M. and K. Arpe, 1997: Evaluation of the hydrological cycle in reanalyses and observations. *Max-Planck-Institut fur Meteorologie Report*, **228**, 52 pp

Stephenson, D. B., K. Rupa Kumar, F. J. Doblas-Reyes, J. F. Royer, F. Chauvin and S. Pezzulli, 1999: Extreme daily rainfall events and their impacts on ensemble forecasts of the Indian monsoon. *Mon. Wea. Rev.*, **127**, 1954 – 1966

Stewart, J. I. and C. T. Hash, 1982: Impact of weather analysis on agricultural production and planning decisions for the semi-arid areas of Kenya. *J. of Appl. Meteoro.*, **21**, 477 – 494

Svensson, C., 1999: Empirical Orthogonal Function Analysis of daily rainfall in the upper reaches of the Huai River Basin, China. *Theor. Appl. Climatol.*, **62**, 147 – 161

Tennant, W. J. and B.C. Hewitson, 2002: Intra-seasonal rainfall characteristics and their importance to the seasonal prediction problem. *Inter. J. Climatol.*, **22**, 1033 – 1048

Tilya, F. F. 2006: Characteristics of wet and dry spells in Tanzania during the wet seasons. *MSc thesis, Department of Meteorology, University of Nairobi, Kenya*

Trenberth, K. E., 1997: The definition of El Niño. *Bull. Amer. Meter. Soc.* **78**, 2771 – 2777

Uppala, S. M., P. W. Kallberg, A. J. Simmons, U. Andrae, V. Da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. Vande Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Holm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P.McNally, J. F. Mahfouf, J. J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo and J. Woollen, 2005: The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, **131**, 2961 – 3012

Usman M. T. and C. J. C. Reason, 2004: Dry spell frequencies and their variability over southern Africa. *Clim. Res.*, **26**, 199 – 211

Van de Boogaard, H. M., 1977: The mean circulation of the tropical and subtropical atmosphere for July. *NCAR/Technical Note 118*, Boulder, Colorado, USA, 48pp

von Storch, H. and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge. 494 pp

Walker, N. D., 1990: Links between South African summer rainfall and temperature variability of the Agulhas and Benguela system. *J. Geophys. Res.* **95**, 3297 – 3319

Ward, N. M. and C. K. Folland, 1991: Prediction of seasonal rainfall in north Nordeste of Brasil using eigenvectors of sea-surface temperature. *Inter. J. Climatol.*, **11**, 711 – 743

Washington, R. and M. Todd, 1999: Tropical-temperate links in southern African and southwest Indian ocean satellite-derived daily rainfall. *Int. J. Climatol.*, **19**, 1601 – 1616

Wilks, D. S., 2006: Statistical methods in the atmospheric sciences. *Associated Press, 2nd edition.* 627pp

Williams, C. J. R., D. R. Kniveton and R. Layberry, 2007: Climatic and oceanic associations with daily rainfall extremes over southern Africa. *Int. J. Climatol.* **27**, 93 – 108

Wyrtki, K., 1973: An equatorial jet in the Indian Ocean. *Science*, **181**, 262 – 264

Xavier, P. K., 2002: Empirical prediction and predictability of dry and wet spells of the Indian summer monsoon. *M.S thesis, India Inst. of Sci., Bangalore, India*

Zhang, H. and T. Casey, 2000: Verification of categorical probability forecasts. *Weather and Forecasting*, **15**, 80 – 89

Zorita, E. and F. F. Tilya, 2002: Rainfall variability in Northern Tanzania in the March-May season (long rains) and its links to large scale climate forcing. *Clim. Res.*, **20**, 31 – 40