UNIVERSITY OF READING

Department of Meteorology

Investigating the Use of Satellite Derived Rainfall for Agricultural Insurance in Tanzania

Emmanuel Lubango Ndetto

A Dissertation Submitted in Partial Fulfilment of the Requirement for the Degree of M.Sc. Atmosphere, Ocean and Climate

August 2009

Dedication

To God,
(from whom knowledge and wisdom come)

&

In memory of my father Lubango Ram and Aunt Suzana Nyakake
(You laid the foundation of my education)

Acknowledgements

I would like to thank Dr. David Grime for his supervision and guidance to my dissertation. I am also grateful to Mr. James Sharpe for his helpful comments and discussions on my work and organising a trip to Tanzania in order to learn how agriculture insurance works as well to collect crop production data.

I would also like to express my sincere gratitude to the Microensure Company who sponsored the trip to Tanzania. Many thanks are to the overall sponsor of my master's degree program, British Foreign and Commonwealth Office through its generous Chevening scholarship programme.

My special thanks go to Dr. Charles William who agreed to read and provide constructive comments to my final draft. In the same way, I thank my course tutor Dr. Andrew Charlton-Perez for his continual effort to give guidance, advice and encouragement in my whole period of study.

Quite a few people have given assistance in my work in different aspects. I would like to mention Rob Chadwick for his assistance in using software for data analysis. Also are George Ameyaw and Timiza Wilbert who agreed to proof-read the write up.

Finally, I want to thank my family (my daughter Faith and my wife Agnes), mother Margret in Tanzania for their moral support during the study. They have contributed greatly to the achievement of my study indirectly. I could not imagine going through it without their patience and love. I dedicate the dissertation to them too.ist of acronyms and abbreviation

List of acronyms and abbreviations

CCD	Cold Cloud Duration
CERES	Clouds and Earth's Radiant Energy System
CHARM	Collaborative Historical African Rainfall Model
ENSO	El Niño Southern Oscillation
FEWS NET	Famine Early Warning Systems Networks
IOD	Indian Ocean Dipole
ITCZ	Inter Tropical Convergence
JAXA	Japan Aerospace Exploration Agency
LIS	Lightning Imaging System
MAE	Mean absolute error
MASIKA	Long rains usually happening during the March – May season
MSIMU	Rain season like MASIKA and VULI but often used to refer to areas having a one continuous season (uni-modal regime)
NASA	National Aeronautics and Space Administration (of United States)
NBS	Tanzanian National Bureau of Statistics responsible for providing statistics in every sector
PR	Precipitation Radar
RMSD	Root mean square difference
TAMSAT	Tropical Applications of Meteorological satellites
TIR	Thermal Infrared Images
TMI	TRMM Microwave Imager
TMPA	TRMM Multisatellite Precipitation Analysis
TRMM	Tropical Rainfall Monitoring Mission
VIRS	Visible and Infrared Scanner
VULI	Short rains usually happening during the October- December season

Abstract

In most African countries, agriculture is mainly rain-fed hence vulnerable to risks associated with weather and climatic hazards. Excess and inadequate rainfall both contribute greatly agriculture losses in these developing countries leaving the rural population poor and lack of food security. Index insurance particularly that uses rainfall as a proxy for loss is envisaged to relieve the rural farmers from the risks. Index insurance thus needs real-time availability of weather data. Satellite based rainfall is therefore sought to be used in order to extend this innovation to other places where weather stations are not closely available. This study investigated the use of satellite derived rainfall in agriculture insurance in two zones of Tanzania (north eastern highlands and south western highlands). The study used satellite products from TAMSAT archive and TRMM products. Rain gauge data and crop yield data for maize for Tanzania, altogether covering a period from 1997 to 2007 were also used. Calibration results for TAMSAT method suggests that -30°C is the optimal threshold temperature for most months except for January, February and October in some areas. Further, correlation results between rain gauge and satellite rainfall shows good results for both north eastern and south-western highlands where the comparison were done. However, poor correlation results are indicated in April and October for north eastern highlands. Co-fluctuation of yield and rainfall data was investigated and it was revealed that there is a good relationship in north-eastern highlands as compared to south-western highlands. These results indicate a high potential of using satellite based rainfall in agriculture insurance. A general conclusion for using satellite rainfall in agriculture insurance will better be reached after doing similar and further analysis on the whole area. The current results can therefore be treated as preliminary to this novel idea of satellite-rainfall application in agriculture weather index insurance.

Table of Contents

Chapter One

Introduction

1.1 Introduction

Many African farmers are vulnerable to rainfall variability, in particular to crop failure caused by inadequate rains in the rainy season. An obvious possibility is to have an insurance scheme which would cushion the effects of losses in dry years. The simplest approach is to trigger payouts based on rainfall amount compared to the amount required for a given crop. In order to be seen to be fair, the rainfall must be an 'official' measurement. Most of the current insurance schemes typically use the local regional Meteorological Office, which has been satisfactory for many of the small pilot schemes that have been in operation to date. However, throughout most of Africa official rain gauges are so sparsely distributed that most farmers would not be able to take part in similar schemes.

An alternative is to use satellite based measurements of rainfall. This would have significant advantages in terms of wide coverage of much of Africa without any infrastructure expense. The MicroEnsure Company based in Cheltenham, United Kingdom (UK) successfully run small scale insurance schemes in developing countries based around rain gauge data.

Currently, MicroEnsure Company is piloting a crop insurance scheme in Tengeru, Arusha Tanzania since 2008. So far the scheme has attracted 339 small scale farmers in that region. The aim of this project is to test the feasibility of using satellite data for a pilot project in Tanzania.

In Tanzania, agriculture has been the main employer since the colonial period. For example, according to the results of the national agriculture census conducted in 2003, the crop sector alone is said to provide income and sustenance to 4,858,810 rural households growing crops (about 99% of the total number of farming households in the rural areas and 95% of the total rural households). Maize dominates the small holder crop production in almost each region, though, there is wide variety of crops (over 95 types), (NBS, 2006)

Further results from the census reveal that there was zero growth in the number of households with access to irrigation over the last inter-censual period (1994 to 2003). It implies that small scale agriculture in Tanzania is mainly rain-fed. Since rain-fed agriculture is prone to weather shocks like drought and floods, then crop weather index insurance can be thought as one of the ways of relieving the majority of rural population against poverty due to crop failure (NBS, 2006).

1.2 Statement of the problem

In order to have sustainable weather index insurance, certain factors should be considered. Two of these factors are: 1) exposure to one or more spatially correlated weather events that can generate catastrophic losses; 2) historical data regarding the weather event(s) of sufficient quantity and quality. The data should also be reliable and available in real-time. This is especially much important in order to evaluate the frequency and magnitude of loss for weather index insurance contracts. Ideally, the time should be at least 30 years of daily or dekadal (10-day) measurements. It is further argued that measurements be conducted by a trusted third party for example a national or international meteorological association (Skees & Collier, 2008).

Skees and Collier (2008) further assert that the lack of available and reliable weather data in Africa is one of the major constraints when considering weather index insurance. In view of the above mentioned factors, it is worthwhile idea to consider using satellite based rainfall in order to help spread agriculture weather index insurance in much to the whole of Africa.

The technology of satellite rainfall estimates has been in use since 1970s and is becoming increasingly more accurate over time (Barrett, E.C., 1970, Ebert et. al., 2007, Lebel et al., 1992, Dinku et al., 2007, Xie et al., 2007). This makes nearly 30 years of data availability in many regions hence meeting one of the basic requirements of weather index insurance (Skees & collier, 2008). Since that time, wide applications of satellite data in various areas have been considered and often feasibility study have also been done in different parts of the globe (Sawunyama & Hughes, 2008, Wardah et. al., 2008, Ebert et.al., 2007, Todd et al., 2001, Sapiano & Arkin, 2009 and Yan & Gebremichael, 2009). In Tanzania, a study of satellite rainfall estimates has been done only over the Rufiji river basin in view of hydrologic application. The study revealed that the use of cold cloud duration (CCD) as a back ground image proved to be very useful in estimating rainfall quantity (Moges et. al., 2007). By definition, Cold Cloud Duration (CCD, usually measured in hours) is the length of time that a pixel is colder than a given threshold temperature (Thorne et al, 2001). The idea of cold cloud duration as used in satellite rainfall estimation is described in detail in section 3.1.2 of this study.

Though, there have been suggestions and at the same time doubts of using satellite based rainfall in index insurance from some of the index insurance community (Skees & Collier, 2008); such doubts are based only on previous literature of satellite estimated rainfall. Such estimates have been done in reference to such application. The same study also narrates that models adding satellite data to rain gauge interpolation and topography estimates, such as the Collaborative Historical African Rainfall Model (CHARM) can contribute to weather risk gridding. Satellite data can then serve as a valuable check, providing actual data values between the rain gauges.

To clarify the situation, investigative studies should be done in different areas of the globe especially in Africa where the trend of rain gauges maintenance is said to be declining (Ali et al., 2005; Sawunyama and Hughes, 2008). This study therefore aims to investigate the use of satellite derived rainfall in agriculture index insurance in Tanzania.

1.3 Area of study

Tanzania is located between latitudes 1°S -12°S and longitudes 29°E - 40°E and about 87 rain gauge stations distributed over the country were available during the study period to provide the ground data (figure 1). Tanzania has varied topography as can be observed in figure 2. It can be seen that the great East African lakes (i.e. Victoria in the north, Tanganyika to the west and Nyasa to the south) together with the Indian Ocean to the east, form a water boundary to the country (Basalirwa et al., 1999). The area of Tanzania is about 945, 100 square kilometres (NBS, 2008).

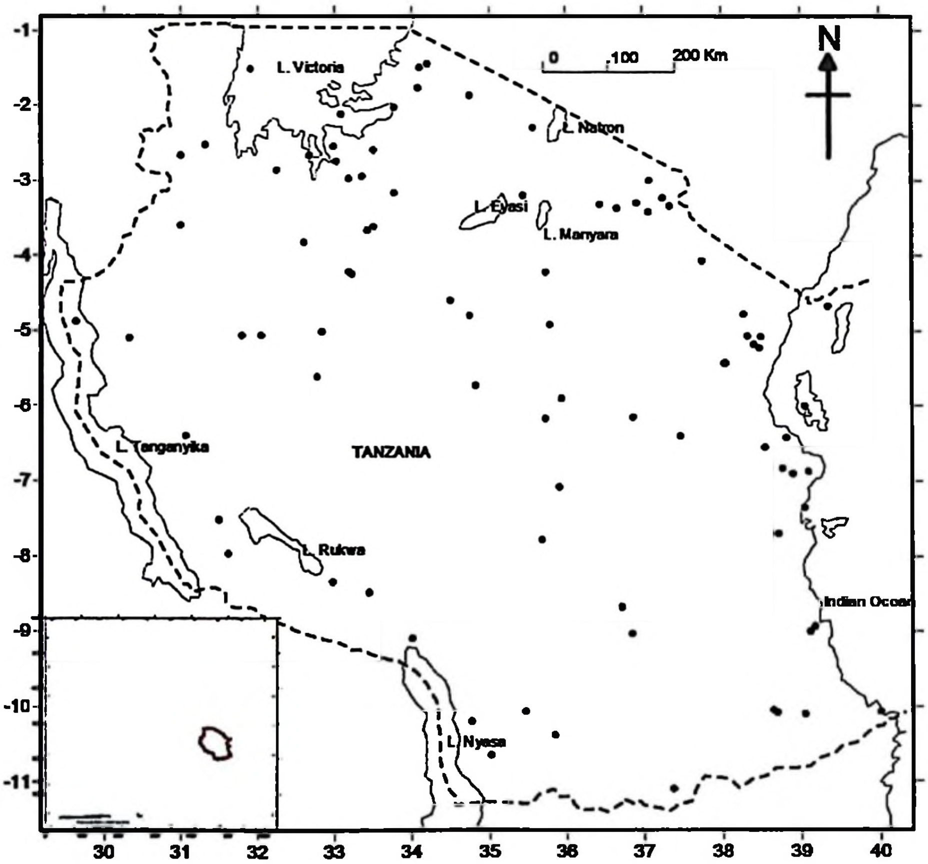


Figure 1: Distribution of rainfall stations used in the study. The insert shows the location of Tanzania on a map of Africa.

Besides, the inland water bodies, the country's topography also include highlands over the north eastern, south western and extreme north western parts of the country. The highest point of Africa, Mount Kilimanjaro at about 5950 m above sea level is located within the north eastern highlands. The highland areas and the coastal region including the Islands of Zanzibar and Pemba receive substantial amount of rainfall of over 1500 mm a year. The rest in the central plateau receives rainfall as low as 500 – 1000 mm per year (Basalirwa et al., 1999; citing Griffiths, 1972).

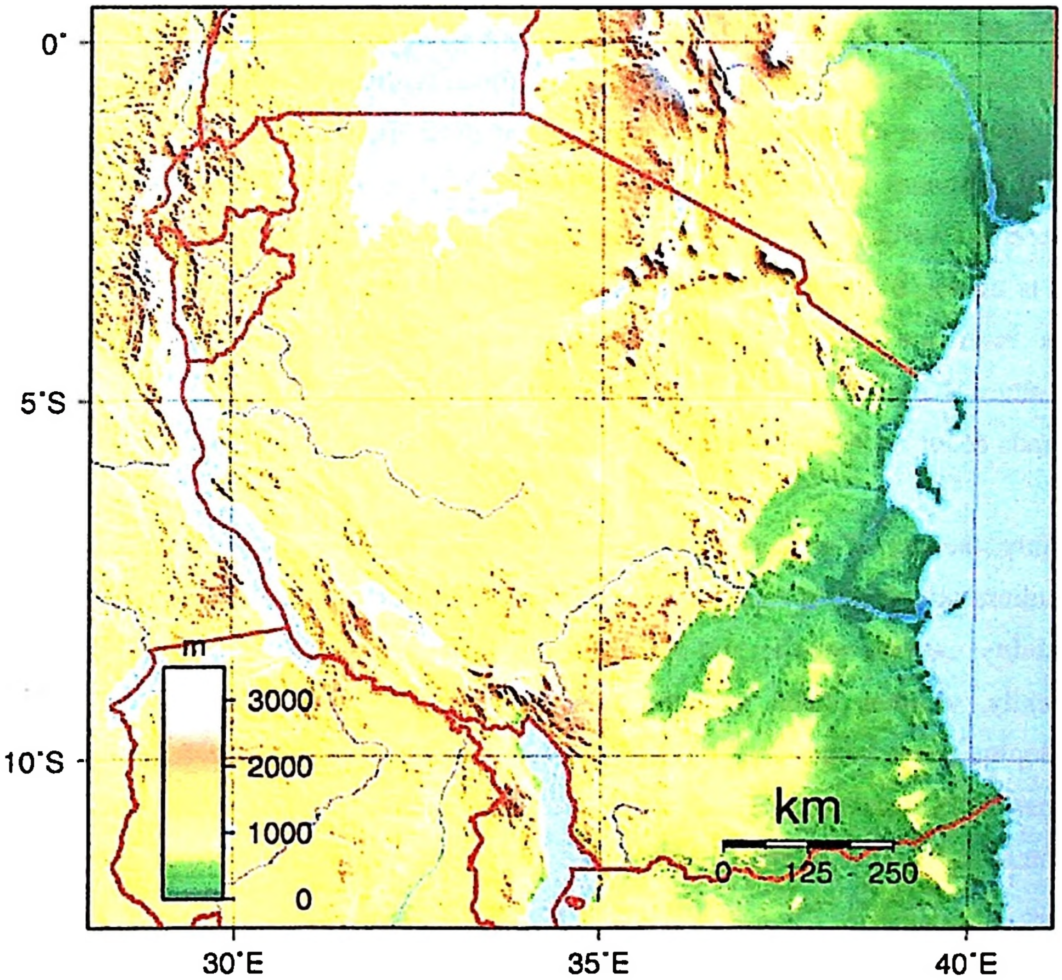


Figure 2: Map of Tanzania indicating topography of the country.

Source: Wikimedia Foundation

The great East African Rift valley passes through the country from Kenya along the north eastern highlands to south western highlands where it divides with one branch running northwards along Lake Tanganyika to Lake Kivu in Uganda. The other branch goes southwards along Lake Nyasa. Due to diversity of physical features and the play of various synoptic conditions, the country experiences both spatial and temporal variation in rainfall. Rainfall climatology is described in more detail in chapter two.

1.4 Motivation

The current wide application of satellite derived rainfall for example in flood forecasting, crop yield modelling, hydrological modelling, food security and early warning presents a high confidence of widening its application in index insurance. TAMSAT method of rainfall estimates has been used in various tropical countries and so far results are impressive (Wardah et al., 2008, Sawunyama & Hughes, 2007 and Grimes, et al, 1999). This is among of the force behind the investigation of satellite rainfall application in index insurance in particular to Tanzania where the economy depends much on agriculture (Kijazi & Reason, 2005, 2009, Zorita & Tilya, 2002, Basalirwa, et al., 1999, Kabanda & Jury, 1999, and Mapande & Reason, 2005).

Possibly, the use of satellite based rainfall in index insurance is potential to stimulate agriculture activities especially among the rural farmers who are vulnerable to rainfall variability, extreme weather and other climatic hazards impacting agriculture production. Currently, weather index agriculture insurance though been established in various developing countries base on gauge rainfall (Hellmuth et al., 2009, Hazell & Skees, 2005, Werner, 2005, Tressler & Mbaka, 2009, and Skees & Collier, 2008). Thus, weather index insurance in Africa is currently benefiting a few farmers around the weather station.

Therefore, the current problems facing agriculture in Tanzania, sparseness of rain gages and the envisaged potential benefits of index insurance especially by incorporating satellite derived rainfall are the motives of this study.

1.5 Objective of the study

The main aim of the study is to test the feasibility of using satellite data for a pilot project on agriculture insurance in Tanzania. In order to make sure that this goal is realizable, the following objectives were put forward:

- To describe the characteristic of rainfall climatology using the rain gauge datasets.
- To assess the validity of satellite derived rainfall in different parts of the country.
- To explore the basic requirements of weather index insurance and how it operates.
- To investigate correlation between satellite-rainfall and gauge rainfall with crop yield over specific areas in Tanzania.

1.6 Justification and significance of the study

In weather index insurance payouts base on a weather index that is highly correlated with actual losses. For example, rainfall is often used since in most crops, production or growth is affected mainly affected by water availability. In fact if farm management is good, other factors that affect crop production have little variation from year to year. Such factors are crop variety, fertilizer use, salinization, pests, disease and mechanization (Senay and Verdin, 2003). Thus, rainfall remains as a single major factor affecting crop production in any rainfed agriculture.

Since losses may be caused by disease, insect infestation, or any number of factors other than the weather variable on which the index is based, then payouts cannot be made basing directly to crop yield. This involves farm-level risk assessment which is a complex process requiring expertise which would not be readily available in rural areas of developing countries. Index insurance contracts are relatively straightforward, operating at low costs relative to traditional insurance products due to the simplicity of sales and loss adjustment, requiring no farm-level risk assessment or loss adjustments. The goal is only to ensure available risk management to the rural poor.

A problem with index contracts is that an individual can suffer a loss and not be paid because the major event triggering a payment has not occurred. For example, a farmer with rainfall insurance could lose a crop to drought at a micro-location, but not receive a payout if the rainfall at the regional weather station remains above the trigger point. It is also possible for an individual to be paid when they suffer no losses. As defined by Barnet and Mahul (2007), "basis risk" is imperfect correlation between the index and losses as experienced by the policyholder. With this problem, index contracts trade off higher "basis risks" for lower transaction costs. Weather index insurance will not be attractive if the basis risk is too high (Hazell and Skees, 2005).

There is also a question of differences in soil moisture holding capacity. This will result in individuals experiencing different crop risks in the same region.

Weather index insurance requires not only accurate weather data, but also should be secure and independent of individual influences. Ground weather stations are rarely very secure and could be tampered with once significant sums of money are dependent upon their readings. New options have been sought like having automatic rain gauges which are evenly distributed in the region. However, an optimal system is to have a variety of verification systems like remote sensing data taken from satellite images, Doppler radar, or soil moisture readings.

It is also worth stating that satellite data have a wider coverage as it relates to an area than the rain gauge data which relates a single station.

Because of the investments on satellite imagery, products using satellite data continue to improve, providing more precise and real-time estimates of weather in remote locations in Africa (Xie et al, 2007 and Todd et al, 2001). Satellite data will also represent a low-cost alternative to weather station data for index insurance and offer a real-time data that can track emerging weather trends as they occur. Furthermore, it has the potential to lower basis risk.

1.7 Structure of this dissertation

This dissertation has been organised into five chapters with the first chapter giving the general introduction of the dissertation. It basically informs about the whole question of weather-index insurance which in a long run is also seen as a means of adaptation to climate change for rural farmers. Chapter two lays the knowledge foundation of the major parts of this work by giving a review of related information about the problem and what has been done elsewhere. This involves a review of the currently available literature.

The methods for analysing data as used in this study are briefly discussed in chapter three while the results and discussion are presented in chapter four. The last chapter concludes the major findings of the study and also giving an outlook for suggested future work on the subject.

Chapter Two

Literature Review

2.1 Overview of agriculture weather index insurance

In this section, relevant literature about the weather index insurance in agriculture is presented. The section begins with highlighting how index insurance differs from traditional weather insurance before discussing how it works. The section finally gives the experience of weather insurance in Tanzania.

2.1.1 Risks in agriculture

The agriculture sector is prone to several weather related risks such as drought, floods, hail, frost and cyclone. One way of dealing with weather risks is to institute agriculture insurance at least to reduce the effects on farmers. Agriculture insurance is also considered as a component of adapting to climate change especially in places where climate projections show increase in floods and/or droughts.

Lack of insurance for weather related risks contribute both directly and indirectly to recurrent chronic poverty especially to rural households. Households that recognise the potential for weather related shocks become reluctant to forego short-term consumption to invest in risky productive assets, resulting to low production and forcing farmers to a cycle of poverty.

2.1.2 Weather index insurance

Two forms of insurance can be considered in agriculture sector. Traditionally, insurance is done by assessing the actual loss of a particular product e.g. crop production after a certain catastrophe. Such kind of assessment can be subjective if not carefully done (Skees & Collier, 2008, Burton, I., & T. Dickinson, 2008).

On the other hand, weather index insurance is a form of insurance where payouts are dependent upon the recording of a specific weather phenomenon captured objectively in an index. It relies on the occurrence of a weather event (for example a rainfall event), rather than crop failure which is the consequence of weather. Briefly, a location with historical index data is chosen. The mean of the index (e.g. average rainfall, temperature, and water levels) is calculated, and a range below the average index (rainfall) is selected as the threshold value. Payments are now based on seasonal rainfall being within this range (threshold). If the rainfall is equal to or greater than the threshold value, no payment is made. If the rainfall is less than the chosen threshold value (e.g. in the event of drought) the index is triggered and payment is automatically delivered to the farmer, regardless of crop outcome (Burton, I., & T. Dickinson, 2008).

The weather data, now serve as a proxy for loss; eliminating the need for costly individual loss assessments. The index is based on an objective measure such as rainfall, livestock mortality, temperature, water levels in a river, etc. These measures should be highly correlated with the economic losses (crop failure, death of livestock, loan defaults, etc.) that might be experienced by an insured entity but (unlike the actual loss) the insured has no ability to affect the index (Skees & Collier, 2008).

Historical weather data is particularly important in order to understand weather risks sufficiently to allow for the design and pricing of weather index insurance contracts. If the future availability of reliable weather data is not assured, no weather index insurance is possible (Barnet & Mahul, 2007). This can be a constraint in many regions of Africa where weather station infrastructure is simply too sparse and poorly maintained (Skees & Collier, 2008).

2.2 Rainfall climatology for Tanzania

In general, weather controls in Eastern Africa are associated with the interplay between synoptic (large scale) and mesoscale (small scale) systems. The later are accentuated by topographical features that include large inland water bodies, mountains and highlands

(Basalirwa, et al., 1999). The synoptic weather systems include Inter Tropical Convergence zone (ITCZ), high pressure cells (mainly Mascarene and St. Helena highs and sometime the Arabian ridge), the Congo air mass and sometimes the teleconnections effects of El Niño Southern Oscillation (ENSO). In this subsection we attempt to give a brief account of how these factors control the rainfall climatology over Tanzania.

2.2.1 Inter Tropical Convergence Zone

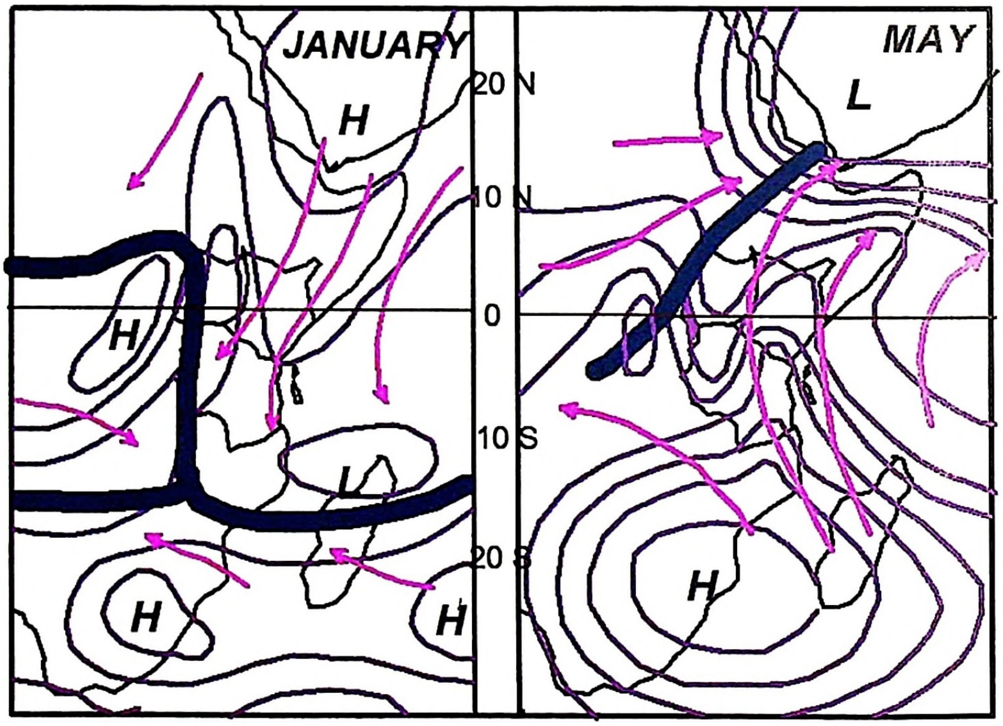


Figure 3: Mean positions of ITCZ during January and May. Left panel shows the splitting of ITCZ into zonal and meridional arms.

The Inter Tropical Convergence Zone (ITCZ) is a zone of low-pressure near the equator (near equatorial trough). It is characterized by maximum surface heating and meeting of the two easterly trade winds that originate from the Northern and Southern hemispheres. Thus, convection, cloudiness, and rainfall are enhanced within the zone.

Over Central Africa, the ITCZ breaks into two components, the zonal and meridional components. The division is a consequence of the topography mainly the Great Rift Valley and the mountain chains of East Africa. It is the zonal arm which delineates the seasonal characteristics over most parts of East Africa. This branch generally lags the overhead sun by 3-4 weeks and migrates seasonally north and south following the overhead sun creating a bimodal nature of rainfall over the most parts of the region.

Figure 3 illustrates the mean position of ITCZ over eastern Africa in January and May. The associated seasonal winds as influenced by the anticyclones are also shown. During December – January (DJF) the meridional arm of ITCZ influence much of rainfall over Tanzania especially on the unimodal areas.

With particular reference to Tanzania, the bimodal rainfall is experienced over the Lake Victoria, North Eastern Highlands and the northern Coast. The bimodal seasons are generally termed as "Short rains" or the October – December (OND) season and "Long rains" or the March – May (MAM) season. The OND is locally known as VULI while MAM is known as MASIKA (Camberlin & Okoola, 2003, Song, et. al., 2004, Mapande & Reason, 2005, Kijazi & Reason, 2009).

Figures 4 and 5 illustrate the rainfall regimes in Tanzania. These illustrations can be simply observed by using the mean monthly rainfall of the particular station. The Arusha station (north eastern highlands) illustrates the bi-modal regime while Songea station (southern) illustrates the unimodal rainfall regime. In the bi-modal regime usually the MAM season more intense than the OND (Zorita and Tilya, 2002). However, in the Arusha plot the situation is contrary and here it is important to note the influence of 1997/98 El Niño event which usually affects the OND season (Kijazi & Reason, 2005, and Kabanda & Jury, 1999).

Over satellite pictures, ITCZ can be easily located as a band of clouds though not always connected as a continuous line. However, over east Africa due to the presence of various topographic features, it is diffuse and therefore sometimes difficult to locate on weather maps (Camberlin & Okoola, 2003).

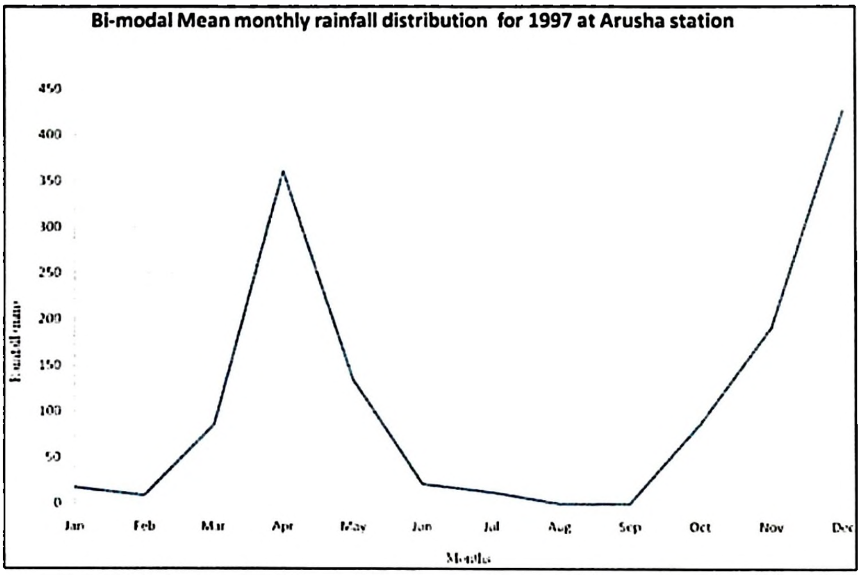


Figure 4: Rainfall distribution for Arusha synoptic station (northern) depicting the bi-modal regime

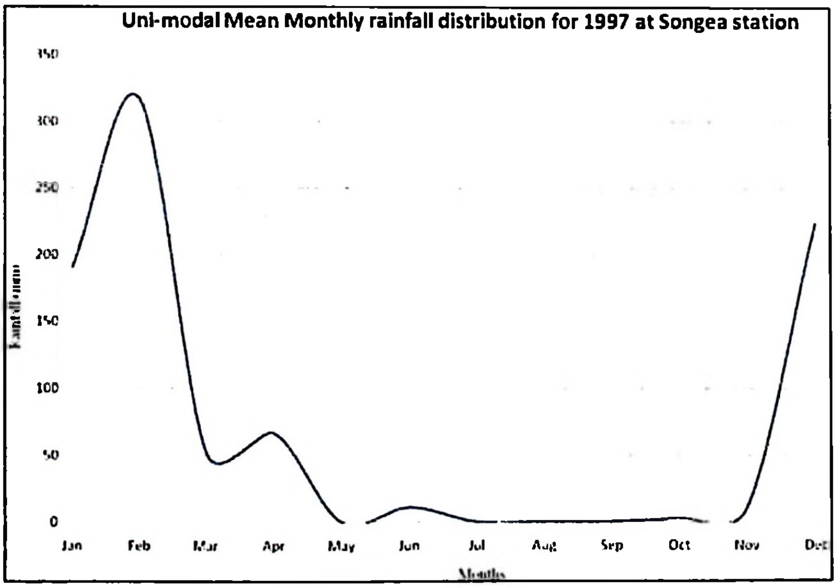


Figure 5: Rainfall distribution for Songea synoptic station (southern) depicting the unimodal regime

The other parts of the country experience one long rainy season starting from October to May. It is also highly influenced by the movement of the ITCZ and the interplay of other factors (Kijazi & Reason, 2009, Basalirwa, 1999). The meridional arm of ITCZ, which has little movement from east to west, influences some parts of western Tanzania when it especially facilitates the penetration of westerly winds further eastwards. It has been observed that it leads to wet conditions especially when influenced by the two southern subtropical highs (Sun, et., al., 1999). The influence of subtropical highs on rainfall over Tanzania is described in the section 2.2.2.

2.2.2 Subtropical anticyclones

In contrast to the ITCZ, the subtropical anticyclones are quasi-stationary centres of high pressure. There are four anticyclones whose position, strength and orientation may influence weather in east Africa generally. These are the Southwest Indian Ocean (Mascarene High), Southeast Atlantic Ocean (St Helena High), North Atlantic Ocean (Azores/Saharan High), and the Arabian High.

High pressure cells influence weather by forcing the seasonal winds which carry moisture from the oceans in different times of the year. The differential strength of the sub-tropical high-pressure systems also determines the relative positions of the ITCZ and the associated weather systems. For example as mentioned in section 2.2.1 the enhancement of St. Helena High and weakening of Mascarene High cause the meridional arm of ITCZ to shift eastwards. These cause wet conditions in most areas over western Tanzania as the region is also influenced by the moist air from the Congo tropical forests (Sun, et., al., 1999). This is a combined effect of the two anticyclones.

Figure 6, indicates the relative positions of the four anticyclones influencing weather over Africa. The Mascarene high is split into two parts possibly due to the presence of Madagascar Island hence being diffused. This makes the Arabian ridge to extend more southwards and cause the zonal arm of ITCZ to be more on the southern part of the country and the meridional arm influencing the western part. The end result is for these parts to experience a long rainy season (locally known as "MSIMU").

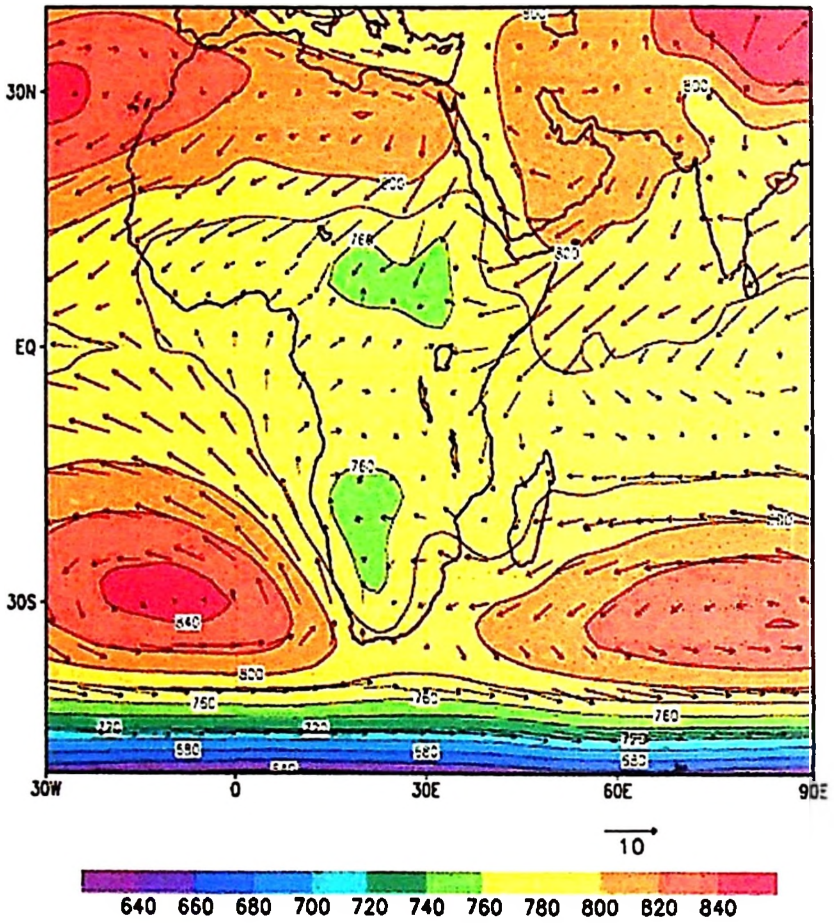


Figure 6: Mean positions of the anticyclones over African continent and the associated wind field. Source: Tilya, 2007.

Individually, the Mascarene high drives the south easterly flow from the Indian Ocean towards the east Africa. While St. Helena high forces the south westerly flow and in some years their influences has been associated with flood like the 1961 floods in the country (Kijazi & Reason, 2009).

In general, the Arabian ridge has not much influence to east Africa in terms of rainfall. The ridge comes into play especially in January and since the north easterly winds are mainly continental by origin, this period is generally dry over most areas.

2.2.3 Other synoptic factors

Besides the St. Helena high being a force for westerly flow, the equatorial rainforest of Congo contributes a lot of moisture through the Congo air mass. This air mass converges with the easterly flow from Indian Ocean resulting in heavy precipitation during April especially over the western part of the country (Mapande & Reason, 2005).

The greatest teleconnection effect which has already been exhaustively studied is the influence of ENSO on the OND short rains. Several studies have been done to account its effects on different parts of the country especially over the northern coast and northern part of the country (Kijazi & Reason, 2005, 2009, Nicholson, and Camberlin & Okoola, 2003). In addition, other teleconnections include the variation of the Walker circulation which Influences Sea surface temperatures (SSTs) over Indian Ocean and consequently enhances (reduces) moisture flux over central East Africa (Ummenhofer, 2009, and Kijazi & Reason 2009). In this list also is the recently observed Indian Ocean Dipole (IOD) which sometimes accompanies El Nino hence influencing the short rains significantly (Kijazi & Reason, 2009).

The mesoscale weather systems over East Africa in general are associated with the presence of prominent physical features. For example, Lake Victoria is said to cause its own circulation system and this causes the significant spatial rainfall variations where the western lake region is wetter than the rest of lake regions (Song et al, 2004). Highlands and mountains, rift valley and Indian Ocean are other sources of mesoscale systems like sea breeze, mountain winds. Effects of such mesoscale systems have been well accounted by several studies in rainfall climatology over the country (Basalirwa et al 1999, Zorita & Tilya, 2002, Kijazi & Reason, 2005, and Mapande & Reason, 2005).

It should be understood that most of the weather systems interact together and there is no single factor which influences weather at a single time and this has also made the climatic classification based on rainfall more complex in Tanzania (East Africa generally). Studies on homogeneous rainfall classification in Tanzania have come up with different regions. Basalirwa et. al. (1999) came up with 15 zones (see figure 7) while Tilya (2007) came up with 11 zones. Earlier studies as cited by Basalirwa et al. (1999) indicated that by using synoptic systems along the country can be divided into 10 or 4 dominant rainfall homogeneous zones as studied by Ogallo (1989) and Nyenzi (1992) respectively.

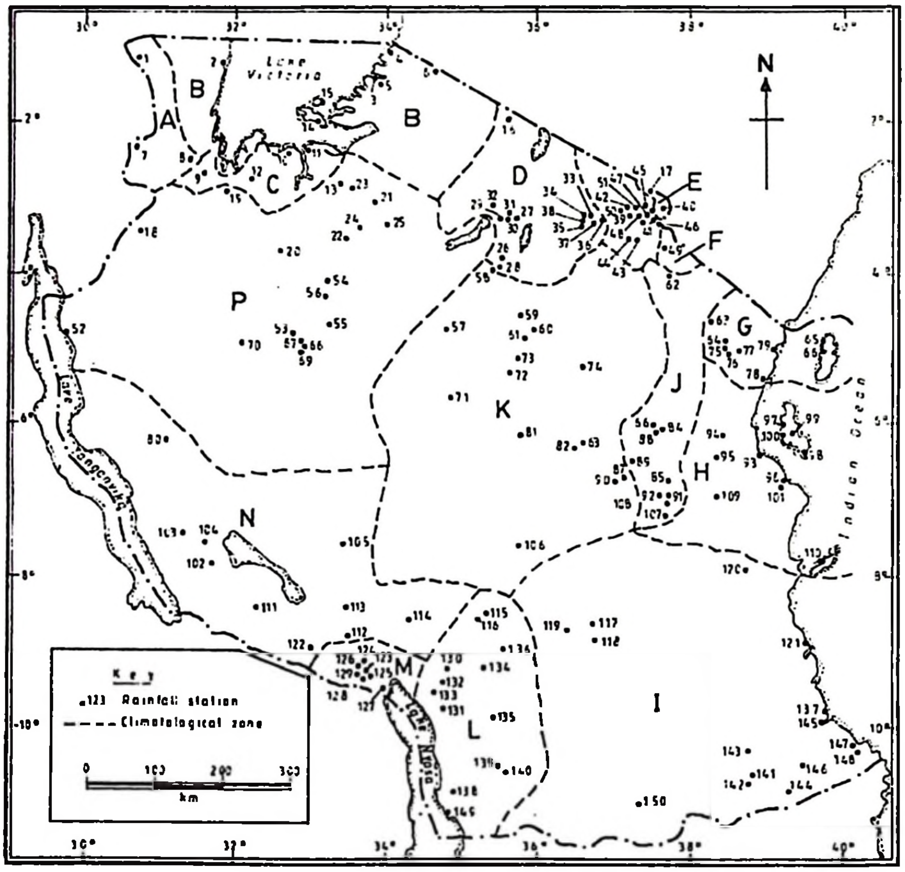


Figure 7: 15 rainfall homogeneous zones as delineated in Basalirwa et al (1999). Zones A – H and J are bimodal while I, K – N and P experience unimodal rainfall regime.

Source: Basalirwa et.al. (1999).

Data and Methods

3.1 Data type and sources

In this section, the data and its associated sources are described. In addition, data quality control is also discussed. Generally, the study involved two main datasets, which are rainfall (both gauge and satellite estimates) and crop yield (maize) for Tanzania.

3.1.1 Gauge rainfall data and data quality control

The rainfall data set involved in this study consists of daily and/or dekadal (ten day) rainfall from a total of 87 gauges. These data were obtained from Tanzania Meteorological Agency for the period ranging from 1970 to 2007 in some of the gauge stations. However, our main focus is the period between 1997 and 2007 corresponding to the period where the satellite information is available. It should be remembered that TAMSAT estimation started from 1989 while TRMM data are available from 1998.

However, the additional range of period was quite useful in studying the rainfall climatology for Tanzania.

In an attempt to do an initial data quality control, the number of missing values for each station was determined within their range of available data for the particular station. Percentages of missing data were also calculated in order to determine the relative quality of gauge data from all stations used in the study. Out of 87 gauge stations, only 21 stations (about a quarter of all gauges) had no missing value. Most of these stations are synoptic making observations every hour. Five stations have more than 50% of missing values and these stations were not considered in analysis (refer to column 6 of table 1). Among these five stations, Biharamulo gauge has highest number of missing data (i.e. 75%). It has data only for the year 2002 and 2008 which is also out of our focus period.

We also put a more stringent condition to calculate the percentage of missing values by considering the period of 1997 – 2007 irrespective of the available range of individual stations. By this condition, any station with more than 60% of missing data should not be considered. Stations within 50% to 60% of missing values were only retained provided it has at least 4 consecutive years of available data. For example, by this condition, Loliondo, Handeni Meteorological and Mpanda Bomani stations were retained while Sikonge Hyrometeorological station was not considered. Table 1 presents the list of stations which were not considered in the analysis after failing to meet the two conditions imposed for data quality control. The full list of stations used is provided in appendices (Table A1) in which the total number of stations used in analysis is 78.

Table 1: List of gauge stations which did not pass the data quality control conditions

Station Name	Lon	Lat	Data Period	Data type	Missing data (%)	1997-2007 missing data (%)
Biharamulo	31.32	-2.63	2002-2008**	Dekadal	75	91
Ibadakuli	33.5	-3.62	2002-2007	Dekadal	29	61
Kaliua	31.8	-5.07	1997-2007	Dekadal	61	61
Magu	33.45	-2.6	1997-2005	Dekadal	43	61
Masasi Girls	38.82	-10.7	1997-2005	Dekadal	70	75
Maswa	33.77	-3.17	1997-2002	Dekadal	60	65
Mbamba Bay	34.77	-11.03	2002-2007	Dekadal	36	65
Mwadui	33.59	-3.53	1997-2002	Dekadal	58	69
Sikonge Hydromet.	32.77	-5.62	1997-2007	Dekadal	41	59

* For Biharamulo, only data at years 2002 and 2008 were available.

** Available data for Gairo were only between 1972 and 1989 inclusive.

3. 1.2 Satellite derived rainfall

Techniques of deriving rainfall estimates from weather satellites have been motivated by early studies and there have been further investigations into the possibilities of deriving acceptable estimates of rainfall distributions over the globe (Barrett, 1970, Diro et al., 2009). The need for estimating rainfall from satellites is due to the fact that precipitation is generally discontinuous and highly variable in both space and time. Thus, estimates could at least give comprehensive coverage of the entire surface of the earth (Sapiano & Arkin, 2009).

Satellite precipitation estimates can be grouped into two major types. One being those based on geosynchronous infrared (IR) measurements and those derived from polar-orbiting microwave observations. The former exhibits a high sampling frequency (full globe scans are currently available every 15 min), however their relationship to precipitation is indirect. It means outgoing longwave radiation is used to locate the position, and cloud-top temperature of cloud masses and precipitation rates are inferred from these attributes (Sapiano & Arkin, 2009 citing Arkin and Meisner, 1987). On the other hand, Sapiano & Arkin, (2009) further narrate that the microwave sounders have a more direct inference of precipitation.

In this study two products of satellite estimation techniques are used to determine the possibility of using satellite derived rainfall in index insurance. Both of the products are developed to be used specifically in tropic regions. The Tropical Application of Meteorology using Satellite (TAMSAT) is an algorithm which has been developed at University of Reading and has been used in different application in various regions of African continent (Thorne, et al., 2001, Nanteza, 2008, Assiri, 2006).

The other product is V6 of TRMM and other satellite merged rainfall (3B42) from Tropical Rainfall Measuring Mission (TRMM). TRMM is a joint mission between the National Aeronautics and Space Administration (NASA) of United States and Japan Aerospace Exploration Agency (JAXA) launched in November 1997. TRMM is designed to monitor tropical rainfall (Nanteza, 2008) and has provided quality rainfall data for more than 9 years. TRMM operations have been recently extended to 2009 hence increasing its potential value for application use (Chokngamwong and Chiu, 2008).

In this study, TAMSAT datasets for the period 1997 to 2007 were obtained from TAMSAT archive and the TRMM data for the period 1998 to 2007 from the NASA archives were used. The brief description of the algorithms for both satellite rainfall products is given in the following sub sections.

3.1.2.1 TAMSAT algorithm for satellite rainfall estimation

Derivation of rainfall from satellite images by the TAMSAT algorithm dates back to 1988 when it was used to produce rainfall estimates for operational purposes in northern Africa, particularly for drought and famine warning (Grimes, et al, 1999). For southern Africa, it was first used since 1993 (Thorne, et. al., 2001, Tucker & Sear, 2001, Dinku, et. al., 2007). Since then, validation studies into various parts of Africa has been done and there is a strong confidence of using it especially in plateau areas where previous studies indicate that it has best performance (Thorne, et. al., 2001, Dinku, et. al., 2007, Nanteza, 2008 and Assiri, 2006).

Basically, the TAMSAT algorithm uses the Cold Cloud Duration (CCD) method with two basic assumptions. In Meteosat Thermal Infrared Images (TIR) cloud-top temperatures are measured and hence a CCD can be determined. TAMSAT algorithm first assumes that the rainfall is produced by clouds (convective) in tropical storms which can be identified by cloud-top temperatures and secondly the CCD is linearly related to the amount of rainfall (Thorne et. al., 2001, Bellerby et. al., 2000, Tucker & Sear, 2001, Dinku, 2007 and Moges et. al., 2007).

Usually, the CCD estimates as produced by TAMSAT method are summed over ten days (dekadal) which is also a significant period for crop growth. Mathematically, the relationship can be given as:

$$Rain = a_0 + (a_1 \times CCD) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (1)$$

The parameters of this relationship are found through calibration using rain gauge data (Thorne et. al., 2001 and Dinku, 2007). Since 1990, dekadal CCD images for southern Africa have been produced with threshold temperatures of -40°C, -50°C and -60°C and it have been extended to include -30°C and -70°C since 1996 (Thorne, et al., 2001). The threshold temperature is defined as that which gives the best agreement between gauges and their containing satellite pixels as to the presence or absence of rain and is done by drawing up contingency tables (Grimes, et al 1999).

In this study the TAMSAT satellite rainfall estimates were generated at the threshold temperatures of -30°C, -40°C, -50°C and -60°C for the period from 1997 to 2007 with a pixel resolution of about 6km × 6km (0.05°×0.05°). Then calibration was done at these threshold temperatures using the CCD value and gauge rainfall in order to determine the optimal threshold temperature for Tanzania.

During calibration, the median of gauge rainfall (mm) was regressed against the CCD (hours) values for each zone containing the gauges (Grimes et al., 1999). This was done using a contingency table to record the regression parameters (see table 2 for an example).

Table 2: Contingency table used to select the optimal threshold

	$CCD = 0$	$CCD = 0CCD$
Rainfall = 0	n_{11}	n_{12}
Rainfall > 0	n_{21}	n_{22}

Then, the optimal threshold temperature is that satisfies the following two conditions.

$$n_{11} + n_{22} \gg n_{12} + n_{21} \quad \ldots\ldots (2)$$

and

$$n_{21} \cong n_{12} \quad \ldots\ldots (3)$$

The relationships between IR - derived cloud indices and rainfall are variable in both space and time and mostly are calibrated using gauge data. Correction is sensitive to the density and spatial distribution of the gauge network.

3.1.2.2 TRMM rainfall products

The TRMM satellite on the other hand is in a non sun-synchronous orbit providing sampling of the complete diurnal cycle of rainfall. Detailed description of TRMM is given in Kummerow et al. (2000) and Chokngamwong & Chiu (2008) give a brief description about the recent state of the mission's products. In brief, the TRMM platform consists of five instruments: the TRMM Microwave Imager (TMI), the Precipitation Radar (PR), the Visible and Infrared Scanner (VIRS), the Clouds and Earth's Radiant Energy System (CERES), and the Lightning Imaging System (LIS). Initially, TRMM was the three-year mission with the objective of measuring rainfall and energy exchange (i.e. latent heat of condensation) of tropical and subtropical regions for a better initialization of global weather and climate models (Levizzani et. al., 2002).

While TRMM Microwave Radiometer (TMI) is passive and the TRMM Precipitation Radar (PR) is active; together they provide high quality, high resolution, instantaneous microwave based estimates (Bellerby et al., 2000 and Feidas et al., 2006). TRMM PR also boosts in the absence of sun-synchronicity and this makes it to be a good source of satellite derived rainfall to be considered in many studies (Bellerby et. al., 2000).

In our study, TRMM data which was extracted from the internet with the use of MicroEnsure in-house software (Perl software) that extracts dekadal estimates make the second set of satellite derived rainfall. However, these dataset is at a much coarser resolution as compared to TAMSAT products. The pixel resolution of TRMM rainfall estimate is $0.25^{o}\times0.25^{o}$. TRMM product is also disadvantaged as it does not have a regular overpass at a particular location which means it cannot be entirely used by itself to give daily, dekadal or monthly totals. It is usually used in conjunction with other data from other satellites. In this study, TRMM data were extracted for the period of 1998 – 2007. This TRMM rainfall product is known as Tropical Rainfall Measuring Mission Multisatellite Precipitation Analysis (TMPA) 3b42 Version 6 channel (Sapiano and Arkin, 2009).

Though, TRMM product seems to be disadvantaged as compared to TAMSAT, the idea to incorporate it is to have a variety of satellite rainfall sources. This can give us some other information which could not have been obtained on one satellite rainfall source.

3.1.3 Crop yield data

The crop yield data for maize were obtained from the Tanzanian Ministry of Agriculture and Food Security as well from the National Bureau of Statistics (in collaboration with the Planning Commission). All these government bodies are based in Dar es Salaam, Tanzania.

Basically, the yield data were derived from the socioeconomic reports of the 21 regions of Tanzania Mainland, the period ranging from 1991 to 2005. These data were supplemented by those from Tanzania National Agriculture Census report. Measurements of maize yield survey done by village agriculture officers who report annually to the district agriculture officer and then to the ministry responsible of agriculture sector.

In addition, the Tanzanian National Bureau of Statistics (NBS) in collaboration with the National Planning Commission and Ministry of Agriculture and Food Security do conduct agriculture census nationwide at predetermined years. The main method of data acquisition in census is the survey and interviews. The last census was done in 2002/2003 year and report for the next census of 2007/2008 is expected to be out at the end of this year (personal communication, and NBS, 2006).

For instance during the 2002/2003 census, large scale farms were completely enumerated while sampling was done for small farms. In the report, crop cultivation referred to agricultural year 2002-03 means the production from October 2002 to September 2003 (NBS, 2006).

Maize was chosen since currently, MicroEnsure Company is piloting index insurance for maize in Tengeru Arusha. In addition, maize has become a dominant crop grown in almost every part of the country.

3.2 Data Analysis

The main goal of this study is to investigate the possibility of using satellite derived rainfall in index insurance for Tanzania. In order to accomplish this goal the study uses a three-step analysis procedure. First, is to study the rainfall climatology for Tanzania through time series analysis and graphical analysis. This is mainly directed to understanding of the nature of temporal and spatial variability of rainfall as it has been pointed out by previous studies (Kijazi and Reason, 2009 and 2005, Sumner, 1983, Mapande & Reason, 2005 and Zorita & Tilya, 2002).

Having reviewed the rainfall climatology, the following analysis is to determine the relationship between satellites derived rainfall and rain gauge measurements by using various statistics. A range of these statistics includes the bias, correlation and root mean square (difference) error.

The last analysis of the study involves trend analysis and correlation to determine the relationship between the rainfall and the maize yield. Such method has been used by Senay and Verdin (2003) on characterizing yield reduction in Ethiopia. This was done by simply determining the year-to-year variation of yield and the corresponding water requirement satisfaction index (WRSI) value. Brief description of each analysis method then follows in the subsequent sub sections.

3.2.1 Descriptive statistics

Empirical descriptive statistics were performed on the gauge rainfall datasets in order to explore the distribution of rainfall in Tanzania. Though, it is not possible to analyse the spatial distribution by empirical statistics, our analysis is geared towards exploring the nature of the datasets that were used in the subsequent analyses like knowing the maximum and minimum rainfall by stations and months.

The datasets used for descriptive statistics are the daily data from 11 synoptic stations. Distribution of these synoptic stations into the delineated zones are; zone 1 (lake), 3 stations and 2 stations for each in zones 2, 4, 5 and 6. For zone 3 (western), we had no any synoptic station during this study.

Also analysis of similar kind was done on the dekadal datasets in which the gauge data were from all stations were arranged into their respective zones. The particular kind of empirical statistics which are done in this study include determining the means, maximum, minimum, standard deviation as well determining the number of missing values in each datasets.

3.2.2 Calibration of TAMSAT estimation for Tanzania

In order to have confidence in the TAMSAT rainfall estimate, the algorithm has to be calibrated for Tanzania as it has been done for other tropical regions where TAMSAT estimates have been used. For this purpose, the country was divided into six zones. The two main criteria for dividing a region for TAMSAT calibration is the homogeneity and a sizable area to contain sufficient number of gauges (Assiri, 2006).

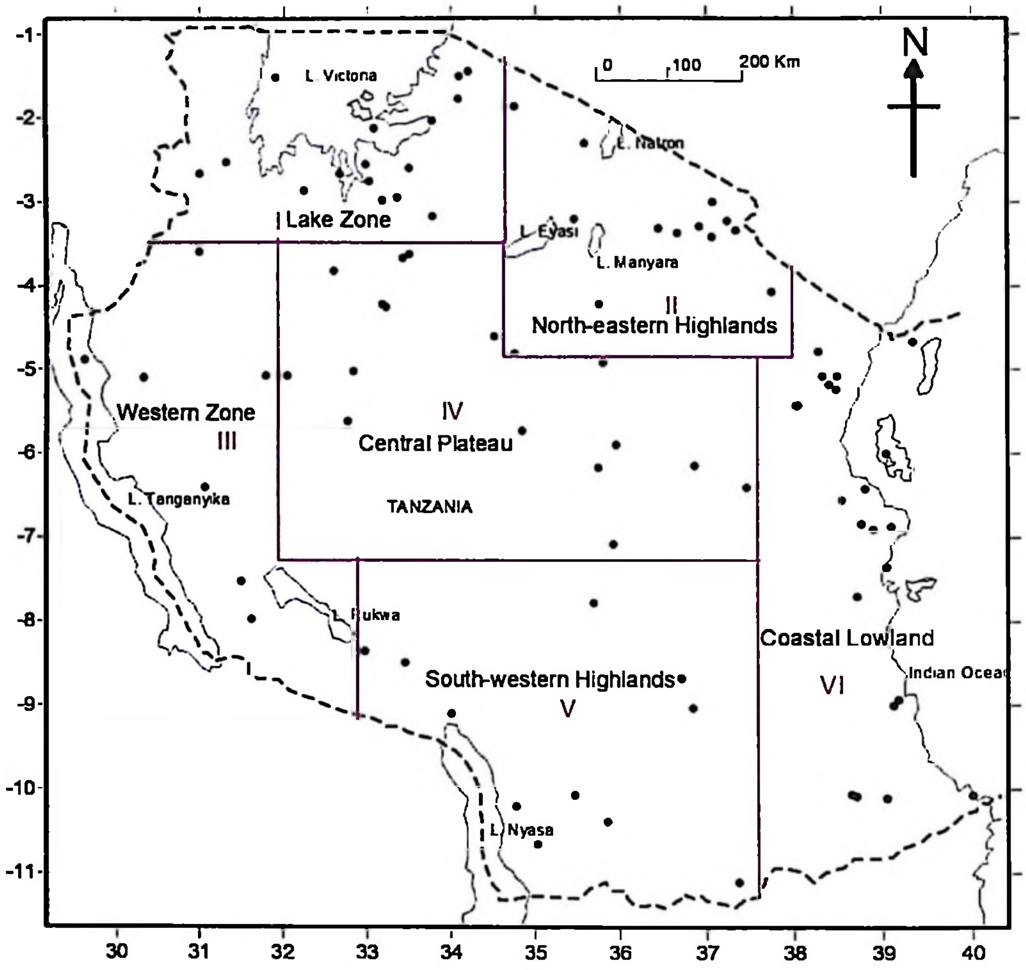


Figure 8: Calibration zones for TAMSAT rainfall estimation as used in this study. The small discs represent distribution of rain gauge in each zone as being used in this study.

The previous delineation of Tanzania rainfall homogeneity zones is 15 zones according to Basalirwa et. al. (1999), where some zones are very small compared to the other. Therefore, in order to have at least sizable and comparable regions, zones were delineated on the basis of rainfall climate, seasonality and topography (see figure 8). For instance, zones 1 and 2 are bi-modal but have different physical features. While zone 1 is mainly influenced by the presence of Lake Victoria, zone 2 is within the north eastern highlands and the northern arm of the rift valley.

Exceptions to the above criteria are on the coastal lowland where the northern coast is bi-modal and the southern coast is uni-modal. But it should be understood that this region is close to the Indian Ocean and is lowland with several rivers draining into the ocean from the South western highlands (southern coast). Other rivers flow from north eastern highlands (northern coast). As the whole coast seems to be influenced by relatively similar physical features, it was decided to make it as one zone. However, over the coast it has been also observed that topography has less control on rainfall variability (Sumner, 1983). The full delineation of calibration zones for TAMSAT is given in figure 8. The delineation has been done for simplicity in reference to figure 7 in order to enable a smooth calibration into relatively large areas than using the zones in figure 7.

3.2.3 Statistical comparison

In modelling studies or any study where a quantity is being estimated from observations, it is imperative to determine the accuracy of the method by using a variety of validation statistics. Such statistics include bias, correlation, mean absolute error (MAE), root mean square difference (RMSD) and skill score. This list is not exhaustive as there are other statistics. In this study, validation has been done by using correlation coefficient, bias (normal and multiplicative), root mean square difference and the Nash Index (NI). These statistics have also been used in previous studies. For example in various studies of comparing observed data and estimated (model) data, Diro et., al. (2009) used the bias, MAE and correlation while Kidd et., al. (2003) and Sapiano and Arkin (2009) used correlation, bias and some area statistics namely probability of detection, and false alarm rate among others.

In general the purpose of using these statistics is to provide a quantitative evaluation of the performance of the method of estimation or model skill (Diro, et., al., 2009). As defined by Symeonakis et al (2009), the following parameters were therefore used in this study to determine the strength of statistical relationship between the estimated or satellite rainfall (E) and measured gauge rainfall value (G) for the number of elements (N):

Bias

$$bias = \frac{\sum_{i=1}^{N}(E_i - G_i)}{N} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Multiplicative bias (biasM)

$$biasM = \frac{\frac{1}{N}\sum_{i=1}^{N} E_i}{\frac{1}{N}\sum_{i=1}^{N} G_i} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

While bias (also may be called additive bias) determines the average difference or error of the estimated rainfall, multiplicative bias is used to give the ratio of the average rainfall estimates to the average rain gauge measurements. For a perfect estimate, bias (additive) is zero and the other is one (Knaff, et al., 2008).

The root mean square difference (RMSD) or root mean square error (RMSE) is a frequently-used measure of the differences between values predicted by a model or an estimator and the values actually observed from the quantity being modelled or estimated. RMSD is a good measure of accuracy.

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N}(E_i - G_i)^2}{N}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

For a perfect estimate and RMSD is zero.

Linear correlation coefficient is given as,

$$cor = \sqrt{\frac{\sum_{i=1}^{N}(E_i - \bar{E})(G_i - \bar{G})}{N\delta_E\delta_G}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

In the above relation, $\bar{E}$ and $\bar{G}$ are the mean values of the estimated satellite rainfall and gauge rainfall at location, i, while δ_E and δ_G the standard deviations of the precipitation parameters respectively. The coefficient is a measure of co-fluctuation between the precipitation parameters. It is not sensitive to a bias.

All of the above statistics were computed after removing any missing data in the two datasets and making sure that the remaining data elements correspond exactly to each other. However, it is also possible to examine the results of different datasets by using the Nash Index (non-dimensional skill score). It is the relative measure of the deviation between estimate and observation (Symeonakis, 2009). This statistic is given as:

$$\mathrm{NI} = 1 - \frac{(\mathrm{RMSD})^2}{\delta_G^2} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (6)$$

Nash Index is computed with respect to the standard deviation of the gauge value and when equals to unit imply a perfect estimate. Therefore, the use of NI and biasM could provide a quick decision for the perfect estimate as both will tend to one.

Validation was done for the months of October to April. May estimates were not available from the TAMSAT archive. The general aim was to validate the rainy season for Tanzania which comprises the months of October to May. Though, the TAMSAT algorithm was calibrated in all zones (six zones), validation was done only in zones 2 and 5. This is because, already there is a pilot project of rainfall index insurance for small scale maize farmers in zone 2 (Tengeru, Arusha) while zone 5 comprises the three of the big four grain producer regions in Tanzania (Mbeya, Iringa and Ruvuma). Time constraints for this paper limited the capacity to do the validation in all zones and hence validation into other zones will be a future work apart from this dissertation.

Though, calibration was done at a fine resolution of 0.05°×0.05°, validation was done at a slight coarse resolution of 0.5°×0.5°. By this resolution, zone 2 has 46 pixels while zone 5 has 65 pixels. Validation was done by taking the means of the estimates and the available rain gauge measurements in each zone. There are 13 gauges in zone 2 and 11 in zone 5. Further study could improve this calibration by taking a more refined resolution up to 0.05°×0.05°.

3.2.4 Comparison of maize crop yield and rainfall data

In determining the relationship between maize yield and rainfall, we first determined the crop calendar in the two zones. The crop calendar is mainly determined by the rainfall regime within the zone. Thus, two main crop calendars are experienced in Tanzania as indicated in figure 9 according to Tanzania FEWSNET office (FEWSNET Food Security Update report, March 2006). In our study, we therefore used this calendar to determine the relationship between the rainfall amount and crop yield for maize.

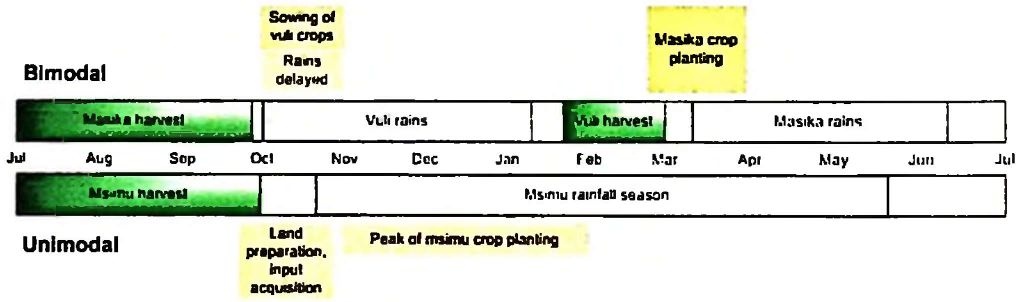


Figure 9: The mean cereal crop calendar in Tanzania as used by FEWSNET in food security.

Comparison between crop yield and rainfall was done by using the trend analysis. The aim here is to determine the relationship between rainfall and crop yield which is vital in deciding the application of satellite derived rainfall for agriculture weather index insurance. Trend analysis has also been used in Ethiopia to determine the association of crop water requirement and yield variation (Senay and Verdin, 2003). In addition, the correlation coefficient is determined though there is limited data for the proper statistical analysis. We had only seven full seasonal years. The rainy season in Tanzania begins in October and ends in May the following year, thus this period is termed as seasonal year in our analysis.

The presence of two types of rainfall regime has attributed two types of crop calendar. In particular over highlands maize growing period is longer as it takes about 6-8months while in lowlands takes about 3-4 months. In most areas especially in unimodal regions (zone 5), maize planting date is in November or early December, harvesting starting in April to June (Kaliba et al., 1998). This also depends on the variety of maize specie used. It is also possible to plant in February and harvest in June or July. In view of these factors, our analysis has been done in two cropping calendar apart from the major rain rainy season of October – May (April for satellite derived rainfall due to unavailable estimates in May). The two cropping calendar are November – April and February – July. Again for the February – July calendar there is no estimates May to July, however, analysis has been done as a formality to visualize the situation. Results under this part will not give a conclusive result regarding the success of the method.

The next chapter describes the results starting with results of descriptive statistics for the overview of the rainfall climatology in Tanzania. Also in the results are the calibrations for TAMSAT algorithm over the whole country followed by its validation results in zone 2 and 5. Lastly in the results is the trend analysis of rainfall and maize yield. Discussion is held on the important points throughout the results.

Chapter Four

Results and Discussion

4.1 Descriptive statistics of gauge rainfall data

We first attempt to explore in detail the rainfall climatology for Tanzania by using descriptive statistics and graphical analysis that include time series. Daily rainfall distribution indicates that high rainfall daily storms are more prevalent to zone I (Lake zone). Between the period of October 1997 and December 2006, the highest daily storm ever recorded is 185 mm at Musoma station (see table 3). Earlier studies have associated the occurrence of intense rainfall in Lake Zone (zone 1) with the availability of moisture supply from Lake Victoria (Basalirwa, et al, 1999). In addition, for the same period, the mean daily precipitation is high for Bukoba at 5.3 mm/day (Lake zone) and the lowest is 1.6 mm/day in Dodoma and Same (Central and North eastern highlands respectively).

Table 3: Descriptive statistics for the synoptic stations

Station	Zone	Max	Sdev	Mean
Bukoba	1	140.4	11.5	5.3
Musoma	1	185	8.6	2.6
Mwanza	1	148	9.1	3.1
Arusha	2	109.5	7.2	2.1
Same	2	88.3	6.0	1.6
Dodoma	4	107	7.6	1.6
Tabora	4	119.5	8.4	2.7
Mbeya	5	87.6	7.4	2.6
Songea	5	83.9	8.1	2.8
NIA Dar	6	135.2	9.6	3.0
Mtwara	6	133.9	9.5	3.0

Considering the dekadal datasets, descriptive statistics reveal that the maximum rainfall for the dekads is 525 mm in the second dekad of November in zone 5. In other zones the maximum rainfall is 325.9 mm (zone 1), 469 mm (zone 2), 204.4 mm (zone 3), 392.6 mm (zone 4) and 515.8 mm (zone 6). The dekads for these maxima are dekad 2 April, dekad 2 November, dekad 3 March, dekad 3 December, and dekad 2 April respectively.

The dekadal maximum rainfall does not necessarily correspond with monthly peak rainfall. In Basalirwa et al. (1999) was revealed that the peak of the mean monthly total for zone M (figure 7) which also forms zone 5 (figure 8) is in April with an amount exceeding 600 mm. It means the dekad can be influenced by few rainstorms within that period of 10 ten days and hence give the maximum rainfall. This structure of rainfall distribution in Tanzania was also revealed in the previous studies. Nieuwolt (1974) confirmed that the rainfall distribution in the country is characterised by individual rainstorms. The storms are also observed to be erratic however; the prevalence area of occurrence is the region around Lake Victoria.

4.2 Results for TAMSAT calibration

From the optimization of calibration parameters, the results suggest that -30°C is the optimal threshold temperature for TAMSAT estimation method in Tanzania for zones 2 and 5 with the exception of some months. The exception is in February for zone 2 where estimation must be done at -40°C while in zone 5 estimations in January and February are done at -50°C and at -40°C in November. The full optimal temperature thresholds are summarized in table 4 and figures 10 and 11 illustrate some of the regression parameters for January. Refer to section 3.1.2.1 for the details about the meaning of threshold temperatures.

The observed differences in optimal temperatures could probably be associated with the different nature of rain seasons and the corresponding synoptic forcings in both zones (Zorita & Tilya, 2003, Kijazi & Reason, 2009 and Basalirwa et al., 1999). Zone 2 follows a 2 rainy seasons (short and long rains) for months of October to December and March to May respectively.

Table 4: Optimal temperature thresholds for zone 2 and 5

Month	January	February	March	April	October	November	December
Zone 2	-30°C	-40°C	-30°C	-30°C	-30°C	-30°C	-30°C
Zone 5	-50°C	-50°C	-30°C	-30°C	-30°C	-40°C	-30°C

On the other hand, zone 5 rainy season is uni-modal starting from late October to May (Zorita & Tilya, 2003). We can also expect some shift in the trigger mechanism to be at a much colder temperature especially on early months of the season. Rainfall peaks in April and is said to be mainly controlled by the movement of the ITCZ. This has been discussed in Basalirwa et al. (1999) as part of their study to divide the country into homogeneous rainfall Climatological zones.

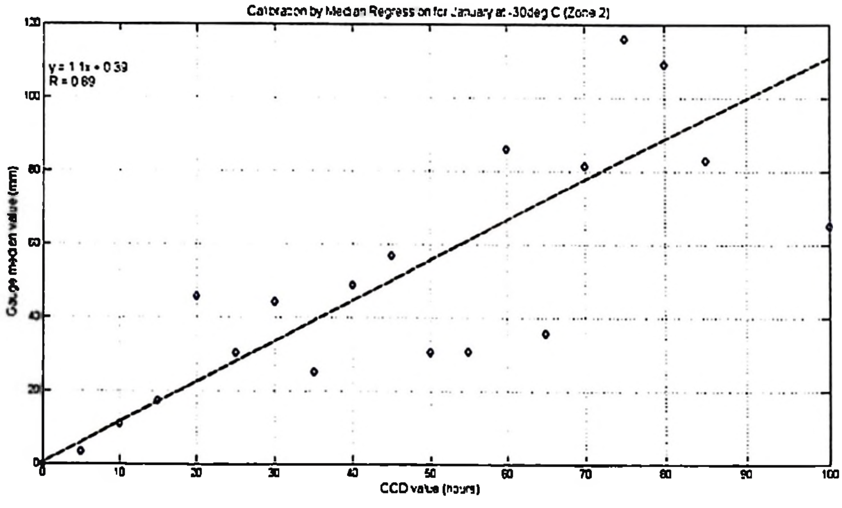


Figure 10: Plots for calibration of TAMSAT method in zone 2 during January

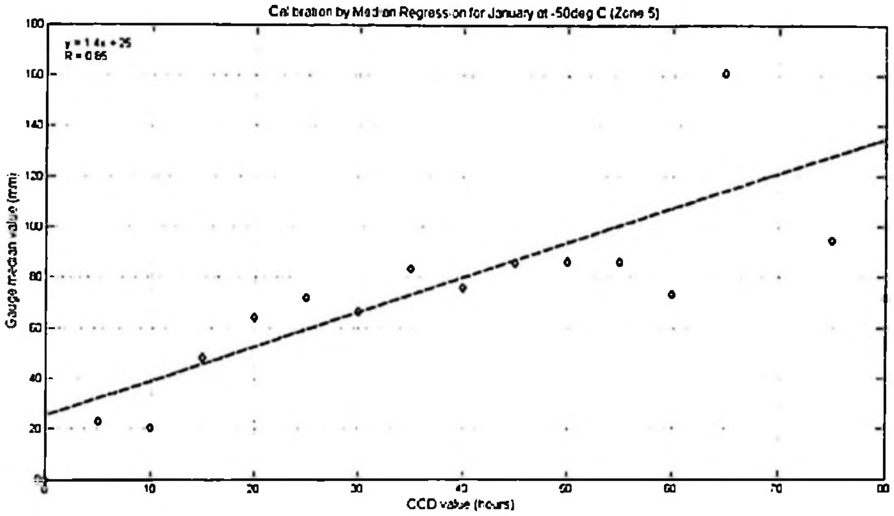


Figure 11: Plots for calibration of TAMSAT method in zone 5 during January

4.3 Statistical comparison of gauge data and satellite estimates

Satellite rainfall estimates and gauge data were compared for the months of October to April in zone 2 and 5. The general picture from the validation results indicate that satellite rainfall estimates tend to follow the general trend of the gauge rainfall in almost all the months in both zones. However, there are some quite large differences from month to month in both zones as is discussed later.

In zone 2, the best month is January as it shows a good correspondence between the estimates and measured value for the years 2001 to 2005 (figure 12). In terms of validation statistics, it scored better in four statistics (bold) out of five (table 5). Other months which indicate a good relationship are November and December while April displays a worst estimate. In fact the months with bad estimate could easily be noted by a negative number with a Nash Index.

Table 5: Comparison statistic parameters for zone 2

Month	Linear Corr.	Bias (mm)	RMSD (mm)	Nash Index	BiasM
January	**0.7413**	**0.5306**	21.9664	**0.5479**	**0.9833**
February	0.4316	8.4178	18.1615	-0.1284	0.5459
March	**0.6471**	5.4305	22.3101	0.3578	**0.8622**
April	0.0998	19.8236	33.0300	-0.0285	0.6793
October	-0.0823	**0.4123**	29.2949	0.0851	**0.9737**
November	**0.7526**	5.3821	16.0731	**0.7274**	**0.8037**
December	**0.7178**	2.4716	22.1702	**0.6387**	**0.9266**

Bolded statistical values indicate some significant relationships.

As observed in figure 12, there is a slight underestimate of high rain especially the second dekad in 1998 for January. This is also observed in February and March. This is contrary to December where there is a slight overestimation of high rain (see appendix B1).

The varied performance of TAMSAT estimation in different months could possibly be attributed to the nature of rainfall regime in the zone. Zone 2 has two rain seasons of OND and MAM as already explained in chapter two. If there is a tendency to underestimate the high rains, then we can expect a bad performance in April since it is normally a peak month in the MAM (Basalirwa, et.al, 1999) and MAM is more intense than the OND (Zorita and Tilya, 2002). However, April is also much affected by missing data for CCD and lack of CCD in May also decrease the confidence in this attribute. The high rain observed in January and February for 1998 is attributed to the 1997/98 El Nino event.

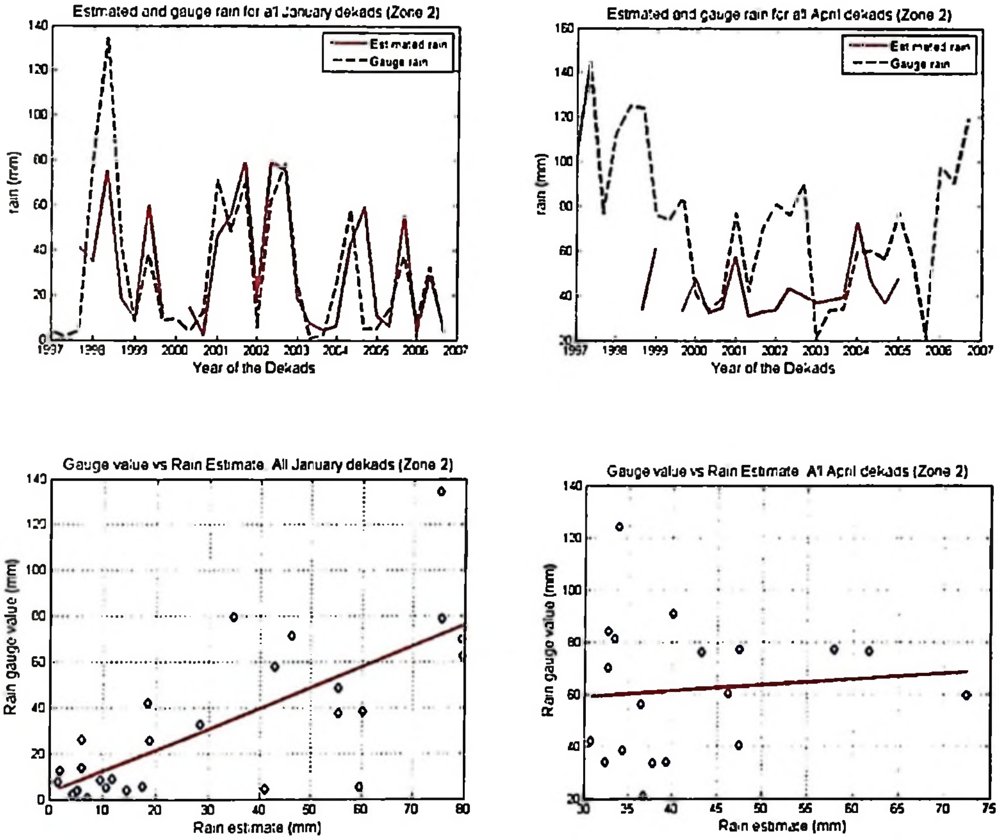


Figure 12: Comparison of measured precipitation to estimate in zone 2 for January and April. Top figures indicate trends for the dekads of the month from 1997 to 2007 while the bottoms are the respective scatter plots with the best fit line.

In zone 5 results indicate general underestimation of rainfall in most months except slightly in December and November. December has scored best in three of the validation statistics. Again, the underestimation is easily identified by the negative Nash Index (table 6). Though, the performance is not good as in zone 2, still the general trend is good as observed that there is a good correspondence of satellite rainfall estimate and gauge rain. This is indicated by the high numbers of the coefficient of linear correlation as also shown by the best fit line (figure 13 and more on figure B2 appendix).

Table 6: Comparison statistic parameters for zone 5

Month	Linear Corr.	Bias (mm)	RMSD (mm)	Nash Index	BiasM
January	0.5148	39.8611	48.9471	-1.1767	0.4830
February	**0.6104**	28.4077	37.1037	-0.5222	0.5522
March	0.5262	47.3877	56.0091	-1.6437	0.4261
April	**0.8976**	49.2593	59.3260	-0.8721	0.2055
October	0.3731	4.1844	7.7256	-0.1822	0.3434
November	**0.7787**	9.9383	19.9289	0.4562	0.5881
December	**0.9043**	8.6738	22.5545	**0.7958**	**0.8710**

Bolded statistical values indicate some significant relationships.

For April, though there is underestimation we notice a very good correlation and it can be observed in figure 13 that the trend of increase or decrease between the satellite rainfall and gauge rainfall correspond to each other. This is contrary to what was observed in zone 2.

We can also attribute this performance to the smaller number of rain gauge used in zone 5 compared to zone 2 apart from the different relief features and rainfall regimes in the two zones. Zone 5 is unimodal with rains from late October to May, however, we had only 11 gauges and the area is larger than that of zone 2 where we had 13 gauges. A number of studies of validating satellite rainfall products to gauge data have indicated the influence of large gauge number on good estimates (Sapiano and Arkin, 2009, Symeonakis et al, 2009, Kidd et al, 2003 and Ali et al, 2005).

In general the correlation results in both zones 2 and 5 are very good except for April and October in zone 2 (i.e. the north-eastern highlands). These suggest a possibility of using the derived satellite in various agricultural applications and if well calibrated possible use in agriculture insurance can be considered. It should be noted that the TAMSAT method has never been properly calibrated for Tanzania and for some months, the calibration parameters for southern Kenya have been used.

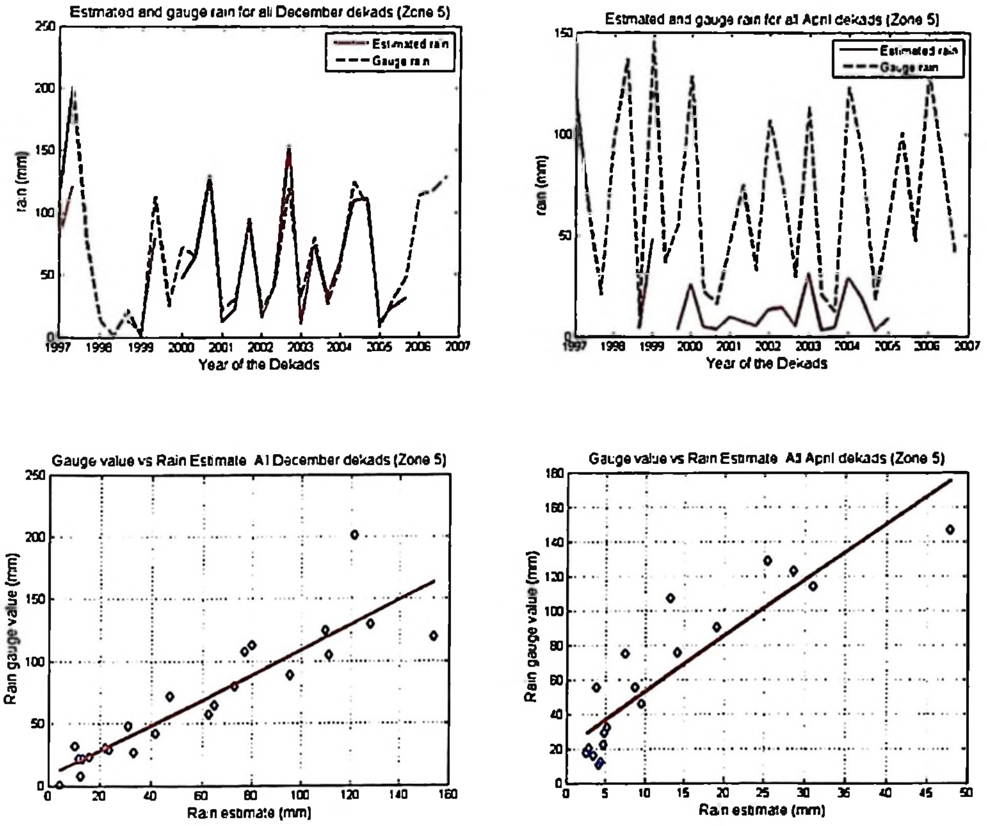


Figure 13: Comparison of measured precipitation to estimate in zone 5 for December and April. Top figures indicate the trends for the dekads of the month from 1997 to 2007 while the bottoms are the respective scatter plots with the best fit line. In April, satellite rainfall is underestimated but the trend is well correlated as indicated on the bottom figure.

4.4 Relationship between rainfall and crop yield

The relationship between the two types of rainfall information we used in this study and the crop yield was done using the trend analysis. Trend analysis was done since the two kinds of datasets to be compared (i.e. rainfall and maize yield) have significant quantitative difference. In addition, the two types of datasets are measured differently then, trend analysis is sufficient to realize the effect of rainfall in yield.

Results discussed under this section can only be regarded as preliminary results towards the possibility of using satellite derived rainfall in agriculture insurance. Our results have been obtained with limited data coverage in spatial and temporal for both rainfall and yield. We start by describing the relationship for the whole rainy period from October to April. This general consideration was done since the yield data does not clearly indicate as to what cropping season was it obtained although there is a little possibility of farmers planting twice in a year. Table 6 therefore summarizes the correlation coefficients for zone 2 and 5 where there is encouraging values in zone 2 especially in the November – April cropping calendar. Further discussion of these results follows in the next two subsections.

Table 7: Correlation coefficients for relationships between yield and rainfall in zone 2 and 5

Rainfall correlated with maize yield	Zone	Whole rainy season (October - April)	November - April cropping calendar	February - July cropping calendar
Gauge rainfall	2	0.17	0.58	0.54
	5	0.01	0.12	0.30
Satellite-rainfall	2	0.49	0.43	0.05
	5	-0.09	-0.09	-0.04

Note: In February to July cropping calendar, the satellite data used are for February to March only.

4.4.1 Results for trend analysis for zone 2

The trends for maize yield and that of rainfall trends in the November – April and February – July for zone 2 is shown in figure 14. The results suggest a good correlation between the yield and the two types of rainfall in both calendars (see table 7). However, there is some mismatch in some seasonal years where rainfall trend is increasing while yield is decreasing and vice versa. Particular case can be seen in the 2002 seasonal year in yield increased to highest amount for the period 1999 – 2005 while the two rainfall types are decreasing.

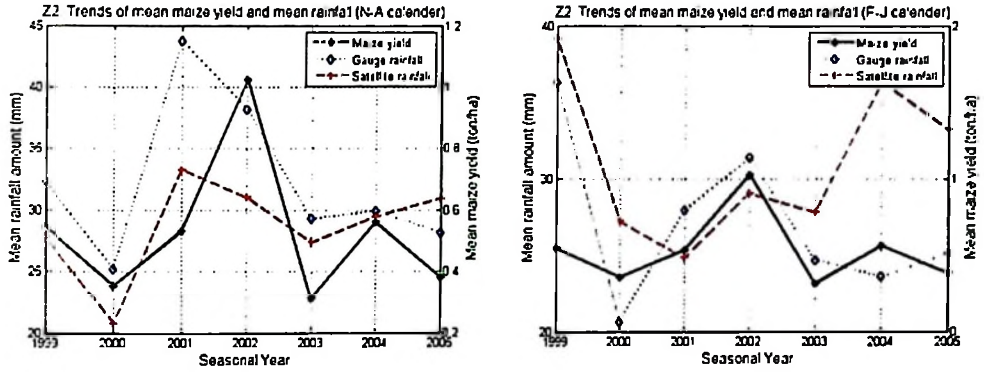


Figure 14: Maize yield, satellite-derived rainfall and gauge rainfall trends for zone 2 during the November – April and February – July cropping calendars.

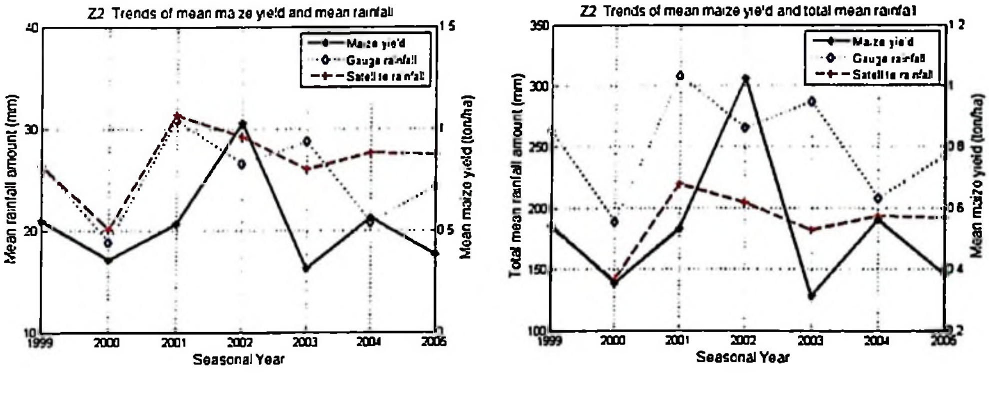


Figure 15: Maize yield, satellite-derived rainfall and gauge rainfall trends for zone 2 during the whole rainy season from October - April.

Similarly, on considering the whole rainy season, it was also observed that on the year 2002, the rainfall trends decreases while the yield increases (figure 15). Further, we noted that the gauge rainfall seems to be uncorrelated with yield beyond the year 2001 and hence making its correlation coefficient as low as 0.17 compared to the other cropping calendar. In the two cropping calendars the coefficient is above 0.5 for the gauge rainfall.

In zone 2, it is important to determine the proper cropping calendar and perhaps the satellite estimates could only be used during the November – April calendar. However, less correlation coefficient could be attributed by other reasons of which this study could not be at a position to establish them. It should be noted also that we had no rainfall estimates for the months of May to July and lack of this could attribute the low correlations observed in this cropping calendar.

4.4.2 Results for trend analysis for zone 5

The trends for maize yield and that of rainfall trends in zone 5 are shown in figure 16 and 17. As indicated on table 6, we have very low correlation coefficients in this zone except for the February- July cropping calendar between maize yield and gauge rainfall where the coefficient is 0.30. The results further demonstrate that satellite rainfall is weakly correlated with the yield. This is perhaps due to missing months of CCD data.

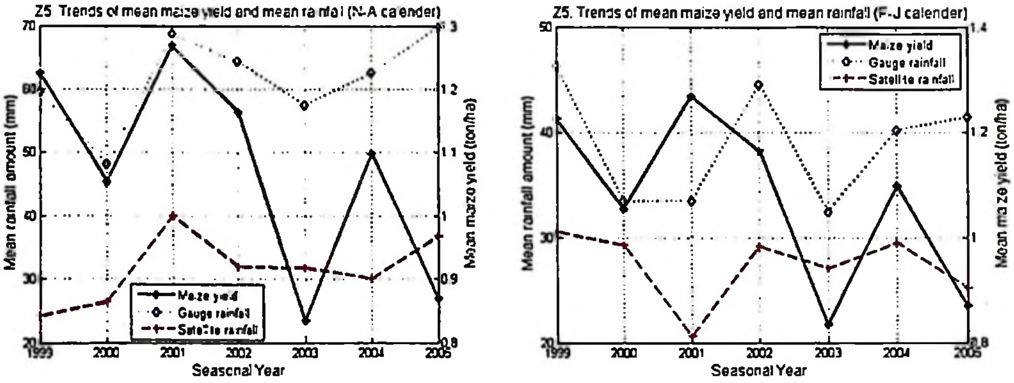


Figure 16: Maize yield, satellite-derived rainfall and gauge rainfall trends for zone 5 during the November – April and February - July.

It might be possible that the lower number of rain gauges used in a relatively large area of zone 5 is one of the reasons behind the weak correlations observed. In zone 5 we have only 11 gauge stations while in zone 2 we have 13 gauge stations, after data quality control of rain gauge data. Together with the fact that TAMSAT method has never been well calibrated to Tanzania instead calibration parameters for Kenya are used, situation can be worse to the southern parts of Tanzania.

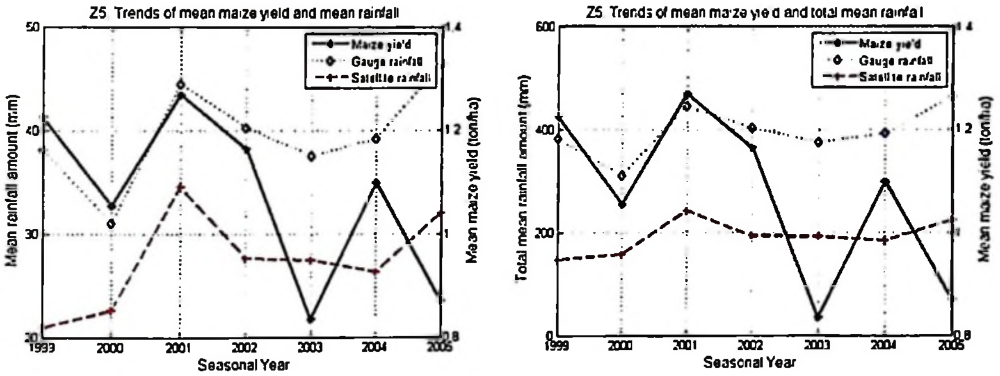


Figure 17: Maize yield, satellite-derived rainfall and gauge rainfall trends for zone 5 during the October – April rainy season.

Similarly the results for trend analysis by considering the whole rainy season of October to April on the available satellite estimates also indicate low correlation coefficients with maize yield. Further, satellite rainfall estimates seem to be negatively correlated to maize yield. On specific years, it seems gauge rainfall slightly correlates well with maize yield for the years 1999 to 2004. Whilst with satellite estimates from 2002 to 2005 it seems to be negatively correlated and hence attributing to the overall negative correlation. Example of these results is indicated in figure 17. More results for trend analysis are attached as appendix C. The analysis includes the trends for means and the total of the means in each zone and particular season.

In further discussion of our results, we can also generally attribute the influence of extreme weather events in some of the seasonal years contributing to the mismatch of the trends. Specifically we can point out that it is possible the decline of yields in the year 2003 and 2005 be attributed to the influence of El Niño in which much increased rainfall may affect the yield especially when floods occur in the fields.

Chapter Five

Conclusion and suggested future work

5.1 Conclusion

The study to investigate the use of satellite derived rainfall on agriculture index insurance was done in Tanzania particularly over the north-eastern and south western highlands areas. The north eastern highlands comprises the regions of Arusha, Manyara and Kilimanjaro and currently, MicroEnsure Company is piloting a project of agriculture insurance in Tengeru, Arusha, basing on rainfall data of one gauge station at Tengeru. These benefit only farmers with farms around the radius of 20 km from the station. On the other hand, the south western highlands of Tanzania comprise three of the regions which are main producer of maize and other cereals. The idea of investigating the use of satellite information in agriculture insurance is thought to be a novel idea where many farmers would benefit and hopefully stimulating agriculture activities in the country. It should be noted that agriculture is the backbone of Tanzania economy as it employ majority of the rural population. Still it suffers a lot from weather and climatic risks.

Analysis was done using gauge and satellite rainfall datasets for the period of 1997-2007. Gauge rainfall data was obtained from Tanzania Meteorological Agency while satellite data were obtained from TAMSAT archives. Similarly crop yield data for maize was used covering the same period. Though there is a problem in administration of crop yield data as there were not sufficient information, the available information were able to give some interesting results, calling further research in the area. Essentially, the main methodology includes comparison and trend analysis which include various statistics.

In the descriptive statistics analysis, the study confirmed that rainfall distribution in Tanzania is influenced by relief features, the passage of ITCZ and some other synoptic and mesoscale features. Further, individual rainstorms which seem to be erratic as confirmed by previous studies influence largely the statistics of the rainfall amount.

From the calibration analysis of TAMSAT method of rainfall estimation in Tanzania, the optimal threshold temperature is -30°C in most of the months where calibration was done. Exception is on February for north eastern highlands and on January, February and November for south western highlands where estimation should be done at a much colder temperature. Calibration into other zones is the next research and possible finding of calibration correction parameters is needed which would improve the accuracy of estimated rainfall. Currently, calibration correction factors for Kenya are used instead.

On the comparison analysis between satellite estimates and gauge rainfall results are very interesting and promising that satellite rainfall estimates could potentially be used in various agriculture applications including index insurance. However, further study is still needed especially in other areas where calibration was not carried out. Specifically, correlation results in both zones 2 and 5 are very good except for April and October in zone 2.

Part of the analysis was to check the relationship between maize yield and the rainfall products. This is vital as it was done in order to give an indication of whether the datasets are good enough to be used in agriculture insurance. Trends of maize yield and that of rainfall products from the seasonal year of 1999 to 2005 were done. Results indicate good correlations in zone 2 especially in the November to April cropping calendar. It is not particularly surprising to see negative correlations in zone 5 for the satellite rainfall as this can be attributed to the fact that proper calibration has not been done to Tanzania.

So far our results seem to be good generally. However, it is important to recognise some of the caveats that if solved out, results would be even better. First, rainfall estimates were calculated using the linear relationship given in equation 1. The estimates did not employ any correction factors that would have improved our results. On the rain gauge data sets it was found that some of the stations have missing data. It is important to note that though the comparisons were done against gauge data, these cannot be treated as the accurate reference. Gauge measurements are also constrained by a number of errors.

The results can still be treated as preliminary till also compared to the insurance contract and see if they will reproduce the relatively similar trends of previous weather risks. Due to constraint of time, this analysis was not done.

In general the results are very promising and it is possible in future to use satellite estimates in agriculture insurance. This calls for further research in the area as calibration was done only in zone 2 and 5. The current results can therefore be regarded as preliminary results. It is therefore concluded that TAMSAT rainfall estimates in Tanzania are estimated at -30°C and are very potential for agriculture weather index insurance use.

5.2 Suggested future work

It has been suggested that the current results are preliminary towards the use of satellite rainfall estimates in agriculture insurance in Tanzania. Some problems have been noted and possibly have affected some of our results. Improvements in the results are required in order to achieve the aim of using satellite information in agriculture insurance. First, there is need to calibrate the TAMSAT method into other areas in Tanzania and obtain the proper correction parameters. In addition, use of TRMM data can be considered. The main analysis to be done for both TRMM rainfall and TAMSAT estimates is to compare them against the krigged gauge data.

Our trend analysis for the yield and rainfall products was done for the period of seven years only from 1999 to 2005 seasonal year. It would be a good idea of using at least ten years of data. This is because most literature regarding agriculture index insurance suggest a long time period of historical data in order to see the relationship between yield and rainfall. Further study should also consider how to separate the loss or increase in yield attributed to other factors like insect, change in farm mechanisation means and other factors.

References

Assiri, M.E. (2006): Merging Satellite and Rain gauge Estimates of Rainfall using Geostatistical Techniques. M.Sc. Thesis. Univ. Of Reading, UK

Barnett, B. J. and O. Mahul (2007): Weather index insurance for agriculture and rural areas in lower-income countries. Amer. J. Agr. Econ. **89** (5), 1241–1247.

Barrett, E.C. (1970): The Estimation of Monthly Rainfall from Satellite Data. Monthly Weather Review **98**, (4), 322-327.

Basalirwa, C.P.K., J.O. Odiyo, R.J. Mngodo and E.J. Mpeta (1999): The Climatological Regions of Tanzania based on the Rainfall Characteristics. Int. J. Climatol. **19**: 69–80

FEWS NET (2006): Tanzania: Food security update, March 2006. < www.fews.net/docs/Publications/Tanzania_200603en.pdf > [Accessed: 02/08/2009]

Bellerby, T., M. Todd, D. Kniveton and C. Kidd (2000): Rainfall Estimation from a Combination of TRMM Precipitation Radar and GOES Multispectral Satellite Imagery through the Use of an Artificial Neural Network. J. of Appl. Meteorology, **39**, 2115-2128.

Ben Kautisire (2007): Climate risk management in Africa: Learning from practice. Available: <http://www.fanrpan.org/documents/d00255/3-Climate-risk/Africa/2007c> [Accessed: 30/04/2009]

Burton, I., and T. Dickinson, 2008: The Potential for Insurance to Help the World's Poor Adapt and Thrive as the Climate Changes. Pages 1-15 in. Development in the Balance: How Will the World's Poor Cope with Climate Change? Brookings Institution Press. Available: <http//:www.brookings.edu/events/2008/~/media/Files/Programs/Global/brookings/blum/roundtable/2008/burton/Dickinson> [Accessed: 30/04/2009]

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

Chokngamwong, R and L. S. Chiu, (2008): Thailand Daily Rainfall and Comparison with TRMM Products. Journal of hydrometeorology, **9**, (2), 256 – 266

Dinku, T., P. Ceccato, E. Grover-Kopec, M. Lemma, S. J. Connor and C. F. Ropelewski (2007): Validation of Satellite Rainfall Products over East Africa's Complex Topography. Int. J. of Remote Sensing, **28**, (7), 1503–1526.

Diro, G. T., D. I. F. Grimes, E. Black, A. O'Neill and E. Pardo-Igúzquiza (2009): Evaluation of reanalysis rainfall estimates over Ethiopia. Int. J. Climatol. **29**, 67–78.

Ebert, E.E., J.E. Janowiak, and C. Kidd, (2007): Comparison of Near-real-time Precipitation Estimates from Satellite Observations and Numerical Models. Bul. of Amer. Met. Soc. 47-64.

Feidas, H., G. Kokolatos, A. Negri, M. Manyin, and N. Chrysoulakis (2006): A TRMM-Calibrated infrared technique for rainfall estimation: Application on rain events over eastern Mediterranean. Advances in Geosciences, **7**, 181–188.

Grimes, D.I.F., E. Pardo-Igúzquiza and R. Bonifacio (1999): Optimal areal rainfall estimation using raingauges and satellite data. Journal of Hydrology **222**, 93–108.

Hazell, P and J. Skees (2005): Insuring against Bad Weather: Recent Thinking. Available: <http://www.globalagrisk.com/Pubs/2005%20Insuring%20Against%20Bad%20Weather%20hazell%20and%20jrs.pdf> [Accessed: 14/05/2009]

Hellmuth M.E., Osgood D.E., Hess U., Moorhead A. and Bhojwani H. (eds) 2009. Index insurance and climate risk: Prospects for development and disaster management. Climate and Society No. 2., International Research Institute for Climate and Society (IRI), Columbia University, New York, USA

Kabanda, T. A and M. R. Jury (1999): Inter-annual variability of short rains over northern Tanzania. Clim. Res. **13**, 231–241.

Kaliba, A.R.M., H. Verkuijl, W. Mwangi, D.A. Byamungu, P. Anandajayasekeram, and A.J. Moshi. (1998): Adoption of Maize Production Technologies in Western Tanzania. Mexico, D.E: International Maize and Wheat Improvement Centre (CIMMY), the United Republic of Tanzania, and the Southern Africa Centre for Cooperation in Agricultural Research (SACCAR). Available: . [Accessed: 04/08/2009]

Kijazi, A. L. and C. J. C. Reason (2005): Relationships between Intraseasonal Rainfall Variability of Coastal Tanzania and ENSO. Theor. Appl. Climatol. **82**, 153–176.

Kijazi, A. L. and C. J. C. Reason (2009): Analysis of the 2006 Floods over Northern Tanzania. Int. J. Climatol. **29**, 955–970.

Knaff, J.A., M. DeMaria and C. Lauer (2008): Verification of the Monte Carlo Tropical Cyclone Wind Speed Probabilities: Project Update. Available: < http://www.nhc.noaa.gov/jht/ihc_07/s8-04Knaff.pp> [Accessed: 10/07/2009]

Kummerow, C., J. Simpson, O. Thiele, W. Barnes, A. T. C. Chang, E. Stocker, R. F. Adler, A. Hou, R. Kakar, F. Wentz, P. Ashcroft, T. Kozu, Y. Hong, K. Okamoto, T. Iguchi, H. Kuroiwa, E. Im, Z. Haddad, G. Huffman, B. Ferrier, W. S. Olson, E. Zipser, E. A. Smith, T. T. Wilheit, G. North, T. Krishnamurti, and K. Nakamura. (2000): The Status of the Tropical Rainfall Measuring Mission (TRMM) after Two Years in Orbit. J. of Appl. Meteorology, **39**, 1965 - 1983

Lebel, T, H. Sauvageot, M. Hoepffner, M. Desbois, B. Guillot and P. Hubert, (1992): Rainfall estimation in the Sahel: The EPSAT-NIGER Experiment. J. of Hydrol. Sciences, **37**, (3), 201-215.

Levizzani, V., R. Amorati and F. Meneguzzo (2002): A Review of Satellite-based Rainfall Estimation Methods. MUSIC – MUltiple-Sensor Precipitation Measurements, Integration, Calibration and Flood Forecasting. Available: www.isao.bo.cnr.it/~meteosat/MUSIC/Rep/Sat/Precip/6.1.pdf > [Accessed: 14/05/2009].

Lomas, P. (Ed), (2009): Micro Insurance Matters. The Quart. Newsletter of Microensure. Available: <http://www.microensure.com/Microinsurancenewsletter/latest.pdf.> [Accessed: 19/05/2009].

Mapande, A. T. and C. J. C. Reason (2005): Interannual Rainfall Variability over Western Tanzania and ENSO. Int. J. Climatol. **25**, 1355–1368.

Morton, J., and S. Anderson: Climate change and agrarian societies in drylands. Available: <http://siteresources.worldbank.org/EXTSOCIALDEVELOPMENT/Resources/244362-1170428243464/3408356-1170428261889/3408359-1202746084138/drylands/022708.pdf> [Accessed on 12 May 2009]

Nanteza, J. (2008): Comparison of Satellite and Rain-gauge Estimation over Uganda. M.Sc. Thesis, Univ. of Reading, UK

National Bureau of Statistics (NBS), (2008): Morogoro Region Socioeconomic Profile.

National Bureau of Statistics (NBS), (2006): National Sample Census of Agriculture 2003-04.

Nicholson, S.E. A Review of Climate Dynamics and Climate Variability in Eastern Africa. Available: <http://www.met.fsu.edu/people/nicholson/papers/clidyn.pdf>. [Accessed on 04/06/2009].

Nieuwolt, S., (1974): Rainstorm distribution in Tanzania. Geogr. Ann. 56A (3-4):241-250

Nyenzi, B.S. (1992): An analysis of the Interannual variability of rainfall over East Africa, J. Afr. Met. Soc. (SMA), 1, 57–79.

Ogallo, L.J. (1989): The spatial and temporal patterns of East African seasonal rainfall derived from principal component analysis. Int. J. Climatol., **9**, 145–167.

Sadalmelik (2007): Topographic map of Tanzania. Available: <http://commons.wikimedia.org/wiki/File:Tanzania_Topography.png> [Accessed on 12 May 2009].

Sapiano, M. R. P. and P. A. Arkin, (2009): An Intercomparison and Validation of High-Resolution Satellite Precipitation Estimates with 3-Hourly Gauge Data. Journal of Hydrometeorology, **10**, 149-166.

Sawunyama, T., and DA Hughes (2007): Application of satellite-derived rainfall estimates to extend water resource simulation modelling in South Africa. Water SA Vol. 34 No. 1 January 2008. Available: <http://www.wrc.org.za> [Accessed: 14/05/2009].

Senay, B. G., and J. Verdin (2003): Characterization of Yield Reduction in Ethiopia using a GIS-based Crop Water Balance Model. Can. J. Remote Sensing, **29**, (6), 687 – 692.

Skees, J.R., and B. Collier (2008): The Potential of Weather Index Insurance for Spurring a Green Revolution in Africa. Available: <http://www.globalagrisk.com/2008/ Skees%20and%20Collier/AGRA%20Paper%20on%20Weather%20Insurance.pdf [Accessed: 18/05/2009].

Song, Y., F. H. M. Semazzi, L. Xie and L. J. Ogallo (2004): A Coupled Regional Climate Model for The Lake Victoria Basin of Eastern Africa. Int. J. Climatol. **24**, 57– 75.

Sumner, G. N., (1983): Daily rainfall variability in coastal Tanzania. Geogr. Ann. 65A (1-2): 53-66.

Sun, L., F. H. M. Semazzi, F. Giorgi, and L. Ogallo (1999): Application of the NCAR Regional Climate Model to eastern Africa 2. Simulation of Interannual Variability of Short Rains, J. Geophys. Res., 104(D6), 6549–6562.

Symeonakis, E., R. Bonifacio and N. Drake (2009): A Comparison of Rainfall Estimation Techniques for Sub-Saharan Africa. Int. J. of App. Earth Obs. & Geoinfo. **11**, 15–26.

Tilya, F. F., (2007): Characteristics of Wet and Dry Spells in Tanzania during the Rainy Seasons, M.Sc. Thesis. University of Nairobi, Kenya.

Todd, M.C., C. Kidd, D. Kniveton, T. J. Bellerby (2001): A Combined Satellite Infrared and Passive Microwave Technique for Estimation of Small-Scale Rainfall. J. Of Atmos. and Oceanic Techn., **18**, 742-755.

Tressler, E.B. and M. Mbaka (2009): Index Based Weather Insurance in Kenya for Crops. Available: <http://www.WeatherInsurance/Kenya/Tressler/16Mar09/Workshop.ppt > [Accessed: 14/05/2009].

Ummenhofer, C.C., Gupta, A.S., M.H. England and C.J.C. Reason (2009): Contributions of Indian Ocean Sea Surface Temperatures to Enhanced East African Rainfall. J. Climate 22, 993 – 1013.

Wardah, T., S.H. Abu Bakar, A. Bardossy, M. Maznorizan (2008): Use of geostationary meteorological satellite images in convective rain estimation for flash-flood forecasting. Journal of Hydrology, **356**, 283– 298

Wenner, M (2005): Agricultural Insurance Revisited: New Developments and Perspectives in Latin America and the Caribbean. Available: < http://www.iadb.org/sds/RUR-Agricultural%20Insurance/ FINALVERSIONENGLISHJULY06> [Accessed: 14/05/2009]

Yan, J., and M. Gebremichael (2009): Estimating actual rainfall from satellite rainfall products. Atmospheric Research **92** 481–488.

Xie, P., A. Yatagai, M. Chen, T. Hayasaka, Y. Fukushima, C. Liu, and S. Yang (2007): A Gauge-Based Analysis of Daily Precipitation over East Asia. J. of Hydromet., **8**, 607-627.

Zorita, E., and Tilya, F.F (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.

Appendices

A: Rain gauge station

Table A1: List of gauge stations showing location, period and percentage of missing data

Station Name	Lon	Lat	Data Period	Data type	Missing data (%)	1997-2007 missing data (%)
Arusha	36.67	-3.37	1970-2007*	Daily/dekadal	1	1
Babati Agriculture	35.75	-4.22	1997-2007	Dekadal	31	31
Bagamoyo Agricult	38.92	-6.42	1997-2007	Dekadal	3	3
Biharamulo	31.32	-2.63	2002-2008**	Dekadal	75	91
Buhemba	34.08	-1.77	1997-2006	Dekadal	21	26
Bukoba	31.81	-1.32	1997-2007	Daily/dekadal	1	1
Bunda Agriculture	33.77	-2.03	1997-2007	Dekadal	32	41
Dodoma	35.74	-6.17	1997-2007	Daily/dekadal	0	0
Engare	37.48	-2.95	1997-2007	Dekadal	20	29
Gairo	36.87	-6.15	1972-1989***	Dekadal	15	15
Geita	32.25	-2.87	1997-2007	Dekadal	6	6
Handeni-Maji	38.05	-5.43	1997-2007	Dekadal	19	20
Handeni Met.	38.03	-5.43	2003-2007	Dekadal	5	57
Hombolo	35.95	-5.9	1997-2007	Dekadal	1	1
Ibadakuli	33.5	-3.62	2002-2007	Dekadal	29	61
Ilonga	36.85	-9.03	1997-2007	Dekadal	2	2
Iringa	35.69	-7.78	1997-2007	Dekadal	0	0
Kahama	32.6	-3.83	1997-2007	Dekadal	4	4
Kaliua	31.8	-5.07	1997-2007	Dekadal	61	61
KIA	37.07	-3.42	1997-2007	Dekadal	0	0
Kibaha	38.97	-6.83	1997-2007	Dekadal	0	0
Kibiti Hospital	38.92	-7.7	1999-2007	Dekadal	13	29
Kibondo	30.72	-3.6	1997-2007	Dekadal	8	8
Kigoma	29.63	-4.88	1997-2007	Dekadal	0	0
Kilwa Masoko	39.52	-8.93	1997-2007	Dekadal	4	4
Kisarawe	39.07	-6.9	1997-2007	Dekadal	2	2

Kizimbani	39.32	-9.06	1997-2007	Dekadal	23	23
Kondoa	35.8	-4.92	1997-2007	Dekadal	2	11
Korogwe	38.47	-5.17	1997-2007	Dekadal	2	2
Loliondo	35.58	-2.02	1997-2002	Daily/Dekadal	13	52
Lushoto	38.28	-4.78	1997-2007	Dekadal	2	2
Lyamungu	37.25	-3.23	1997-2007	Dekadal	1	1
Mabuki Livestock	33.18	-2.98	1997-2007	Dekadal	8	8
Mafia	39.67	-7.92	1997-2007	Dekadal	6	6
Magu	33.45	-2.6	1997-2005	Dekadal	43	61
Mahenge Hydro.	36.72	-8.68	1997-2007	Dekadal	0	0
Mahenge Met.	36.72	-8.66	1997-2007	Dekadal	0	0
Manyoni	34.83	-5.73	1997-2007	Dekadal	23	23
Masasi Agriculture	38.8	-10.73	1997-2007	Dekadal	4	4
Masasi Girls	38.82	-10.7	1997-2005	Dekadal	70	75
Maswa	33.77	-3.17	1997-2002	Dekadal	60	65
Mbamba Bay	34.77	-11.03	2002-2007	Dekadal	36	65
Mbeya	33.44	-8.89	1997-2007	Daily/dekadal	0	0
Mbinga	35.02	-10.93	1998-2007	Dekadal	10	18
Mbozi	32.97	-9.03	1998-2007	Dekadal	10	18
Monduli	36.45	-3.32	1997-2007	Dekadal	2	2
Morogoro	38.41	-5.18	1997-2007	Dekadal	1	1
Moshi	37.34	-3.34	1997-2007	Dekadal	0	0
Mpanda Bomani	31.07	-6.4	1997-2002	Dekadal	22	58
Mtera	35.92	-7.08	1997-2007	Dekadal	6	6
Mtwara	40.18	-10.27	1997-2007	Daily/dekadal	0	0
Mugumu	34.72	-1.87	1997-2007	Dekadal	10	10
Musoma	33.8	-1.5	1997-2007	Daily/dekadal	0	0
Mwadui	33.59	-3.53	1997-2002	Dekadal	58	69
Mwanza	32.89	-2.52	1997-2007	Daily/dekadal	0	0
Nachingwea	38.75	-10.35	1997-2007	Dekadal	2	2
Nansio	33.08	-2.12	1997-2007	Dekadal	7	7
Newala	39.3	-10.95	1997-2007	Dekadal	8	8
Ngara	30.63	-2.47	1997-2007	Dekadal	5	5
Ngorongoro	35.45	-3.2	1999-2007	Dekadal	29	47
Ngudu	33.35	-2.95	1997-2007	Dekadal	43	43
NIA Dar es Salaam	39.2	-6.87	1997-2007	Daily/dekadal	0	0
Nkasi Namanyere	31.50	-7.52	1997-2007	Dekadal	5	5
Nyanza Salt mine	30.33	-5.1	1997-2007	Dekadal	22	40
Nzega	33.18	-4.22	1997-2007	Dekadal	0	0
Nzega Hydromet.	33.22	-4.25	1997-2007	Dekadal	2	2
Pangani	38.98	-5.43	1997-2007	Dekadal	3	3

Pemba	39.75	-5.2	1997-2007	Dekadal	0	0
Peramiho	35.47	-10.57	1997-2007	Dekadal	3	21
Ruvu Farm	38.75	-6.75	1997-2005	Dekadal	37	48
Same	37.75	-4.07	1997-2007	Dekadal	0	0
Sengerema	32.67	-2.67	1997-2007	Dekadal	5	5
Shirati	33.98	-1.13	1997-2007	Dekadal	14	20
Shinyanga-Maji	33.42	-3.67	1997-2007	Dekadal	2	14
Sikonge Hydromet.	32.77	-5.62	1997-2007	Dekadal	41	59
Singida	34.75	-4.8	1997-2007	Dekadal	11	20
Songea	35.65	-10.68	1997-2007	Daily/dekadal	0	0
Sumbawanga	31.62	-7.97	1997-2007	Dekadal	12	12
Tabora	32.83	-5.02	1997-2007	Daily/dekadal	0	0
Tanga	39.07	-5.08	1997-2007	Dekadal	0	0
Tengeru	36.92	-3.3	1997-2007	Dekadal	5	5
Tukuyu	33.63	-9.25	1997-2007	Dekadal	0	0
Tunduru	37.37	-11.1	1997-2007	Dekadal	28	28
Ukiriguru Agriculture	33.02	-2.75	1997-2007	Dekadal	23	33
Urambo	32.05	-5.07	1997-2007	Dekadal	14	14
Wami Prison	37.47	-6.4	1997-2007	Dekadal	9	9
Zanzibar	39.2	-6.17	1997-2007	Dekadal	0	0

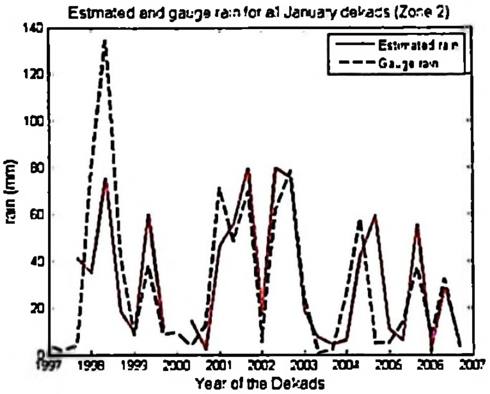
* Daily data for Arusha were available from 1970 to 2006.

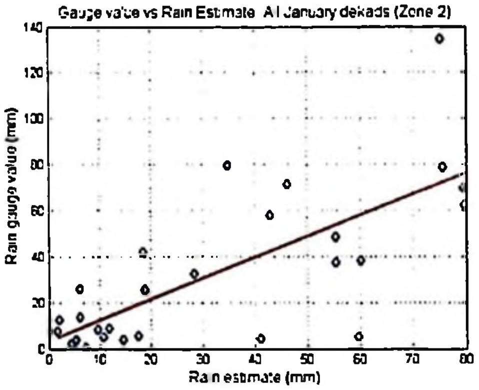
** For Biharamulo, only data at years 2002 and 2008 were available.

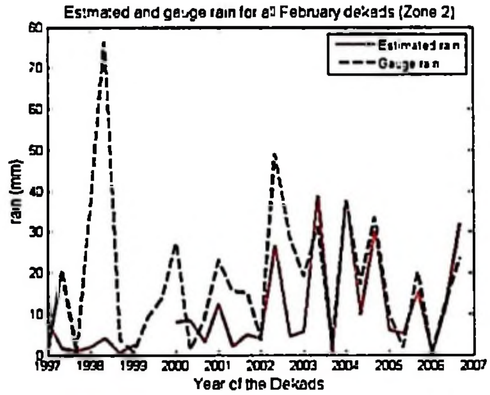
*** Available data for Gairo were only between 1972 and 1989 inclusive.

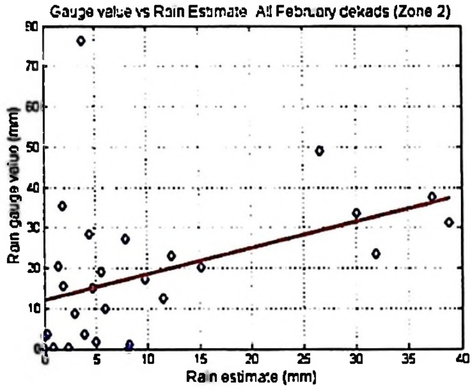
Stations with bolded values of missing percentages were not retained for analysis.

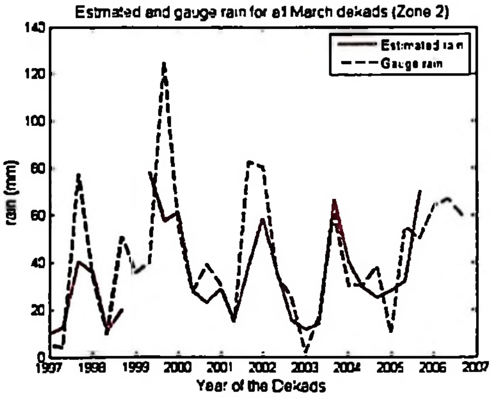
B: Validation

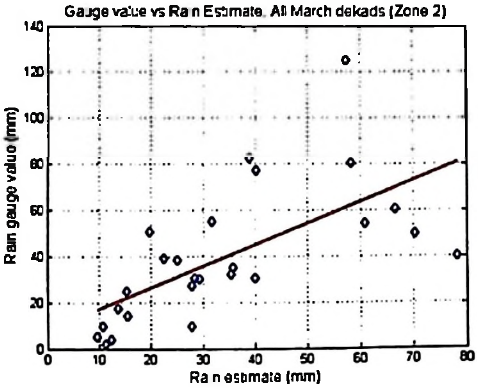












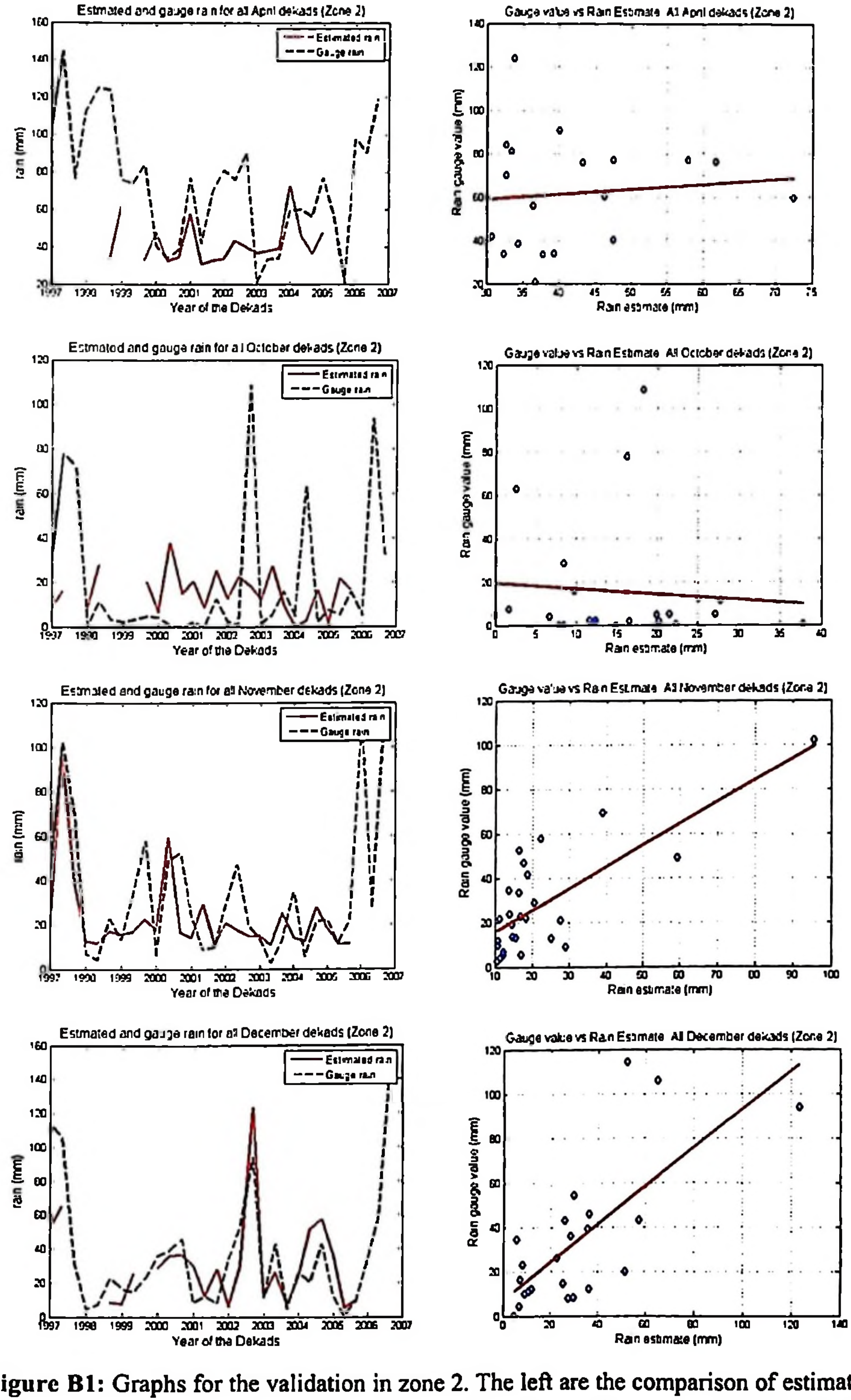
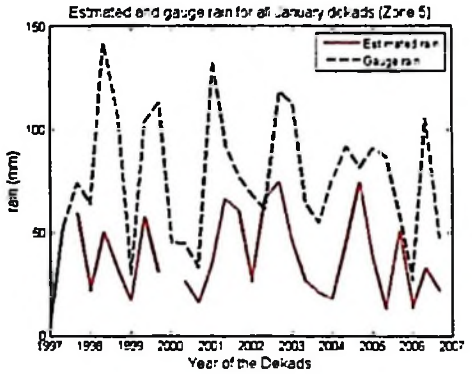
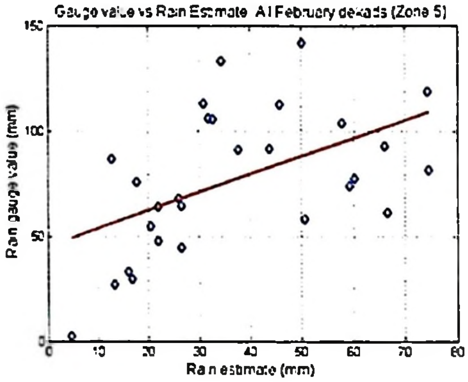
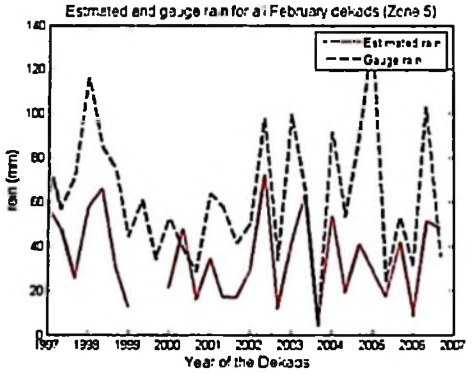
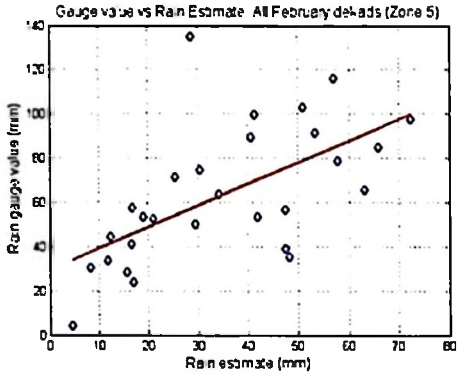


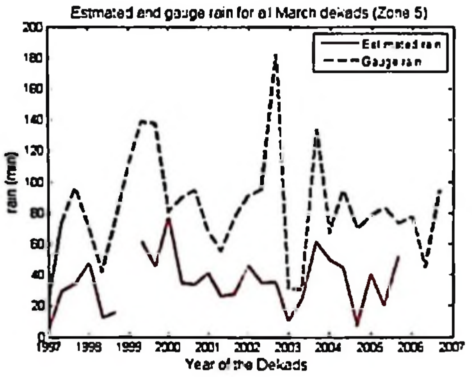
Figure B1: Graphs for the validation in zone 2. The left are the comparison of estimated rain and gauge rain while the right are their corresponding scatter plots with the best fit lines.

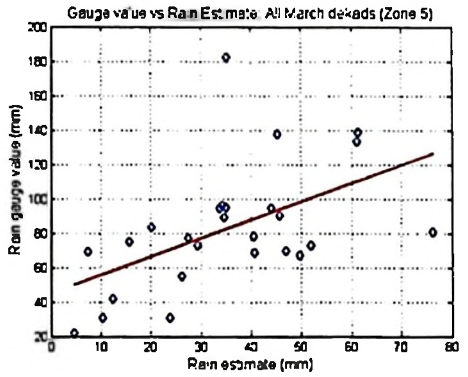

Estimated and gauge rain for all January dekads (Zone 5)
Estimated rain
Gauge rain
rain (mm)
150
100
50
0
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
Year of the Dekads

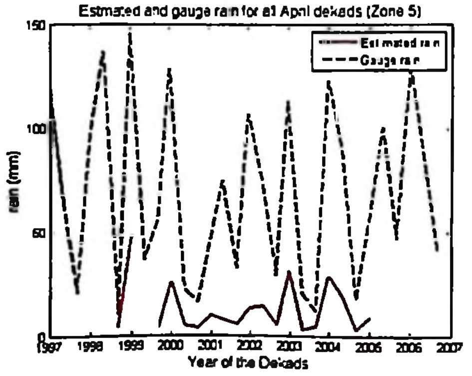

Gauge value vs Rain Estimate All February dekads (Zone 5)
Rain gauge value (mm)
150
100
50
0
0
10
20
30
40
50
60
70
80
Rain estimate (mm)

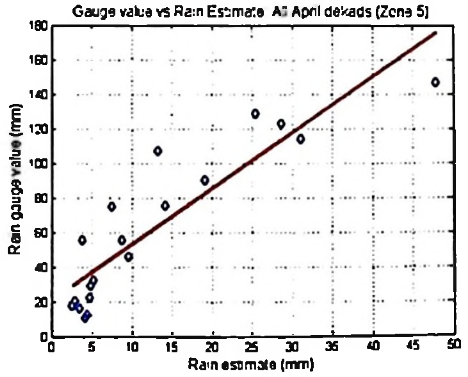

Estimated and gauge rain for all February dekads (Zone 5)
Estimated rain
Gauge rain
rain (mm)
140
120
100
80
60
40
20
0
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
Year of the Dekads


Gauge value vs Rain Estimate All February dekads (Zone 5)
Rain gauge value (mm)
140
120
100
80
60
40
20
0
0
10
20
30
40
50
60
70
80
Rain estimate (mm)


Estimated and gauge rain for all March dekads (Zone 5)
Estimated rain
Gauge rain
rain (mm)
200
180
160
140
120
100
80
60
40
20
0
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
Year of the Dekads


Gauge value vs Rain Estimate: All March dekads (Zone 5)
Rain gauge value (mm)
200
180
160
140
120
100
80
60
40
20
0
10
20
30
40
50
60
70
80
Rain estimate (mm)


Estimated and gauge rain for all April dekads (Zone 5)
Estimated rain
Gauge rain
rain (mm)
150
100
50
0
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
Year of the Dekads


Gauge value vs Rain Estimate All April dekads (Zone 5)
Rain gauge value (mm)
180
160
140
120
100
80
60
40
20
0
0
5
10
15
20
25
30
35
40
45
50
Rain estimate (mm)

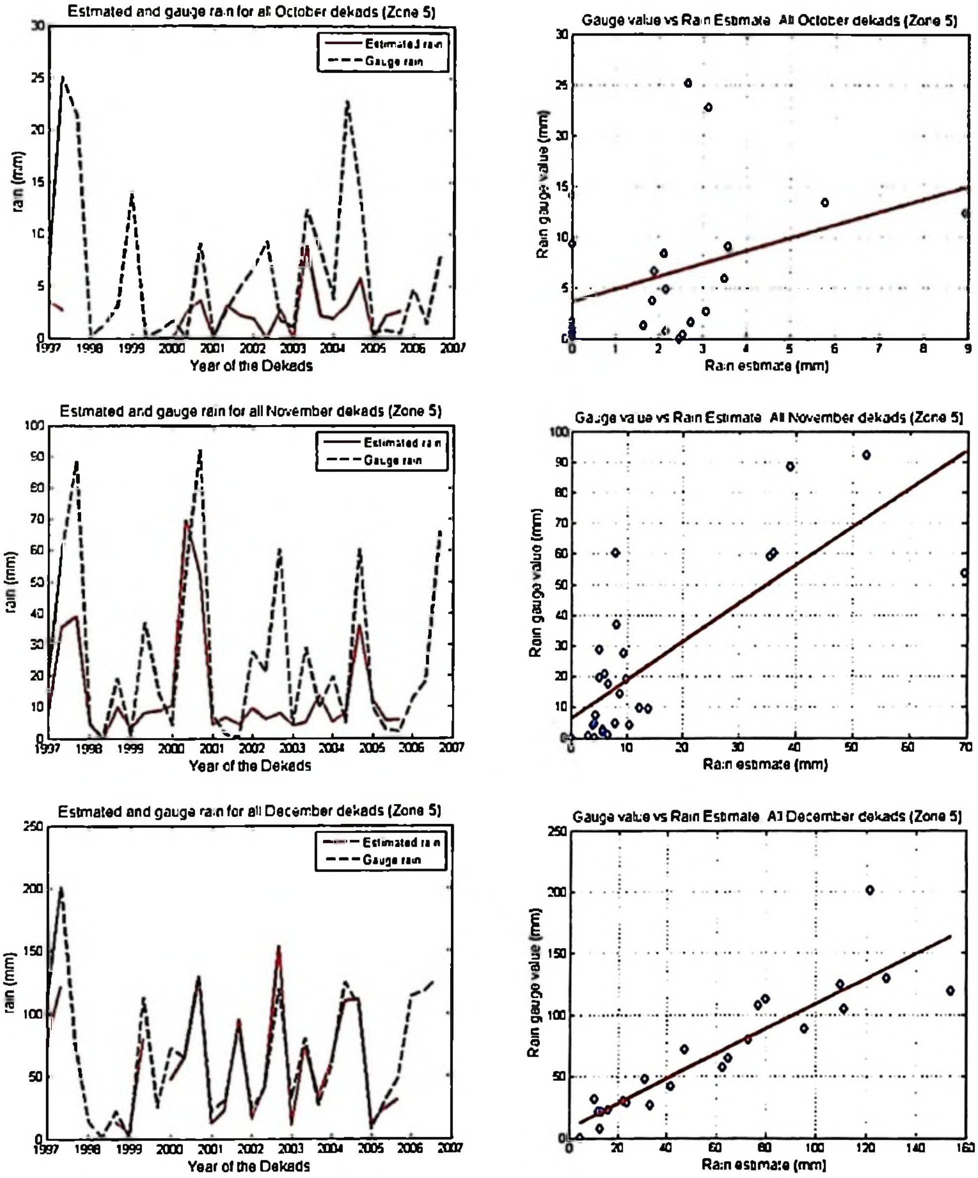


Figure B2: Graphs for the validation in zone 5. The left are the comparison of estimated rain and gauge rain while the right are their corresponding scatter plots with the best fit lines.

C: Trend analysis

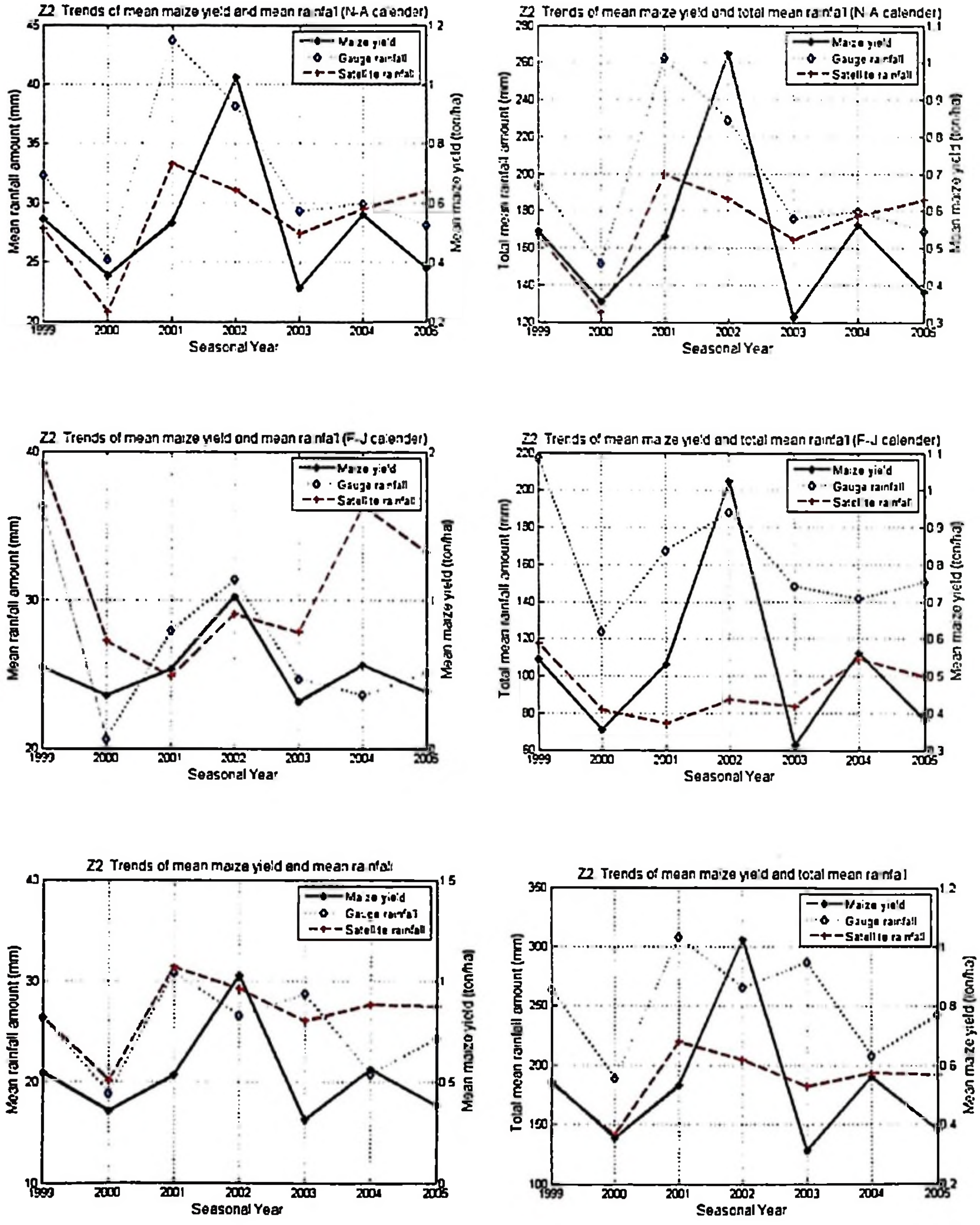


Figure C1: Trends of maize yield, satellite-derived rainfall and gauge rainfall for zone 2 by considering the means of the yield, satellite rainfall and gauge rainfall (left panel). Right panel are those of mean maize yield and the total of the means of the rainfall in the particular period.

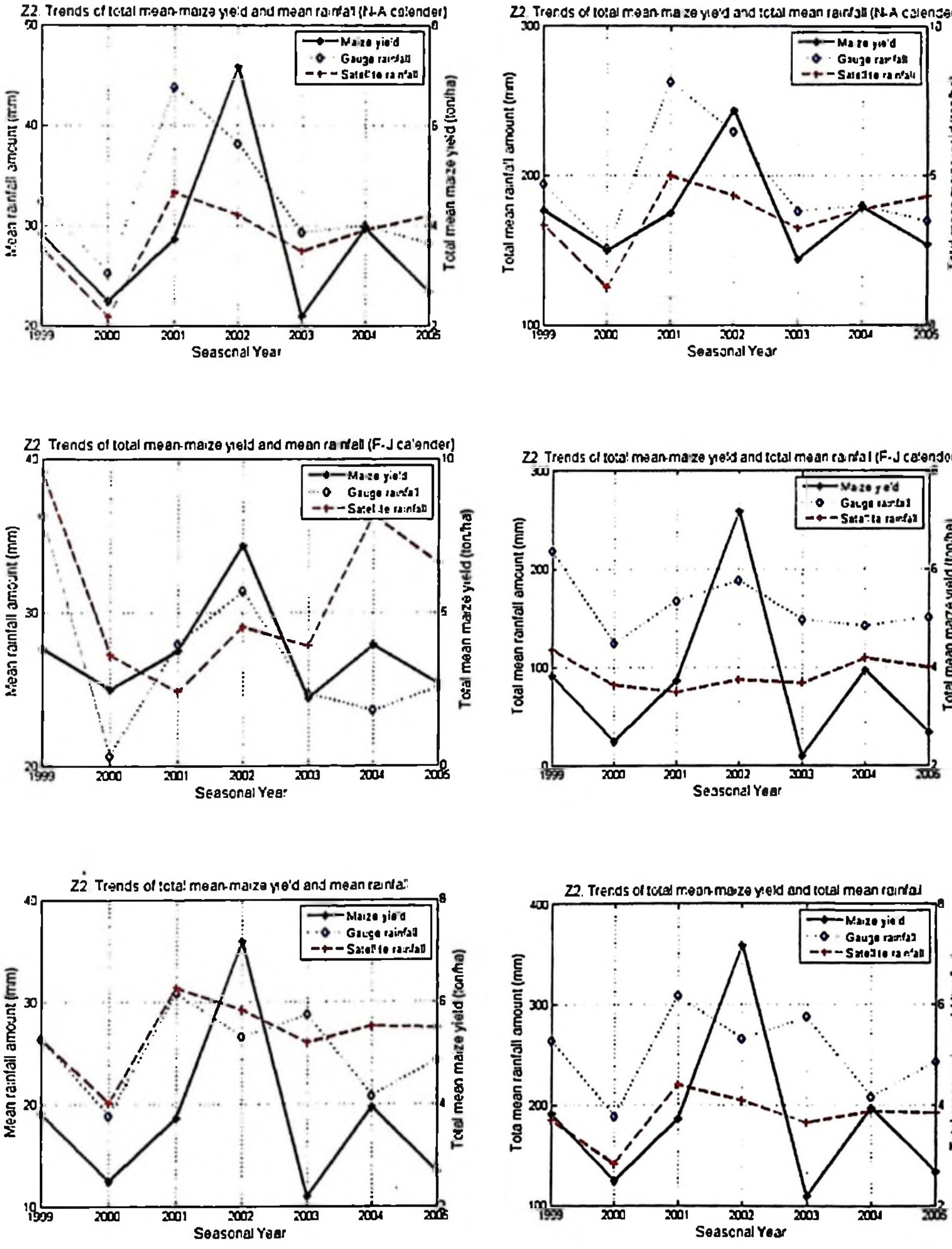


Figure C2: Trends of maize yield, satellite-derived rainfall and gauge rainfall for zone 2 by considering the total of the means of the yield and means of the satellite rainfall and gauge rainfall (left panel). Right panel are those of total of the means of maize yield and means of the rainfalls in the particular period.

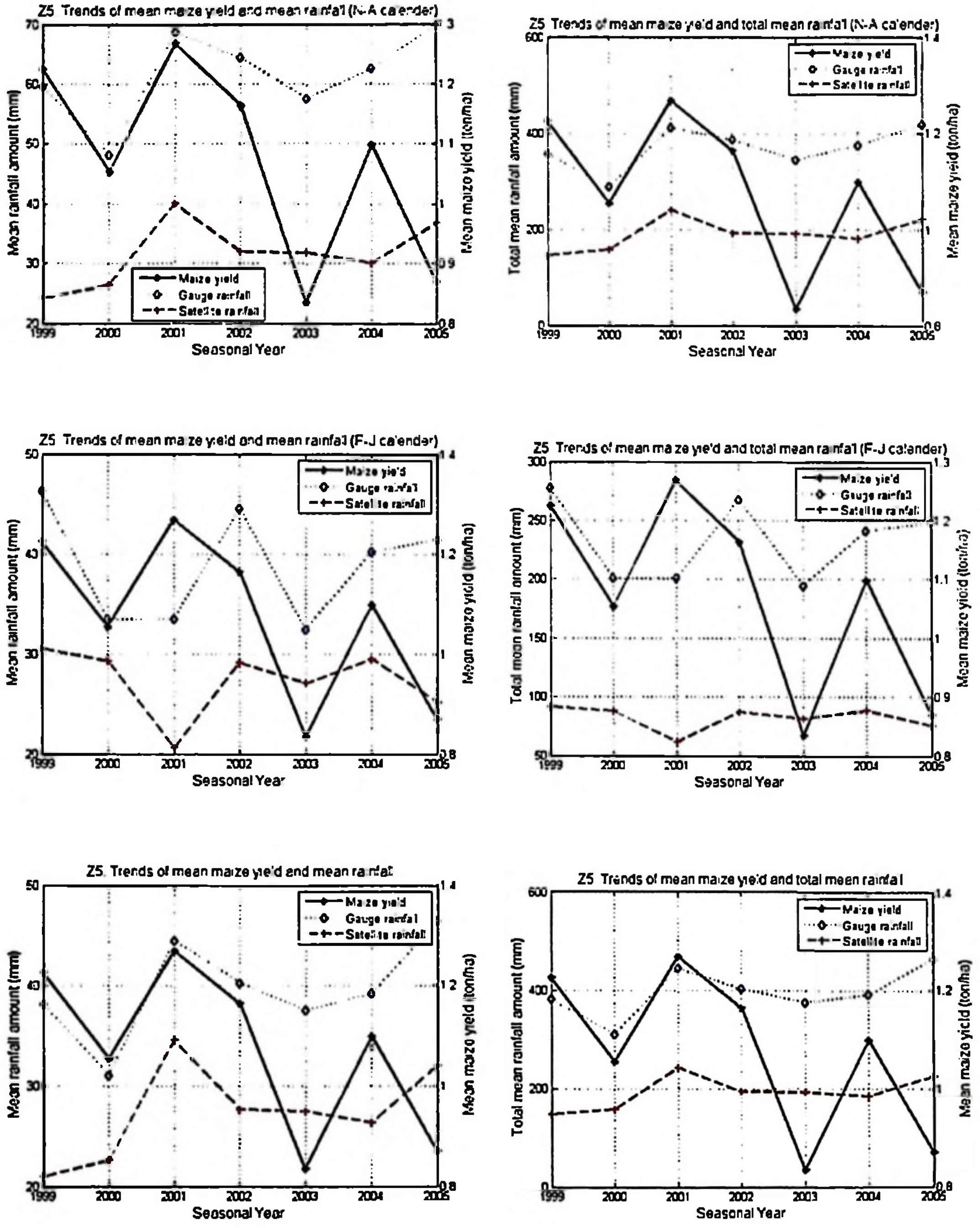


Figure C3: Trends of maize yield, satellite-derived rainfall and gauge rainfall for zone 5 by considering the means of the yield, satellite rainfall and gauge rainfall (left panel). Right panel are those of mean maize yield and the total of the means of the rainfall in the particular period.

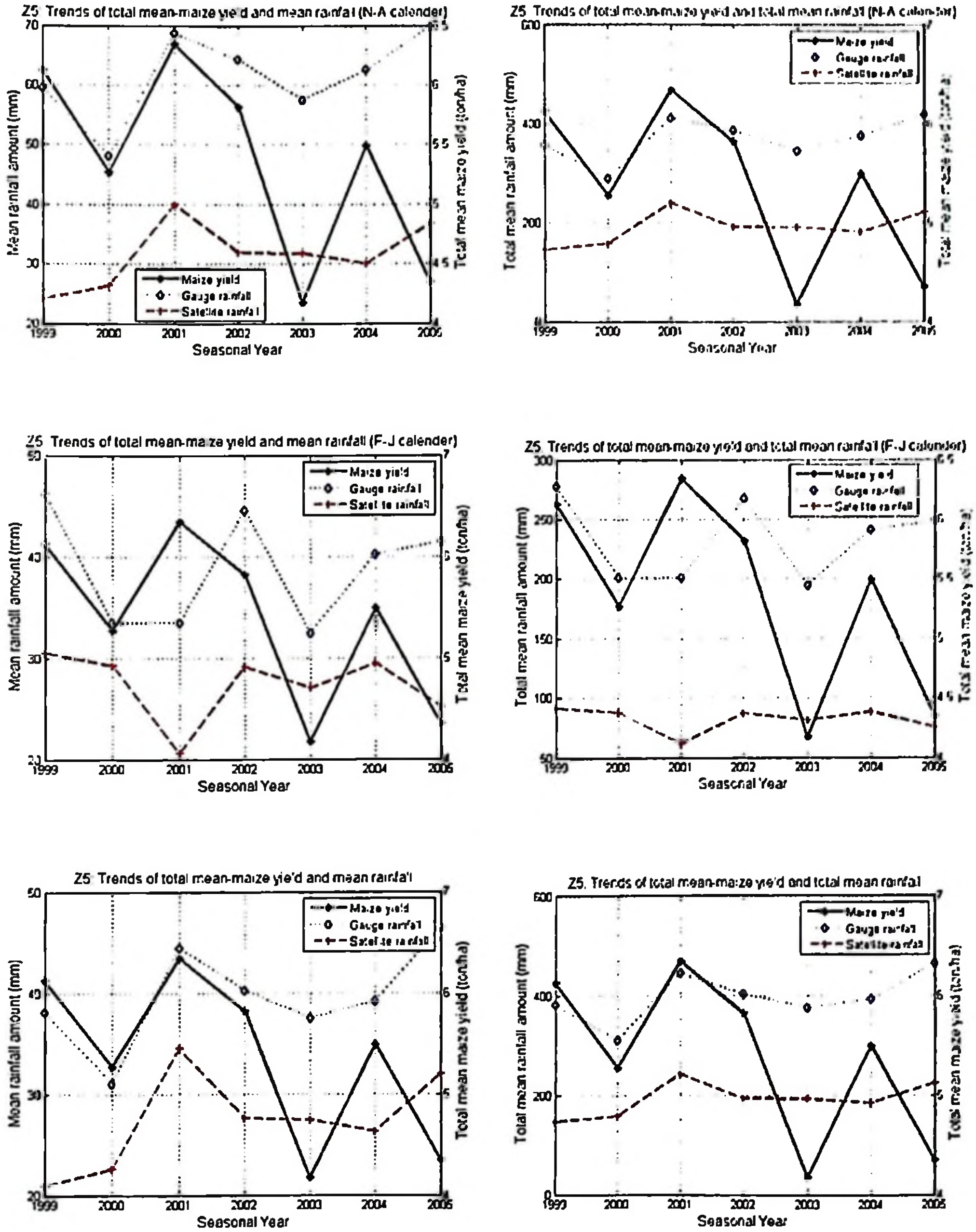


Figure C4: Trends of maize yield, satellite-derived rainfall and gauge rainfall for zone 5 by considering the total of the means of the yield and means of the satellite rainfall and gauge rainfall (left panel). Right panel are those of total of the means of maize yield and means of the rainfalls in the particular period.