ADJUSTING HISTORICAL RAINFALL DATA FOR AGRICULTURAL RESEARCH: A CASE STUDY OF MAKINDU

BY

MAWORA THOMAS MWAKUDISA

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN APPLIED STATISTICS

DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE

MASENO UNIVERSITY

Abstract

There is much concern worldwide about how climate change would impact rain fed agriculture in developing countries, since many in their farming population are small scale and depend on rain. Farmers get information about the risk of cultivating in the expected weather from meteorologists. Their governments also contribute by advising on farming practices and subsidizing the cost of fertilizers. The governments use credible information from researchers, one such is the International Crop Rresearch Institute for the Semi-Arid Tropics (ICRISAT). In 2008, ICRISAT researchers conducted a research project on how to mitigate and adapt to climate change and concluded that with improved practices, there was hope for rain fed agriculture even under climate change. One of their Kenyan sites used in the ICRISAT research was Makindu, approximately 170 km South East of Nairobi. In this thesis, using climate data for Makindu, a test was conducted to see the difference in rainfall amount, start of rains and resultant yield when rainfall amount and pattern were changed. Since climate was the variable under investigation, fifty years of rainfall data was collected from the Kenya Meteorological Department for use. After cleaning up the data, GenStat was used to create climate change scenarios by adjusting the number of rainy days, spells and rainfall amount by 10%. The analysis showed that changing pattern would result in varying rainfall but delayed the start in rains. The four climate scenarios were then entered into APSIM (Agricultural Production System sIMulator) which simulated crop yield. Overall, change in amount of rain had the most effect on yield. However when looking at the long rains and short rains, change in patterns had more effect than when the change was affecting only the amount.

Chapter 1: Introduction

1.1: Overview

There are many factors that usually affect the agricultural output including phenotypic, genotypic and external factors like slope of land, soil fertility, crop management and climate. Statistical analysis usually helps to assess and filter factors that "truly" influence the productivity. Because of this, it has been applied a lot to agricultural research over the years. This in turn has helped to improve on agricultural land management and the phenotype of the seeds used.

The role climate plays in the variability of agricultural production is quite significant, [1]. Lack of adequate rainfall could result in complete crop failure. Statistical analysis has been used to understand the climatic behaviour of certain regions and help inform farmers on the risks involved from historical data [2–4].

In many countries, there are National Meteorological Departments that collect climate data for different centres. This meteorological data can also be analysed for climate trends, hence the possibility of its impact on crops grown within that region. Steiner mentions that a pressing need is to "change our data-rich to an information-rich environment"[5]. An example of this might be to use the analysis of climatic data to inform farmers on when to plant.

Over the last quarter of the 20th century, development of decision support tools has helped to shorten the time span needed to conduct research projects that would initially take years. These tools use mathematical models that when input with certain conditions, they simulate climate data stochastically[6] or agricultural outputs deterministically [6,7]. With them, the possibilities of future climate[8] can all be analysed to help in planning and informing farmers. They have been used to analyse the effect of a variety of land and crop management options on the agricultural output in as little as a week[9].

In this thesis three different mathematical and statistical models were used. They each had a different role to play; they were selected because they had free license, had online tutorials and had been used before as decision support tools to fill in missing weather data or generating yield values. Stochastic climatic models were used to fill in missing values in the climate data. Deterministic models were used to simulate changes in rainfall amount and pattern. A mathematical crop model was used to simulate yields for Maize under the different climatic scenarios.

Fifty years of historical climate data for Makindu were first analysed for trends. Temperature data gave statistically significant trends but there was not enough evidence to support a change in rainfall. The results from the temperature were also questionable since only short term data was available (data for twenty years), with some days having missing data. In addition, to conduct climate analysis, at least thirty years of data was needed. Therefore, Weatherman [6] was used to fill the missing temperature and radiation data.

Climate change scenarios were then created by changing the rain by 10% as in [9] and using other methods. The rainfall changes of 10% were simulated for rainfall amount, rain days and rain spells using GenStat [10] and then analysed for the differences on the total annual rainfall.

The data from the different climate scenarios were then input to Agricultural Production Systems sIMulator (APSIM) [7], a crop model, in order to simulate yields for maize. Analysis was conducted for overall yields and for seasonal yields. The results have shown that there are differences in yields with a similar change in proportion of rainfall amount or pattern.

1.2: Statement of the problem

There are many studies currently looking at the effect of future climatic changes on crop production. This study has been left to Agronomists with limited statistical methodology in the generation of the climate change scenarios. To help these analyses, models that simulate crop performance have been developed. These models can also be used to simulate crop performance under a climate change scenario. APSIM for instance can change temperature or rainfall experienced on a given day by a certain proportion, without affecting the pattern. The problem is that should climate change, in particular rainfall, it is the pattern that will be affected most. Therefore this thesis sought to look at how changing historical rainfall patterns by the same proportion will compare with when only the amount was adjusted (as in the models) and also how this may affect the yield.

1.3: Justification

The analysis of climate data depends upon getting the actual information on the trends of rainfall and temperature from historical data. Though a lot of evidence indicates climate change, it is not clear how rainfall will change in the future. Models can enable analysis of this and were used to simulate a couple of possible rainfall change scenarios in this thesis.

The ICRISAT study was conducted in 2008[9]. This thesis is timely to provide more ideas to further the ICRISAT research mentioned by incorporating the change in rainfall distributions in them. By doing this, the results from the decision support tools may be deemed more dependable. The decision support tools are software that can be used to simulate possible occurrences of either climate or agricultural production. The two Decision support tools used in this thesis are Weatherman and APSIM (Agricultural Production System sIMulator).

Makindu is classified as a Semi-Arid Tropic site and was used in the ICRISAT project in 2008[9]. Its climate data was readily available from the Kenya Meteorological Department. In addition, APSIM had the maize module for the cultivar that is predominantly grown there – *katumani*. For the two reasons above, Makindu was the site of choice for this thesis.

1.4: Objectives

The main objective was to compare different ways of creating rainfall change scenarios. In particular, how a change in pattern differs from just changing the amount. The Specific objectives are:

- To compare differences in rainfall amount when 10% of the number of rain days are removed and when 10% of rainfall amount is reduced on all rain days
- To compare differences in rainfall amount when 10% of the number of rain spells are removed and when 10% of rainfall amount is reduced on all rain days
- To compare yield differences in case when there are 10% less rain days and when 10% of rainfall amount is reduced on all rain days
- To compare yield differences in case when there are 10% less rain spells and when 10% of rainfall amount is reduced on all rain days

1.5: Significance of the study

The work contained in this thesis is important to both farmers and students. Historical climate data gives a good complement to the intuitive knowledge of farmers on climate patterns. Its analysis also helps to confirm farmer perceptions on climate trends.

This thesis contributes the element of including a change in the distribution of rainfall when studying the effect of climate change. The change was based on historical data using a deterministic approach which can be done in many statistical packages. This means that this concept is transferrable and can be applied in different locations and contexts hence contribute to more informed analysis of effects of climate change on crop yields.

The use of models helps simulate effects of different climate change scenarios on crop yield also give them information on the risks involved.

1.6: Outline of the thesis

Chapter 2: gives a review of what has been done in this area of research. The review looks at what different researchers are saying concerning climate change, and also shows the growing use of simulation models to fill missing rainfall data and also to simulate crop yields. The focus is on the two models Weatherman and APSIM respectively.

The methodology chapter (Chapter 3:) gives a detail on how the whole process was conducted. It explains more on the models and how they work.

The chapter on results (Chapter 4:) gives the tabulation of the data and explanation of analysis resulting from the data. The first section analyses the Makindu temperature and rainfall data by exploration and testing for a statistically significant trend over the years. The data used for this initial analysis was the raw data from Kenya Meteorological Department. This chapter also contains the analysis of the comparison of the simulated yield results.

The last chapter (Chapter 5: gives a conclusion on the comparison between the different ways of simulating climate change and their effect on yield.

Chapter 2: Literature review

In 2008 the ICRISAT researchers, led by Peter Cooper, analysed potential effect of changes in climate on rain fed agriculture in Semi-Arid Tropics (SAT) sites within the Sub-Saharan Africa and Asia. The experiment was a 3 ×5 factorial (three rainfall factors each having five levels of increase in temperature) with climate as the experimental unit. To create the climate change scenarios, they used guidelines from the Intergovernmental Panel for Climate Change (IPCC. They then created fifteen climate change scenarios from available historical climate data by changing rainfall amounts by +10%, 0% and -10% and increasing temperature by 1, 2, 3, 4 and 5°C. This data was then fed into APSIM which simulated yields that were analysed to check for significant differences in the climate change scenarios[9].

This section shares work done building up to the various components employed in the research above. First we shall look at what has been done with regard to experiments in Makindu, then have a closer look at different research over the years on climate variability and change. Finally, the development and uses of the crop simulation model will be discussed.

2.1: Makindu

Makindu town is located in the South Eastern part of Kenya, latitude 2° 16' 30.00" S and longitude 37° 49' 12.00" E (see figure 1) with an altitude of 990 meters. A website by Makindu elders' society, mentioned that agriculture contributes to 70% of their economy. This agriculture is mostly rainfall dependent and uses traditional techniques[11].

Makindu has over the years been used as a study site by various researchers. One reason for its preference is the availability of long term historical climate data. For instance, F. A. Mutere who was interested in studying bats chose Makindu due to the availability of a nearby Meteorological station which had adequate rainfall data[12].

Makindu is also close to the Kenya Agricultural Research Institution (KARI) at Katumani which has been partnering with several research institutions like the ICRISAT. ICRISAT is an international institution that has been working with several partner countries and testing different interventions across their sites. One of their key research areas is resilience of the Semi-Arid tropics on climate variability. Hansen[13], for instance, linked rainfall forecasts to yield predictability for maize crop in 2004. In his work, he re-introduced the use of crop models which are discussed in detail in a subsequent section.

Figure 1: Map showing the location of Makindu (Wikipedia 2011)

2.2: Climate change analyses

The world over, there is general consensus that temperature is increasing. However, rainfall data from different areas exhibit more variability and the direction of climate change is then not clear. This has also been true for Uganda[3] and Zambia[2]. It is no wonder when looking at impact of climate in a Asian Semi-Arid Tropics, Naveen P. Singh mentioned that their rainfall pattern is characterised by a lot of variability[14].

Historically there are several projects that looked at the impacts of climate change, like M V K Sivakumar in 2005[15]. What featured in this work was the change in the frequency of extreme events in the Semi-Arid tropics in Asia. In 2011, C W Recha looked at the variability in a district close to Makindu, Tharaka[16].Recha compared the within season rainfall variability, onset and cessation of rainfall between the March-April-May season and the October-November-December season. The variability is more in the former, he said. This thesis therefore looked at changing the rainfall variability which was not included in the 2008 ICRISAT project[17].However, what is not known completely is exactly how the climate will change (especially rainfall). In the section below, we look at how these studies are currently being facilitated.

2.3: Future climate change projections

These General Circulation Models (GCMs) are particularly important in down scaling the climate change information. In 2009, James W Hansen investigated how the output from these models can be useful to inform farmers on the management options for the next season[18]. His results contributed to "knowledge of seasonal forecast value in a relatively high-risk, high-predictability context; utility and value of forecasts derived from a GCM; and risk implications of smallholder farmers responding to forecasts". His analysis was more on the value addition of the GCM models; in this thesis however, the value was on how they can be used for long term climate forecasting.

The Intergovernmental Panel on Climate Change (IPCC) uses the General Circulation Models (GCMs) to simulate probable climate change scenarios in the short and long term. These simulations are based on the Special Report on Emission Scenarios (SRES), and result in four main classifications B1, B2, A1 and A2 which do not consider a change in emission policies. The scenario B1 is combined with a low climate sensitivity, A1 and B2 with medium climate sensitivity and A2 with high climate sensitivity. Predictions under the three sensitivity classes are predicted to result possibly in 1.5°C, 2.5°C and 4.5°C increase in temperatures respectively in the year 2050 [8]. The results from different working groups are synthesised into one, like the one for 2007[19]. IPCC in addition provided a report on mitigation[20]. These reports are then subjected to analysis, like done by Terry Barker in 2007[21] who assessed the report for the fourth working group.

D. A. Stone in 2008 gave caution on using the future climate scenarios as predicted by the models[22].He mentioned robustness in the predictions for the short term period whereas the uncertainty in climate model structure and the unknown future levels of greenhouse gas emissions should be considered. It is not surprising that the ICRISAT team of researchers, rather than settle for a distinct rainfall change scenario for a site, considered several options[17]. They opted to either leave the rainfall as it is or change by ±10% of the rainfall amount[9]. In this thesis, only the 10% adjustment was applied to both rainfall amount and distribution.

The reports from the IPCC patterns have shown more on the accumulated change in amount[21]. However the need to look at the change in rainfall pattern cannot be over emphasised. The use of available statistical packages to create a deterministic change in pattern is one option used in this thesis.

Weatherman[23], which is a constituent of Decision Support System for Agrotechnology Transfer (DSSAT)[6], has been under development for more than 20 years. The current version uses Windows, unlike the first version that used DOS.

The Windows version creates a more user friendly interface. Data can be imported into Weatherman from a variety of files including Comma Separated files (CSV), Excel files (XLS), text files (TXT) and also database file (DBF).

The ICRISAT team used Weatherman (in 2008) to fill missing rainfall data. The climate change scenarios were created directly in the Crop Simulation Model (APSIM). Apart from Weatherman, Cooper in 2008 also used MarkSim[24]to generate daily weather data for use.

The next section looks at the increasing use of crop simulation models to aid in Agricultural research.

2.4: Using models for research

Crop Simulation Models have been developed and improved for over three decades, and can be used to 'grow crops' within a matter of minutes. The models simulate crop development and production using what is already known about crop phenology. For instance, in 1996, W S Mollah used them in an advanced research cutting across 103 seasons to study rainfall variability in the Northern Territory of Australia[25]. More recently, P. D. Jamieson in 2008 used them to study how they respond to water shortages[26]. In his study, he used both simple and more complex models.

The Agricultural Production System sIMulator (APSIM)[7] is a deterministic crop simulation software that was developed by the Agricultural Productions Systems Research Unit (APSRU) in1991. This was a collaboration of Commonwealth Scientific and Industrial Research Organisation (CSIRO) and Queensland State Government agencies, Australia[7].

APSIM modelling framework constitutes three important modules that can be edited to user preference and a fourth module that is a simulation engine which drives "drives the simulation process and controls all the messages passing between the independent modules", figure 4. The independent modules constitute:

1) Modules that simulate biological and physical processess in plants and the whole farm system. The user has very little control over this module.

2) Modules that allow the user to input rules for crop management like planting dates, irrigation, length of growing period and planting date.

3) Module that allows input and output of data. The data input can be meteorological while the output can be yield or biomass[7].

APSIM requires meteorological and crop management input to generate yield for a paddock. This creates a good platform for comparing different treatments. APSIM is an engine driven by models that capture a variety of factors that affect crop production. For instance, rainfall received is balanced between surface runoffs, transpiration, redistribution and drainage. The water is used at different rates during different stages of the plant growth.

In 2002, H. Meinke used APSIM to develop and validate it when used to simulate crop growth under environmental changes influenced by windbreaks[27]. In 2011, Hongtao Xing compared predictions of Carbon dioxide emissions between APSIM and another model called DAYCENT and had favourable remark for APSIM[28].

Since APSIM captures a realistic development of crop growth, it was the suitable model for use in this project.

Chapter 3: Methodology

3.1: Introduction

This section describes the different activities that were done in cleaning, analysing, simulating climate change scenarios and the subsequent yields. The first section gives an overview of all. The following section shows how Markov models were used for analysis and the last section shows how the models were used.

3.2: Sequence of work on project

Since this project was dealing with climate, at least thirty years of data was needed. The Kenya Meteorological Department (KMD) provided fifty years of daily climate data from Makindu, Kenya upon request. KMD provided available daily climate data for between January 1961 to December 2010 for rainfall, temperature and radiation.

Exploration was then done using boxplots, line plots and other summary statistics to check for outliers. Outliers located using box plots had their context examined. Values for the preceding and succeeding days were compared with the outlier to see if it could have been an error in recording. Corrective measures were then decided. These were; leave outlier as it is, edit it if it followed a pattern, or delete it.

An analysis of the corrected raw data formed the initial analysis. This analysis sought to check for any trends of a change in climate, and also compare it with what some researchers have published. Minimum and maximum temperatures and rainfall were subjected to regression analysis to test for trends.

The exploration and initial analysis were done using GenStat[10]and Instat[29]. Instat was selected mainly because of its climate menu while GenStat was used for its more powerful features for regression analysis. Markov modelling was also used to find the probability of rain on any given day of the year. Both zero-order and first-order models were used.

Since APSIM requires complete daily climate data, Weatherman [6] was used to fill the missing data. Weatherman was chosen since it generated data from the means and variances of the available data. Furthermore, the data was stochastically generated.

To simulate a 10% reduction in rainfall amount, every tenth rainy day was removed starting from the first rainy day in 1961. GenStat syntax was developed to do this. The resultant change in the Markov probability of rainfall was also analysed.

Crop yields were then simulated using APSIM. The crop management options were left as default while the climate data was input for the different rainfall change scenarios.

Finally, the maize yields from the APSIM simulations were analysed using GenStat and Excel. Differences in successful harvests were discussed while those in the maize yields were tested using Generalised Linear Models where the factor variable was the rainfall change scenario.

3.3: Markov modelling for rainfall

Markov models are used to model the probabilities of a phenomena changing from one state to another. In this thesis, the two states were either "a rainy day" or not. None of the states were absorbing. The probability of raining on a given day of the year in the year was calculated from historical data using the formula:

$$probability\ of\ raining\ on\ a\ given\ date = \frac{Number\ of\ days\ with\ rain}{Total\ occurrences\ of\ date\ in\ the\ data}$$

The probabilities were plotted and smoothed using the spline function in GenStat.

3.4: Analysis of climate trends

In order to analyse for a trend in climate elements, regression models were used. Two different sets of models were used; 1) simple linear model where either temperature or rainfall was regressed on year and 2) simple linear model where either temperature or rainfall was regressed on year, but the model factored the months separately. The two were both used in the same order for temperature data.

The models sought to test for a significant slope (whether positive or negative) over the years. This did not indicate that 'year' was a cause of the change. It was simply quantifying the average change in temperature or rain per year.

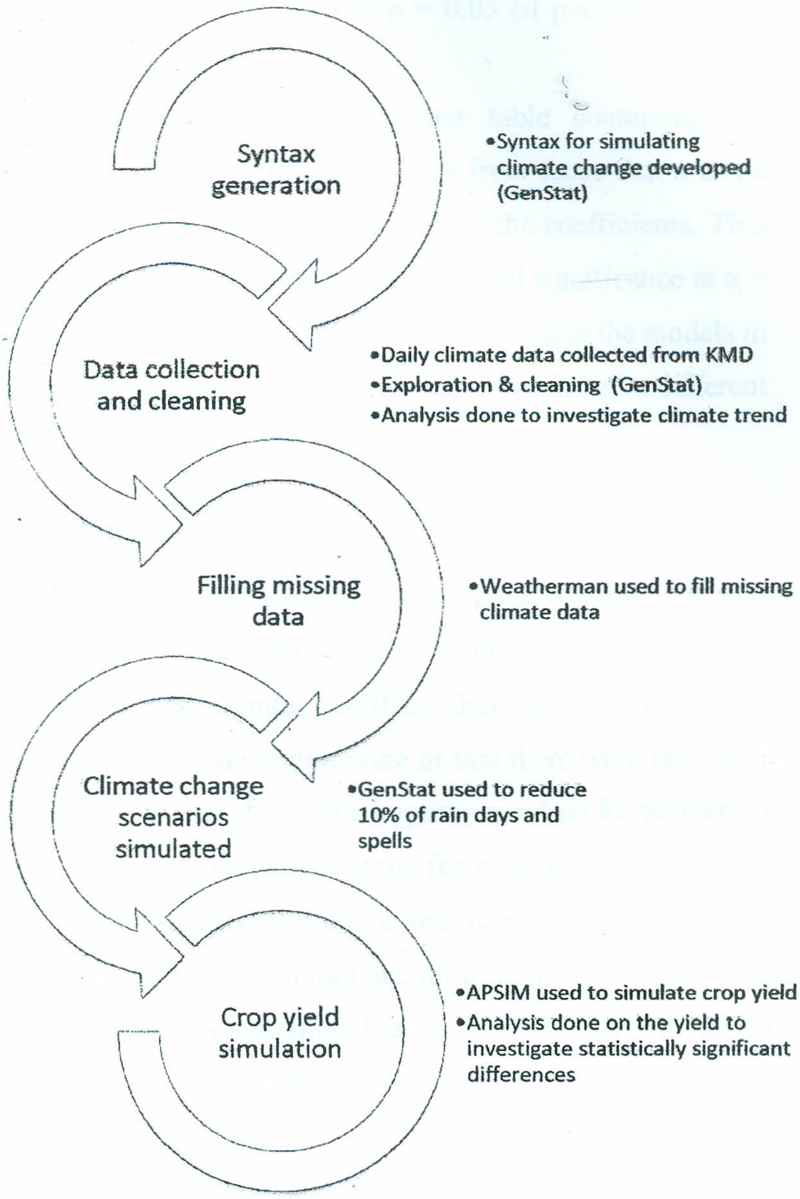


Figure 2: The chronological process as followed in this project

GenStat provided an accumulated ANOVA table for all models used. The accumulated ANOVA tested all the elements in the model without giving the actual estimate of the coefficient. However, it provided the p-values for which one could check whether they were statistically significant at $\alpha = 0.05$ (if p-value was less than 0.05) or not.

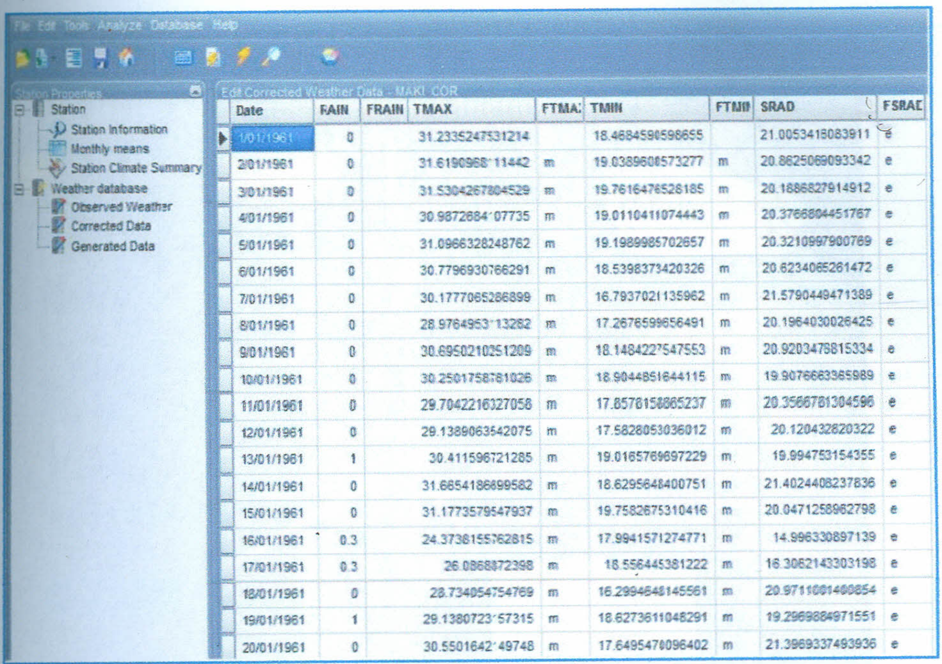
The GenStat output thereafter produced another table containing the estimates for the different coefficients of the model. It is from these that a slope could be calculated. These tables also did a test for each of the coefficients. This meant that some coefficients were not different (with statistical significance at $\alpha =$ 0.05 level) from a 'base' coefficient value. This especially applied to the models in which month was put into the model as a factor hence each month had a different intercept.

3.5: Cleaning and filling weather data

In order to use APSIM to simulate yield, good complete weather data was required. Weatherman was used to check for the quality of the data; like rainfall or radiation values that were less than 0 mm, re-affirm that the minimum daily temperature did not exceed the maximum temperature or that there were any other outliers present. It flagged irregular data during the import procedure for correction purposes. After cleaning the raw data, the data was run for in-filling.

The filling did not affect rainfall data since it was complete from 1961 to 2010. It only generated data for missing maximum and minimum temperature data. There was very little radiation data. Weatherman has an in-built algorithm which calculated the missing radiation data from the available minimum and maximum temperatures.

The complete data was stored in the Weatherman database and exported to the crop simulation software, Decision Support System for Agrotechnology Transfer (DSSAT)[6]. This outputs was saved as a WTG (Weatherman generated), ready for input to crop simulation models.



Date	RAIN	FRAIN	TMAX	FTMA:	TMIN	FTMI	SRAD	FSRAD
1/01/1961	0		31.2335247531214		18.4684590598655		21.0053418083911	e
2/01/1961	0		31.6190968'11442	m	19.0389608573277	m	20.8625069093342	e
3/01/1961	0		31.5304267804529	m	19.7616476528185	m	20.1886827914912	e
4/01/1961	0		30.9872684'07735	m	19.0110411074443	m	20.3766894451767	e
5/01/1961	0		31.0966328248762	m	19.1989985702657	m	20.3210997900769	e
6/01/1961	0		30.7796930766291	m	18.5398373420326	m	20.6234065261472	e
7/01/1961	0		30.1777065286899	m	16.7937021135962	m	21.5790449471389	e
8/01/1961	0		28.9764953'13282	m	17.2676599656491	m	20.1964030026425	e
9/01/1961	0		30.6950210251209	m	18.1484227547553	m	20.9203478815334	e
10/01/1961	0		30.2501758781026	m	18.9044851644115	m	19.9076663365989	e
11/01/1961	0		29.7042216327058	m	17.8578158865237	m	20.3566781304596	e
12/01/1961	0		29.1389063542075	m	17.5828053036012	m	20.120432820322	e
13/01/1961	1		30.411596721285	m	19.0165769697229	m	19.994753154355	e
14/01/1961	0		31.6654186899582	m	18.6295648400751	m	21.4024408237836	e
15/01/1961	0		31.1773579547937	m	19.7582675310416	m	20.0471258962798	e
16/01/1961	0.3		24.3738155762815	m	17.9941571274771	m	14.996330897139	e
17/01/1961	0.3		26.0868872398	m	18.556445381222	m	16.3062143303198	e
18/01/1961	0		28.734054754769	m	16.2994648145561	m	20.9711081400854	e
19/01/1961	1		29.1380723'57315	m	18.6273611048291	m	19.2969884971551	e
20/01/1961	0		30.5501642'49748	m	17.6495478096402	m	21.3969337493936	e

Figure 3: A pictorial view of corrected data in Weatherman. 'm' in the TMAX and TMIN columns is for missing data that was corrected. In the SRAD column, values with 'e' are those that were estimated from the temperature data.

3.6: Using APSIM for crop growth simulations

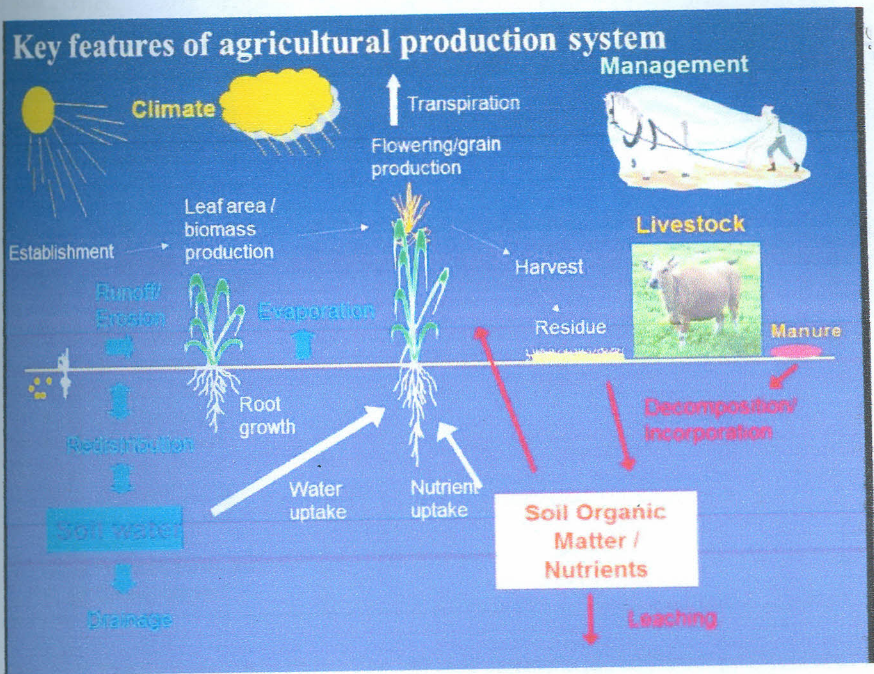


Figure 4: The water and other nutrient balance processes in APSIM (Lisson, 2012)

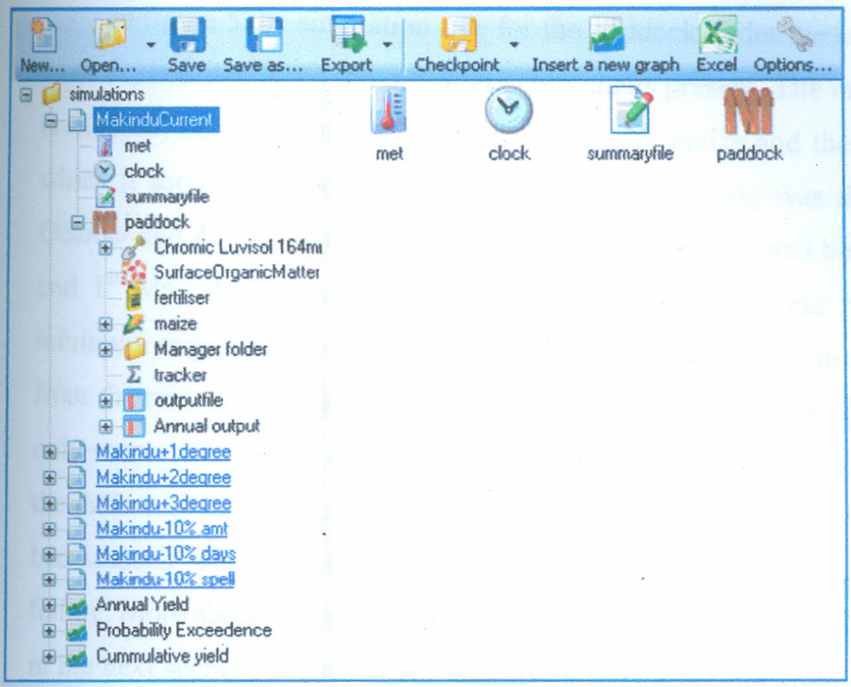


Figure 5: The face view of APSIM. The left side contains the different projects that were simulated using the trials

In this thesis, crop management options and soil properties were used in different paddocks; only the meteorological data was different. For one paddock the input was the recorded (current) climate while for the others the inputs were the different climate change scenarios. In figure 5, it shows simulations done for 1˚C, 2˚C and 3˚C increases in temperature and 10% rainfall change in amount, number of rain days and number of rain spells.

APSIM has some default simulations adapted for different climate patterns. In this project the default was bimodal rainfall patterns which matched the rainfall at Makindu. Since climate was the experimental unit, most of the default settings, like fertilizer application and crop density, remained unchanged.

In figure 5 the simulation tree for the paddock under the current climate has been expanded to show the different components present. The management folder (collapsed) contains details on when to plant the maize and the conditions under which it should be planted for the two seasons. Sowing was done between 10th October and 20th November during the short rains season and between 10th March and 1st May during the long rains season. These were the periods when the minimum requirement of water, 20mm over two days, was more likely to come from the rain as no irrigation was done; the model was set up to have rain as the only source of water. In order to investigate the magnitude of a drop in the yield for the different climate scenarios, APSIM was set up to "force" sowing even in very bad seasons. This meant that lack of output for that season meant an entire crop failure. More about outputs for the different climate change scenarios are discussed in the next section. Management options on the harvesting and preparation is also defined there.

The soil used was Chromic Luvisol 164 mm. It was a soil type found in Katumani which is close to Makindu geographically. The defaults of the soil and fertilizer were not adjusted. The crop grown was maize of the cultivar "*Katumani*" which is adapted to be grown regions in Kenya with low precipitation, including Makindu. The output was set to be produced once at the end of the season. The outputs produced were the yield, biomass and the total rainfall in the season and are analysed in the next chapter.

Chapter 4: Results from analysis

4.1: Introduction

In this section, the analysis results for both the raw data, cleaning procedure, comparison of cumulative rainfall for among the different climate scenarios and comparison of yield results from the different climate scenarios. Statistically, exploratory analysis usually precedes the confirmatory analysis. For this reason, the immediate section explored the Makindu raw data.

4.2: Exploring and cleaning Makindu data

Before doing any analysis, the available data was checked for irregularities so that a decision can be reached on whether to use it as it is or after modifications. Already, it was evident that daily temperature and radiation data would need filling for at least thirty years (see table 1). The temperature and rainfall data was subjected to further investigation for anomalies by manually scrolling through the datasets and using graphs. This section discusses what was found, and the corrective step taken.

4.2.1: The available climate data

The Kenya Meteorological Department (KMD) provided the daily climate data from 1961 (rainfall) and from 1990 (temperatures and radiation). Table 1 gives a pictorial summary of the data that was provided. The first column gives the data that was requested for and the shaded region in a row specify the duration of data available. Rainfall was well recorded for the fifty years; the same could not be said about temperature and radiation.

Climate element	1961 – 1985	1986	1987 – 1989	1990	1991	1992	1993 – 1995	1996	1997 – 2010
Rainfall (mm)									
Max temp (°C)									
Min temp (°C)									
Radiation(Mj/m^2)									

Table 1: A pictorial view of availability of Makindu climate data for the four elements

The rainfall was recorded in millimetres (mm). Days without rainfall were recorded as zero. There was no specified code for a trace value or non-existent data (like for February 29 on non-leap years).

Temperature was recorded in degrees Celsius (°C); both minimum and maximum temperatures were recorded daily. In addition to the years that lacked temperature data, there were forty one days of unrecorded temperature in 2005 which occurred mostly in November. In 2006, there were twenty-nine days with no temperatures recorded with only one of them not in May.

Radiation was measured in mega joules per metre square. Data for it was available between January and October 1990, then between January 1993 and November 1996. In 1994, however, there were fifteen days with missing data spread across the year.

4.2.2: Makindu temperature data

The daily maximum and minimum temperature data were explored to identify causes of noise in the data since they would affect the quality of output from the simulation models to an unknown scale. Initially, the data was checked for instances where the minimum temperature exceeded the maximum. Five days were noted and shared with KMD who responded by giving corrected values.

Further exploration on the temperature used box plots separately for each month, figure 6. The box plots were produced from the 7564 recorded values of daily maximum temperature and 7221 values of daily minimum temperature; the two counts differed since some days did not have both minimum and maximum temperature values recorded. These summaries highlighted a couple of outlier temperature readings for the different months. The outliers exceeding a 3°C gap from the next extreme have been circled. On 17th June 1993, the maximum temperature was 7°C higher than for any other values for June.

There were six overall values for each of the maximum and minimum temperatures where a decision had to be made. This was either to 'edit' the value if a logic was seen in the value or treat the value as missing, and hence to simulate an estimated value for that day. The rationales used are bulleted below:

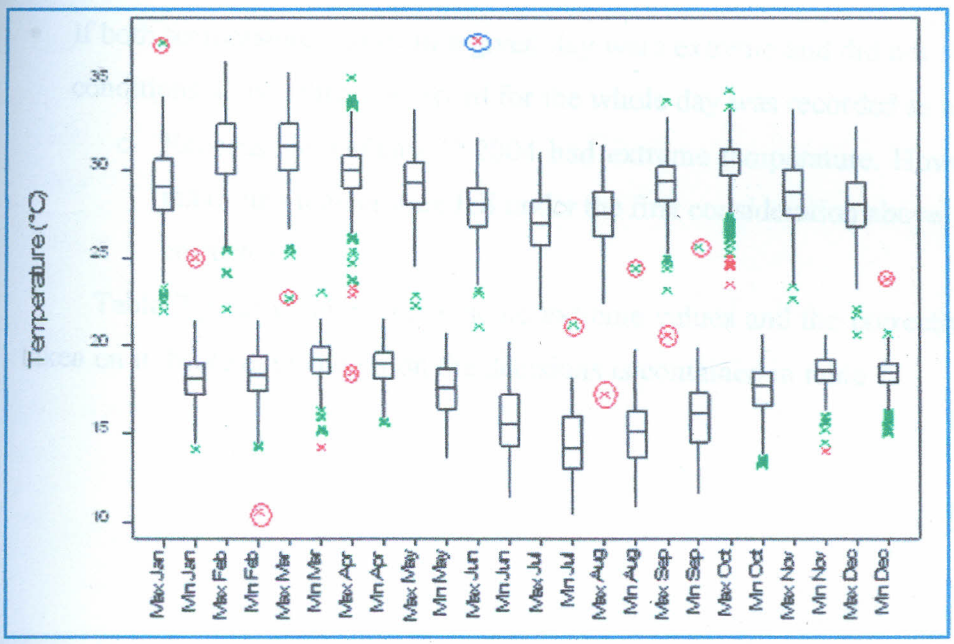


Figure 6: The distribution of recorded daily temperature data for different months (all temperature data included 1990, 1992 - 2010). The box plots are divided for the minimum and maximum daily temperatures.

- If the outlier was 10°C higher or lower than the previous and following days, then adjustment was made by either subtracting or adding 10°C respectively.

For example, the recorded temperature for 9th August 1999 was 17.2°C. The previous and subsequent days recorded 28.6°C and 27.4°C respectively. It is highly unlikely that temperature drops by 10°C in one day and increases again by 10°C in one day. It would seem only logical that the temperature was wrongly entered. Therefore, the correction was set to 27.2°C.

- In case of differences of other magnitudes, a decision was made to either delete it or leave it as it is depending on rainfall observance.
    - Presence of rain on the day in question could be a reason for lower temperatures if the previous and following days didn't rain. In such a case, the record would not be edited. Otherwise the record would be recorded as missing.
- If both temperature values in a given day were extreme and did not fit the two conditions above, then the record for the whole day was recorded as missing.
    - Records for January 1st 2004 had extreme temperature. However, the maximum temperature fell under the first consideration above, so it was corrected.

Table 2 gives a summary of some extreme values and the corrective action taken on it. More explanation on the decisions is contained in table 2.

Temp	Date	Recorded	Previous day	Following day	Decision taken
Max	1-Jan-2004	37.2°C	29°C	31.5°C	Delete
Max	7-Mar-1990	22.7°C	27.8°C	29.2°C	Delete
Max	25-Apr-1997	18.5°C	28.3°C	27.5°C	Delete
Max	17-Jun-1993	37.3°C	26.2°C	27.2°C	27.3°C
Max	9-Aug-1999	17.2°C	28.6°C	27.4°C	27.2°C
Max	17-Sep-2002	20.6°C	28.5°C	24.4°C	Delete
Min	1-Jan-2004	25°C	19.7°C	18.1°C	Delete
Min	18-Feb-1995	10.6°C	20.4°C	18.2°C	20.6°C
Min	22-Jul-2002	21.2°C	17.3°C	12.7°C	Delete
Min	19-Aug-2003	24.4°C	15°C	14.4°C	14.4°C
Min	29-Sep-1990	25.6°C	13.3°C	17.6°C	15.6°C
Min	7-Dec-2002	23.8°C	18.2°C	19.4°C	Delete

Table 2: Context of the extreme values observed and the corrective action taken

The edited data set was again used to produce box plots as shown in figure 7. Even though outliers were present, they were within 2°C of the rest of the data.

4.2.3: Makindu rainfall data

A rainy day was defined as one which had a minimum of 0.85 mm recorded [29]. Days which had less than 0.85mm of rainfall were treated as dry days.

Makindu on average receives 600mm of rainfall from 47 rain days in the full year. The year is itself divided into the short and long rains that receive an average of 370 mm and 220 mm of rainfall from 28 and 19 days respectively. The short rains usually start in October and peak in November while for the long rains it starts in February and peaks in April.

November and December received more rainfall frequency and amount than other months. In half the years recorded, the number of rainy days in November exceeded 15, with the total rainfall exceeding 200 mm for one out of four years, figure 8.

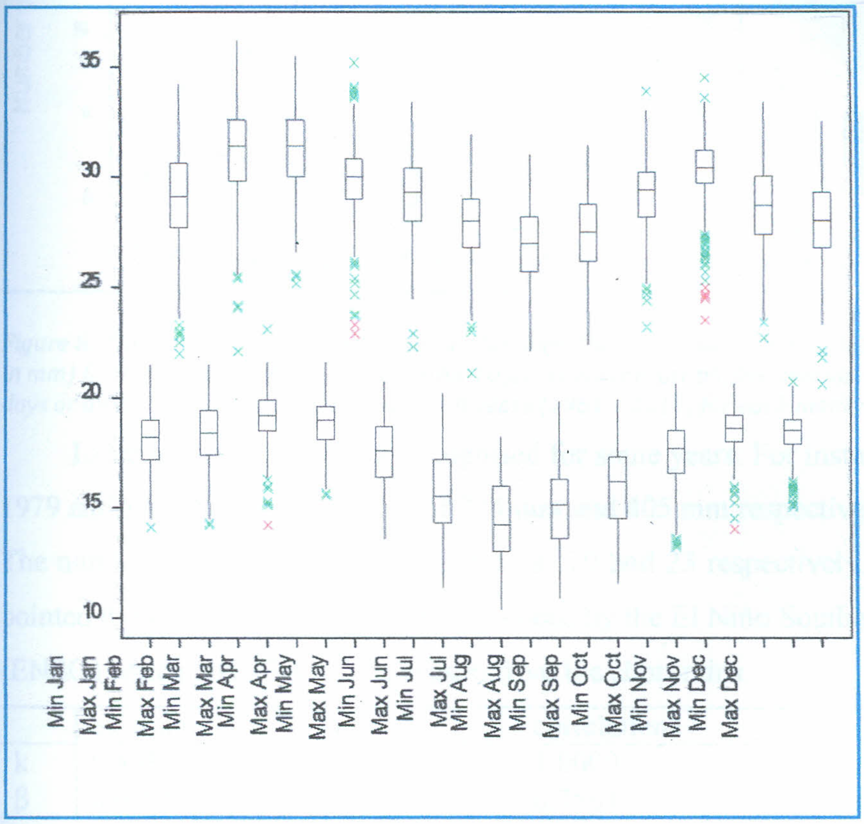


Figure 7: Box plots for the edited Makindu temperature data. In this case, there are no extreme values exceeding 7°C.

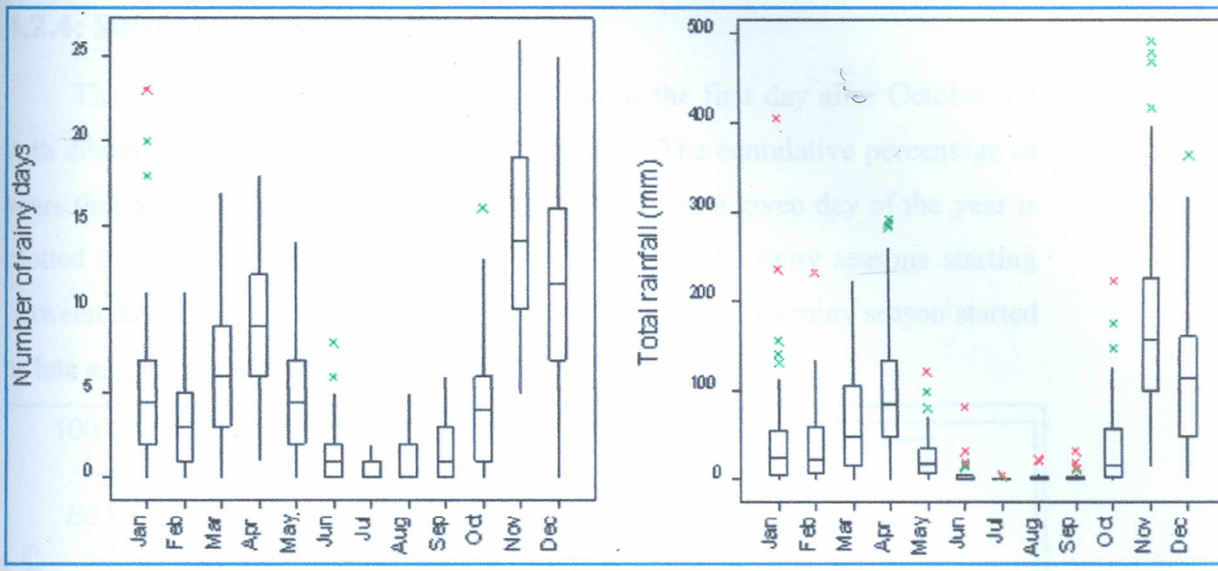


Figure 8: The distribution of number of rainfall days (left) and the total monthly rainfall (right in mm) for Makindu for the twelve months. To generate this graph, the average number or rain days or amount were first calculated for all years (1961 – 2010) for each month

Extreme events were also recognised for some years. For instance, the years 1979 and 1998 had rainfall totals of 234 mm and 405 mm respectively in January. The numbers of rainy days were also high; 20 and 23 respectively. Usher [31] has pointed these years as having been influenced by the El Niño Southern Oscillation (ENSO) which had affected the season from the short rains.

	estimate	s.e.	correlations	
k	0.8818	0.0222	1.0000	
β	0.0705	0.0023	0.7564	1.0000

Table 3: Estimates of the shape and scale parameters to fit the gamma distributed Makindu rainfall data

Since gamma distribution could be used to model rainfall. The shape parameter (k) and the scale parameter (β) were calculated from the available Makindu rainfall data, table 3, and used to calculate the mean rainfall as shown below.

$$mean\ rainfall = \frac{k}{\beta} = \frac{0.8818}{0.0705} = 12.508 \text{ mm}$$

4.2.4: Start of rainfall season

The start of the rainy season was defined as the first day after October 1st with 20mm of rainfall accumulated over 2 days [29]. The cumulative percentage of years that had the start of the rainy season before or on a given day of the year is plotted in figure 9. More than half of the years had the rainy seasons starting between day 301 (October 27th) and 319 (14th November). The rainy season started as late as December in four years.

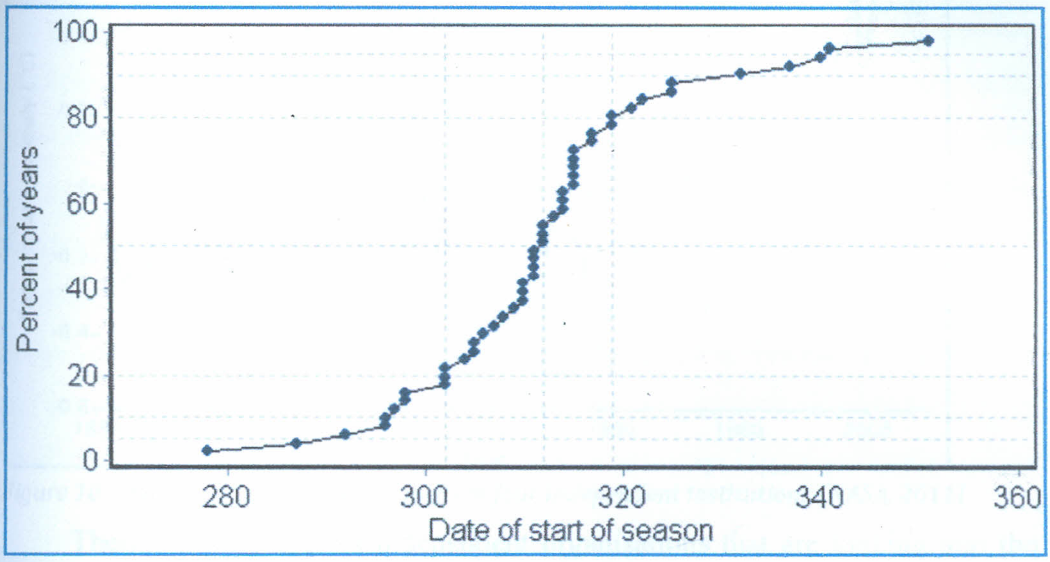


Figure 9: Cumulative percentages of years that had the start of rainy season on the indicated day. The start of rainy season was any day after October 1st (day 275) which accumulated a total of 20 mm in two days.

The effect of El Niño / Southern Oscillations (ENSO) on this event gave contrasting outcomes. Rains started on 7th November 1997 which was not different from other years during the 1997/1998 EI Niño. Due to its magnitude, this event was termed as "The Climate Event of the Century" [31] in the 1998 La Niña summit. In another El Niño season, 1982/1983, rains started on 13th October 1982, which was among the earliest starts of rainy seasons.

Rains started on 6th November 1988 and 28th October 1995 for the 1988/1989 and 1995/1996 La Niña seasons respectively. These were not outlier dates when compared with the rest.

4.3: Long term climate behaviour for temperature

4.3.1: Global warming

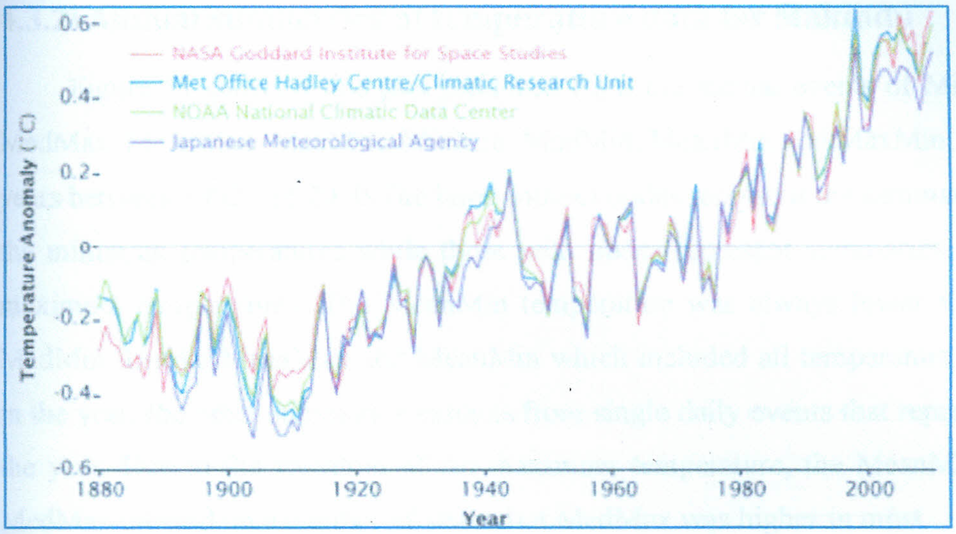


Figure 10: Global temperature anomaly from four independent institutions (NASA, 2011)

There are a number of independent organisations that are looking into the behaviour of temperature over the years from a global perspective. NASA has presented the findings of four such organisations, National Aeronautics and Space Administration (NASA), Met Office Hadley Centre, National Oceanic and Atmospheric Administration (NOAA) and Japanese Meteorological Agency, by plotting annual temperature anomalies for over 100 years, figure 10.

The institutions collected temperature data from different Met Stations across the world and averaged to get a single summary for the globe. A mean of the annual global temperature means for the years 1880 to 2000 was then calculated. The difference of the global annual mean from this is what was plotted in figure 10.

Though there has been an increase in temperature over the 20th century, its rate was higher in the last three decades. In January 2011, NASA, based on a study conducted by Goddard Institute for Space Studies (GISS), mentioned 2010 and 2005 as the two hottest years on record[32]. Figure 10 shows that the hottest seasons have been after the year 2000. NOAA recorded the highest anomaly, of 0.6°C, in the years 1998 and 2005[33].

4.3.2: Annual summaries of temperature data for Makindu

Figure 11 shows a line plot produced from the annual events of MinMax, MedMax, MeanMax, MaxMax, MinMin, MedMin, MeanMin and MaxMin, for the years between 1992 and 2010.The lines without nodes represent the summaries for the minimum temperatures while those with nodes represent summaries for the maximum temperatures. The MeanMin temperature was always lower than the MedMin. Unlike MeanMax and MeanMin which included all temperature values in the year, the other summaries extracts from single daily events that represented the year. Due to the variation of the maximum temperature, the MeanMax and MedMax crossed on a number of years, but MedMax was higher in most.

Figure 11 was produced to check whether temperature has been increasing over the years. From the figure, summaries for the minimum temperature did not visually increase over the years. However, the MinMax increased from 20.6°C in 1992 to 24.6°C in 2010. The MeanMax values, which included all yearly maximum temperatures, crossed the 29°C point in 2003 and has never since gone below that. For Makindu, both MeanMin (17.92°C) and MeanMax (29.96°C) were the highest in 2009. NOAA had mentioned that 2005 and 2010 were hottest [33].

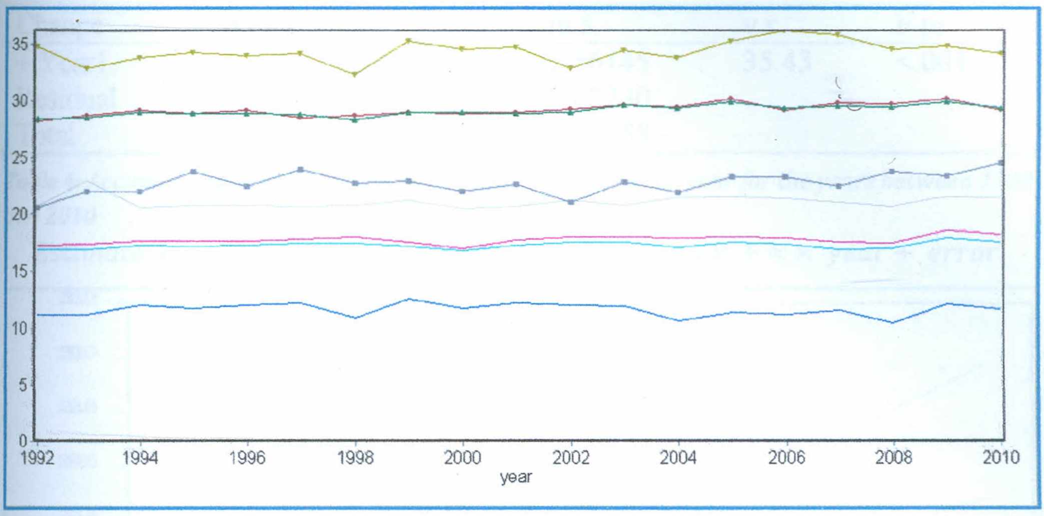


Figure 11: The behaviour of the annual Minimum, Median, Mean and Maximum of minimum and maximum daily temperatures between 1992 and 2010. The lines with points are for the MinMax, MeanMax and MaxMax. The continuous lines from the bottom are MinMin, MeanMin and MaxMin

4.3.3: Confirming trend in Makindu annual temperature

The above section considered the visual investigation of the temperature from a graph. More analysis was done by fitting a line of best fit and testing it for statistical significance. Two lines were fitted. For both, the independent variable was 'year' while the dependent variables were the MeanMin(Mean of the Minimum temperatures) and MeanMax(Mean of the Maximum temperatures). The annual means were chosen since they included all the values for the year. GenStat's simple linear regression function was used. This applies only for years between 1992 and 2010. The simple linear regression equation was:

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ Year1	1	2.56145	2.56145	35.43	<.001
Residual	17	1.22903	0.07230		
Total	18	3.79048	0.21058		

Table 4: Accumulated ANOVA for the regression of MeanMax on year for the years between 1992 and 2010

$$Estimate\ of\ mean\ maximum\ temperature = constant + \propto \times\ year + error$$

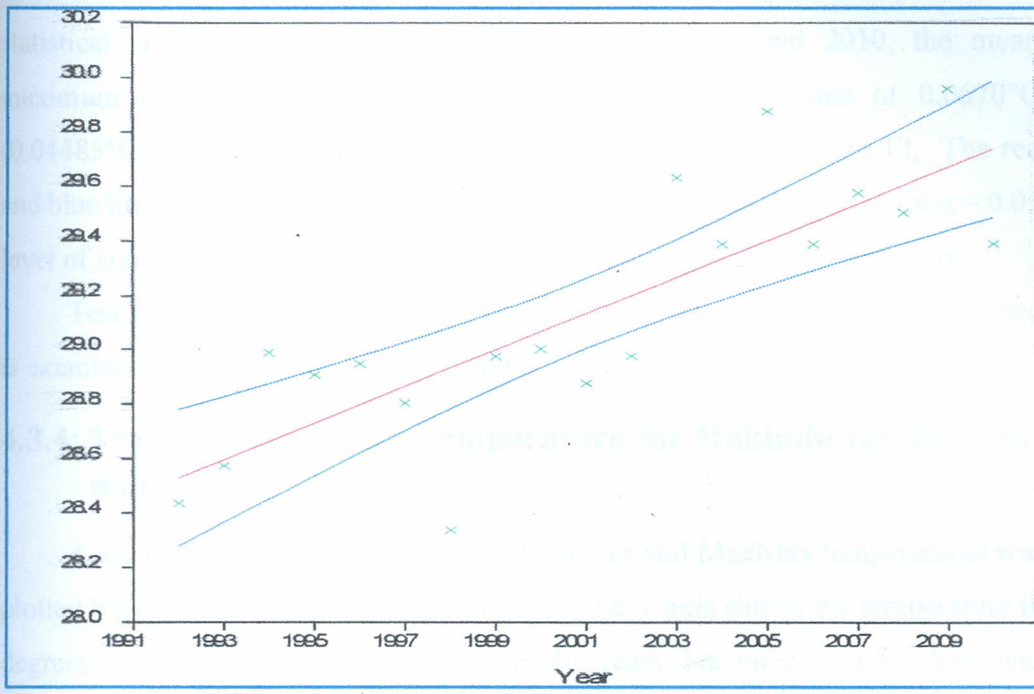


Figure 12: Fitted and observed values for the annual mean maximum temperature with 95% confidence interval

The ANOVA table in table 4 explains the amount of variability in temperature explained by the different factors in the simple regression model above. The regression model had nineteen summary values for each of the years (1992 – 2010). Because of this, the total degrees of freedom are eighteen. There is one explanatory variable in the model and the error term. This makes the total degrees of freedom for the regression equation to be 1.

The regression equation, of MeanMax on year, accounted for 65.7% of the total, table 4. R^2 is calculated by the equation:

$$R^2 = \left(1 - \frac{SS_{residual}}{SS_{Total}}\right) \times 100 = \left(1 - \frac{1.22903}{3.79048}\right) \times 100 = 65.7\%$$

Since the p-value was less than 0.001, therefore the regression equation had statistical significance. This means that between 1992 and 2010, the mean maximum temperature had been increasing at an annual rate of 0.0670°C (0.04485°C, 0.08915°C). A graphical illustration is shown in figure 12. The red and blue lines represent fitted values and their 95% confidence interval (at $\alpha = 0.05$ level of significance) while the crosses are the observed values for the years.

Test for regression of MeanMin on year was statistically not significant and is examined in detail in the section below.

4.3.4: Trends of maximum temperature for Makindu for different months

A summary of monthly MinMax, MeanMax and MaxMax temperatures was plotted for the years 1990 to 2010, figure 13. The Y axis shows the temperature in degrees centigrade while the X axis gives the years. For most months, there was some increase in the later years, despite the variability. December shows this clearly. The MeanMax can also be seen to increase over the years for June and September but it is not the case for MinMax and MaxMax of the same months.

Again a regression line was fitted where the maximum temperature was the dependent variable and year independent. This was for the years between 1992 and 2010. In this instance the variable 'months' was factored out. This means that the y-intercept and slopes were thus different for different months. This regression could best be summarised in the linear model equation:

$$\text{Estimate of Max temperature} = \text{Constant} + \text{year effect} + \text{month effect} + \text{year.month effect} + \text{error effect}$$

In the equation above, the variable month was set to be categorical while the one for year was left as numeric. This means that in the equation, the different months will have different y-intercepts and slope.

To do the regression analysis, 6834 days between January 1992 and December 2010 were used. The regression accounted for 39.889% of the total variation, see the calculation below. Most of this variation was however due to the difference in monthly temperatures (36%).

$$R^2 = \left(1 - \frac{SS_{residual}}{SS_{Total}}\right) \times 100 = \left(1 - \frac{19606.179}{32616.650}\right) \times 100 = 39.889\%$$

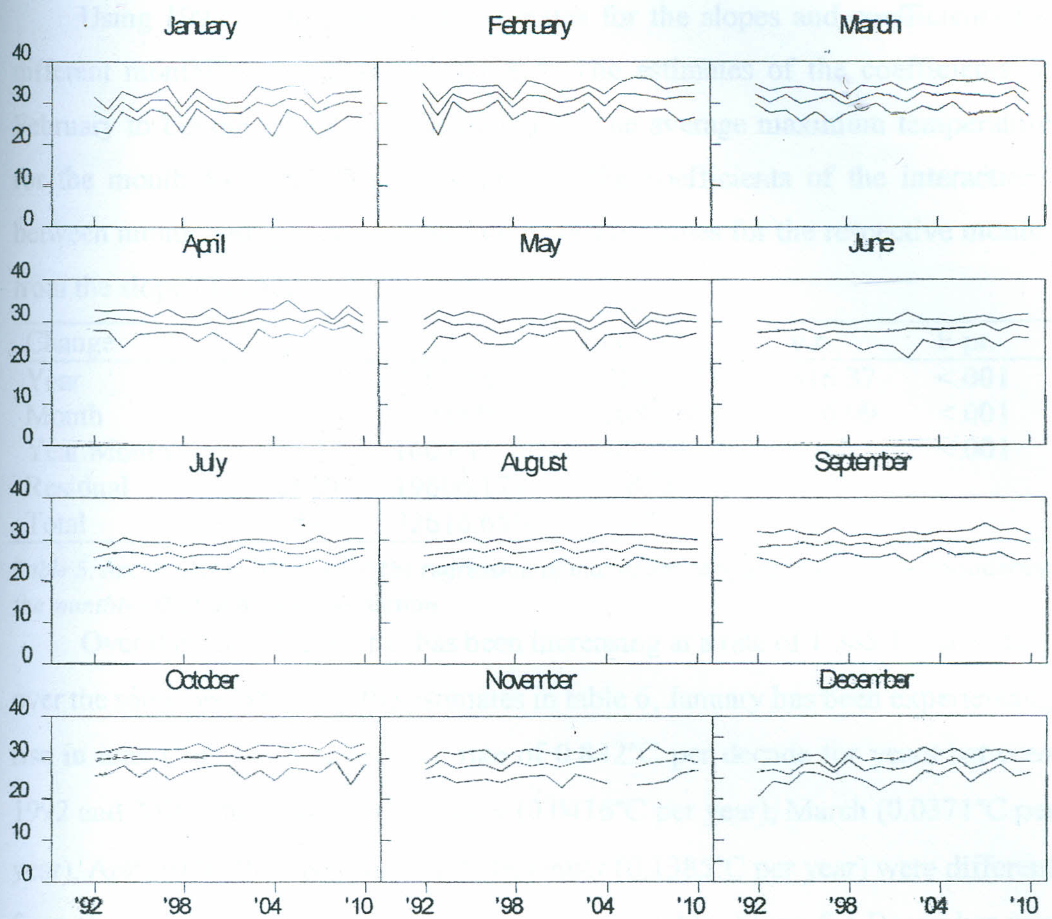


Figure 13: Trellis showing the behaviour of the minimum, mean and maximum of maximum daily temperatures (in degrees centigrade) over the years for the different months.

In order to test for slope for different months, the variable - year – was left as numeric while that for months was converted to nominal categorical. There was statistical significance for year, month and their interaction, table 5. The statistical significance for year implies that between 1992 and 2010, temperature for January had a slope. The statistical significance for month means that the y-intercept for the different months were different and the statistical significance for their interaction means that in that period, the different months had slopes that were different from that of January, table 5.

Using 1992 as the zero year, estimates for the slopes and coefficients for different months were calculated, table 5. The estimates of the coefficients of February to December gave the deviation for the average maximum temperature for the month for 1992. The estimates for the coefficients of the interactions between month and year are the deviations of the slopes for the respective month, from the slope of January.

Change	d.f.	s.s.	m.s.	v.r.	F pr.
Year	1	910.848	910.848	316.37	<.001
Month	11	11938.978	1085.362	376.99	<.001
Year.Month	11	160.645	14.604	5.07	<.001
Residual	6810	19606.179	2.879		
Total	6833	32616.650	4.773		

Table 5: Accumulated ANOVA for the regression of maximum temperature on year considering the monthly effect and their interaction

Over the years, December has been increasing at a rate of 1.385°C per decade over the same period. From the estimates in table 6, January has been experiencing rise in maximum temperature at a rate of 0.842°C per decade for years between 1992 and 2010.The slopes for February (0.0416°C per year), March (0.0371°C per year), April (0.0236°C per year) and December (0.1385°C per year) were different from that of January with statistical significance, table. Slope for December was higher than that of January while the rest had slightly lower slope.

When conducting the analysis, it was important to separate the rate at which temperature is increasing for different months. This was important to consider since the crops are grown during the rainy seasons only. As was expected, the rate of temperature increase for different months was different as shown in the paragraph above. This may be due to the differences in cloud cover in different months.

The above estimates are high when compared with the global trends shown in figure 10. Since only eighteen years of temperature data was used for this regression, it would be important to first consider using more data before generalising the slope of maximum temperature for Makindu.

Parameter	y-intercept	Slope	Deviation from Jan	s.e.	t(6810)	t pr.
January Intercept	28.418			0.134	211.31	<.001
February Intercept	30.736		2.318	0.195	11.91	<.001
March Intercept	30.984		2.566	0.190	13.49	<.001
April Intercept	29.839		1.421	0.192	7.40	<.001
May Intercept			0.212	0.190	1.12	0.264
June Intercept	27.168		-1.250	0.195	-6.41	<.001
July Intercept	26.339		-2.079	0.196	-10.62	<.001
August Intercept	26.901		-1.517	0.190	-7.98	<.001
September Intercept			0.172	0.192	0.89	0.371
October Intercept	29.744		1.326	0.190	6.97	<.001
November Intercept	27.933		-0.485	0.192	-2.53	0.012
December Intercept	26.934		-1.484	0.190	-7.80	<.001
January slope		0.0831		0.0128	6.51	<.001
February Slope		0.0416	-0.0415	0.0185	-2.24	0.025
March Slope		0.0371	-0.0460	0.0181	-2.54	0.011
April Slope		0.0236	-0.0595	0.0182	-3.27	0.001
May Slope			-0.0099	0.0183	-0.54	0.586
June Slope			-0.0023	0.0184	-0.12	0.901
July Slope			-0.0135	0.0184	-0.73	0.464
August Slope			-0.0181	0.0181	-1.00	0.316
September Slope			-0.0113	0.0182	-0.62	0.537
October Slope			-0.0235	0.0181	-1.30	0.193
November Slope			-0.0025	0.0184	-0.14	0.891
December Slope		0.1385	0.0554	0.0181	3.07	0.002

Table 6 shows the estimates of the regression parameters for monthly maximum temperatures for the period between January 1992 and December 2010. The shaded rows are for coefficient estimates that have statistical significance

To get the output in table 6, GenStat automatically calculated the estimates using a base month - which was January. The other estimates were calculated as deviation from January. The standard error, t-statistic and the p-value were calculated for the deviation from January. To get the intercepts for the other months, their estimates were added to that of January. The same was done with the slopes for the other months. In table 6 only the intercepts and slopes which had a statistical significance deviation from that of January have been included and highlighted.

In order to get the estimated mean temperature for a certain month in a given year between 1992 and 2010, the equation should include the base value for January, and the deviation arising due to the difference in months and the slope for that month. From the model, the estimate maximum temperature for December 2010 would be 24.559°C, calculated by:

{***28.418***+ ***0.0831****(2010-1992) - ***1.484****11+***0.0554****(2010-1992)*11} °C.

This regression equation is for months of years between January 1992 and December 2010.

4.3.5: Trends of minimum temperature for Makindu for different months

In figure 14 the y-axis gives the temperature in degrees centigrade while the x-axis gives the years from 1992 to 2010. Plotted are the MinMin, MedMin and MaxMin summary values for each month over the years. July and August were the coolest months.

There was no minimum temperature data for November in 2005 and in May 2006, only three values were recorded. These values were 19.1°C, 19.6°C and 18.6°C. This explains why there is a break in line in figure 14.

The ANOVA in table 7 is again testing the regression equation:

$$\text{Estimate of Min temperature} = \text{Constant} + \text{year effect} + \text{month effect} + \text{year.month effect} + \text{error}$$

Though the term 'year effect' has been used in the model above (and the model for the estimate of maximum temperatures), year was not considered as a cause for a unit change in temperature. There are other causes which lead to this, a common one is the greenhouse gas emissions and other land management practices which have short term and long term effects on the atmosphere.

In the equation above, month was a nominal categorical variable. This means that there were different y-intercept and slopes for the months. The ANOVA in table 7 shows that year, month and their interaction had statistical significance. This means that there was a slope over the years, that the different months had different values for the base year (1992) and that their slopes were different over the years. To get this, data from 6856 days starting January 1st 1992 were used. They explained 54.634% of the variation as explained below.

$$R^2 = \left(1 - \frac{SS_{residual}}{SS_{Total}}\right) \times 100 = \left(1 - \frac{13492.844}{29742.012}\right) \times 100 = 54.634\%$$

Change	d.f.	s.s.	m.s.	v.r.	F pr.
Year	1	81.138	81.138	41.07	<.001
Month	11	16109.185	1464.471	741.20	<.001
Year.Month	11	58.844	5.349	2.71	0.002
Residual	6829	13492.844	1.976		
Total	6852	29742.012	4.341		

Table 7: Accumulated ANOVA for the regression of minimum temperature on year considering the monthly effect and their interaction.

Parameter	y-intercept	slope	Deviation from Jan	s.e.	t(6829)	t pr.
January Intercept	17.747			0.111	159.29	<.001
February Intercept			0.131	0.161	0.81	0.417
March Intercept	18.769		1.022	0.158	6.49	<.001
April Intercept	18.761		1.014	0.159	6.38	<.001
May Intercept	17.327		-0.420	0.158	-2.66	0.008
June Intercept	15.713		-2.034	0.160	-12.75	<.001
July Intercept	14.551		-3.196	0.158	-20.20	<.001
August Intercept	14.832		-2.915	0.158	-18.50	<.001
September Intercept	15.828		-1.919	0.159	-12.08	<.001
October Intercept	17.254		-0.493	0.158	-3.12	0.002
November Intercept	18.161		0.414	0.159	2.61	0.009
December Intercept	18.21		0.463	0.158	2.94	0.003
January slope		0.0336		0.0106	3.17	0.002
February Slope			0.0130	0.0153	0.85	0.396
March Slope			0.0071	0.0150	0.48	0.634
April Slope			-0.0258	0.0151	-1.71	0.087
May Slope			-0.0132	0.0151	-0.87	0.383
June Slope		0.0004	-0.0332	0.0151	-2.20	0.028
July Slope		-0.0105	-0.0441	0.0150	-2.94	0.003
August Slope			-0.0156	0.0150	-1.04	0.296
September Slope			-0.0210	0.0151	-1.39	0.163
October Slope			-0.0134	0.0150	-0.89	0.372
November Slope			0.0115	0.0152	0.76	0.448
December Slope			-0.0101	0.0150	-0.67	0.501

Table 8: Estimates of regression parameters for monthly minimum temperatures when data is grouped by month for the period between January 1992 and December 2010

Between 1992 and 2010, the minimum temperature for January has been increasing at a rate of 0.0336°C (0.012824°C, 0.054376°C) every year (see table 8). Even though the accumulated ANOVA showed that the interaction of months and year had statistical significance, only two months had slopes that were different from that of January, table 8.

In 1992, only February did not have different temperature from January. This could be attributed to its closeness to January (see third row in table 8, the p-value is 0.417).

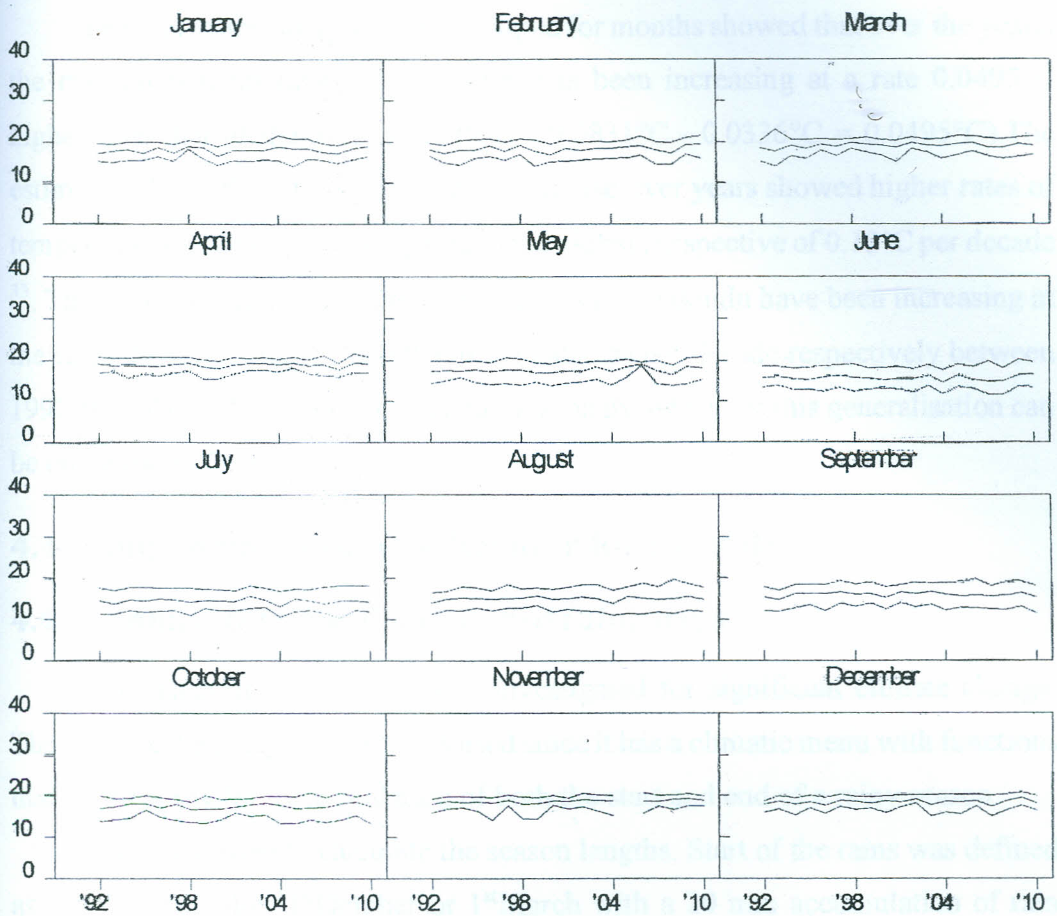


Figure 14: Trellis showing the behaviour of the minimum, mean and maximum of daily minimum temperatures (in degrees centigrade) over the years for the different months.

4.3.6: Summary on temperature behaviour

From the Makindu temperature data between 1992 and 2010, there was statistical evidence to indicate increasing trends for both minimum and maximum temperature. This rate of increase differed for the different months, but most of them did not have statistical significance. However, a lot of variability remained unexplained in both regressions for maximum (39.889%) and minimum (54.634%) and temperatures.

The regression analysis when grouped for months showed that over the years, the maximum temperatures for January has been increasing at a rate 0.0495°C higher than that of the minimum (that is 0.0831°C - 0.0336°C = 0.0495°C).The estimates of coefficients of temperature increase over years showed higher rates of temperature increase when compared to the global perspective of 0.13°C per decade [33]. The maximum and minimum temperatures for Makindu have been increasing at the rates 0.831°C (table 6) and 0.336°C (Table 8) per decade respectively between 1992 and 2010. More data would need to be analysed before this generalisation can be made for all Makindu data.

4.4: Long term Climate behaviour for rainfall

4.4.1: Annual totals and number of rainy days

The seasonal rainfall was also investigated for significant climate change. The statistical package Instat[29] was used since it has a climatic menu with functions that could allow for the calculation of both the start and end of a rainy season.

Instat was used to calculate the season lengths. Start of the rains was defined as the first day after 1^{st}October or 1^{st}March with a 20 mm accumulation of rain water in at most two days, for the short and long rains respectively. To get the end of the rains, the water balance function was used. These were the earliest days after 1^{st} February and 1^{st} June which had less than 0.05 mm of water for the short and long rains respectively.

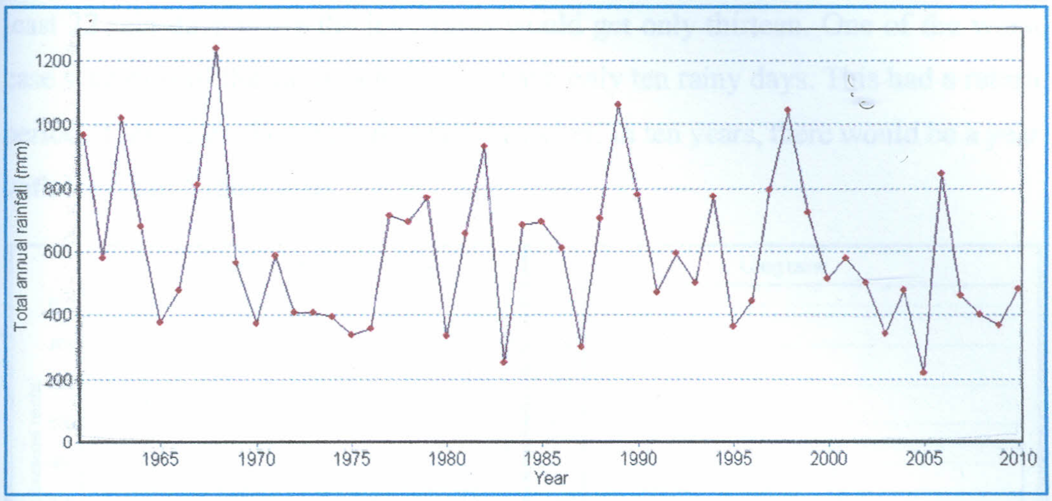


Figure 15: Total annual rainfall for years between October 1961 and June 2010

This was from a total of 1852 rainy days. Figure 15 distinguishes the total rainfall for long and short rains respectively. In the long rains of the years 1972, 1993 and 2009, the total rainfall recorded were 22 mm, 17.9 mm and 18.6 mm respectively (figures 18, 19 and 20). This rainfall was distributed over the season and no two consecutive rain days totalled 20 mm to which was the start of the season. Because of this, the total rainfall was calculated from 1st March of the three years. The Long rains for 1972 had only two rain days whose total gave the 17.9 mm while the long rains of 2009 had three rain days. In 1984, the total rain days for the Long rains were only four; however, the total rainfall recorded for that season was 77.5 mm.

The short rains received a mean of 351.0 mm with a standard deviation of 198.5 mm while the long rains on average received 170.8 mm of rainfall with a standard deviation of 101.9mm.No clear trend of either increase or decrease was evident for the annual and the seasonal rainfall. There was also a lot of variation in the number of rainy days. The mean number of rain days for the short and long rains were 23.92 and 13.44 respectively. The two had standard deviations of 10.03 and 6.437 respectively. In one out of every two years, the short rains would get at least 22 rain days while the long rains would get only thirteen. One of the worst case scenarios of the short rains was to have only ten rainy days. This had a return period of ten years. For the long rains, however, in ten years, there would be a year with less than 5 rainy days.

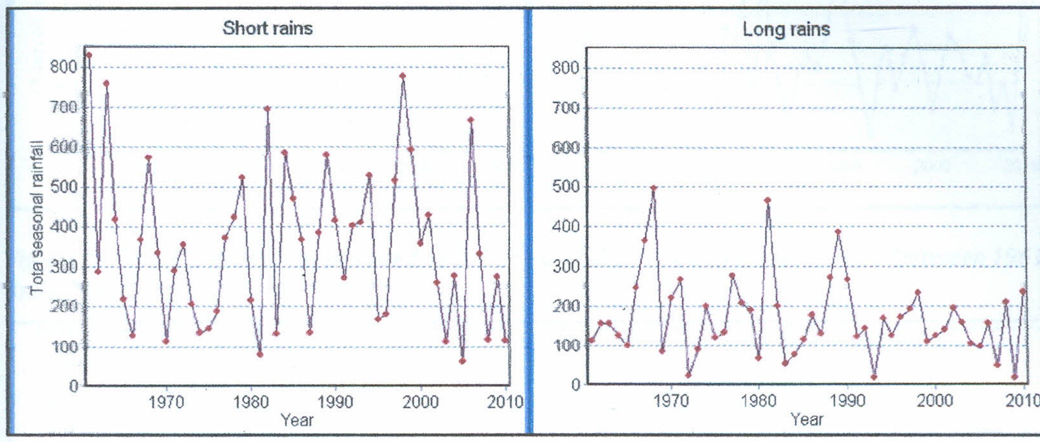


Figure 16: Total rainfall for the long and short rains for the years between 1961 and 2010

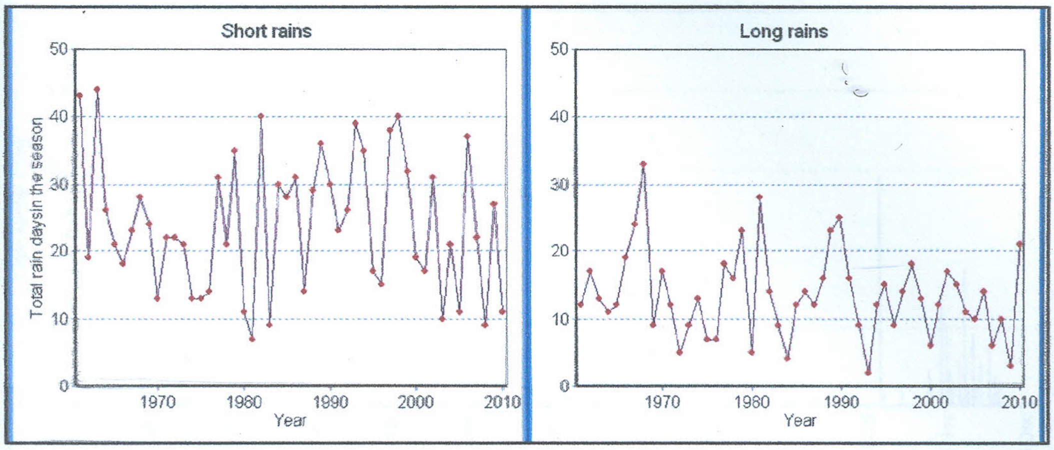


Figure 17: Total rain days experienced in the long and short rains periods for years between 1961 and 2010

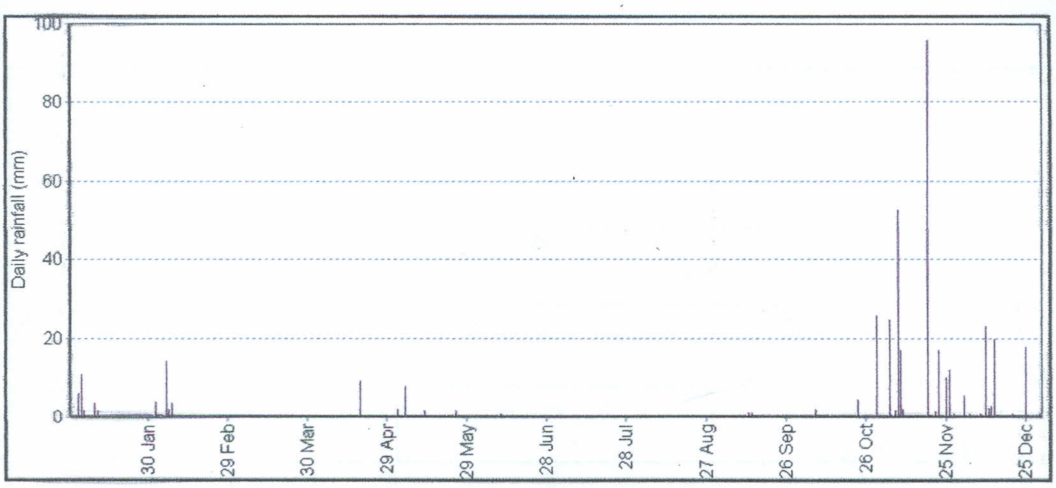


Figure 18: Recorded rainfall for 1972.

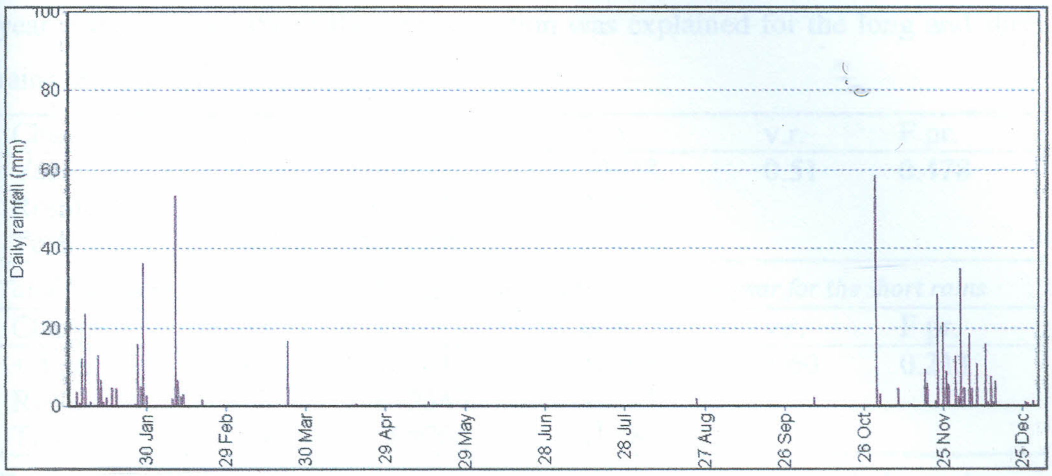


Figure 19: Recorded rainfall for 1993.

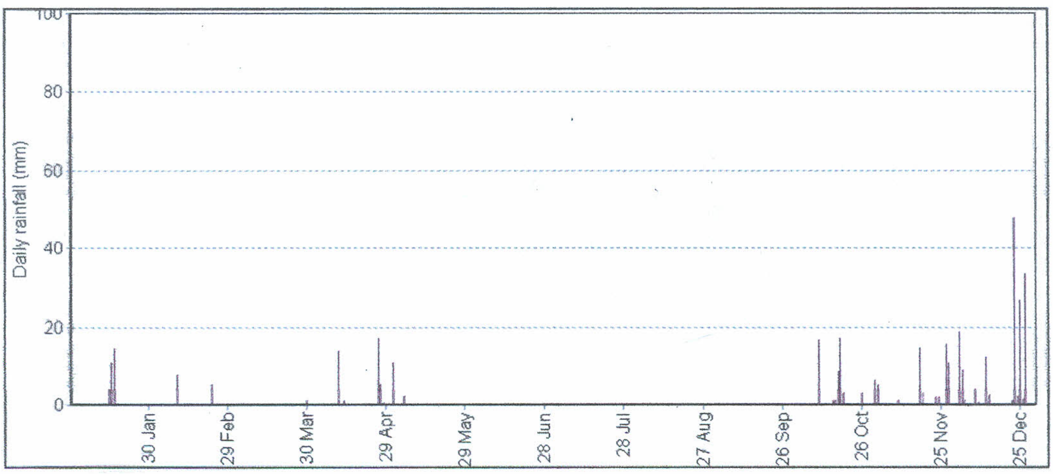


Figure 20: Recorded rainfall for 2009

Regression analysis was done for the short and long rains separately to check for statistically significant slope, tables 9 and 10 respectively. The data used were the rainfall totals for the period between 1961 and 2010. The p-values in tables 9 and 10 were greater than 0.05. This means that the slope over the years was not statistically significant at the 0.05 level. By regressing total seasonal rainfall on year, only 1% and 3% of the total variation was explained for the long and short rains respectively.

Change	d.f.	s.s.	m.s.	v.r.	F pr.
Year	1	20338.	20338.	0.51	0.478
Residual	48	1910251.	39797.		
Total	49	1930589.	39400.		

Table 9: Accumulated ANOVA for the regression of total rainfall on year for the short rains

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ Year	1	16361.	16361.	1.60	0.213
Residual	48	492345.	10257.		
Total	49	508706.	10382.		

Table 10: Accumulated ANOVA for the regression of total rainfall on year for the long rains

4.4.2: Behaviour of rainfall totals over the years for different months

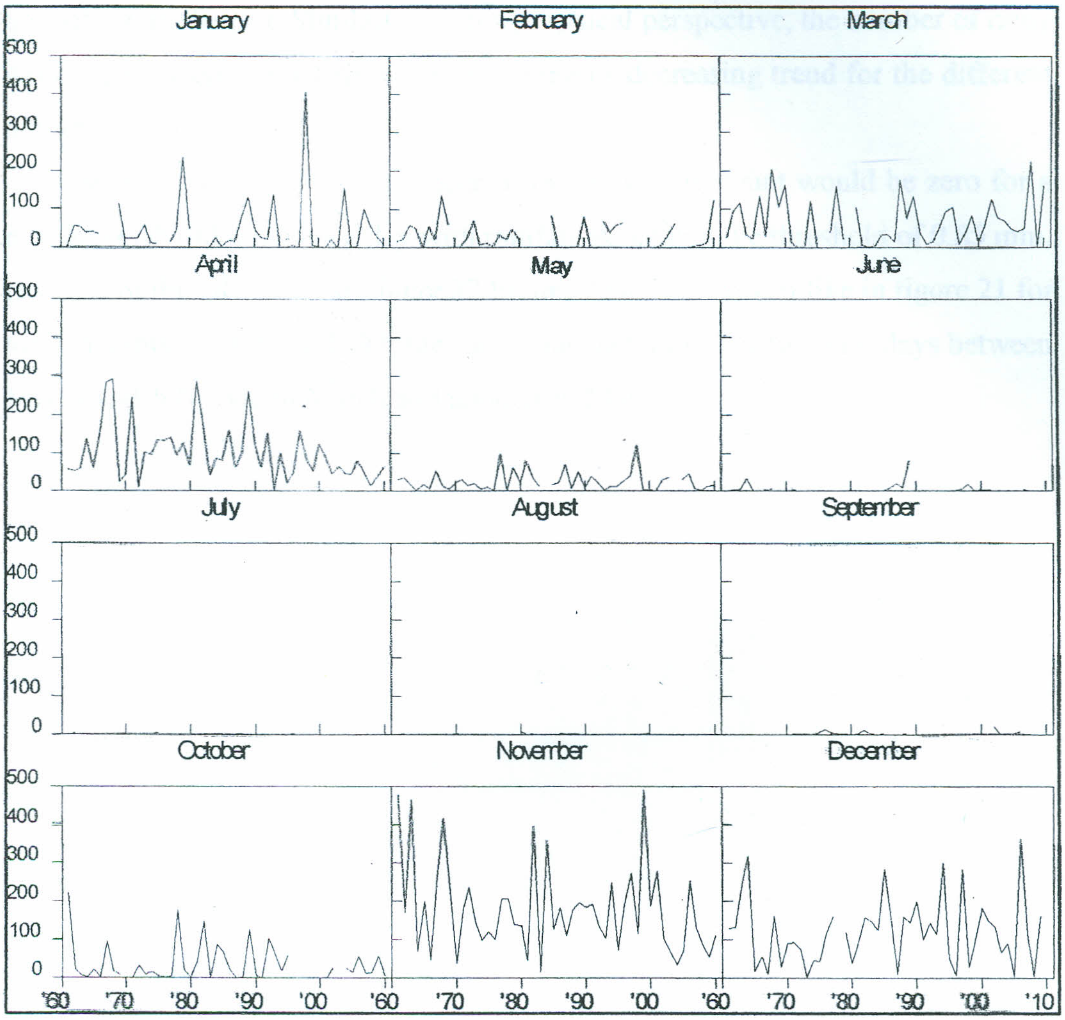


Figure 21: Trellis plot for the total monthly rainfall over the years. The rain data was available for the period between 1961 and 2010 (unlike temperature data).

Analysis was also conducted to see how rainfall totals has been varying over the years. Neither oscillations nor slope could visibly figure 21. Variability in April appeared to have decreased over the years since mid-90s; this is arrived at by looking at the line graph in figure 21. The lines towards the end show there is no sudden high increase and sudden decrease in the values. They are all low. The effect of the 1997/1998 El Niño on the January rainfall for 1998 distinguishes its high quantity from the rest. Similarly, from a graphical perspective, the number of rainy days experienced did not show an increasing or decreasing trend for the different months, figure 22.

When summarising the number of rain days, the count would be zero for a month which did not have a day with rainfall exceeding the threshold of 0.85 mm. Because of this, the lines in Figure 22 Figure 22are not broken like in figure 21 for total monthly rainfall. In 1993, the long rains had a total of two rain days between March and July; one in March and another in May.

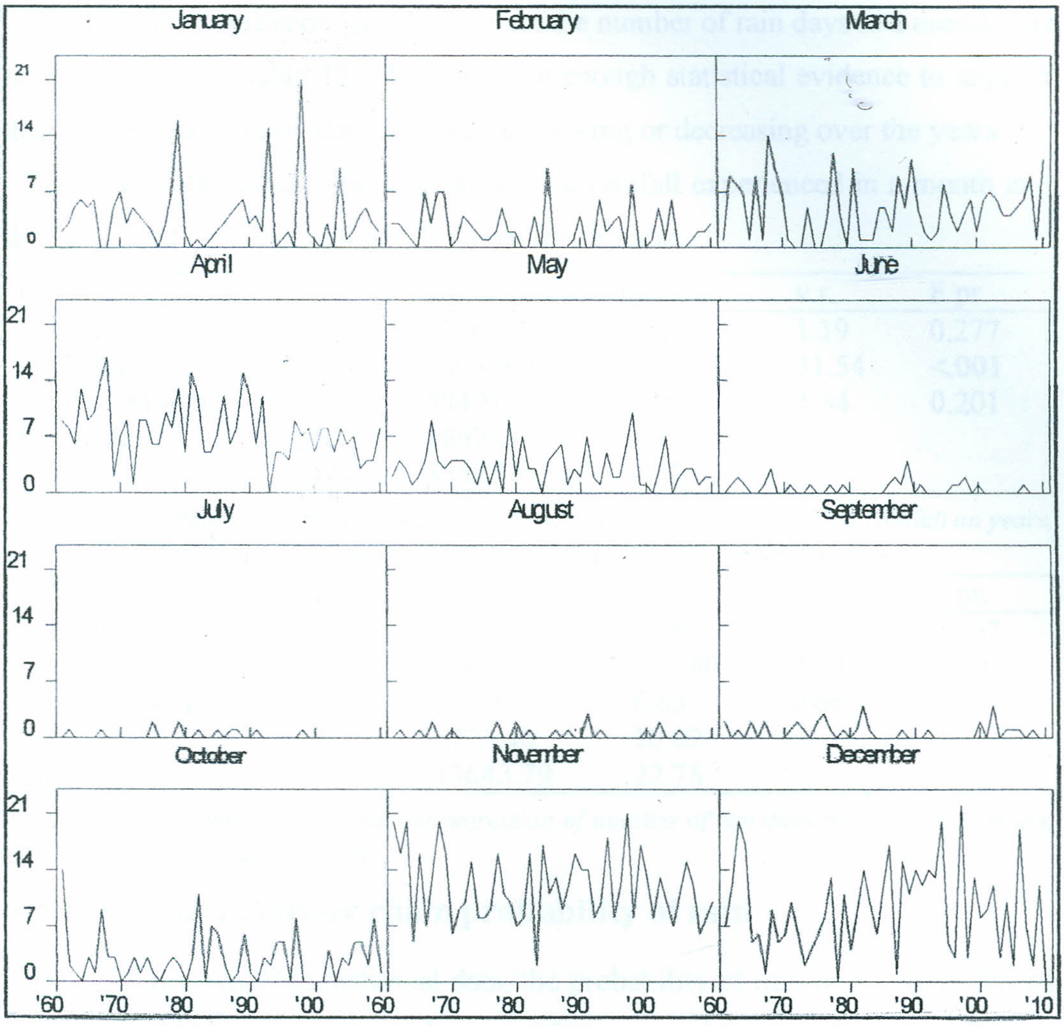


Figure 22: Trellis plot for the total rain days for different months over the years

A linear regression equation for the total monthly rainfall was again fitted and tested for statistical significance. The independent variable was year, table 11. This was done from a total of 486 months that received rainfall between the January 1961 and December 2010. The p-value for year as an explanatory variable was greater than 0.05. This means that slope was not statistically significance at the 0.05 level of significance. Nearly half of the variation (42%) was attributed to month.

A similar regression was done where the number of rain days in a month was regressed on year, table 12. There was not enough statistical evidence to suggest that the number of rainy days has been increasing or decreasing over the years.

From both outputs, the total amount of rainfall experienced in a month and the number of rain days differed between the months.

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ Years	1	4796.	4796.	1.19	0.277
+ Months	11	1403890.	127626.	31.54	<.001
+ Years.Months	11	59497.	5409.	1.34	0.201
Residual	462	1869394.	4046.		
Total	485	3337578.	6882.		

Table 11: This table shows the accumulated ANOVA for the regression of total rainfall on years, grouped by month. This regression equation explains 41.2% of the total variation

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ Years	1	33.65	33.65	3.36	0.067
+ Months	11	7763.81	705.80	70.42	<.001
+ Years.Months	11	72.93	6.63	0.66	0.775
Residual	576	5773.40	10.02		
Total	599	13643.79	22.78		

Table 12: Accumulated ANOVA for the regression of number of ran days in a month on years, grouped by month. The R2 is 56%

4.4.3: Fitting a Markov chain probability of rain

With the available historical data, the probability of rain on a certain day of the year could be calculated. Markov models were used for this. Only the presence of rainfall was considered, not its amount when it rained. From the probabilities, the bimodal pattern was evident.

In figure 23a the probabilities of rain have been plotted while in figure 23b the smoothed probabilities are shown. The x-axis gives the day of the year from 1st January while the y-axis gives the probability between 0 and 1. The middle dash-dot line plot in figure 23a gives the probability of raining without considering memory from the preceding day. The other two line plots consider if it rained before (p_rr) or it didn't (p_rd). The same line scheme has been used for the smoothed Markov plots in figure 23b.

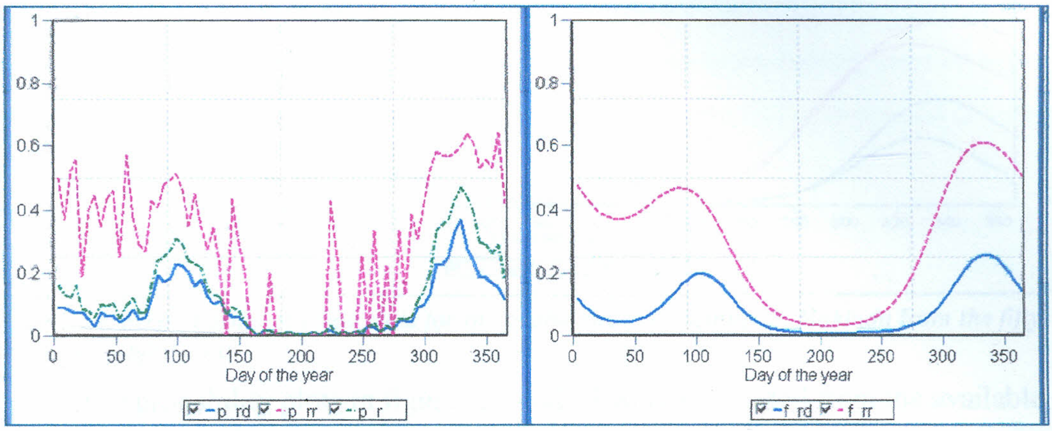


Figure 23: First order Markov plot of the probabilities or raining. The top line (f_rr) shows the chance of raining today when it had rained the previous day while the bottom one shows the probability of raining today when the previous day was dry

Only a few days in November experienced rainfall more than 50% of the times when there was rainfall the previous day. To check if there was that much difference if a two day memory was considered, a second order Markov model for the probabilities was produced, figure 24. The top broken line gives the probability of raining given it rained in the past two days (p_rrr). The top continuous gives the probability of raining given that it rained yesterday but not the day before (p_rrd). The lower continuous line gives the probability of raining given it did not rain yesterday but rained the day before (p_rdr) and the lower broken line gives the probability of raining given that it did not rain in the past two days (p_rdd).

There was not much difference in the probabilities during the dry parts of the year, which is between 1st July and 31st August. This memory had more effect on days in the long rains. For instance, more than 50% of the years that had two rainy days on any day between January 1st and 10th experienced rainfall on that day.

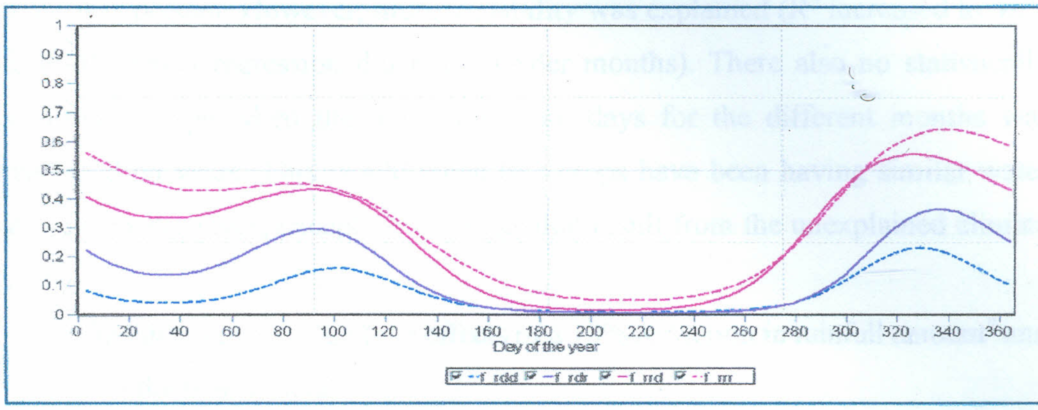


Figure 24: Second order Markov plots for the probabilities or raining in Makindu from the fifty years of rainfall data

The probability plots in figures 23 and 24 were calculated from the available fifty years of daily rainfall data for Makindu. These probabilities can be used to simulate a dataset for the occurrence of rainfall. By using Markov probability models, the simulations would be stochastic, but would not be representative of a climate (rainfall) change scenario. This is because the probabilities have been derived from what has been experienced historically. More parameters would be required in order to simulate the rainfall amount, for instance, a similar curve for rainfall amounts can be used.

One way that climate change can be simulated is by moving the probability curve either up or down, then simulating rain occurrence in a year using the new probabilities. However, this thesis only looked at the deterministic methods of adjusting rainfall for climate change scenarios and analysed their effect on the crop yield.

4.4.4: Conclusion on rainfall

A lot of variation is evident in the Makindu rainfall data. Regression analysis did not show any evidence of increasing or reducing rainfall totals. When the total monthly rainfall was regressed on year, for the different months, no significant trend was present. However, more variability was explained (R^2 increased to 42% from 4% when regression did not consider months). There also no statistically significant slope when the number of rain days for the different months was regressed on years. This would mean that crops have been having similar water stress over the years, except for the risks that result from the unexplained climate variability.

This thesis investigated the effects of a 10% reduction in rainfall amount[9] and pattern on the maize yield.

4.5: Filling for missing data

4.5.1: Missing data

The crop model, APSIM [7], require daily climate data for it to simulate. As table 1 showed a number of years had missing temperature and radiation data. At least thirty years data was required since this project focused on climate, and the data needed to be daily in order to use the crop simulation model, APSIM. Only twenty years had complete temperature data, therefore a specialised software, Weatherman[6], was used to fill the missing data.

4.5.2: Filling Makindu missing data

Data from Makindu was corrected using the observed means and variances. No correction was done for rainfall since the recorded values covered the entire period, 1961 to 2010. This correction did not affect the days that had recorded data. The total days for maximum and minimum daily temperatures that were corrected were 58.6% and 60.5% respectively – the Kenya Meteorological Department was initially not collecting the temperature data until recently for Makindu. The percentage of corrected radiation data was 88.7% of the available 18262 days. The

With the corrections made, fifty years of daily climate data were available to be used as APSIM inputs. Climate change scenarios were created using GenStat for different rainfall change scenarios as discussed in the next section.

4.6: Creating the climate change scenarios

GenStat syntax was developed to create the climate change scenarios by adjusting the number of rain days and spells, 0. To reduce 10%, every tenth rain day was removed respectively for the entire fifty years of Makindu data. The count was not restarted each year. It continued from where it stopped the previous year.

There were at least ten different to ways in which 10% of rainy days or spells could be removed deterministically. This would depend on whether one started the count at 1, 2…9. Because of this, the simulation syntax was set to loop creating ten different climate change scenarios. Using the uniform distribution, one of the ten climate change scenarios was selected at random. This selection was done using uniform distribution since it was unbiased; each of the ten scenarios had 0.1 chance of being selected. This procedure was done for both the climate change scenario affecting the rain days and for the one affecting rain spells.

Climate change Scenario	Software used to simulate
10% reduction in rainfall amount	APSIM during crop simulation
10% reduction in rain days	GenStat
10% reduction in rain spells	GenStat

Table 14: Software used to create different climate change scenarios.

In order to remove the tenth rain day or tenth rain spell, the recorded rainfall for the selected days or spells were edited to read 0 mm.

4.6.1: Effect of 10% change in distribution on rainfall amounts

The ANOVA conducted in table 15 was done to test how much the climate change scenarios that were simulated explained in the differences in rainfall amounts. This was also important since it would be a measure of whether there was statistical significance in the climate data. The season, year, and month are other factors that may affect the cumulative rainfall experienced in a place. For this reason, they were included in the ANOVA table. All these factors had p-values that were less than 0.05 which meant they were statistically significant at the 0.05 level of significance. The ANOVA table was not used to indicate causation, but just to explain why the rainfall amounts were different. For instance, it is not uncommon for January and February rainfalls to differ, but it does not mean that February is the cause.

Reducing 10% of the rain days resulted to10.32% reduction in total rainfall for the fifty years of data. The reduction in the total annual rainfall ranged between 2% and, on a worst case scenario, 28% (experienced in 1995). In one out of ten years the reduction was 16.62% whereas the median reduction was 9.985%.

Overall, by removing 10% of the rain spells, the fifty years of data had 10.83% less rainfall amount. The minimum percentage reduction in a year was 0.55% of total annual rainfall and in one out of every ten years had an 18% reduction. The maximum reduction experienced was a 44.42% reduction (in 1973). The median reduction experienced in a year was 9.01%.

The variation in rainfall was tested for different sources, table 15. The test indicated strong evidence for differing rainfall according to years, months and Seasons. Season was defined as a combination of year and month. The grand mean for the rainfall received in a day was 1.534 mm. This calculation included only rain days which recorded 0.85 mm and above. The three climate scenarios had statistically significant differences in this mean rainfall. The mean daily rainfall under current climate, 10% less rainfall amount, 10% less rain days and 10% less rain spells were 1.664 mm, 1.498mm, 1.492mm and 1.484mm respectively.

The complete table containing the reduction of total annual rainfall after removing 10% rain days and spells have been tabulated in 0.

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Climate	3	410.77	136.92	3.33	0.019
Year	49	24490.49	499.81	12.16	<.001
Month	11	192212.01	17473.82	424.97	<.001
Seasons	599	201669.86	336.68	8.19	<.001
Residual	72385	2976312.13	41.12		
Total	73047	3395095.26			

Table 15: ANOVA testing for the different sources of variation for rainfall

4.6.2: Effect of the 10% reduction of rainfall on start of the rain seasons

Reducing 10% of the rainfall affected the start of the rainfall spells. This effect was different for the different ways rainfall was changed. Tables 16 and 17 show summaries of the starts of the short and long rains respectively. The start of rains was defined as the first day after 1st October or 1st March that received a minimum of 20 mm of rainfall over a period of two days for short and long rains respectively.

The asterisks (*) are indicative of a season that did not meet this criterion and therefore dates for the next season was selected. The next earliest or latest date that fell under the season was taken as representative.

	Earliest	Lower Quartile	Median	Upper Quartile	Latest
Short Rains (Current)	1st Oct	17th Oct	29th Oct	5th Nov	20th Nov
Short Rains (-10% amount)	4th Oct	31st Oct	7th Nov	14th Nov	16th Dec
Short Rains (-10% days)	4th Oct*	31st Oct	7th Nov	14th Nov	16th Dec
Short Rains (-10% spells)	13th Oct*	30th Oct	7th Nov	14th Nov	16th Dec

Table 16: The differences in the start of short rains under current climate ad when rainfall is reduced by 10%

Under current climate, there was no season that did not meet the criterion for a start of the season. Despite short rains for 1973 receiving 73.4 mm with 10% less rain spells, there were no two days that accumulated 20 mm rainfall.

As can also be seen from the summary of the start of short season (table 16), one in four years resulted in the date of planting being pushed to December. This leads to shortening of the length of growing period which would affect the biomass development and eventual maize harvest. One extreme start was of the short season was in the 2008/2009 season. The start of rains was delayed till 17th January 2009,andwas followed by only 13.4 mm received in two day before 1st March. However, the cumulative rainfall for that season was 87.7 mm which could have been used to result in a harvest.

	Earliest	Lower Quartile	Median	Upper Quartile	Latest
Long Rains (Current)	1st Mar	3rd March	19th Mar	25th Mar	23rd Apr*
Long Rains (-10% amount)	1st Mar	19th Mar	26th Mar	10th Apr	27th Apr*
Long Rains (-10% days)	1st Mar	20th Mar	27th Mar	7th Apr	27th Apr*
Long Rains (-10% spells)	1st Mar	19th Mar	26th Mar	7th Apr	28th Apr*

Table 17: The differences in the start of long rains under current climate ad when rainfall is reduced by 10%

The long rains for 1972, 1983 and 1993 did not meet criteria for the start of rains under current climate and also under climate change. Though some days received rainfall above 10 mm rainfall, a one day break in between resulted to the start of the rains criterion not being met.

4.6.3: Effect of the 10% reduction of rainfall on Markov probabilities

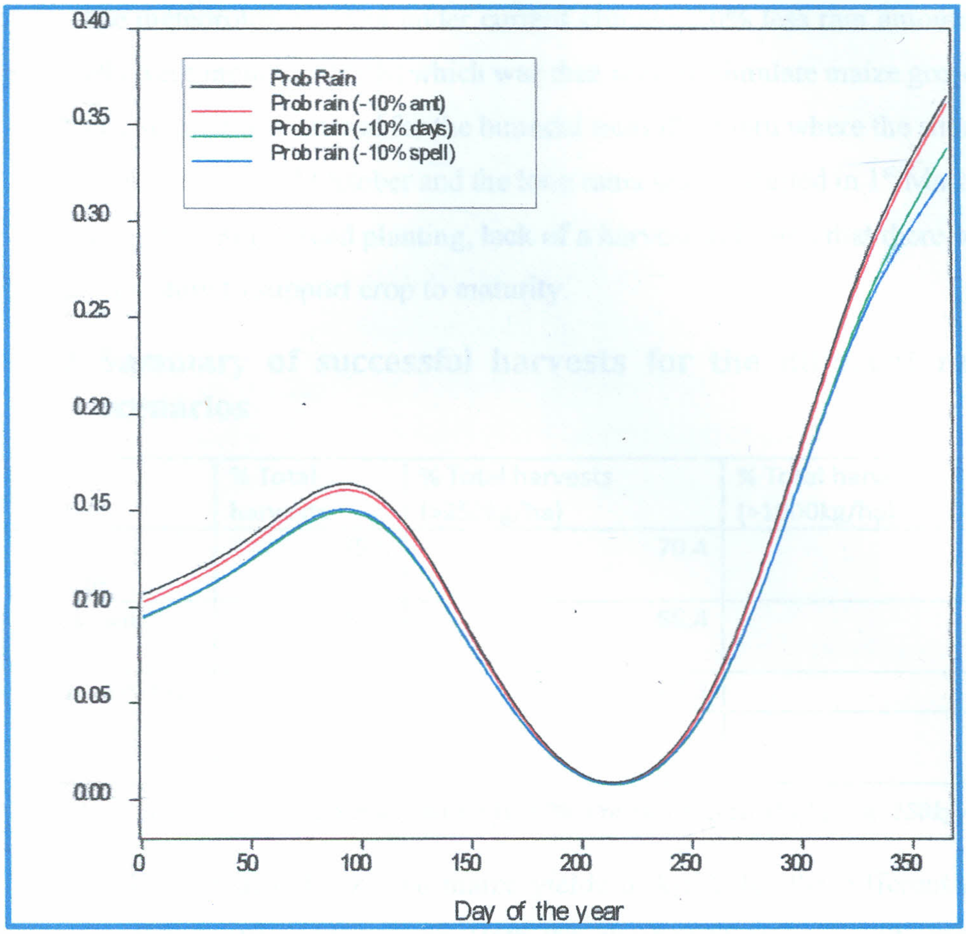


Figure 25: The effect of the three changes on the probability of rain occurring on a given day of the year

A Markov model of the probability of rainfall occuring on a given date in the year was also calculated from the years for the differentclimate change scenarios. The probabilities are plotted in figure 25. As the figure shows, the probabilities of raining under current climate and when only the amount had been changed were very close. However, affecting the pattern (days and spell) resulted in reduction of these probabilities, especially during the rainy season.

4.7: Results from simulation of crop yield under the different climates

The meteorological data under current climate, 10% less rain amount, days and spells were input to APSIM which was then used to simulate maize growth and yield. The simulations catered for the bimodal rainfall pattern where the short rains season was started in 1st October and the long rains season started in 1st March each year. Since the setup forced planting, lack of a harvest indicated that there was not enough moisture to support crop to maturity.

4.7.1: Summary of successful harvests for the different rainfall scenarios

	% Total harvests	% Total harvests (>250kg/ha)	% Total harvests (>1000kg/ha)
Current climate	75.5	70.4	52.0
- 10% rain amount	74.5	68.4	45.9
- 10% rain days	70.4	67.3	48.0
- 10% rain spells	71.4	66.3	49.0

Table 18: Percentages of successful harvests with the thresholds of 0kg/ha, 250kg/ha and 1000kg/ha

APSIM simulated seasonal maize yields in kg/ha for the different rainfall scenarios for ninety eight seasons. Both the number of seasons with successful harvest and the amount variation in yield amount were analysed. The percentage of successful harvests for three different thresholds has been given in table 18. The average yields were 1180.47Kg/ha, 1059.25Kg/ha, 1079.9Kg/ha and 1067.6Kg/ha for the current rainfall and 10% reduction in rainfall amount, rain days and rain spells respectively. This section however discusses more about counts of successful harvests.

In the short rains seasons for 1974, 1987 and 1998, there was 10% more yield under a 10% reduction in rainfall amount and days than under current rainfall. This was 3% of the seasons. Yields for the short rains of 1974 and 1998 also were 10% more when10% of rain spells had been removed when compared to yield in the current climate. In 1974, the rainfall recorded was 179mm which reduced by 11.5% and 16.9% when 10% of rainy days and spells were removed respectively.

Even though the 1987 short rains, 200.3mm of rainfall was recorded; the maize failed under current climate and with a 10% reduction in rain spells. The reduction in rain days and spells reduced the rainfall amount by 24.3% and 6.5% respectively. Under current climate and a 10% reduction in spells did not affect the record of November 5th 1987, which was 42mm of rainfall followed by a nine day dry spell that could result to crop failure. By removing the days, this planting date was shifted to November 22nd which ended with a successful harvest since no immediate long dry spell was present.

One of the objectives for this thesis was to determine whether significant differences occur when rain days are reduced by adjusting the distribution and not just the amount, albeit with the same percentage. A season was considered to have a different yield if it had 10% more or less yield under either 10% reduction in rain days or spells, when compared to rain amount. More than one in three seasons had different yields when compared to a 10% decrease in rainfall amount. This was 32% and 44% of seasons under 10% less rain days or spells respectively.

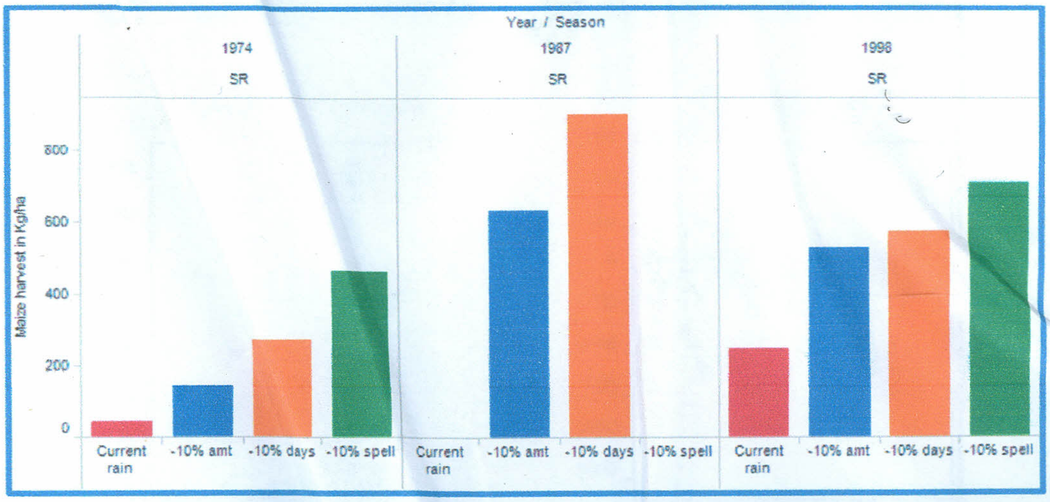


Figure 26: The three seasons in which the yield harvested was 10% higher in under rainfall change when compared to under current climate. Plotting of all seasons resulted in a very crowded graph hence the seasons in which the current climate did had more yield were excluded.

4.7.2: Testing for the different sources of variation in yields

A number of sources explained the variation that existed in the yields. They included the season, year and also the climate scenario present. The crop management with regard to fertilizer use were left as default and were not changed in order to contain any variation that they might cause and remain with the three mentioned.

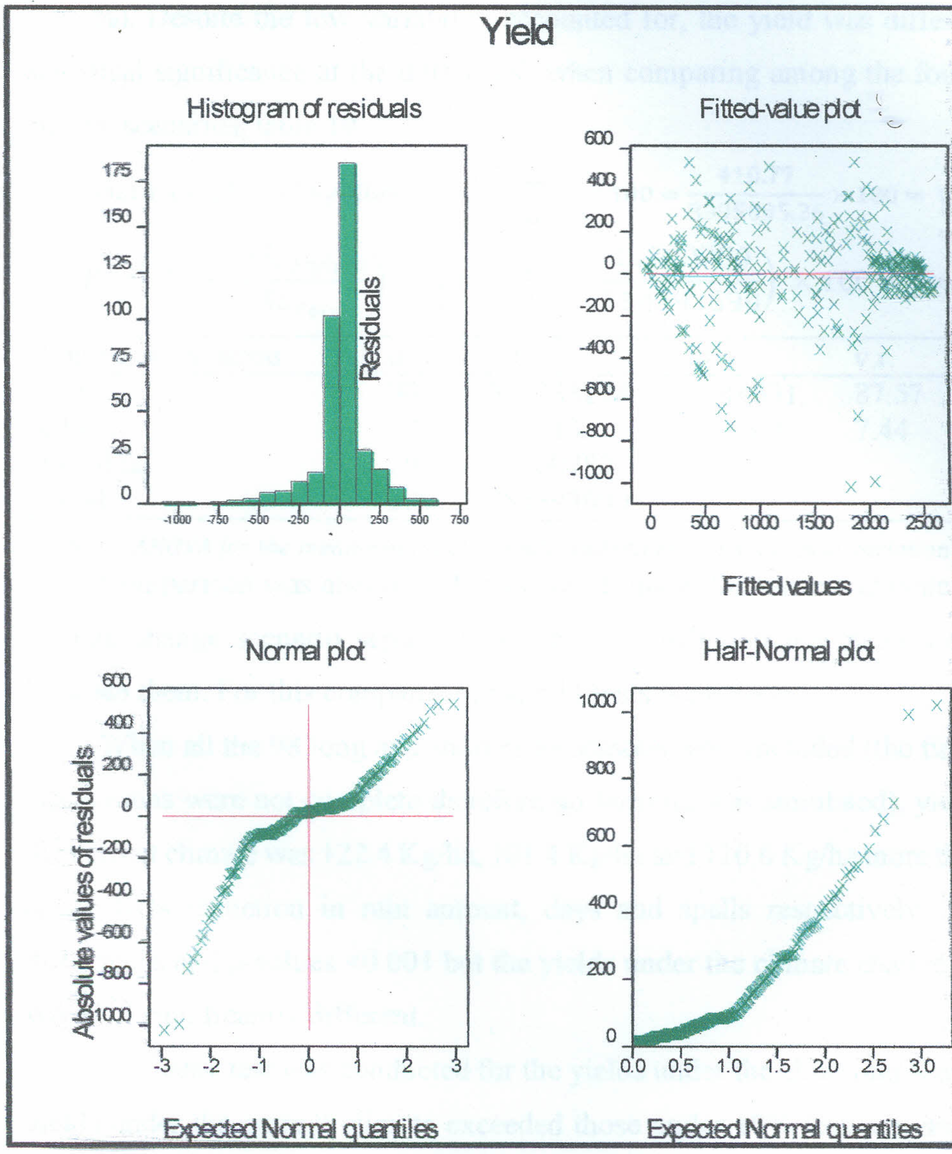


Figure 27: Residual plots to test the normality assumption for the Maize outputs

An ANOVA was conducted to test the different sources of variation in yields, table 15. To test the normality assumptions, a residual plot was also produced for visual analysis, figure 27. The histogram of the residuals had a normal distribution pattern. Climate explained only 0.12% of the total variation explained by the model

(96.7%). Despite the low variability accounted for, the yield was different (with statistical significance at the 0.05 level) when comparing among the four climate change scenarios, table 19.

$$Variation\ explained\ by\ climate = \frac{SS_{Climate}}{SS_{Total}} \times 100 = \frac{410.77}{3395095.26} \times 100 = 0.012\%$$

$$R^2 = \left(1 - \frac{SS_{residual}}{SS_{Total}}\right) \times 100 = \left(1 - \frac{12700780}{384393690}\right) \times 100 = 96.7\%$$

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Seasons	99	370738298.	3744831.	87.57	<.001
Climate	3	954613.	318204.	7.44	<.001
Residual	297	12700780.	42764.		
Total	399	384393690.			

Table 19: ANOVA for the mean yield with climate and seasons as sources of variation

Comparison was also done for the yields under the current climate and each climate change scenario separately to see how different the mean yields were between them. For this comparison, paired t-tests were done.

When all the 98 long and short rains seasons were included (the first and the last seasons were not complete therefore no farming was simulated), yields under the current climate was 122.4 Kg/ha, 101.4 Kg/ha and 110.6 Kg/ha more than yields under 10% reduction in rain amount, days and spells respectively. The three differences had p-values <0.001 but the yields under the climate change scenarios were not significantly different.

The same test was conducted for the yields under the short rain seasons. The yields under the current climate exceeded those under 10% less rain amount and spell by 88.83 Kg/ha and 143.7 Kg/ha respectively. The two differences had statistical significance (p-values were0.014 and 0.009). The yields under the current climate and those under 10% less rain days were not statistically significantly different.

For the long rains, the yield under the current climate was higher than under the different climate change scenarios with strong statistical significance. It was 155.4 Kg/ha, 162.7 Kg/ha and 78.79 Kg/ha more than yields when there was 10% less rain amount, days and spells.

The tests were also conducted for the different climate change scenarios for the short and long rains seasons. The difference in yield during short rains seasons under 10% change in rain amount and days was -50.67 Kg/ha (-91.00 Kg/ha, -10.34 g/ha). This was the only significant difference in yield under the climate change scenarios.

Chapter 5: Summary, Conclusions and Recommendations

5.1: Summary

This thesis sought to analyse how changing rainfall pattern would result in different rainfall amount when compared to change in just the amount. The same proportion of rainfall change (-10%) was used for both pattern and amount. This change was applied on the historical rainfall for Makindu.

Since the thesis considered a possible climate change scenario (for rainfall), raw data from Makindu was analysed for a trend in climate that may suggest a changing climate. Simple linear models (without and with factoring for months) were used to test for a statistically significant slope for both temperatures and rainfall. The results gave statistically significant slope for temperature but not for rainfall. The rate for temperature was more than three times higher than that given of the global rise in temperature, and this might be due to the few number of years involved (only 19 years).

Again, in order to use the Crop Simulation Model (APSIM) good complete daily climate data was needed. The rainfall data was available for the whole period of interest (1961 – 2010). Temperature data only covered 1992 – 2010 and radiation even less time. The four climate elements were needed in order to successfully run crop models. Therefore, a weather generator (Weatherman) was used to fill the missing temperature and calculate radiation from the temperature data. The temperatures and radiation were not adjusted for climate change, and their purpose was solely for simulating yield.

The climate change scenarios were generated by 1) reducing the amount of rainfall on a rainy day by 10%, 2) reducing the number of rain days systematically by 10% and 3) reducing the number of rainfall spells systematically by 10%. After this the effect of these scenarios were analysed for differences using ANOVA. The climate scenarios, though having statistical significance, did not explain a lot of the variability. Seasons were then included as factors affecting the variability in cumulative rainfall amounts and they explained more.

Finally, the four climate scenarios were fed into the APSIM model. Only rainfall was affected but temperatures were left as they were. The ANOVA of the yield results showed that when you consider the seasons, a change in rainfall pattern has more effect than simply changing the rainfall amount. Overall, however, the change in amount had more effect on the yield.

5.2: Conclusion

Since the procedure used by ICRISAT in 2008 was not considering the seasons independently, then it was sound for that analysis. This is because overall, the change in overall yield for fifty years was not significantly different between the climate change scenarios.

One limitation experienced is the lack of enough quality temperature and radiation data for Makindu. Because of this, some temperature data was corrected and a stochastic weather generator was used to in-fill the missing temperature data and calculates radiation data from the temperature data. Analysis of climate trends for was done after the correction but before in-filling. This thesis holds that this was representative of Makindu since the corrected data were outliers which were affecting the trend negatively, and the context in which they were corrected they appeared as errors in recording. Therefore, between 1992 and 2010, both minimum and maximum temperatures for Makindu have been increasing at the rates of 0.0831°C and 0.0336°C respectively (at the $\alpha = 0.05$ level of significance). This rate was very high when compared with the rate of global temperature increase; but this may be due to the fact that the data only covered 20 years. The rainfall amount and days has, however, not been changing with any statistical significance.

The main climate element under investigation was rainfall, how different ways of changing it will result in differences in amount and eventual maize yield. For this reason when simulating climate change scenarios, temperature and radiation were left unaffected. As it emerged, the changes had varying effects on the start of the rains and cumulative rainfall for different seasons. Delay in start of rains shortened the length of growing period for the maize crop while the reduction in amount increased the moisture stress for the crops. Since it is more likely that the pattern will be the one which will change in future, the thesis created preview of the 'what if' rainfall scenarios and the effect on crop yield. A 10% decrease on rain amount, days and spells resulted in 1) approximately 5% reduction in successful harvests and 2) lower yields. However, by changing the distribution, one year had the start of rains delayed ad resulted in better yields. This was because planting occurred in time for that season and it did not have any moisture stress due to in-season long dry spells. Given these variations, it is important to consider changing the rainfall pattern when analysis rainfall-change scenarios.

5.3: Recommendations

This study could be improved if there was more climate (temperature and radiation) data for Makindu. This would eradicate the use of stochastic models to simulate the climate data. In addition, having better quality data which would again not necessitate a reason for correction would very much improve on a follow up on this research project.

Another recommendation for further study is look at all possible rainfall change scenarios (of 10%) and analyse the effect on yields. The analysis conducted in this thesis looked at only one of the many possible scenarios of change in rain days and spells each. It is almost certain that the results found in this thesis could be different in case rain days or spells were changed differently.

This thesis limited itself to investigating the effects of reducing only 10% of rain amount, days and spells. More work can be done by also incorporating a 10% increase in rainfall amount, days or spells. This is since the IPCC has shown that the rainfall change may be either negative (as done here) or positive. In addition, the percentage change can also be increased to 20% in accordance with the IPCC Special Report on Emission Scenarios for 2007.

Another recommendation is the use of stochastic processes when simulating climate change scenarios; in this thesis rainfall change scenarios were all deterministic. This can be done by shifting the Markov probabilities (figures 23 and 24) and using them to stochastically create rain days, then after finding suitable parameters, the rainfall amount can be simulated using gamma distribution or other distributions used to model rainfall amount.

References

1. Chaudhari K, Oza M. Impact of climate change on yields of major food crops in India. *isprs.org*. 2007. Available at: http://www.isprs.org/proceedings/XXXVIII/8-W3/b2/3–B3-92.pdf. Accessed April 29, 2012.

2. Stern RD, Cooper PJM. Assessing Climate Risk and Climate Change Using Rainfall Data – a Case Study From Zambia. *Experimental Agriculture*. 2011;47(02):241–266. Available at: http://www.journals.cambridge.org/abstract_S0014479711000081. Accessed December 22, 2011.

3. Osbahr H, Dorward P, Stern RD, Cooper S. Supporting Agricultural Innovation in Uganda To Respond To Climate Risk: Linking Climate Change and Variability With Farmer Perceptions. *Experimental Agriculture*. 2011;47(02):293–316. Available at: http://www.journals.cambridge.org/abstract_S0014479710000785. Accessed December 27, 2011.

4. Dixit P, Cooper P, Dimes J, Rao K. Adding value to field-based agronomic research through climate risk assessment: A case study of maize production in Kitale, Kenya. *Experimental Agriculture*. 2011;47(02):317–338.

5. Steiner JL, Hatfield JL. Winds of Change: A Century of Agroclimate Research. *Agronomy Journal*. 2008;100(Supplement_3):S–132–152. Available at: https://www.agronomy.org/publications/aj/abstracts/100/Supplement_3/S-132.

6. Hoogenboom G, Jones JW, Wilkens PW, et al. Decision Support System for Agrotechnology Transfer (DSSAT) v4.5 CD-ROM. 2011.

7. Keating BA, Carberry PS, Hammer GL, et al. An overview of APSIM , a model designed for farming systems simulation. *Science*. 2003;18.

8. Hulme M, Doherty R, Ngara T, New M, Lister D. African climate change : 1900 – 2100. 2001;17:145–168.

9. Cooper P, Rao K, Singh P. Farming with current and future climate risk: Advancing a'Hypothesis of Hope'for rainfed agriculture in the semi-arid tropics. *Jounal of SAT Agricultural Ressearch*. 2009;7(December):1–19.

10. VSN. VSN International. 2012. Available at: http://www.vsni.co.uk/. Accessed August 12, 2012.

11. Makindu Elders Society. Agriculture. 2012. Available at: http://www.makinduawake.org/index.php?option=com_content&view=article&id=55&Itemid=58. Accessed February 6, 2012.

12. Mutere FA. A comparative study of reproduction in the populations of the insectivorous bats, Otomops martiensseni, at latitudes 1� 5' S and 2� 30' S. *Journal of Zoology*. 1973;171:79–92.

13. Hansen J. Linking dynamic seasonal climate forecasts with crop simulation for maize yield prediction in semi-arid Kenya. *Agricultural and Forest Meteorology*. 2004;125(1-2):143–157. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0168192304000450.

14. Singh NP, S Bantilan MC, Kumar AA, Janila P, R Hassan AW. Climate Change Impact in Agriculture: Vulnerability and Adaptation Concerns of Semiarid Tropics in Asia. In: *Crop Adaptation to Climate Change*. Wiley-Blackwell; 2011:107–130. Available at: http://dx.doi.org/10.1002/9780470960929.ch9.

15. Sivakumar MVK, Das HP, Brunini O. Impacts of Present and Future Climate Variability and Change on Agriculture and Forestry in the Arid and Semi-Arid Tropics. Salinger J, Sivakumar MVK, Motha RP, eds. *Climatic Change*. 2005;70(1-2):31–72. Available at: http://www.springerlink.com/index/R8167645734QXH72.pdf.

16. Recha CW, Makokha GL, Traore PS, Shisanya C, Lodoun T, Sako A. Determination of seasonal rainfall variability, onset and cessation in semi-arid Tharaka district, Kenya. *Theoretical and Applied Climatology*. 2011;(2000):1–16. Available at: http://www.springerlink.com/index/10.1007/s00704-011-0544-3.

17. Cooper PJM, Dimes JP, Rao KPC, Shapiro B, Shiferaw B, Twomlow S. Coping better with current climatic variability in the rain-fed farming systems of sub-Saharan Africa: An essential first step in adapting to future climate change? *Agriculture, Ecosystems & Environment*. 2008;126(1-2):24–35.

18. Hansen JW, Mishra A, Rao KPC, Indeje M, Ngugi RK. Potential value of GCM-based seasonal rainfall forecasts for maize management in semi-arid Kenya. *Agricultural*

Systems. 2009;101(1-2):80–90. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0308521X09000468.

19. Ipcc IPOCC. *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. (Pachauri RK, Reisinger A, eds.). IPCC; 2007:104. Available at: http://www.ipcc.ch/publications_and_data/ar4/syr/en/contents.html.

20. Ipcc CWT. *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. (Metz B, Davidson OR, Bosch PR, Dave R, Meyer LA, eds.). Cambridge University Press; 2007:x, 851 p. Available at: http://www.ipcc.ch/ipccreports/ar4-wg3.htm.

21. Barker T. Climate Change 2007 : An Assessment of the Intergovernmental Panel on Climate Change. Pachauri RK, Reisinger A, eds. *Change*. 2007;446(November):12–17. Available at: http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf.

22. Stone DA. Predicted climate changes for the years to come Predicting 21st Century weather. *Change*. 2008;27(2):319–330.

23. Pickering NB, Hansen JW, Jones JW, Wells CM, Chan VK, Godwin DC. Weatherman - A utility for managing and generating daily weather data. *Agronomy Journal*. 1994;86(2):332–337.

24. Jones PG, Thornton PK. A rainfall generator for agricultural applications in the tropics. *Agricultural and Forest Meteorology*. 1993;63:1 – 19.

25. Mollah WS, Cook IM. Rainfall variability and agriculture in the semi-arid tropics - The Northern Territory, Australia. *Agricultural and Forest Meteorology*. 1996;79(1-2):39–60. Available at: http://www.sciencedirect.com/science/article/B6V8W-3Y3YRYW-8/2/af69992575158255d32c550bbed87320.

26. Jamieson PD. Crop Responses to Water Shortages Crop Responses to Water Shortages. *Journal of Crop Production*. 2008;2(2):71–83. Available at: http://www.scopus.com/scopus/inward/record.url?eid=2-s2.0-0033390820&partner=40&rel=R4.0.0.

27. Meinke H, Carberry PS, Cleugh HA, Poulton PL, Hargreaves JNG. Modelling crop growth and yield under the environmental changes induced by windbreaks. 1. Model

development and validation. *Australian Journal of Experimental Agriculture*. 2002;42(6):875–885.

28. Xing H, Wang E, Smith CJ, Rolston D, Yu Q. Modelling nitrous oxide and carbon dioxide emission from soil in an incubation experiment. *Geoderma*. 2011;167-168(3):328–339. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0016706111001923.

29. Stern RD, Rijks D, Dale IC, Knock J. Instat Climatic Guide. 2006;(January).

30. Stern RD, Coe R. The use of rainfall models in agricultural planning. *Agricultural Meteorology*. 1982;26:35–50.

31. Usher PEO. Kenya and ENSO: An Observation and La Niña Prediction. 2011.

32. NASA. NASA Earth Observatory: Different Records, Same Warming Trend. 2011. Available at: http://earthobservatory.nasa.gov/IOTD/view.php?id=48574. Accessed March 19, 2012.

33. NOAA. National Oceanic and Atmospheric Administration Climate Services. 2012. Available at: http://www.climate.gov/#dataServices. Accessed March 11, 2012.