Journal of Mathematics and Statistics 9 (1): 1-11, 2013
ISSN 1549-3644

doi:10.3844/jmssp.2013.1.11 Published Online 9 (1) 2013 (http://www.thescipub.com/jmss.toc)

TREND ANALYSIS OF MONTHLY RAINFALL DATA IN CENTRAL ZONE

Telemu Kassile

Department of Biometry and Mathematics, Faculty of Science, Sokoine University of Agriculture, Tanzania

Received 2012-12-17, Revised 2013-02-01; Accepted 2013-03-15

ABSTRACT

Proper trend analysis of rainfall data is necessary for social and economic planning especially in the tropics where rainfall is a major limiting factor for crop production. In spite of the growing number of studies on climatic variables including rainfall in recent years, in the context of Tanzania little is known as to whether or not the observed time trend in rainfall in recent years is statistically significant. This study examines the evolution of rainfall in central zone focusing on Dodoma region and ascertains whether the observed time trend in rainfall is significant in statistical terms. Also, the study establishes whether there was a shift in the distribution of rainfall for both short and long rainy seasons in the study area. The data for the analysis consists of monthly rainfall (mm) records covering the period from January, 1981 to December, 2010. Linear regression analysis through OLS estimation and the Kruskal-Wallis and Mann-Kendall's tau test statistics were employed. The data demonstrate that rainfall in the study area is unpredictable both in terms of onset and cessation. The mean annual rainfall during the years 1981-2010 displayed a decreasing trend with the years 1981-1989 recording the highest mean annual rainfall as opposed to the succeeding two periods: 1990-1999 and 2000-2010. Overall, the mean annual rainfall has declined by about 4% between the periods 1981-1989 and 2000-2010. Parametric and nonparametric analysis results did not find any statistically significant evidence of a trend in the amount of rainfall in Dodoma region over the last 30 years. Nevertheless, there was a shift in the distribution of rainfall for both short and long rainy seasons. The study recommends for further research that will include rainfall data from the various zones of Mainland Tanzania in order to permit a comparative analysis of the issues examined in this study.

Keywords: Trend Analysis, Rainfall Data, Central Zone

1. INTRODUCTION

In recent years, climate change is one of the issues, which have attracted the attention of policy makers and advisors, scholars, directors of several research institutes, among others around the world. In Africa, a great part of the continent's economies and the livelihood of Africa's poor depend on agriculture. In Africa, especially in sub-Saharan Africa, one cannot speak about economic growth without addressing agriculture and one cannot talk about addressing agriculture without taking into account the issue of climate change according to African Economic Research Consortium (Lyakurwa, 2009). While the numerous climatic variables interact with the crop in complex ways, rainfall is the most important limiting factor for crop production in most parts of the tropics (Stern and Coe, 1984; Perera *et al*., 2002). In this linkage, variations and trends in rainfall have significant social and economic impacts on agriculture and in turn, people's livelihood in these areas. Consequently, describing changes in the patterns of rainfall (Pennycuick and Norton-Griffiths, 1976) is one of the areas, which have received particular attention by researchers across the world.

The undesirable impacts of climate change are apparent nearly all over the world. In Tanzania, the impacts of climate change have been manifest both at the micro and macro-levels. Severe and repeated droughts especially during the past recent years have resulted into food shortages and overwhelming power crisis in many parts of the country. Temperature measurements from 21 meteorological stations in the country revealed a stable increase in temperature as a result of global warming.

Fig. 1. Map of Tanzania showing Dodoma region

Because of the rising temperature, it is expected that the whole glacier of Mount Kilimanjaro will be gone by 2025 (URT, 2007).

The amount of rainfall and its distribution over the years largely affect the productivity of agriculture in semi-arid regions of Africa (Yengoh *et al*., 2010). Thus, appropriate trend analysis of rainfall data is essential for social and economic planning to measure the effects of global warming (Ghosh *et al*., 2009). However, as early noted by Stern and Coe (1984), the rate of collecting rainfall data does not match with proper analysis to permit agricultural planning decisions. In recent years, the number of studies on climatic variables has been growing. In the context of Tanzania, relatively little is known in statistical terms about the impact of climate change on climate indices including rainfall. Moreover, the few studies (Prins and Loth, 1988; Pennycuick and Norton-Griffiths, 1976) which have been done on rainfall variability have for the most part been focusing on the northern part of the country.

This study examines the evolution of rainfall in central zone over the last 30 years and ascertains whether the observed time trend in rainfall during this period is statistically significant. In addition, the study attempts to establish whether there was a change or shift in the distribution of rainfall for both short and long rainy seasons in recent years relative to 20 years ago.

1.1. Data Source and Description of the Study Area

The data for analysis in this study were obtained from the Tanzania Meteorological Agency Headquarters in Dar as Salaam, Tanzania. The data consist of monthly rainfall (mm) records covering the period from January, 1981 to December, 2010. Dodoma region is located in the central zone in Mainland Tanzania and lies between latitude 4, 7’ and 7°C21’ south of the equator and between longitude 36, 43’ and 35°C5’ east of Greenwich (**Fig. 1**). The region has a total land area of 41,311 square kilometers and is the 12th largest region in Mainland Tanzania in terms of land area (in square kilometers). The region is largely semi-arid with an annual rainfall ranging from 400 mm to 900 mm. The temperature in the region varies depending on altitude and season. During the period from October to December, the minimum and maximum temperatures are 18 and 31°C respectively whereas for the cool dry season (June-August) the minimum and maximum temperatures range from 10 to 11 and 27 to 28°C respectively (RCO, 2003).

In Dodoma region, as elsewhere in Tanzania, a large proportion of the population lives in rural areas.

Table 1. Dodoma region: Selected indicators

Indicator	Dodoma	National average	Percent ± national average
Population			
Population (%), 2002[a]	5.1	4.8	+6.3
Annual (%) growth rate (1988-2002), 2002[a]	2.3	2.9	-20.7
Education			
Males 15-49 years with no education (%), 2005[b]	17.3	12.7	+36.2
Females 15-49 years with no education (%), 2005[b]	29.7	24.7	+20.2
Employment			
Males 15-49 years unemployed in last 12 months (%), 2008[c]	18.5	18.4	+0.5
Females 15-49 years unemployed in last 12 months (%), 2008[c]	25.4	19.1	+33.0
Economy			
Household population in the lowest wealth quintile (%), 2008[c]	24.2	21.5	12.6
Household population in the highest wealth quintile (%), 2008[c]	11.1	17.9	-38.0
GDP (%) at current prices, 2000[a]	3.4	5.0	-32.0
Nutritional status of under-five children			
<5 children stunted-low height-for-age (-2SD) (%), 2005[b]	44.4	38.0	+16.8
<5 children underweight-low weight-for-age (-2SD) (%), 2005[b]	30.4	21.9	+38.8

[*]Regional population as a percentage of national (Mainland Tanzania) population. **Source:** a: RCO (2003); b: ORCM (2005); c: THMIS (2008)

Approximately 85% of the total 2,214,657 projected (2012 estimates) population live in rural areas and depend mainly on rain-fed agriculture and livestock keeping for their livelihood. It is estimated that agriculture employs over 90% of the workforce. Major food crops grown in the region are maize, sorghum, bulrush millet and cassava. Per capita GDP is generally low compared to most of the regions in Mainland Tanzania. For example, in 1995, 1997, 1999 and 2000 Dodoma region was ranked 18, 18, 18 and 19th out of the then total 20 regions in Mainland Tanzania in terms of GDP per capita. Underdevelopment of the natural resources sector, which includes forestry, beekeeping, fishing and mining, greatly contributes to the region's small input to national GDP and poor quality of life of the majority of the population (RCO, 2003). Comparative analysis by the International Food Policy Research Institute (IFPRI, 2006) reveals that Dodoma region features predominantly among the poorest regions in Mainland Tanzania in many aspects. For example, in terms of incidence of poverty, IFPRI (2006) found that in 1991/92, 1996, 1999 and 2003 Dodoma was the 11, 18, 16 and 18th poorest region respectively out of 20 regions considered in the analysis. **Table 1** gives a succinct comparison between Dodoma region and the national (Mainland Tanzania) average on the selected indicators in which a positive (+) or a negative (-) sign represents the percentage by which the regional characteristic is above or below the national indicator respectively.

2. MATERIALS AND METHODS

As already mentioned, the data for analysis in this study were monthly records for the periods 1981-2010. Scrutiny of the data revealed that the record for October, 2009 was missing. A decision was made not to discard the records for the entire year but to estimate the missing value. For simplicity, the last observation carried forward (Verbeke and Molenberghs, 2000 for details) procedure was applied in estimating the missing observation. No attempt was made to investigate the influence of the said observation on the results.

2.1. Exploratory Analysis

As a prelude to the linear regression analysis of the data, descriptive analysis (both numerical and graphical) was first carried out in order to explore the distribution of rainfall in the region. The arithmetic mean was used for numerical summary measure, whereas scatter plots and trend lines of rainfall against time were generated for graphical presentation of the data. For the purpose of understanding changes in the amount of rainfall over time, the reference period (1981-2010) was grouped into three separate time periods: 1981-1989, 1990-1999 and 2000-2010. For both numerical and graphical measures, the analysis was done separately for the three time periods and overall (1981-2010). In order to have a complete understanding of the changes in the amount of rainfall over the last 30 years, the monthly amount of rainfall for each year were reduced to a single value. From the annual rainfall series obtained, plots of departure from a long term mean during the periods 1981-1989, 1990-1999 and 2000-2010 and overall (1981-2010) were constructed in order to discern years of anomaly in the amount of rainfall. Furthermore, the twelve months (Jan-Dec) were categorized according to major rainy and dry seasons in the study area. That is, October-January (short rain season), February-May (long rain season) and June-September (dry season). Rainfall (mm/season) was calculated and profiles of rainfall for each season were created separately for each of the three time periods (1981-1989, 1990-1999 and 2000-2010).

2.2. Testing for Trend

Tests for trend in a data set can be done with both parametric and nonparametric methods. The former tests are more powerful when the data are normally distributed than is the case when it is not (Onoz and Bayazit, 2003). When the data include outliers or are severely non-normally distributed, the use of parametric methods can give incorrect results hence invalid inference (DAPD, 1999). Moreover, nonparametric methods relax the parametric assumption imposed on the data generating process and allow the data to determine a suitable functional form (Racine, 2008). For comparative purposes of the results in this study, both parametric and nonparametric methods were employed. In particular, linear regression analysis (for parametric) and the Mann-Kendall or Kendall's tau (τ) statistic (for nonparametric) methods were employed to ascertain in statistical terms, whether there was a trend in the amount of rainfall over time. Furthermore, in order to gain more insights into the impact of time on the amount of rainfall, the time periods 1891-1989, 1990-1999 and 2000-2010 were treated as three independent time points. That is, t_1, t_2 and t_3 then the Kruskal-Wallis rank sum nonparametric test statistic, H equivalent of the one-was analysis of variance (Zar, 1984) was used.

2.3. Linear Regression Analysis

Based on the graphical representation of the data discussed in the exploratory analysis section, a straight line regression model was hypothesized to describe the relationship between the amount of rainfall and time, hence characterize the prediction of any amount of rainfall y (mm), given any time t (in year). The equation relating the amount of rainfall with time was as follows:

$$y = \gamma_0 + \gamma_1 t + \varepsilon \quad (1)$$

In which y is the dependent or response variable representing the amount of rainfall (mm), t is the covariate or explanatory variable and ε is the unobserved error or disturbance. The goal is to estimate the regression parameters γ_0 the intercept and γ_1 the slope. A familiar assumption in linear regression is that the error has a mean of zero and that each explanatory variable is uncorrelated with the error term (Wooldridge, 2001). In the structure of the model in Equation 1 this assumption is equivalent to $E(\varepsilon) = 0$, $E(t, \varepsilon) = 0$.

In the perspective of the present analysis, γ_1 was interpreted as representing the average rate of change of rainfall throughout every one year time period. Significant γ_1 ($p<0.05$) was interpreted as indicating that there was a trend in the amount of rainfall and was equal to the magnitude of γ_1. Otherwise, insignificant γ_1 signifies absence of a trend in the amount of rainfall over time. On the other hand, the direction (increasing or decreasing) was represented by the sign of the slope parameter γ_1 in which a negative sign indicates decreasing trend while a positive implies an increasing trend with time.

The model in Equation 1 was fitted by Ordinary Least Squares (OLS), a parametric method. Prior to estimation of the model parameters, the outcome measure of interest was subjected to a normality test, which as mentioned above, is one of the requirements for classical methods including the OLS. This was achieved using the Shapiro-Wilk test of normality at the 5% (or 0.05) level of significance. Under this test, the null, H_0 and alternative, H_a hypotheses were H_0: The annual amount of rainfall was normal and H_a: The annual amount of rainfall was non-normally distributed. The univariate procedure (Cody and Smith, 1997 for details) was carried out in the SAS system software version 9.2 (SAS Institute Inc., Cary, NC, USA). By specifying the option normal the univariate procedure computes the Shapiro-Wilk statistic, W together with the corresponding probability (or p) value. At the pre-specified level of significance, the null hypothesis of normally distributed annual rainfall was not rejected ($p>0.05$). Therefore, the amounts of rainfall y were modeled in their natural occurring scale of measurement. The model in Equation 1 was fitted using the reg procedure in SAS.

2.4. Kruskal-Wallis Test

Following Zar (1984) the Kruskal-Wallis test statistic, H was calculated as in Equation 2:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1) \quad (2)$$

where, n_i is the number of observations in group i; $N = \sum_{i=1}^{k} n_i$ is the total number of observations in all K (=3 in the present case) groups; and R_i is the sum of the ranks of the n_i observations in group i (=1, 2, 3). Under this (H) test, the null and alternative hypotheses were H_0: The annual amounts of rainfall were the same in all three time (t_1, t_2 and t_3) periods and H_a: The annual amounts of rainfall were not the same in all three time periods. The npar1way procedure in SAS was used to compute H in Equation 2.

2.5. Kendall's Tau Test

The Kendall's τ statistic is one of the nonparametric trend tests that have been frequently used in the literature. Kahya and Kalayci (2004) is an excellent reference for numerous other trend test techniques. The Kendall's τ statistic has been described in much detail by Onoz and Bayazit (2003); Kahya and Kalayci (2004) and Partal and Kahya (2006), among others. In brief, following Kahya and Kalayci (2004); Onoz and Bayazit (2003) and Partal and Kahya (2006), Kendal's τ test first ranks all observations by date order, then the difference between each consecutive value is calculated and the sum of the signs of these differences is calculated as the Kendall sum, S statistic given as in Equation 3:

$$S = \sum_{J=k}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}\left(x_j - x_k\right) \tag{3}$$

In which $\operatorname{sign}\left(x_j - x_k\right) = \begin{cases} 1 & \text{if} \quad \left(x_j - x_k\right) > 0 \\ 0 & \text{if} \quad \left(x_j - x_k\right) = 0 \\ -1 & \text{if} \quad \left(x_j - x_k\right) < 0 \end{cases}$

The expected value and variance of S are E (S) = 0 and Var (S) = $\left[n(n-1)(2n+5) - \sum_t t(t-1)(2t+5)\right]/18$ respectively with t indicating the extent of any given time and $\sum_t$ denotes the sum across all the ties in the rainfall data. For n>0, the standard normal variate is calculated as in Equation 4:

$$z = \begin{cases} \dfrac{S-1}{\sqrt{\operatorname{Var}(S)}} & \text{if} \quad S > 0 \\ 0 & \text{if} \quad S = 0 \\ \dfrac{S+1}{\sqrt{\operatorname{Var}(S)}} & \text{if} \quad S < 0 \end{cases} \tag{4}$$

Under the Kendall's τ test, a positive value of S in Equation 3 indicates an increasing trend whereas a negative value indicates a decreasing trend (Kahya and Kalayci, 2004). The null and alternative hypotheses under S were H_0: There was no trend in the annual amounts of rainfall and H_a: There was a trend in the annual amounts of rainfall.

As already pointed out, in all tests of significance, the decisions to reject the null hypothesis proceeded in the same fashion as standard tests of hypotheses or significance. That is, if the p-value of the test was less than the level of significance, the hull hypothesis was rejected, but if the reverse was the case, then the null hypothesis was not rejected. Kendall's τ test of significance was carried out in S-PLUS 2000 Professional Edition for Windows using the cor. test function with method=K for Kendall's τ statistic.

All tests of significance were two-sided and considered significant at the 0.05 level. All figures were created in R version 2.7.2 (The R Foundation for Statistical Computing).

3. RESULTS

3.1. Distribution of rainfall

Figure 2 gives monthly distribution of rainfall during the years 1981-1989, 1990-1999 and 2000-2010. As evident from the figure, there are divergent erratic variations in the amount of rainfall across several months and years during the periods 1981-2010 in the area. For example, a comparison across several months reveal that January was characterized by the highest rainfall during the period 1981-1989 followed by the periods 1990-1999 and 2000-2010. In contrast, in February and in March, the highest monthly rainfall were recorded during the period 1990-1999 with the periods 2000-2010 and 1981-1989 representing the second and the last positions respectively. The month of December in the period 2000-2010 recorded the highest monthly rainfall followed by the period 1981-1989 and lastly the period 1990-1999.

The irregularity of monthly rainfall across these year periods is also depicted in **Fig. 3** in which during the period 1981-1989, January of 1987 recorded maximum rainfall (330.8 mm) while during the periods 1990-1999 and 2000-2010 it is December 1997 and 2009 which recorded maximum rainfall of 433.6 mm and 336.9 mm respectively. Moreover, as revealed in **Fig. 3**, maximum monthly rainfall varied between periods (1981-1989, 1990-1999 and 2000-2010) during the years 1981-2010.

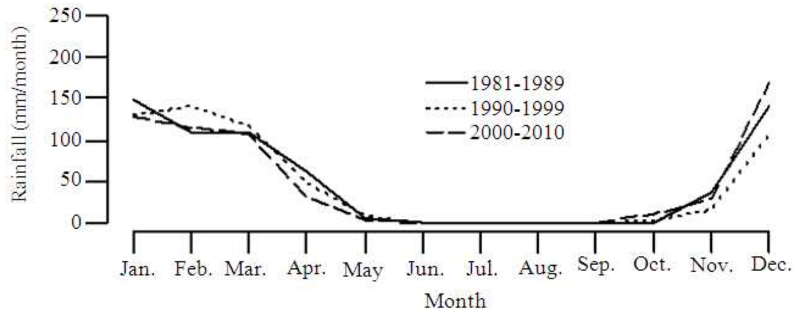

Fig. 2. Distribution of rainfall in Dodoma region (1981-2010)

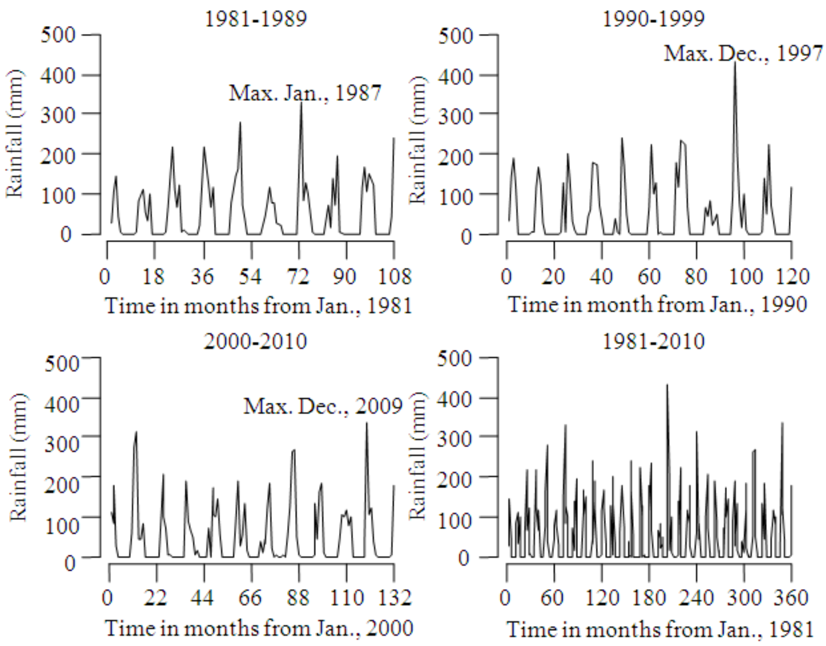

Fig. 3. Observed monthly rainfall in Dodoma region (1981-2010)

As marked in **Fig. 4**, there appears to be a shift in the distribution of rainfall for both short and long seasons. During the years 1981-1989, short rain season was concentrated between November and January. In contrast, during the years 1990-1999, the concentration of rainfall was between December and January. Likewise, there is an upward shift in the distribution of rainfall for both long and dry seasons during the years 1990-1999. However, the shift which is observed during the years 1990-1999 is transitory in that the profile of rainfall observed in recent years (2000-2010) is rather similar across seasons to that of the years 1981-1989.

The change or shift in the amount of rainfall between the years 1981-1989 and 1990-1999 is more pronounced in the short season than in the long season. In the former season, the change (decrease) was about 60% while in the later season the change (increase) was approximately 54%. Also, the difference between the amounts of rainfall during the years 1990-1999 and 2000-2010 was higher in the short season (51.2) than in the long season (46.6). On the contrary, the magnitude of the gap in the amounts of rain was much higher in the long season (8.3) than in the short season (1.6) (**Table 2**)

3.2. Variations in Annual Rainfall

As seen in **Fig. 5**, the annual amount of rainfall varied both within and between periods during the years 1981-1989, 1990-1999 and 2000-2010.

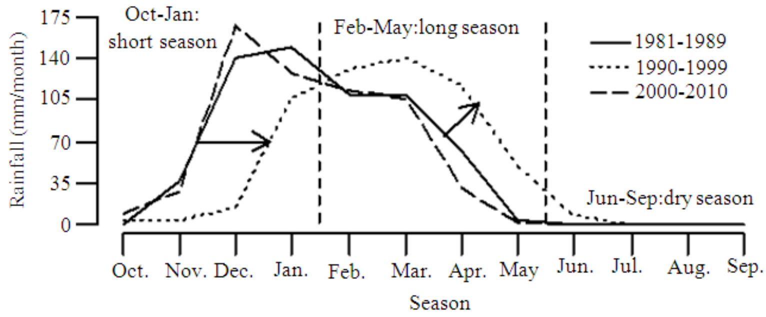

Fig. 4. Seasonal variations of rainfall in Dodoma region (1981-2010)

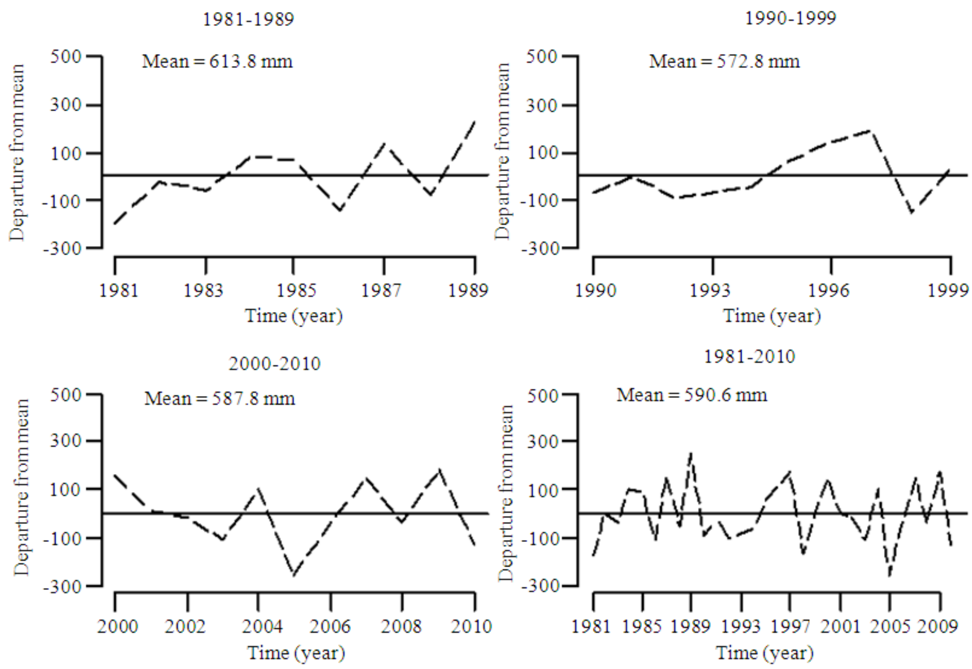

Fig. 5. Trends in departure from mean annual rainfall (1981-2010)

Table 2. Changes in mean seasonal rainfall

	Rainfall (mm/season)	
Period	Long season	Short season
1981-1989	71.4	82.0
1990-1999	109.7	32.4
2000-2010	63.1	83.6

Over the periods 1981-1989, 1990-1999 and 2000-2010, the average annual rainfall (mm) was 613.8, 572.8 and 587.8 respectively and during the periods 1981-2010 the average was 590.6 mm. During the three time periods (1981-1989, 1990-1999 and 2000-2010), the annual rainfall value of the highest negative departures from mean were recorded in 1981 (411.9 mm), 1998 (420.8 mm) and 2005 (329.7 mm). On the other hand, the highest positive departures were recorded in the years 1989, 1997 and 2009 with corresponding annual rainfall values of 836.0 mm, 763.1 mm and 768.4 mm respectively. In general, there appears to be an even distribution of annual rainfall within the three time periods; that is, about equal numbers of years lie in both sides of the zero line. For instance, during the time period 1981-1989, five years (1981, 1982, 1983, 1986 and 1988) lie below the zero reference line while the remaining four years (1984, 1985, 1987 and 1989) are above the zero line.

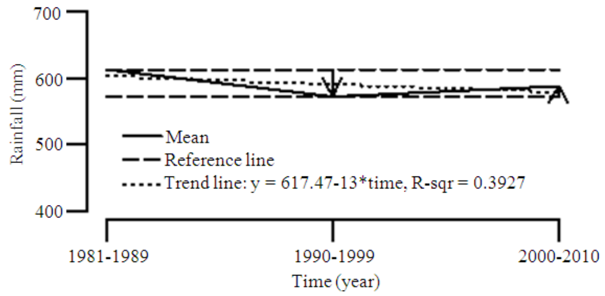

Fig. 6. Profile of mean annual rainfall (1981-2010)

Table 3. Kruskal-Wallis rank sum test

Quantity	SAS	S-PLUS
Chi-Square	0.3465	0.3465
DF	2	2
p-value	0.8409*	0.8409*

*; Not significant at 0.05 level

Table 4. Parameter estimates for model in Equation (1)

Parameter	Estimate (standard error)	p-value
γ_0	590.7191 (47.1473)	<0.0001
γ_1	-0.0085 (2.6558)	0.9975*

*; Not significant at 0.05 level

Table 5. Parameter estimates for model in Equation (1) including a quadratic term

Parameter	Estimate (standard error)	p-value
γ_0	560.4704 (74.8004)	<0.0001
γ_1	5.6631 (11.1230)	0.6148*
γ_2	-0.1830 (0.3482)	0.6035*

*; Not significant at 0.05 level

Figure 6 plots mean annual rainfall. Between the periods 1981-1989 and 1990-1999, the mean annual rainfall decreased by about 7% but increased by approximately 3% between the periods 1990-1999 and 2000-2010. Between the periods 1981-1989 and 2000-2010, the average annual rainfall decreased by about 4%; that is, decreasing from 613.8 mm during the period 1981-1989 to 587.8 mm during the period 2000-2010. This trend of increase and decrease in the mean annual rainfall is depicted by the area between the solid (mean) and dotted (reference) lines as in **Fig. 6**. In general, the mean annual rainfall observed during the period 1981-1989 exceeds the one which is observed in the succeeding two periods.

Figure 7 displays the trends in the observed annual rainfall during the three time periods: 1981-1989, 1990-1999 and 2000-2010. As evidenced in the figure, the first two time periods display increasing linear trends denoted by positive slopes in the trend line equations. However, the years 1990-1999 exhibit a lower rate of increase than the rate during the years 1981-1989. In contrast, the years 2000-2010 demonstrate a decreasing trend, which is represented by a negative slope in the corresponding trend line equation. On the whole, the years 1981-2010 indicate a decreasing linear trend, though, at a much lower rate (0.008) than the one (3.727) observed during the years 2000-2010.

3.3. Relationship between Rainfall and Time

Table 3 gives Kruskal-Wallis results of the test of the effect of time on annual rainfall. As indicated in the table, both SAS and S-PLUS show the same results, thereby the same conclusion; that is, there are equal calculated test statistic values (Chi-Square), Degrees of Freedom (DF) and p-values. Based on the p-value, there is not enough evidence to reject the null hypothesis of the same annual rainfall recorded in the three time periods: 1981-1989, 1990-1999 and 2000-2010. This is equivalent to saying that, the annual amounts of rainfall were not statistically different for the three time periods.

3.4. Trend

Table 4 presents estimation results for the model in Equation 1. The slope parameter γ_1 is negative suggesting a declining trend in annual rainfall. However, the parameter (γ_1) is not statistically significant ($p>0.05$) signifying that there is no trend in the amount of rainfall over time. Moreover, the overall fit of the model was not statistically significant ($p>0.05$) with R-square value equal to zero, indicating that the assumed linear model does not explain any of the observed variability in y.

To examine the impact of misspecification of the model, the model in Equation 1 was fitted including a quadratic term. That is, $y = \gamma_0 + \gamma_1{}^t + \gamma_2 t^2 + \varepsilon$ in which γ_0 is the intercept; γ_1 is the linear term; γ_2 is the quadratic term; and ε is the residual. An OLS solution gave the results in **Table 5** in which both the linear and quadratic terms were not significantly different from zero at the 0.05 level. Notable from the results in **Table 5** is a change in both the sign and magnitude of the linear term (γ_1). Inclusion of a quadratic term in the model changed the sign of the linear term from negative (**Table 4**) to positive (**Table 5**). Moreover, the magnitude of the estimate of the term (γ_1) changed from 0.0085 to 5.6631. Accordingly, the p-value for testing for the effect of the linear term changed from 0.9975 of that of the model without the quadratic term in **Table 4** to 0.6148 of the model with the quadratic term in **Table 5**.

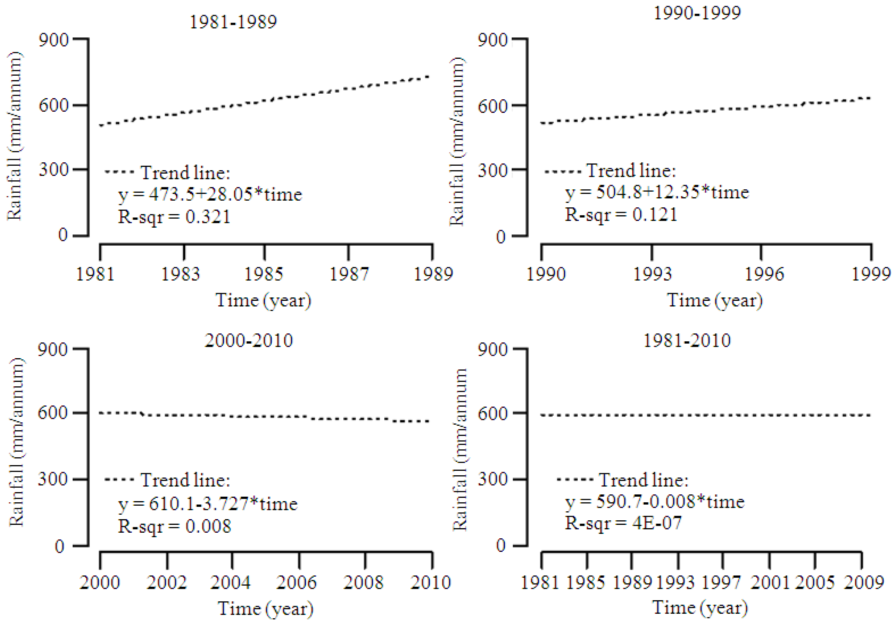

Fig. 7. Trends in observed annual rainfall (1981-2010)

Table 6. Kendall's tau trend test

Quantity	Value
Kendall's tau	0.0207
Normal-z	0.1606
p-value	0.4362*

*; Not significant at 0.05 level

Like the model without the quadratic term, the overall fit of the model with the quadratic term was also not good ($p>0.05$) though R-square improved slightly reaching a value of 0.0101 for the later model compared to 0.0000 for the former model.

The results from the Mann-Kendall's trend test (**Table 6**) fail to reject ($p>0.05$) the null hypothesis of no trend in the rainfall series at the 0.05 level. Kendall sum, S statistic is positive indicating that the sum of the differences between each consecutive value exceed zero.

4. DISCUSSION

In this study, the trend analysis of rainfall data in the central zone in Tanzania was carried out first to understand the evolution of rainfall over time and to ascertain in statistical terms whether the observed time trend in rainfall was increasing or decreasing over the last 30 years. The data for analysis were obtained from Tanzania Meteorological Agency and comprised of monthly records of rainfall covering the period from January, 1981 to December, 2010. To test for trend in the data, both parametric and nonparametric approaches were employed. The impact of time on rainfall was also examined in a nonparametric framework disaggregating the years 1981-2010 into three distinct time periods: 1981-1989, 1990-1999 and 2000-2010.

The results demonstrate that rainfall in the study area is unpredictable both in terms of onset and cessation. For the years 1981-1989 major rainfall during the short season (October-January) began in October while during the years 1990-1999 it began in December. While cessation of rainfall during the long season was in May during the years 1981-1989, it extended to June during the years 1990-1999. However, during the years 2000-2010, the onset and cessation of rainfall demonstrate a similar pattern to that of the years 1981-1989 (**Fig. 3**). This finding is consistent with the results by Prins and Loth (1988) for northern Tanzania. The authors found that the amount of rain in the long rainy season was more predictable than the amount in the short rainy season.

Mean annual rainfall during the years 1981-2010 displayed a decreasing trend with the years 1981-1989 recording the highest mean annual rainfall than the succeeding two periods of 1990-1999 and 2000-2010. However, the annual rainfall recorded during the years 2000-2010 was slightly higher than that recorded during the periods 1990-1990. During the periods 1981-

1989 and 1990-1999, the annual rainfall displayed an increasing trend at a higher rate during the former than during the later period. In contrast, during the years 2000-2010 and over the entire reference period (1981-2010) the annual rainfall demonstrated a decreasing trend. Overall, the mean annual rainfall has declined by about 4% between the periods 1981-1989 and 2000-2010. The declining trend in the amount of rainfall observed in the present analysis is consistent with the observations in many studies. Long rains in central Kenya, for instance, have declined by more than 100 mm since the mid 1970s. This decline is most likely due to warming in the Indian Ocean and is expected to continue (Funk, 2010). In Turkey, Partal and Kahya (2006) found manifest decrease in the annual mean rainfall in some parts of the country particularly in western and southern, as well as along the coasts of the Black Sea.

Parametric analysis of trend by OLS gave a negative slope parameter suggesting decreasing annual rainfall over time. Nonetheless, the parameter was not different from zero in statistical terms implying absence of a linear trend in the rainfall data over a 30-year time period. The Mann-Kendall's tau nonparametric test statistic yielded the same conclusion of not enough evidence to suggest the existence of trend (increasing or decreasing) in the rainfall data in Dodoma. The impact of time on the amount of rainfall was investigated in a nonparametric framework employing the Kruskal-Wallis equivalence of the one-way analysis of variance. Also, the null hypothesis that the annual rainfall was the same in all three time periods was not rejected.

While the decline in the amount of rainfall between the periods 1981-1989 and 2000-2010 seems to be statistically insignificant, its impact may be devastating to the livelihood of the people in areas such as Dodoma. As already noted, rainfall in Dodoma region is low and unpredictable for crop production. In 2001, potential areas for irrigation in Dodoma region were estimated to be 16,152 hectares. Of these, only 3,756 hectares (about 23%) were under irrigation in 2001. Lack of resources to develop irrigation schemes was one of the reasons for low coverage of irrigation schemes in the region (RCO, 2003).

5. CONCLUSSION

The purpose of analyzing the rainfall data in this study was to determine whether or not the observed time trend in rainfall during the last 30 years in the central zone in Tanzania focusing on Dodoma region is statistically significant. Additionally, the study examined whether there was a shift in the distribution of rainfall for both short and long rainy seasons in the study area in the past recent years.

All three methods employed in the present study: linear regression analysis by OLS, Kruskal-Wallis and Mann-Kendall's tau nonparametric test statistics gave the same conclusion of none existence of a meaningful statistical trend in the amount of rainfall over the last 30 years. Relative to the years 1981-1989, there appears to be a shift in the amount of rainfall for both long and short rainy seasons during the years 1990-1999. The study recommends for further research that will involve rainfall data from the various zones in Tanzania in order to provide a comparative analysis of the issues examined in this study.

6. REFERENCES

Cody, R.P. and J.K. Smith, 1997. Applied Statistics and the SAS Programming Language. 4th Ed. Prentice-Hall, Inc. ISBN-10: 0137436424, pp: 445.

DAPD, 1999. S-PLUS 2000 Guide to Statistics. 1st Edn., Data Analysis Products Division, Math Soft, Inc., pp: 23.

Funk, C., 2010. A Climate Trend Analysis of Kenya. U.S. Geological Survey Fact Sheet 2010-3074.

Ghosh, S., V. Luniya and A. Gupta, 2009. Trend analysis of Indian summer monsoon rainfall at different spatial scales. Atmospheric Sci. Lett., 10: 285-290. DOI: 10.1002/asl.235

IFPRI, 2006. Poverty and malnutrition in Tanzania: New approaches for examining trends and spatial patterns. International Food Policy Research Institute.

Kahya, E. and S. Kalayci, 2004. Trend analysis of stream flow in Turkey. J. Hydrol., 289: 128-144. DOI: 10.1016/j.jhydrol.2003.11.006.

Lyakurwa, W., 2009. Climate Change and Economic Development in sub-Saharan Africa. South African Institute of International Affairs.

Onoz, B. and M. Bayazit, 2003. The power of statistical tests for trend detection. Turkish J. Eng. Environ. Sci., 27: 247-251.

ORCM, 2005. Tanzania: 2004-05 Demographic and Health Survey. 1st Edn., National Bureau of Statistics Tanzania, pp: 34.

Partal, T. and E. Kahya, 2006. Trend analysis in Turkish precipitation data. Hydrol. Proc., 20: 2011-2026. DOI: 10.1002/hyp.5993.

Pennycuick, L. and M. Norton-Griffiths, 1976. Fluctuations in the rainfall of the Serengeti ecosystem, Tanzania. J. Biogeogr., 3: 125-140. DOI: 10.2307/3038141

Perera, H.K.W.I., D.U.J. Sonnadara and D.R. Jayewardene, 2002. Forecasting the occurrence of rainfall in selected weather stations in the wet and dry zones of Sri Lanka. Sri Lankan J. Phys., 3: 39-52. DOI:10.4038/sljp.v3i0.184

Prins, H.H.T. and P.E. Loth, 1988. Rainfall patterns as background to plant phonology in northern Tanzania. J. Biogeogr., 15: 451-463. DOI: 10.2307/2845275

Racine, J.S., 2008. Nonparametric econometrics: A primer. Foundat. Trends Econ., 3: 1-88. DOI: 10.1561/0800000009

RCO, 2003. Dodoma Region: Socio-Economic Profile. 2nd Edn., National Bureau of Statistics, Tanzania, pp: 219.

Stern, R.D. and R. Coe, 1984. A model fitting analysis of daily rainfall data. J. Royal Stat. Soc. Seri. General, 147: 1-34.

THMIS, 2008. Tanzania HIV/AIDS and Malaria Indicator Survey 2007-08. University of Southampton.

URT, 2007. National Adaptation Program of Action. Vice President's Office, United Republic of Tanzania.

Verbeke, G. and G. Molenberghs, 2000. Linear Mixed Models for Longitudinal Data. 1st Edn., Springer-Verlag, New York, ISBN-10: 0387950273, pp: 568.

Wooldridge, J.M., 2001. Applications of generalized method of moments estimation. J. Econ. Perspectives, 15: 87-100.

Yengoh, G.T., F.A. Armah, E.E. Onumah and J.O. Odoi, 2010. Trends in agriculturally-relevant rainfall characteristics for small-scale agriculture in Northern Ghana. J. Agric. Sci., 2: 3-16.

Zar, J.H., 1984. Biostatistical Analysis. 2nd Edn., Prentice-Hall, ISBN-10: 0130779253, pp: 718.