Research Article **Open Access**

Prediction of Long-term Pattern and its Extreme Event Frequency of Rainfall in Dire Dawa Region, Eastern Ethiopia

Rediat Takele* and Solomon Gebretsidik

Jigjiga University, Jigjiga, Somali regional state, Ethiopia

Abstract

Rainfall is the most critical and key variable both in atmospheric and hydrological cycle. Its patterns usually have spatial and temporal variability. Its variability is assumed to be the main cause for the frequently occurring climate extreme events such as drought and flood. In this study, spectrum analysis, cross-spectral analysis as well as seasonal auto-regressive integrated moving average (SARIMA) model were used in intending to predict the pattern and its extreme event frequency of rainfall in Dire Dawa Region based on data obtained from Dire Dawa and adjacent stations: Dengego and Haramaya. The result indicated that the amount of rainfall at Dengego and Haramaya are more or less the same on average in all seasons and much higher than that of Dire Dawa during last 30 year study period. The variability of annual rainfall in Dire Dawa during the study period is a bit larger than neighboring station's rainfall (Dengego and Haramaya), indicating that climate instability is high in Dire Dawa than other nearby stations. The result also indicates that relatively there is a tendency of increasing pattern in average annual rainfall of Dire Dawa in forecasted period. In the region, the rainfall extreme event like flood predicts or inferred to be recurring at about 4.17 years. Moreover, the rainfall periodicities of Dengego and Dire Dawa are found to be more likely associated, which implies that rainfall extreme event frequency in two districts is statistical significantly associated. It is proposed that assessing the factors that cause the fluctuation in the pattern and frequency of rainfall distribution in the region is needed to be study.

Keywords: Extreme event; Climate variability; Periodicities; SARIMA; Spectral analysis

Introduction

Located within the tropics, Ethiopia has great geographical diversity with high and rugged mountainous, flat-topped plateaus, deep gorges, etc. Its altitudinal range lies between 120 m below sea level and 4600 m above sea level [1]. The differences in altitude and relief create a large variation in climate in various regions of the country [2]. Several regions receive rainfall throughout the year, but in some regions rainfall is seasonal and low making irrigation necessary [3]. Particularly, Semi-arid regions receive very small, irregular, and unreliable rainfall. Rainfall is the most critical and key variable both in atmospheric and hydrological cycle. Its patterns usually have spatial and temporal variability. These variabilities affect agricultural production, water supply, transportation, environment and urban planning and the existence of its people. Its variability assumed to be the main cause for the frequently occurring climate extreme events such as drought and flood.

Ethiopia is one of the countries whose economy is highly dependent on rain-fed agriculture and also facing recurring cycles of flood and drought. Current climate variability is imposing a significant challenge to Ethiopia in general and Dire Dawa in particular. Metrologically, Dire Dawa Administration is characterized by an arid and semi-arid climate, thus, receives low and erratic rainfall [4]. Despite the fact that, sloppy topography of the Dire Dawa administration are surrounded by the neighboring mountainous areas (Haramaya and Dengego). These are the main catchments for contributing to the disaster flood and drought event in Dire Dawa. Flood disasters are occurring more frequently in the urban area, and are having an ever more dramatic impact in terms of the costs on lives, livelihoods and environmental resources. The May 1984 catastrophic flood claimed 42 people, and property worth 10 million birr was lost. On August 2006 claimed 650 people, displaced 35,000 people, and caused direct damage of huge infrastructure estimated at 100 million Ethiopian Birr and indirect damage of similar magnitude. On April 2010- though there was the property of worth 28 million birr had been lost. On the other hand, the area has often experienced and affected by frequent and prolonged drought claiming the lives of millions of people. The impacts of the 2004 and 2005 droughts incident posed food shortage up to 85% of the rural population [5].

Methods of prediction of rainfall extreme events have often been based on studies of physical effects of rainfall or on statistical studies of rainfall time series [2,3,6]. The study in various region of African shows that rainfall has Periodic tendencies [7,8]. A study made by Mersha and Seifu [9,10] on rainfall cyclicity over selected stations in Ethiopia shows that there appears to be periodic tendency in the annual rainfall series, particularly, Gode, Dire Dawa, Jigjiga, Negelle, and Debre Zeit station. Haile [11] found that drought occur at every 6-8 years in the semi-arid regions of Ethiopia. Tsegay [12] on his study the occurrences of drought and frequencies of rainfall suggest that drought occur every 3-5 years in Eastern Ethiopia. On the other side, the study by Yilma and Alamerew et al. [2,3] reach the conclusion of rainfall extreme event recurs in Addis Ababa region between 10 to 11 years.

Awareness about the characteristics of the rainfall over an area such as the source, quantity, variability, distribution and the frequency of rainfall is essential for efficient control and management of water resources, the implication in utilization and associated problems.

The impacts of climate variability and climate change are evident on almost all socio-economic sectors. The output of this study can play an important role in the design of hydrological structures in the area of study and help the policy makers in improve their decisions by taking into consideration the available and future water resources.

Materials and Methods

Data and variable of the study

A time series of monthly 30 years rainfall data in mm for the period January, 1984 to January, 2014 collected by the National Meteorological Agency of Ethiopia were used in the study. The data were collected from the synoptic stations of Dire Dawa and adjacent station: Dengego and Haramaya for further investigation.

Methodology

The development of climatology as a science has given rise to growing statistical applications on climatic information. Conducting investigations using standard statistical methodologies is an essential step in the development of climatology [13]. For instance, time series analysis is used in order to evaluate the temporal and spatial behavior of rainfall [14]. In this study a univariate Box-Jenkins Methods, in particular, Seasonal Autoregressive Integrated Moving Average (SARIMA) methods and spectral analysis and cross-spectral analysis are employed [15].

Seasonal autoregressive integrated moving average (SARIMA) modeling is one of the most widely implemented methods for analyzing univariate time series data that exhibits a seasonal variation such as temperature and rainfall, which have strong components corresponding to seasons. Hence, the natural variability of many physical, biological, and economic processes tends to match with seasonal fluctuations [16]. The seasonal autoregressive integrated moving average model, denoted by SARIMA $(p,d,q)*(P,D,Q)_s$, as in Shumay and Stoffer [17] and is given by:

$$\Phi_P(B^s)\phi(B)\nabla_S^D x_t = \alpha + \Theta_Q(B^s)\theta(B) w_t \quad (1)$$

where w_t is the usual Gaussian white noise process. The ordinary autoregressive and moving average components are represented by polynomials $\phi(B)$ and θ (B) of orders p and q, respectively and the seasonal autoregressive and moving average components by Φ_P (B^s) and Θ_Q(B^s) of orders P and Q. The ordinary and seasonal difference components can be written as: $\nabla^d = (1-B)^d$ and $\nabla_S^D = (1-B^S)^D$

Test for stationary assumption

Before developing a Box-Jenkins modeling process, it is important to check whether the data under study meets basic assumptions such as stationary. A time series is said to be stationary if there is no systematic change in mean (no trend), if there is no systematic change in variance and if periodic variations have been removed. Many testing procedures for testing stationary are employed in the literature.

Time plot: This procedure may reveals seasonality, trends either in the mean level or the variance of the series, long- term cycles, and so on. If any such patterns are present, then these are signs of non-stationary.

The correlagram test: test based on the plot its sample autocorrelation function and sample partial autocorrelation function. Both function used to reveal important information regarding given series stationary. Wei [18] states that if the sample ACF decays very slowly, which indicates that series is not stationary.

The augmented dickey-fuller (ADF) test: This test first poses the null hypothesis that the given time series has a unit root, which means that the time series is non-stationary. That is:

H$_0$: ϕ=1 (has a unit root) Vs H$_1$: ϕ< 1 (has root outside the unit circle)

To use the Augmented Dickey-Fuller test the following test equation will be used:

$$\nabla x_t = \theta + \beta t + \varphi x_{t-1} + \theta_1\nabla x_{t-1} + ... + \theta_p\nabla x_{t-p} + w_t \quad (2)$$

Where: ∇x_t is the first differenced value of series (x_t), w_t is the error term, ∇x_{t-j} is the jth lagged first differenced of values of x_t, $\theta_o, \beta, \Phi = \phi - 1$, θ_1, θ_2,..., θ_p are parameters to be estimated

Variance comparisons: The characteristics of the sample variances associated with different orders of differencing can provide a useful means of deciding the appropriate order of differencing [19] cited also in Alamarew et al. [3]. The sample variances will decrease until a stationary sequence has been found.

Building seasonal -ARIMA models

To identify a perfect ARIMA model for a particular time series data, Box and Jenkins [16] proposed a methodology that consists of four phases: i) Model identification; ii) parameters Estimation; iii) Diagnostic checking for the identified model and iv) Forecasting.

Model identification: The purpose of the identification stage is to determine the differencing required achieving stationarity and also the order of both the seasonal and the non- seasonal AR and MA operators for the residual series. The autocorrelations function (ACF) and the partial autocorrelation functions (PACF) are the two most useful tools in any attempt at time series model identification. To identify the order of model parameters, we examine the ACF and PACF based on the theoretical pattern using Tables 1 and 2 as summarized in Shumway and Stoffer [17].

Parameter estimation: After choosing the most appropriate order of the model, the next step is to estimate the model parameters $(\phi_1,...,\phi_p, \Phi_1,...\Phi_P, \theta_1,...,\theta_q, \Theta_1,...,\Theta_Q, \delta_k^2)$ by using several estimation procedures like maximum likelihood methods and least square methods.

In time series analysis, there may be several adequate models that can be used to represent a given data set, and hence, information criteria are used for model comparison. The AIC and SBC is a mathematical selection criterion of model building by select the model that gives the minimum of the AIC and SBC defined by Shumway and Stoffer [17]:

	AR(P)	MA(Q)	ARMA(P,Q)
ACF	Tails off	Cuts off after lags Q	Tails off
PACF	Cuts off after lags P	Tails off	Tails off

Source: Shumway and Stoffer, [17].

Table 1: Behavior of the ACF and PACF for ARMA Models.

	AR(P)s	MA(Q)s	ARMA(P,Q)$_s$
ACF	Tails off at lags Ks, K=1, 2 ,….	Cuts off after lags Qs	Tails off at lags Ks
PACF	Cuts off after lags Ps	Tails off at lags ks	Tails off at lag Ps

Source: Shumway and Stoffer, [17].

Table 2: Behavior of the ACF and PACF for Pure SARMA Models.

$$AIC = \log \hat{\delta}_k + \frac{n+2k}{n} \text{ and } SBC = \log \hat{\delta}_k^2 + \frac{k \log n}{n} \qquad (3)$$

Where: $\hat{\delta}_k^2 = \dfrac{SSE_k}{n}$ denotes the maximum likelihood estimator for the error variance and k is the order of seasonal and non-seasonal autoregressive and moving average parameters to be estimated in the model, that is, according to Wei [18], k=p + q + P+ Q + 1 and n is the number of observations in the series.

Diagnostic checking: After fitting a provisional time series model, we can assess its adequacy in various ways. Most diagnostic tests deal with the residual assumptions in order to determine whether the residuals from fitted model are independent, have a constant variance, and are normally distributed. Several diagnostic statistics and plots of the residuals can be used to examine the goodness of fit of the tentative model to the time series data.

The first approaches that can be used to evaluate the adequacy of a model are the plot of the errors over time, which can be written Shumway and Stoffer [17]:

$$w_t = \left(x_t - x_t^{t-1}\right) / \sqrt{p_t^{t-1}} \qquad (4)$$

Where $x_t - x_t^{t-1}$ is the one-step-ahead prediction of (x_t) based on the fitted model and $\sqrt{p_t^{t-1}}$ is the estimated one-step-ahead error variance. If visual inspections of the errors reveal that they are randomly distributed over time, then we have a good model.

The autocorrelations function (ACF) of the series can also be used to examine whether the residual of the fitted model is white noise or not. Under the null hypothesis that residual follows a white noise process, roughly 95% of the autocorrelation coefficient should fall within the range $\pm 1.96 / \sqrt{T}$ [20].

Ljung-Box Q (LBQ) test: It can be employed to check independence of residual instead of visual inspection of the sample autocorrelations. This test uses the null hypothesis that

Ho: $\rho_1 = \rho_2 = \dots = \rho_k = 0$ and the test statistic is calculated as Ljung and Box [21].

$$Q(r) = n'(n'+2) \sum_{j=1}^{k} \frac{\mu_j^2}{n-j} \qquad (5)$$

Where: $n' = (n-d)$, n is the number of observations in the original time series, μ_j^2 is the sample autocorrelation of the residuals at lag j and d is the degree of non- seasonal differencing used to transform the original time series values into stationary time series. If a model is correctly specified, residuals should be uncorrelated and Q(r) should be small (p- value should be large).

Runs test: The runs test can be used to decide if a data set is from a random process. The first step in the runs test is to compute the sequential differences $(Y_i - Y_{i-1})$. If $Y_i > Y_{i-1}$ a 1 (one) is assigned for an observation and a 0 (zero) otherwise. let T be the number of observations, T_a be the number above the mean, T_b be the number below the mean and R be the observed number of runs. Then the test statistic for this test is given as Cromwell et al. [22]:

$$Z_{N=} \frac{R - E(R)}{\sqrt{V(R)}} \approx N(0,1) \text{ where:} \begin{array}{l} E(R) = T + 2T_a T_b / T \\ V(R) = 2T_a T_b (2T_a T_b - T) \Big/ 2T(T-1) \end{array} \qquad (6)$$

The test of series randomness is rejected if the calculated Z_N value exceeds the selected critical value obtained from the standard normal distribution table.

Turning point test: A turning point means when the series changes from increasing to decreasing or vice versa. That is, $X_{t-1} < X_t > X_{t+1}$ or $X_{t-1} > X_t < X_{t+1}$. Let T=the number of turning points in an n- period series. The hypothesis of residual whitenoise should be rejected if the absolute value of N_T(eqn.(7)) $> N_{T(1-\alpha/2)}$, where $N_{T(1-\alpha/2)}$ is the (1- α /2) quartile of standard normal distribution Cromwell et al. [22]:

$$N_T = \frac{|T - \mu_T|}{\sqrt{Var(T)}} \approx N(0,1) \text{ Where:} \begin{array}{l} \mu_T = (2/3)(n-2) \\ Var(T) = (16n - 29) / 90 \end{array} \qquad (7)$$

Test for normality of the residuals: The Shapiro-Wilk Test, histogram and Q-Q plot of a data set give information related to normality. The null hypothesis for any test of normality is that the data are normally distributed. The Shapiro-Wilk statistic "W", which used as a base for the test of normality of residual in Shapiro-Wilks test, is given as [23]:

$$W = \frac{\left(\sum_{i=1}^{n} a_{i,n} x_i\right)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \qquad (8)$$

Procedure for checking of homosecdasticity

Goldfeld-Quandt test: If there is a change in variance (heteroscedasticity) of residuals, a transformation is necessary for the data. For this test:

If test statistic: $F_{cal} = S_2^2 \Big/ S_1^2$ is smaller than the critical value $(F_\alpha(n_1 - 2, n_2 - 2))$, the residuals are assumed to be homoscedastic.

Breusch-Pagan test: This involves applying ordinary least-square (OLS) to:

$\dfrac{\hat{\varepsilon}_i^2}{\hat{\sigma}^2} = \gamma_0 + \gamma_1 X_{1i} + \gamma_2 X_{2i} + \dots + \gamma_K X_{Ki} + u_i$ and calculating the regression sum of squares (RSS). The test statistic is: $\chi_{cal}^2 = \dfrac{RSS}{2}$. Reject the null hypothesis of homoscedasticity if: $\chi_{cal}^2 > \chi_\alpha^2(K)$.

If the constant variance and normality assumptions are not true, they are often reasonably well satisfied when the observations are transformed by a Box-Cox transformation [17]:

$$x_t^{(\lambda)} = \begin{cases} \dfrac{x_t^\lambda - 1}{\lambda}, \lambda \neq 0 \\ \ln x_t, \lambda = 0 \end{cases} \text{ Choose the value of } \lambda \text{ that maximizes:}$$

$$l(\lambda) = -\frac{n}{2} h\left[\frac{1}{n} \sum_{j=1}^{n} (x_j^{(\lambda)} - \bar{x}^{(\lambda)})^2\right] + (\lambda - 1) \sum_{j=1}^{n} h \; x_j \qquad (9)$$

If the selected model is inadequate, the three-step model building process is typically repeated several times until a satisfactory model is finally obtained. The final selected model can then be used for prediction purposes [18].

Forecasting: The last step in time series modeling is forecasting.

There are two kinds of forecasts: sample period forecasts and post-sample period forecasts. The former are used to develop confidence in the model and the latter to generate genuine desired forecasts. In forecasting, the goal is to predict future values of a time series, x_{t+m}, $m =1, 2,...$ based on the data collected to the present, $x=\{x_t, x_{t-1},..., x_1\}$. Then explicit forms of the model for the observation x_{t+m} generated by the ARIMA process may be expressed as follows (Box and Jenkins) [16]:

$$x_{t+m} = \phi_1 x_{t+m-1} + ... + \phi_{p+d} x_{t+m-p-d} - \theta_1 w_{t+m-1} - ... - \theta_q w_{t+m-q} \qquad (10)$$

Forecasting accuracy measures

Once forecasts are made they can be evaluated if the actual values of the series to be forecasted are observed. To assess the precision of the forecasts, prediction interval can be written as [17]:

$$x^t_{t+m} + C_{\frac{\alpha}{2}} \sqrt{P^t_{t+m}} \text{ , Where } C_{\frac{\alpha}{2}} \text{ is chosen to get the desired}$$

degree of confidence. Where: Mean square prediction error,

$$P^t_{t+m} = E(x_{t+m} - \tilde{x}_{t+m})^2 = \delta_w^2 \sum_{j=0}^{m-1} \psi_j^2$$

There are also some measurements of the accuracy of forecasts, which evaluate the accuracy by comparing with observed value are Mean Square Error (MSE); Mean Absolute Error (MAE); Mean Absolute Percentage Error (MAPE); Theil's inequality coefficient (U-Statistics). The scaling of Uis such that it will always lie between 0 and 1. If U=0, there is a perfect fit; if U=1 the predictive performance is as bad as it possibly could be. Moreover, in evaluating the performance of the forecasting models, the lower the RMSE, MAE, MAPE, the better the forecasting accuracy [24].

Spectral analysis: Quite often, hydrologic phenomenon depicts cyclic and stochastic processes. The Spectral analysis based on periodogram method has dominated hydrology for years because of underlying periodicities in hydrologic processes [25]. The main objective of spectral analysis is to estimate and study the spectrum. It is therefore concerned with estimating the unknown spectrum of the process from the data and quantifying the relative importance of different frequency bands to the variance of the process [26]. The spectrum of a time series is the distribution of variance of the series as a function of frequency. The area under the curve in spectrum is proportional to the variance. Hence, we can be estimating the power spectrum by the raw periodogram method as [17]:

$$\bar{f}(\omega) = n[R(\omega)]^2 = n|d(\omega)|^2 \qquad (11)$$

$$\text{Where: } |d(\omega)|^2 = \frac{1}{n}\left[\left(\sum_{t=1}^{n}(x_t - \bar{x})\cos(2\pi\omega t)\right)^2 + \left(\sum_{t=1}^{n}(x_t - \bar{x})\sin(2\pi\omega t)\right)^2\right]$$

and the frequencies (ω_j) are related to the size (n) by: $\omega_j = \frac{j}{n}$, $1 < j < \frac{n}{2}$

By considering the fact that the periodogram estimates (eqn.11) are independent and identically approximated is χ^2-distributed with v-degree of freedom, the approximate $100(1-\alpha)\%$ confidence interval for the spectrum:

$$\frac{2I(\omega_j)}{x_v^2(1-\frac{\alpha}{2})} \leq f(\omega) \leq \frac{2I(\omega_j)}{x_v^2(\frac{\alpha}{2})} \qquad (12)$$

The raw periodogram is a wildly fluctuating estimate of the spectrum with high variance. For a stable estimate, the periodogram must be smoothed. Bloomfield [27] recommends the Daniell window as a smoothing filter for generating an estimated spectrum from the periodogram. The smoothed periodogram is defined as [17]:

$$\hat{f}(\omega) = \sum_{k=-m}^{m} h_k I(\omega_j + k/n); \text{ Where the weights } h_k > 0 \text{ satisfy,}$$

$$\sum_{k=-m}^{m} h_k = 1 \qquad (13)$$

The result in eqn. (13) can be rearranged to obtain an approximate $100(1 - \alpha)$ % confidence interval for the best estimator of spectrum, $f(\omega)$ of the form

$$\frac{2L_h \hat{f}(\omega)}{x_{2Lh}^2(1-\alpha/2)} \leq f(\omega) \leq \frac{2L_h \hat{f}(\omega)}{x_{2Lh}^2(\alpha/2)} \qquad (14)$$

Cross- spectral Analysis: Cross-spectrum analysis examines the relationship between a pair of series. The main objective of cross-spectrum analysis is to determine the existence of correlation between two variables and to simultaneously separate each of two time series in to its harmonic component. The principal tool that measures the strength of relation between two series in cross-spectral analysis is squared coherence function defined as [17]:

$$\rho_{xy}^2(\omega) = \frac{\left|f_{xy}(\omega)\right|^2}{f_{xx}(\omega)f_{yy}(\omega)}, \quad -\frac{1}{2} \leq \omega \leq \frac{1}{2} \qquad (15)$$

Then spectral Matrix of a Bivariate Process is given by:

$$f(\omega) = \begin{pmatrix} f_{xx}(\omega) & f_{xy}(\omega) \\ f_{yx}(\omega) & f_{yy}(\omega) \end{pmatrix} \qquad (16)$$

An estimate of the above spectral matrix is similar to the eqn. (13) defined in power spectra. Then an estimate of the squared coherence between two series, x_t and y_t is given by:

$$\hat{\rho}_{xy}^2 = \frac{\left|\hat{f}_{xy}(\omega)\right|^2}{\hat{f}_{xx}(\omega)\hat{f}_{yy}(\omega)} \qquad (17)$$

At a particular significance level (α), the approximate critical value that must be exceeded for the original squared coherence to be able to reject $\rho_{xy}^2 = 0$ at a specified frequency is [17]:

$$C_\alpha = \frac{F_{2,2L-2}(\alpha)}{L-1+F_{2,2L-2}(\alpha)} \qquad (18)$$

Result and Discussions

The statistical software package used for most of the analysis is by the help of R-15.10.1 and for some test and graphs; however, SAS-9.2 was used.

Descriptive analysis

Here it can be described rainfall of Dire Dawa, Haramaya and Dengego station and mainly focused on analyzing monthly as well as annually rainfall data of Dire Dawa station recorded by the National

Meteorological Agency of Ethiopia (NMAE). From descriptive statistics result, it can be observed that the mean annual rainfall of Dire Dawa, Dengego and Haramaya are 611 mm, 774 mm and 772 mm, respectively. In Dire Dawa, the minimum annual rainfall of 228.3 mm was recorded during the year 1984 and the maximum of 719.7 mm in the year 1997. The amount of rainfall at Dengego and Haramaya are more or less the same on average in all seasons, and is much higher than that of Dire Dawa over the study period. NMA [28] documented that a series with coefficient of variation(CV) less than 0.20 can be considered as less variable, between 0.20 and 0.30 is moderately variable, and greater than 0.30 is highly variable. The monthly rainfall data in all stations show that there is high variability since their CV is greater than 0.30. However, relatively the monthly variability of rainfall of Dire Dawa and Dengego are similar. When we see the overall variability, the high variability occurs during months of dry seasons while the low variability is observed during the rainy season in all stations. The variability of annual rainfall in Dire Dawa during the last 30 year period is a bit larger than neighboring stations. This may indicate that climate instability is high in Dire Dawa than other stations.

Testing stationary assumption

Visual Inspection: Regardless of which technique is used, the first step in any time series analysis is to plot the observed values against time. The pattern of the time series plot in Figure 1 does not show any systematic upward and downward change about the mean. This indicates that the series is non-seasonally stationary. Moreover, the sample autocorrelation function (SACF) plot in Figure 2, the presence of seasonality behavior and seasonally non-stationarity of the mean monthly rainfall series is clear. Because it shows strong seasonal wave pattern at the multiple of seasonal intervals(s=12) and declining slowly while non-seasonal lags are relatively decaying quite rapidly. This can be interpreted as a 6-month seasonal pattern that cycles between summer when there is little to no rainfall, and winter when rain is at its peak.

Therefore, from the actual series plot and sample autocorrelation plot (Figures 1 and 2) we observed that our series has seasonal variation, and seasonal differencing is needed to make it stationary.

Unit root test: This test first impose that the mean monthly rainfall
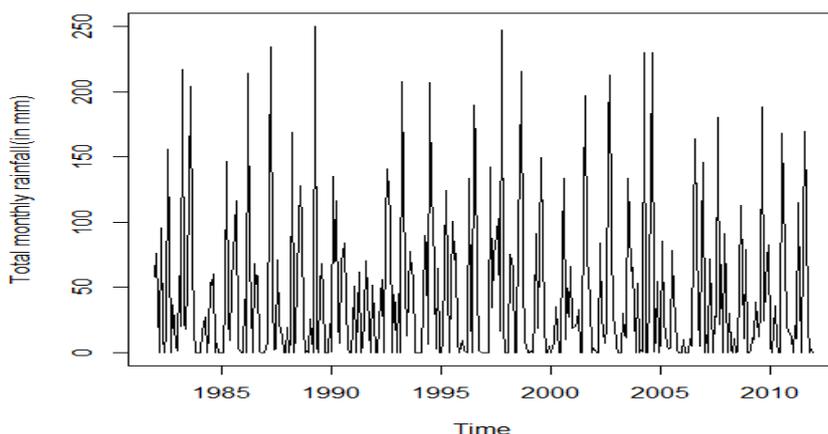
Figure 1: Time plot for total monthly rainfall in mm of Dire Dawa.
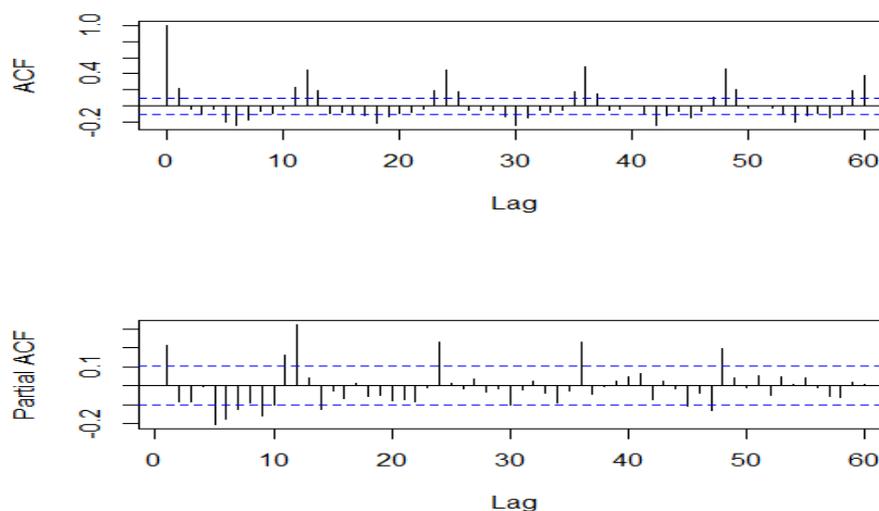
Figure 2: Sample ACF and PACF plots for the original rainfall series of Dire Dawa.

series has a unit root, versus the series is stationary. We test the null hypothesis using the Augmented Dickey-Fuller (ADF) testto know whether our series are stationary or not at level or after differencing. From Figure 1, we observe that the rainfall series of Dire Dawa doesn't have a trend and potentially slow- turn around zero. Therefore, we can use the test equation given in equation (2). For our series, maximum lag lengths (P) of the ADF test to remove serial correlation from the residuals of the regressions were selected as 12. The results of ADF test for the mean monthly rainfall series of Dire Dawa at level (without differencing), after first regular and seasonal differencing shows that, the ADF test statistic for the original and first regular differencing monthly rainfall series are greater than the critical values at 1%, 5% and 10% significance levels. This implies that, the null hypothesis, which has a unit root, for the series should not be rejected at all significance levels. This figure further confirms that original series as well as series obtained after first regular differencing are not stationary. However, as also suggested by visual analysis, after first seasonal differencing, the computed ADF test statistic are smaller than the critical values at 1%, 5% and 10% significant levels. This leads to the rejection of the test that there is a unit-root problem.

Variance Comparison: Increase in the differencing order tends to increase the variance indicating over differencing [29]. In this study the following results were obtained:

Var(∇x_t)=5127.7, Var(x_t)=3252.7, Var($\nabla_{12} x_t$)=2704.76, and $Var(\nabla_{12}^2 x_t)$ =7981.4 values. From these it is clear that: Var(∇x_t) >Var(x_t) and Var ($\nabla_{12}^2 x_t$) >Var(x_t) >Var($\nabla_{12} x_t$). These results again suggest that non-seasonal first differencing (∇x_t) has been over-differenced and hence the original series is non-seasonally stationary. The first seasonal differencing would rather be important (Figure 3).

Consequently, all tests for stationarity seem to agree and suggest that the first-seasonal differencing of the series make stationarity around a constant mean, which is approximately zero and calculated its standard deviation of 52.05 mm (Figure 4).

Moreover, to determine whether stationarity has been achieved, either by trend removal or by differencing, one may examine the autocorrelation function (ACF) and partial autocorrelation function (PACF) of series [30]. The sequence corresponding to a stationary process should converge quite rapidly to zero as the value of the lag increases. From this point of view, Sample autocorrelation and partial autocorrelation plot shown in Figure 5 are in support of monthly rainfall series stationary after having first seasonal difference.

To test whether residual from fitted model come from normally distributed series, we use histogram, QQ-plot of the residual and Shapro-wilks test. The histogram and QQ-plot of the residual shown in Figure 6a, show skewed rather than normally distribution. Shipiro-Wilks test also confirms this fact since it results in P-value of 7.03e-09, which is smaller than 0.05. This implies that the residuals from the fitted models do not come from a normal distribution. This therefore, suggests some kind of transformation for our monthly rainfall series

Figure 3: Plot for total monthly square root transformed rainfall series of Dire Dawa.

Figure 4: Plot for first seasonal differenced square root transformed rainfall series of Dire Dawa.
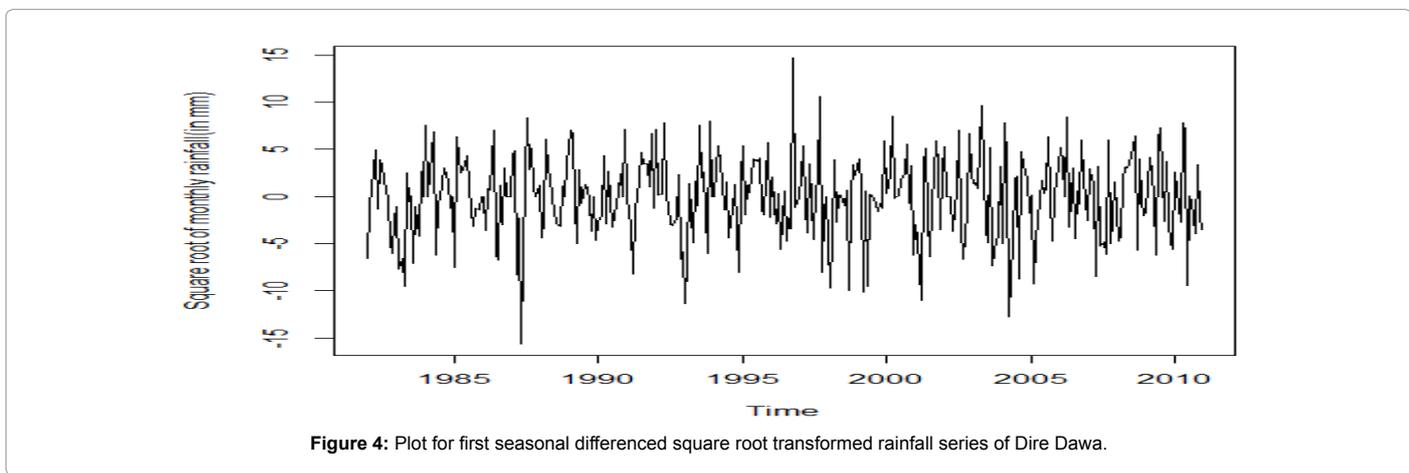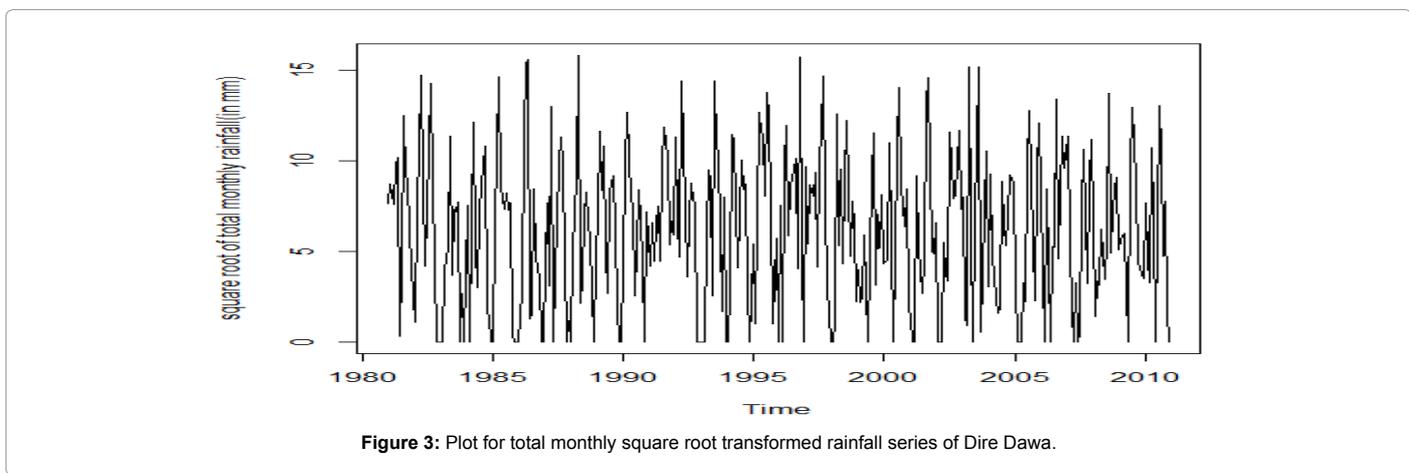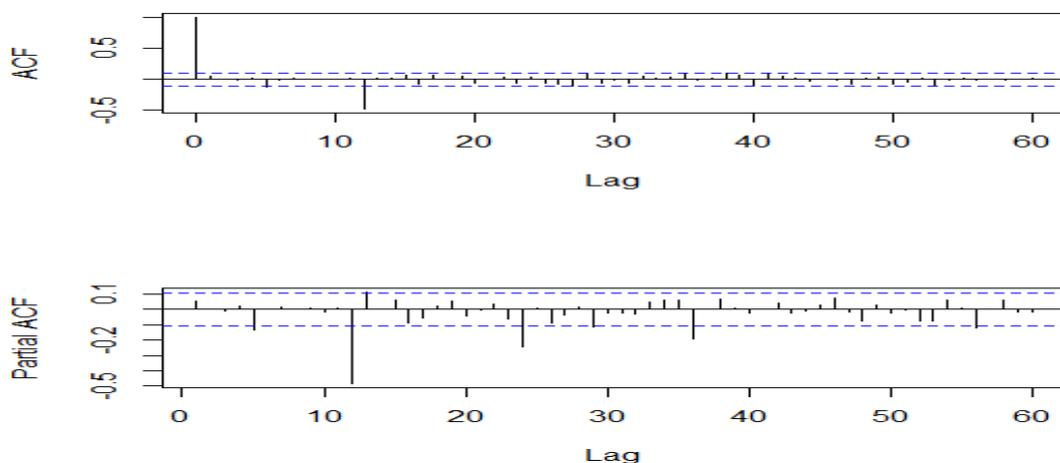
Figure 5: Sample ACF and PACF for first seasonal differenced transformed rainfall of Dire Dawa.
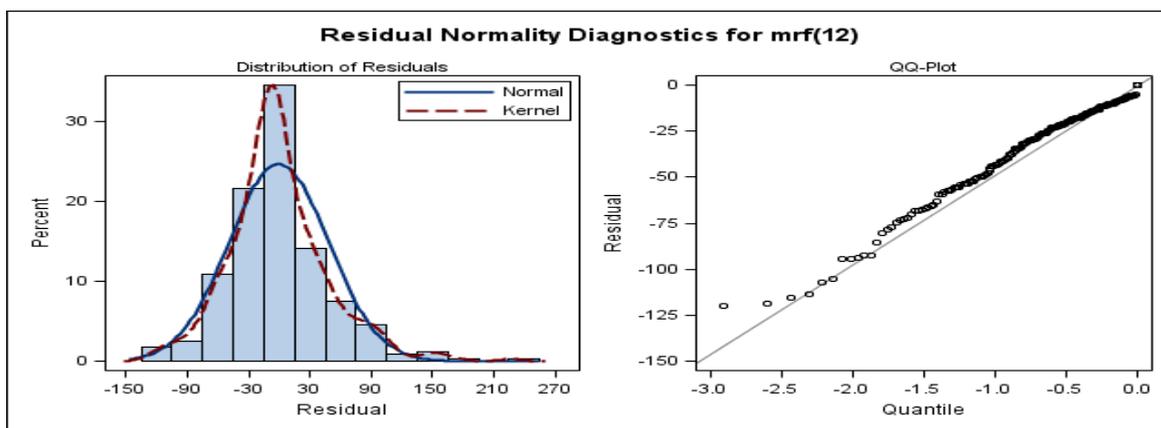
(a)
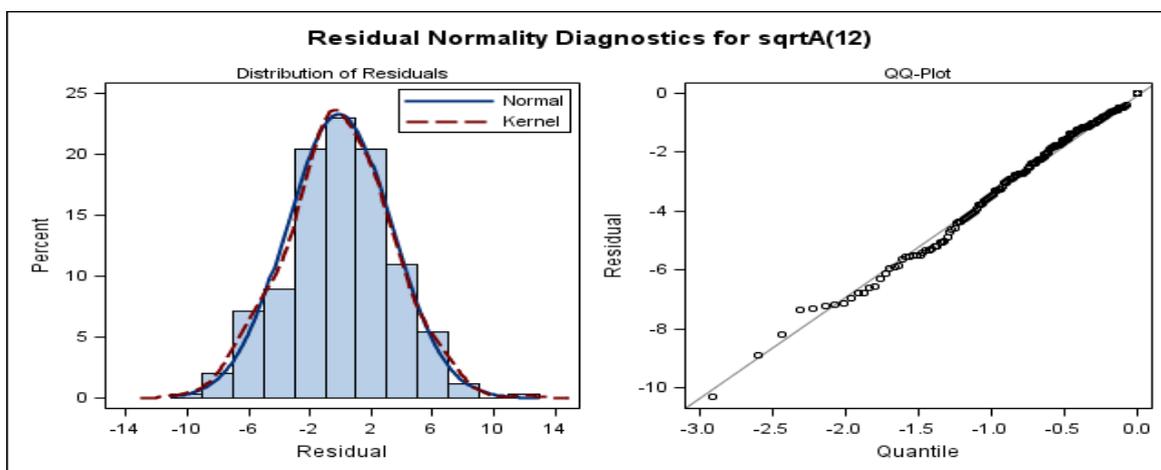
(b)

Figure 6: Normality Diagnostics plot for: (a) untransformed series model, (b) square root transformed series.

to achieve normality. By applying Box-Cox transformation defined in eqn.(9) to the monthly rainfall series, the normality assumptions of residual will be achieved using optimum value λ = 0.5, which maximizes the likelihood function with various iteration by the help of SAS 9.2-software. Figure 6b reveals that after this transformation the problem of non-normality seems dealt with. The Shapiro–Wilk test result based on test statistic defined on eqn.(8) (p-value=0.397) is also in support of the normality of transformed series.

Since normality is not fulfilled for the original series, three stages model building is performed with help of Sample ACF and PACF plot for the square root transformed rainfall data of Dire Dawa.

Model building for monthly rainfall of dire dawa

Since monthly rainfall data has seasonality variation, the univariate Seasonal Authoregressive Integrated Moving Average methodology is used to model rainfall series of Dire Dawa region.

Model identification: Once the degree of differencing has been determined, the autoregressive and moving-average orders are selected by examining the sample autocorrelations and sample partial autocorrelations function. Thus, it considered the SACF and SPACF shown in Figure 5 above and Tables 1 and 2 [17] as a guide and preliminary values of p, q, P and Q are chosen. Because we are dealing with estimates, it will not always be clear whether the sample ACF or PACF is tailing off or cutting off. Also, two models that are seemingly different can actually be very similar. With this in mind, we should not worry about being so precise at this stage of the model fitting.

At this stage, a few preliminary values of p, q, P and Q should be at hand, and we can start estimating the parameters. First, concentrating on the seasonal lags, the characteristics of the ACF and PACF of our transformed series in Figure 2 tend to show a strong peak at h=12 in the autocorrelation function, combined with peaks at h=12, 24, 36 in the partial autocorrelation function. Hence it appears that either: the ACF is cutting off after lag 12 and the PACF is tailing off in the seasonal lags, or the ACF and PACF are both tailing off in the seasonal lags. Consequently, either (i) a seasonal moving average of order $Q=1$, or (ii) due to the fact that both the ACF and PACF may be tailing off at the seasonal lags, perhaps both components, $P=1$ and $Q=1$, are needed. To identify the between-season model, we focus the lags h=1, 2…11 and identify order based on Table 1. First, we set the ACF to be tailing-off and the PACF to be cut-off after lag 5, we identify p=5 and q=0. Also it is possible to think of the PACF to be tailing-off and the ACF to cut-off after lag 5, leading to identify P=0 and Q=5. Fitting the models suggested by these observations, it can be obtained:

SARIMA $(0, 0, 5) \times (0, 1, 1)_{12}$; SARIMA $(5, 0, 0) \times (0, 1, 1)_{12}$

SARIMA $(0, 0, 5) \times (1, 1, 1)_{12}$; SARIMA $(5, 0, 0) \times (1, 1, 1)_{12}$

Parameter estimation: Here the maximum likelihood estimation methods for transformed total monthly rainfall are used to estimate the model parameters and the results are summarized in Table 3, which displays the list of the parameters for each temporally entertained model. For each model parameter, the table presents the estimated value, standard error, t- value, AIC, SBC and variance $(\hat{\delta}_k^2)$ for the estimate. As indicated by Shumway et al. [17], parameters must differ

Model	Parameter	Estimate	Std.Error	T-ratio	P-value	Criteria
(i)	μ	-0.002	0.039	-0.05	0.959	AIC=1863.206
	θ_5	0.165	0.053	3.09	0.002	SBC=1874.763
	Θ_{12}	0.775	0.037	20.88	<.0001	$\delta_k^2 = 11.81$
(ii)	μ	-0.002	0.039	-0.05	0.959	AIC=1860.17
	ϕ_5	-0.181	0.053	-3.41	0.0006	SBC=1870.73
	Θ_{12}	0.777	0.037	20.99	<.0001	$\delta_k^2 = 10.71$
(iii)	μ	-0.002	0.039	-0.05	0.959	AIC=1865.19
	Θ_{12}	0.165	0.053	3.09	0.002	SBC=1879.61
	Θ_{12}	0.772	0.037	20.88	<.0001	
	Θ_{12}	-0.007	0.071	0.09	0.925	$\delta_k^2 = 12.19$
(iv)	μ	-0.002	0.039	-0.05	0.959	AIC=1864.13
	ϕ_5	-0.181	0.053	-3.41	0.0006	SBC=1878.54
	Θ_{12}	0.775	0.037	20.88	<.0001	
	Θ_{12}	-0.011	0.071	-0.16	0.877	$\delta_k^2 = 11.85$

Table 3: Summary of model Parameter Estimates and model selection measures.

significantly from zero and all significant parameters must be included in the model. The result indicated that the seasonal and non-seasonal moving averages as well as non-seasonal autoregressive parameters are all significant since their p-values is smaller than 0.05 and should be retained in the model. However, the constant (μ), non-seasonal autoregressive parameters $(\phi_1, \phi_2, \phi_3, \phi_4)$ and seasonal autoregressive parameters (ϕ_2) in all selected models are insignificant; hence it should be omitted from the model. Moreover, the correlation of the parameter estimates is also used to assess the extent to which collinearity may have influenced the results. If two parameter estimates are very highly correlated, you might consider dropping one of them from the model. However, in our case, correlation between seasonal moving average and non-seasonal autoregressive parameter is -0.079, which indicate that less correlation is observed between parameter estimates.

Finally, the AIC, SBC and variance of the estimate $(\hat{\delta}_k^2)$ in Table 3 are computed according to eqn. (3) and used to compare selected models fit best to the monthly rainfall series. From this, SARIMA (5, 0, 0)*(0, 1, 1)$_{12}$ model has lower AIC, SBC and variance of estimate than other model, so it fits the transformed monthly rainfall series of Dire Dawa better than other, and can be used for further analysis.

Diagnostic checking: If the model fits the rainfall data well, the residuals of the fitted model are random or white noise [31].

Inspection of the time plot of the standardized residuals for square root transformed rainfall series in Figure 7, which can be obtain by help of eqn.(4) above, shows no clear patterns (trend or seasonality behavior).

Under the test that residual follows a white noise process, roughly 95% of the residual autocorrelations should fall within the range of $\pm 1.96/\sqrt{360}$. It is clear, as shown in Figure 7 that there is no pattern in the residuals autocorrelation plot for the selected model, which means there is no autocorrelation coefficient which lies outside the two standard errors significantly for the fitted models. The Ljung Q-statistic again for each group of six lags are computed using Eq. (5) and the results are summarized in Table 4 below and Figure 6, indicate that the p-value is greater than 0.05 for all lags, which implies that the white noise hypothesis is not rejected. Moreover, two alternative methods, namely runs and turning point tests, are applied to check the independence assumption of residuals of the fitted model. The fitted model was found to be consistent with the independence assumption for both methods since their test statistic is smaller than the critical values. Thus, those results also closely follow with the result plot of the standardized residuals in Figure 7 and Table 4.

The results related to the normality of residuals using Shapiro–Wilk tests indicate the residuals of the fitted models is normally distributed since p-value is greater than 0.05. In addition to these tests, Figure 6b shows the histograms and QQ-plot of the residuals. As expected,

Figure 7: Independence diagnostics of the residuals from the fitted SARIMA (5,0,0) x (0,1,1)$_{12}$ model.
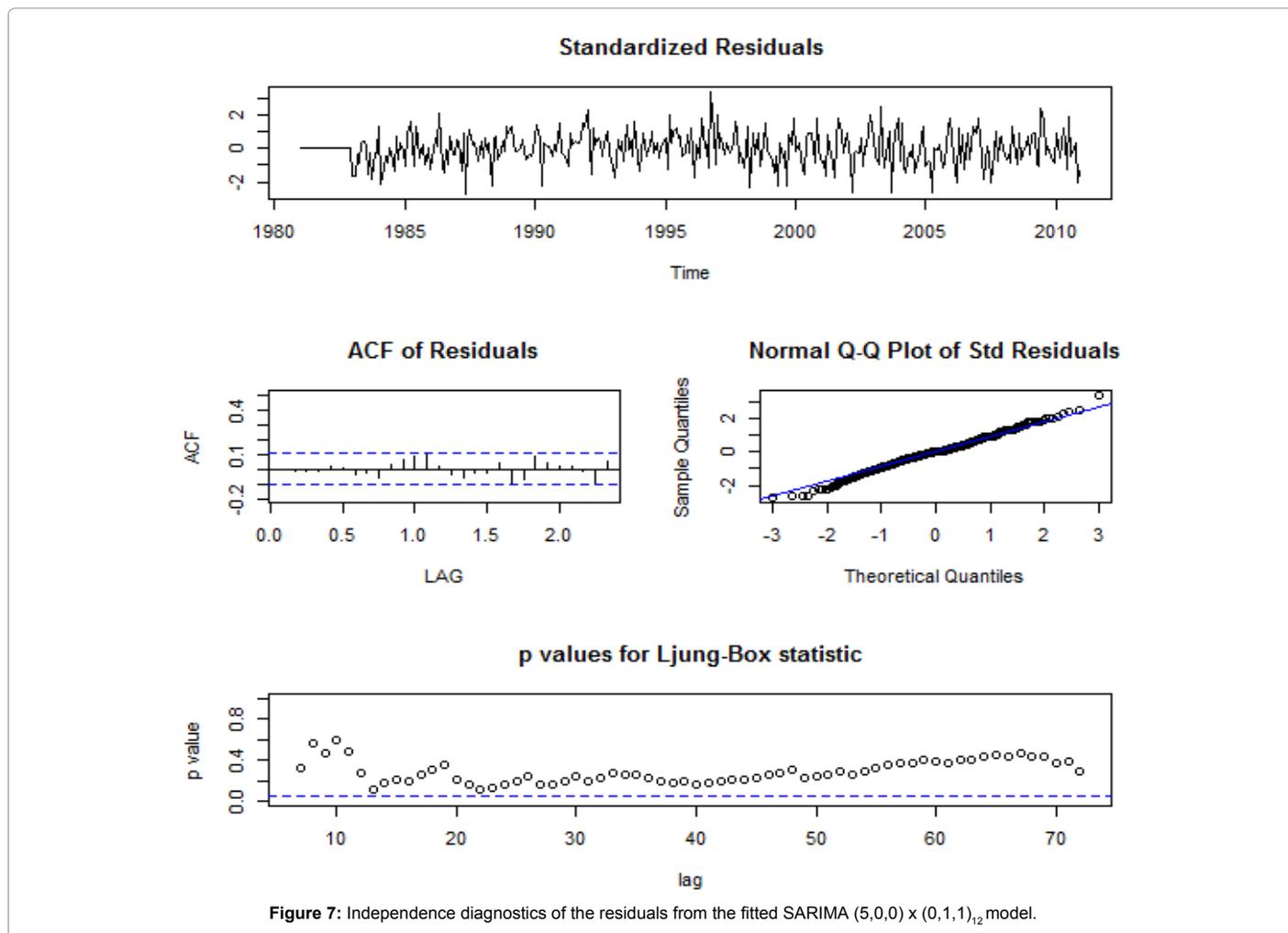
To Lag	Chi- Square	DF	P-Value	Autocorrelations					
6	7.60	4	0.1072	0.03	0.033	-0.021	-0.017	0.008	-0.032
12	12.24	10	0.2695	-0.064	-0.034	-0.023	0.040	0.073	-0.006
18	17.67	16	0.3435	0.095	0.024	-0.007	-0.063	-0.025	-0.025
24	22.56	22	0.4267	0.014	-0.074	-0.033	0.069	0.030	0.024
30	27.21	28	0.5069	-0.021	-0.036	-0.084	0.042	-0.038	-0.012
36	32.48	34	0.5420	-0.062	-0.008	0.016	0.070	0.066	0.012
42	39.07	40	0.5120	0.055	0.079	0.043	-0.055	0.042	0.030
48	40.61	46	0.6970	0.019	-0.024	-0.027	-0.007	-0.042	0.018
54	45.06	52	0.7411	0.059	-0.039	0.009	-0.006	-0.073	-0.020
60	47.08	58	0.8469	-0.000	-0.019	-0.029	-0.055	-0.023	-0.001
66	50.60	64	0.8885	0.026	0.012	0.065	0.029	0.036	-0.032
72	57.42	70	0.8594	-0.019	0.051	-0.044	-0.078	-0.013	-0.067

Table 4: White noises check of SARIMA (5, 0, 0) × (0, 1, 1)12 model using Ljung-Box.
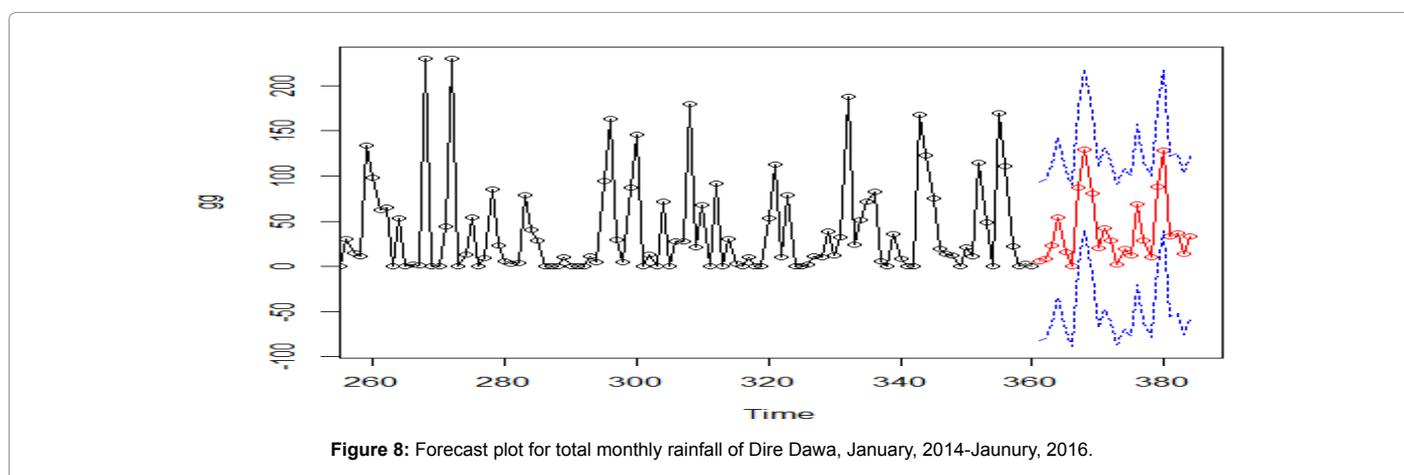
Figure 8: Forecast plot for total monthly rainfall of Dire Dawa, January, 2014-Jaunury, 2016.

the curves significantly reflect a normal distribution. Moreover, all calculated values are found to be smaller than the respective critical values for Goldfeld-Quandt (G-Q) Test and Breusch and Pagan(B-P) test, in the test of homoscedasticity of the residual, which indicate that the residuals has constant variance.

Therefore, from above all result it can be concluded that the hypothesis that the residuals are white noise cannot be rejected indicating that the fitted model is adequate. That is, SARIMA (5,0,0) × (0,1,1)$_{12}$ model is adequate for modeling the square root transformed monthly rainfall series of Dire Dawa region.

Forecasting: Since the model diagnostic tests show that all the parameter estimates are significant and the residual series is white noise, the estimation and diagnostic checking stages of the modeling process are complete. We can now proceed to forecasting the rainfall series with fitted SARIMA (5, 0, 0)*(0, 1, 1)$_{12}$ model. According to eqn. (1) above, the SARIMA (5, 0, 0) × (0, 1, 1)$_{12}$ model can be written as:

$$(1 - \phi_5 B^5)(1 - B^{12})y_t = (1 + \theta_{12} B^{12})e_t \qquad (19)$$

This equation (19) can also be multiplied out and rewritten in a form that is used in forecasting as:

$$y_t = y_{t-12} + \phi_5(y_{t-5} - y_{t-17}) + e + \theta_{12}e_{t-12} \qquad (20)$$

Where $\beta^5 y_t = y_{t-5}$ and $(1 - \beta^{12})y_t = y_t - y_{t-12}$. This equation further re-expressed as:

$$y_{t+m} = y_{t+m-12} + \phi_5(y_{t+m-5} - y_{t+m-17}) + e_{t+m} + \theta_{12}e_{t+m-12} \qquad (21)$$

After substituting the estimated parameter values in Eq. (21) above,

we obtain the following equation:

$$\hat{y}_{t+m} = \hat{y}_{t+m-12} - 0.181(\hat{y}_{t+m-5} - \hat{y}_{t+m-17}) \\ + \hat{e}_{t+m} + 0.777\hat{e}_{t+m-12} \qquad (22)$$

Forecasts are to be made at the origin, *t*=January, 2014 for lead times *m*= 1, …, 24 for the total monthly rainfall series in the coming 24-month. Then the one-step ahead forecast at the origin, t=Jan, 2014, which give us forecast of actual series for the month of January, 2012 are given by:

$$\hat{x}_{Jan,2014+m} = \hat{x}_{Jan,2011+m-12} - 0.033(x_{Jan,2014+m-5} - \hat{x}_{Jan,2014+m-17}) \\ + \hat{e}_{Jan,2014+m} + 0.604\hat{e}_{Jan,2014+m-12} \qquad (23)$$

By the help of eqn.(23), it can be obtain 24 month(2-years)-steps a head forecast and prediction interval for total monthly rainfall series of Dire Dawa, as presented in Figure 8. Results reveal that there is a tendency of relatively increasing in average pattern of monthly and annual rainfall over the forecast period from January 2014 to January, 2016.

Forecasting accuracy evaluation: If the fitted SARIMA (5, 0, 0)*(0, 1, 1)$_{12}$ model has to perform well in forecasting, the forecast error will be relatively small. The accuracy of forecasts is usually measured using root mean square error (RMSE), mean absolute error (MAE), Mean absolute percentage error (MAPE) and Theil's inequality coefficient (Theil-U). The result shows that the Mean Absolute Percentage Error (MAPE) turn out to be 3.56%, which is relatively less than 4% and Theil's inequality coefficient (U-statistic) turn out to be 0.018, which

is relatively close to zero. Besides this result, the bias and variance proportion are also very small, which are 0.047 and 0.001, respectively. Thus, measures indicate that the forecasting inaccuracy is low.

To assess the out-of-sample forecasting ability of the model it is advisable to retain some observations at the end of the sample period. Therefore, with the hold-out last monthly rainfall values from Jan, 2013 to Jan, 2014 for evaluating forecasting accuracy and model validation, The RMSE, MAE, MAPE values are less than 5% and Theil-U statistics is close to zero, which indicates the difference between the actual value andthe predicted value is very small. That is, the predictive powers of the models are better and suitable for m-step ahead forecasting.

Spectral analysis result

Power spectral Analysis is applied for annual and monthly rainfall series of Dire Dawa region. The first step in estimation of the spectrum by the periodogram method is subtraction of the sample mean and removing any obvious trend. Figure 9 show that raw periodogram of annual rainfall of Dire Dawa, in each point of it represents the variance of the series contributed by a frequency range centered at the point. Since raw periodogram is a wildly fluctuating estimate of the spectrum with high variance, the periodogram must be smoothed to ensure stable estimate. Proper amount of smoothing is somewhat subjective and depends on the characteristics of the data, generally, it is a good idea to try several bandwidths that seem to be compatible with the general overall shape of the spectrum, as suggested by the periodogram [17]. Considering the tradeoff in smoothness, stability and resolution in selecting widths of modified Daniel filters, which leads to span of length, c (1,1) and 15 as a suggested value for smoothing the annual and monthly rainfall of Dire Dawa station respectively (Figure 10). The result is displayed in Figure 2a-2c, which provide a sensible compromise between the noisy version, shown in Figure 1a-1c, and a more heavily

Figure 9: Raw periodogram of annual rainfall of Dire Dawa.

Figure 10: Raw periodogram of monthly rainfall of Dire Dawa.
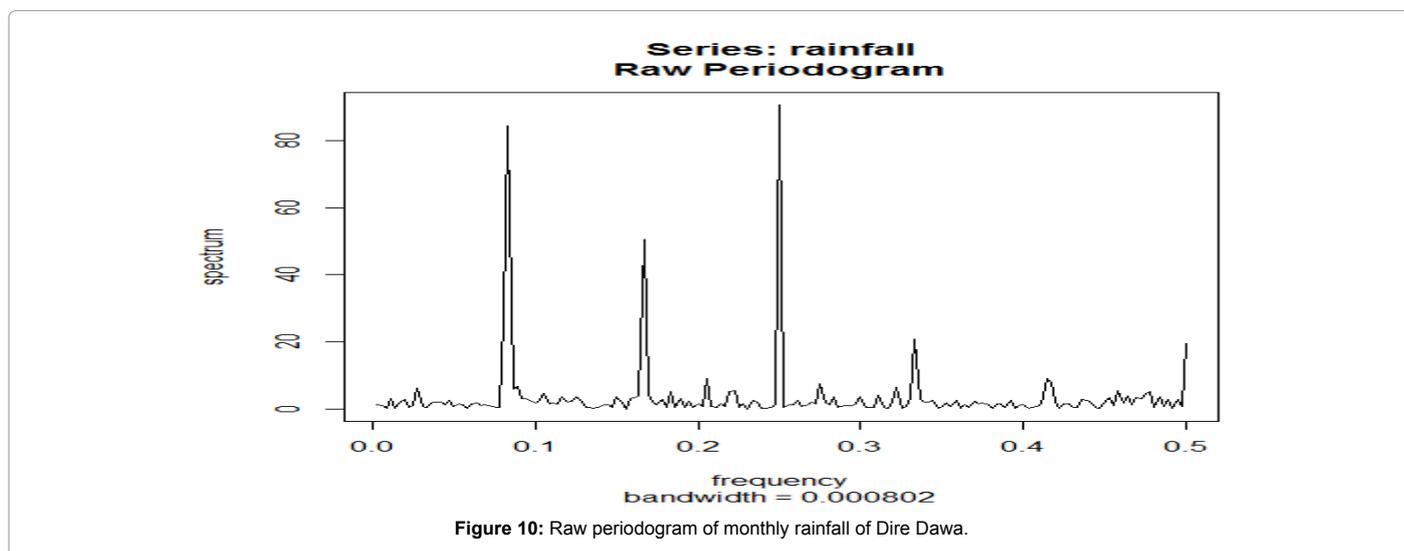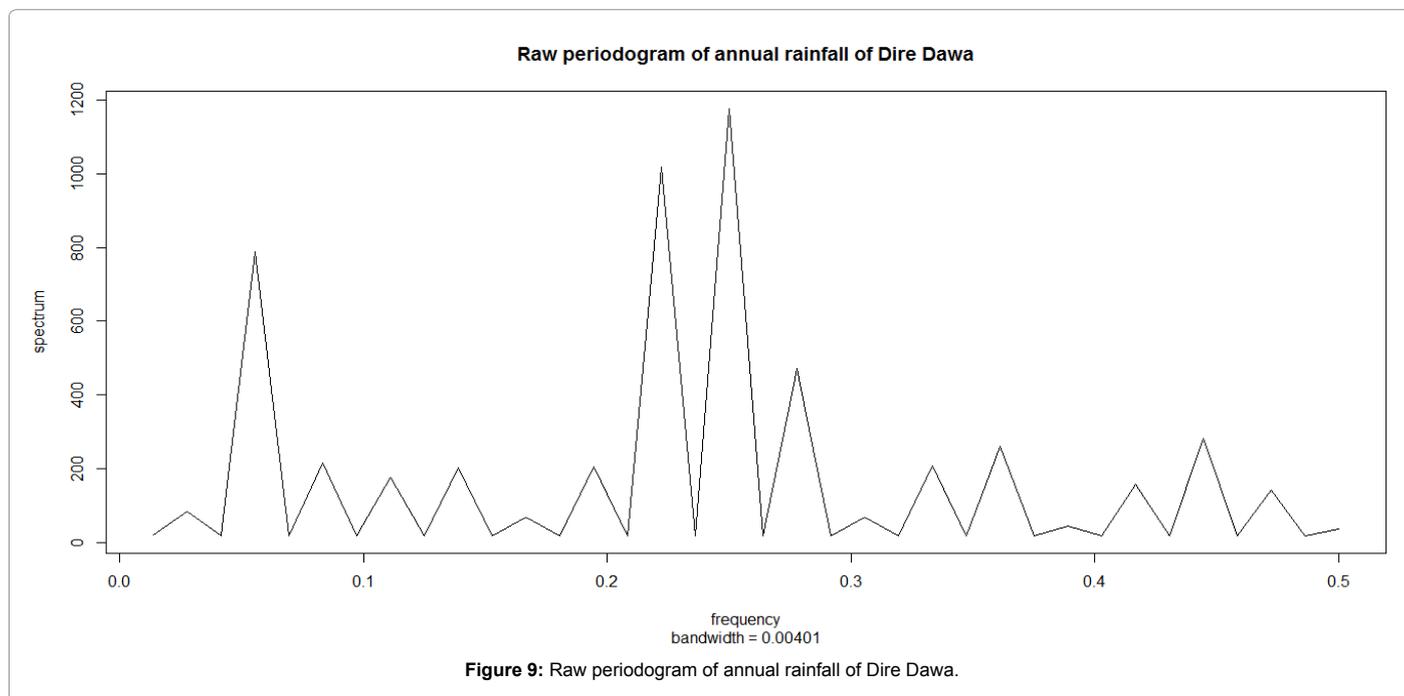
smoothed spectrum, which might lose some of unnecessary peaks in estimation of spectrum. The width of the center mark on the 95% confidence interval indicator indicates the band width [32] cited also in Alemrew et al. [3]. The bandwidth, is equal to 0.1 cycles per year and 0.0417 cycles per month for the spectral estimator. This bandwidth means we are assuming a relatively constant spectrum over about 20% and 8.3% for monthly and annual rainfall with entire frequency interval (0, 0.5), respectively. By using degrees of freedom (d_f) of 6 and 30 for annual and monthly rainfall respectively, it can constructed approximate 100(1-α) %confidence intervals (eqn.(14)) for spectral density for the frequency bands identified as having the maximum power, as shown in Table 5 above. If the lower confidence limit for the spectral value is greater than the baseline level at some predetermined level of significance, we may claim that frequency value as a statistically significant peak [17]. As we observed from Table 5, an approximate 95% confidence interval for the spectrum $f_S(0.37)$ is [39.34, 255.75], which is again too wide to be of much use, but we do notice that the lower value 39.34 is higher than any other periodogram ordinate, so it is safe to say that this value is statistically significant. Similarly, an approximate 95% confidence interval for the spectrum $f_S(0.083)$ is [6.17, 17.25].

The smoothed spectra shown in Figure 11 and summarized in Table 3 provide a sensible compromise between the noisy version, shown in Figure 9, and a more heavily smoothed spectrum, which might lose some of the peaks. As displayed in the Figure, the peak with an average period of 4.17 years, contributing the highest percent to total variance

of the annual rainfall series of Dire Dawa, which corresponds to the existence of "strong" peak at a frequency band centered at 0.24 cycles per year. This indicates that the periodicity of 4.17 years the most dominant in the annual rainfall of Dire Dawa. This can be attributed to the El Nino phenomenon. By looking for the other large spectral ordinate, it can be identified peaks of 16.67, 2.86 and 2.27 year periodicities.

Moreover, the monthly rainfall data are also used and their spectrum will exhibit narrow and sharp peaks at seasonal frequencies, as shown in Figure 12. For such type of data, commonly occurred periodic oscillation: Semi-Annual Oscillation (SAO), Annual Oscillation (AO), Quasi-Biennial Oscillation (QBO) and El-Nino/Southern Oscillation (ENSO) are approximately obtained by 4-7 months, 10-14 months, 22-32 months, and 40-66 months, respectively [33]. The peak with an average period of 12 month contributes a higher percent to total variance of the monthly rainfall series. This indicates that the most dominant peak observed in monthly rainfall of Dire Dawa is annual Oscillation (AO). There is also a subsidiary peak corresponding to a periodicity of 4 month and 6 month with associated frequency band centered around 0.25 and 0.167 cycles per month, respectively. These can also be attributed to the Semi-Annual Oscillation (SAO).

Using spectral analysis there have been observed prominent periodicities of 1.8-2.6 years [34], 2.7-3.3 years [35] and 2.5-8 years [36]. The study result by Perterbough [37] various regions of African rainfall indicate that there is no apparent statistically significant spectral peak in most cases. However, opposite of this result obtained by Nicholson [38], who discovered a statistically significant spectral

Series	Frequency	Period	Power (variance)	Lower limit	Upper limit
Annual rainfall	0.37	2.5 years	62.3	39.34	255.75
	0.2	5 years	38.2	16.76	156.74
Monthly rainfall	0.083	12 month	9.66	6.17	17.25
	0.167	6 month	5.51	3.52	9.85
	0.250	4 month	7.61	4.86	13.59

Table 5: The Smoothed Spectra of the annual and monthly rainfall Series of Dire Dawa.
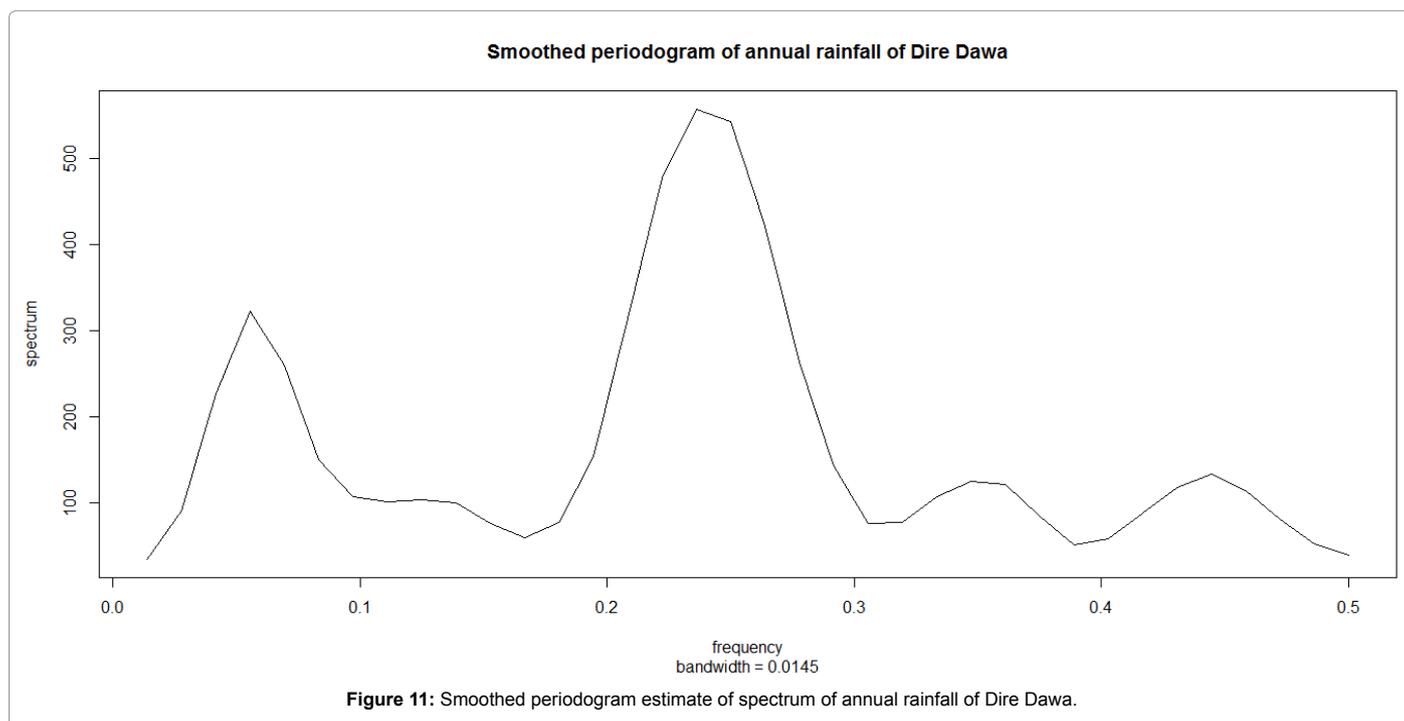
Figure 11: Smoothed periodogram estimate of spectrum of annual rainfall of Dire Dawa.

Figure 12: Smoothed periodogram estimate of spectrum of monthly rainfall of Dire Dawa.
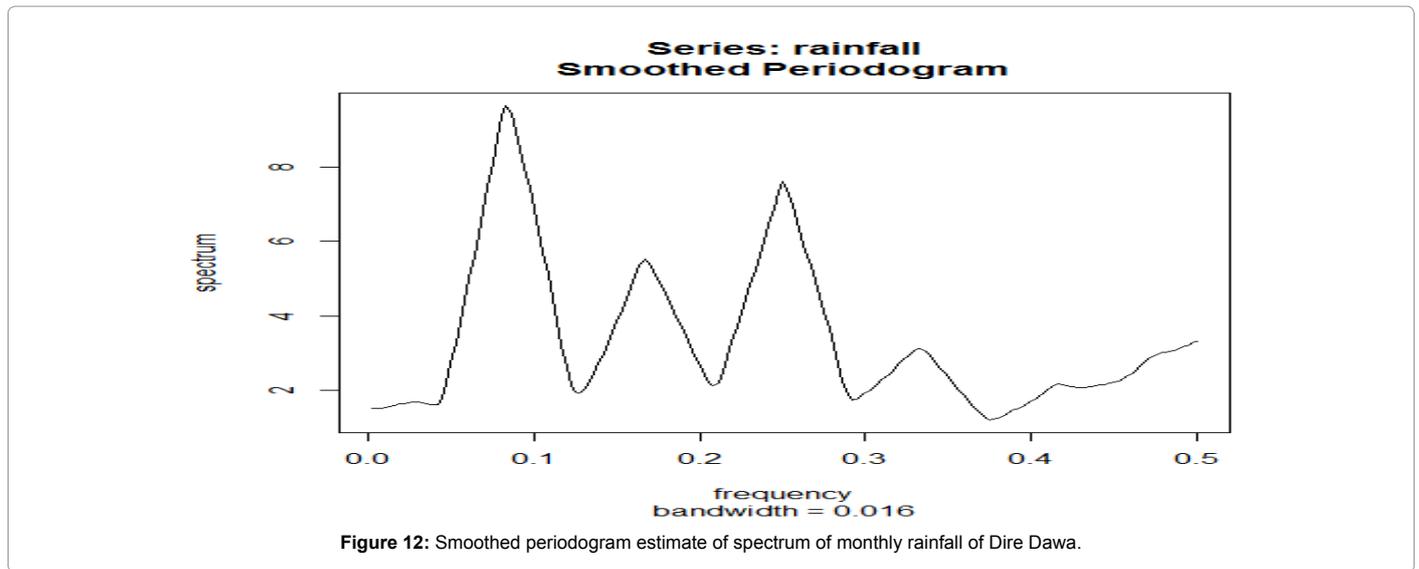
peaks of approximately 2.3, 2.8, 3.6 and 5–6 years is observed in different regions of African rainfall. Moreover, Perterbough [37] study and our result opposed each other. By conducted a detailed power spectrum analysis of African rainfall series, Nicholson et al. [39] found that quasi-periodicities clustered in four bands at 2.2–2.4, 2.6–2.8, 3.3–3.8 and 5.0-6.3 years, common throughout equatorial and southern Africa. Ayoade [26] detected periodicities of 2-6 years while Adejuwon et al. [40] detected prominent periodicities of between 2 and 8 years cycles and less prominent periodicities of 16, 18 and 21 years cycles both in study of precipitation pattern of Southern Nigeria. Alamerew et al. [3] using spectral analysis method found that the dominant cycle with periodicity of 11.24 year for the annual rainfall of Addis Ababa, but they didn't test the statistical significance of spectral peak.

Many study shows that Ethiopian rainfall strongly related to the ENSO phenomena via rainfall extreme event like drought and flood [4,11,41]. The ocean–atmosphere interaction is believed to be a dominant factor in the weather processes in Eastern Ethiopia and contributes the largest share of the total annual rainfall the region. With this inter-action, many studies were used for estimating possible future rainfall frequency of extreme events of the area. Yilma [2] found that a dominant cycle for the annual rainfall of Addis Ababa is 10.22 years. Alemrew et al. [3] inferred that drought recurs in Addis Ababa between 10 to 11 years. Haile [11] also found that drought occur every 6-8 years in the semi-arid regions of Ethiopia. The study by Tsegay [12] based on the frequencies of rainfall deviation from the average suggest that drought occur every three to five years in Eastern Ethiopia. The possibility of using prominent peaks in the spectrum to predict the long-range behavior of the rainfall is attractive [42]. Oduro et al. [6] predict that a rainfall extremes event in Ghana recurs every 5.6 years. Stringer [43] also estimated that the climatic events recur every 2 to 2.5 year in Africa.

Cross-spectra analysis result

The cross-spectrum analysis is used to provide a means of determining the contributions of fluctuations in various frequency bands to covariance quantities such as the fluctuations in rainfall series at different nearby stations.

Figure 13 shows the squared coherence between rainfall series of Dire Dawa station with Haramaya and Dengego stations with L_h =15

and 30 degree of freedom and $F_{2,30}(0.05)=3.34$ at 5% significance level. Hence, the hypothesis of no coherence is rejected for the values of estimated square coherence (eqn.17) that exceed $C_{0.05}=0.16$ (eqn.18). As it can be observed in the same Figure that, at lower seasonal frequency, the Dire Dawa rainfall characteristics are significantly associated with both Haramaya and Dengego rainfall, which mean there is strongly coherent. However, the degree of relatedness or coherence is higher in Dengego than Haramaya rainfall to the rainfall of Dire Dawa.

Conclusion of the Study

In this study, 30 years annual and monthly rainfall records were analyzed, using data from Dire Dawa region and neighboring area in the Eastern Ethiopia, mainly to study the rainfall pattern and its extreme event frequency of Dire Dawa. Based on the overall results of the research, the following conclusions could be drawn:

➢ A time series model for monthly rainfall series of Dire Dawa was adjusted, processed, diagnostically checked and lastly SARIMA(5,0,0)*(0,1,1)$_{12}$ model is established and adequately be used to forecast 2 years monthly rainfall values. Further results reveal that there is a tendency of relatively increasing in average pattern of monthly and annual rainfall over the forecast period from January 2014 to January, 2016.

➢ Rainfall extreme event like flood predict or inferred to be recurring at about 4.17 years in Dire Dawa region.

➢ The annual rainfall of Dire Dawa region found to dominate with El Nino phenomenon. On other hand, Annual oscillations dominate monthly rainfall of the area.

➢ Rainfall extreme event like flood and drought highly association between Dire Dawa and Dengego area.

According to the result in this study, the pattern and frequency in rainfall extreme event is increasing. So, every effort should be made to minimize such risks as deaths, loss of property and damage to infrastructure, associated with rainfall extreme event like drought and flood, crop failure by pre-planning and make decision against adverse effects of it on the region. Moreover, more study need to assess the factors that cause the change in the pattern of rainfall distribution in the region.

(a)

(b)
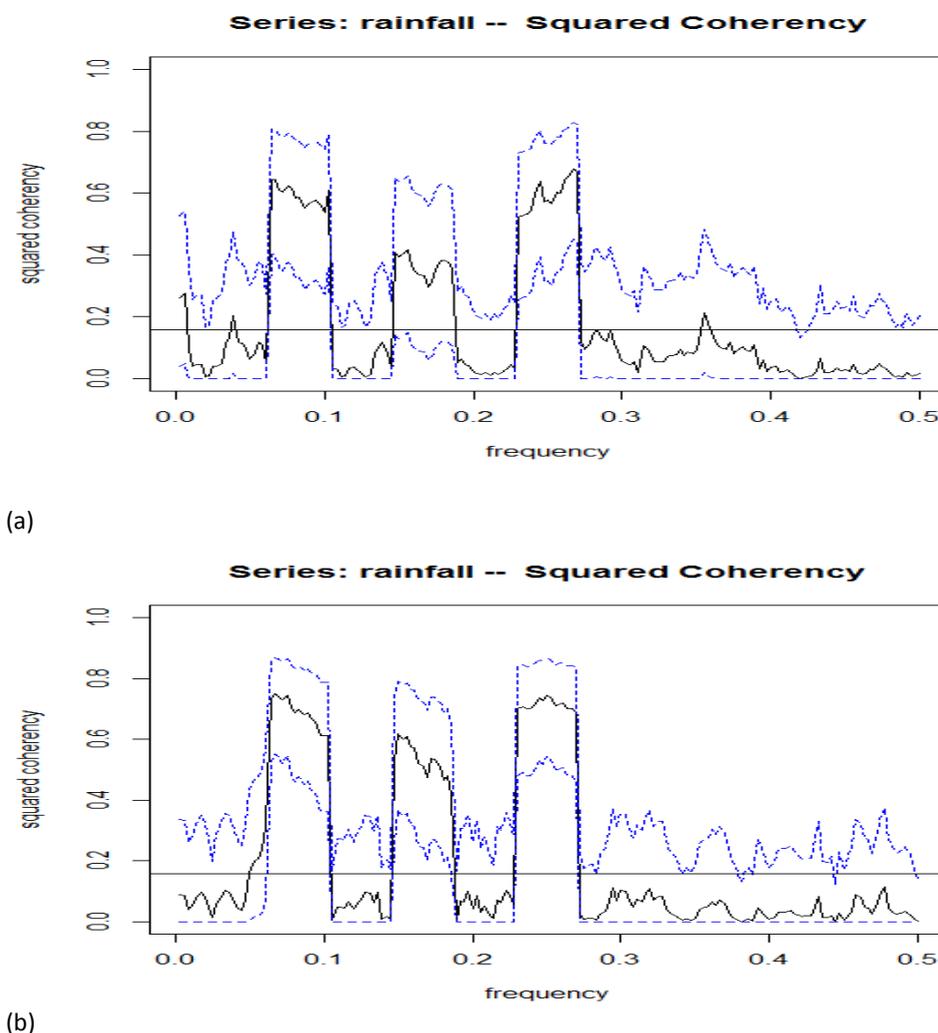
Figure 13: Squared coherency between rainfall series of: (a) Dire Dawa and Haramaya station; (b) Dire Dawa and Dengego stations.

References

1. Admasu G (1989) Regional flood frequency Analysis. Technical report, Royal institute of technology stockholm, Sweden 46-84.

2. Yilma S, Demaree GR, Delleur JW (1994) Sunspot numbers as a possible indicator of annual rainfall at Addis Ababa, Ethiopia. International Journal of Climatology 14: 911-923.

3. Alamerew B, Eshetu W (2009) Assessment of local climate in Addis Ababa. Journal of Ethiopian Statistical Association 18: 46-84.

4. Bekele F (1997) Ethiopian use of ENSO information in its seasonal forecast. Internet Journal for African Studies (IJAS).

5. Dire Dawa Environmental Protection Authority (DDAEPA) (2011) Dire Dawa administration program of adaptation to climate change.

6. Oduro-AfriyieK, Adukpo DC (2006) Spectral characteristics of the annual mean rainfall series in Ghana. West Africa Journal of Applied Ecology 9: 15-18.

7. Wood CA, Loveth PR (1974) Rainfall, drought and the solar cycles. Nature 251: 594-596.

8. Dyer TGJ, Tyson PH (1977) Estimating above and below-normal rainfall periods over South Africa 1972-2000. Journal of Applied Meteorology 16: 145-147.

9. Mersha E (2002) Determination of rainfall cyclicity over selected location in Ethiopia. Ethiopian J Water Science and Technology.

10. Seifu A (2004) Rainfall variation and its effect on crop production in Ethiopia. Addis Ababa University, unpublished MSC thesis, Ethiopia.

11. Haile T (1988) Causes and characters of drought in Ethiopia. Ethiopian Journal of Agricultural Sciences 10: 85-97.

12. Tsegay W (1998) El Nino and drought early warning in Ethiopia. In Using Science Against Famine: Food Security Early Warning System and El Nino, Internet Journal of African Studies.

13. Polyak I (1996) Computational statistics in climatology, Oxford University Press, Oxford.

14. Bewket W, Conway D (2007) A note on the temporal and spatial variability of rainfall in the drought-prone Amhara region of Ethiopia. International Journal of Climatology 27: 1467-1477.

15. Gibbons RD (1994) Statistical methods for groundwater monitoring. John Wiley & Sons, New York.

16. Box GEP, Jenkins GM (1976) Time series analysis: forecasting and control, Holden day.

17. Shumway RH, Stoffer DS (2010) Time series analysis and its applications with R Examples. Springer (3rdedn).

18. Wei WWS (1990) Time series analysis: univariate and multivariate methods. Addison-Wesley Publishing Company, Inc, New York, USA, p. 478

19. Mills T (1999) The econometric modeling of financial time series. Cambridge University Press, Cambridge.

20. Lehmann A, Rode M (2001) Long-term behaviour and cross-correlation water quality analysis of the River Elbe, Germany. Water Res 35: 2153-2160.

21. Ljung GM, Box GEP (1978) On a measure of lack of fit in time series models. Biometrika 65: 297-303.

22. Cromwell JB, Labys WC, Terraza M (1994) Univariate tests for time series models. A Sage Publications, London, 96: 07-99.

23. Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52: 591-611.

24. Makridakis S, Wheelwright SC, Hyndman RJ (1998) Forecasting methods and application. John Wiley & Sons, New York.

25. Yevjevich V (1972) Stochastic processes in hydrology, Water Resources Publications, Fort Collins.

26. Ayoade JO (1973) Annual rainfall trends and periodicities in Nigeria. Nigeria Geographical Journal 16: 167-176.

27. Bloomfield P (2000) Fourier analysis of time series: an Introduction. (2ndedn) John Wiley & Sons, Inc., New York.

28. NMA (1996) Climate & agro climate resources of Ethiopia. NMSA, Meteorological Research Report Series, Addis Ababa.

29. Hamilton J (1994) Time series analysis. Princeton University Press, New Jersey.

30. Janacek G, Swift L (1993) Time series forecasting, Simulation, Application, Ellis Horwood, New York, USA.

31. Chatfield C (1991) The Analysis of time series - an introduction (4thedn.), Chapman and Hall: London.

32. Venables W, Ripley B (1999) Modern applied statistics with S-PLUS. Springer-Verlag, New York.

33. Sinta BS, Hariadi TE (2003) The spectrum analysis of rainfall in Indonesia. Indonesian Journal of Physics 14.

34. Klaus B (1977) Spatial distribution and periodicities of mean annual precipitation south of the Sahara. Archives for Meteorology Geophysics and Bioclimatology Series 26: 17-27.

35. Ogallo L (1979) Rainfall variability in Africa. Monthly weather review 107: 1133-1139.

36. Farmer G, Wigley TML (1985) Climatic trends for tropical Africa. A Research Report for the Overseas Development Administration, England.

37. Peterbough TL (1983) A regional comparison of seasonal African rainfall anomalies. In Proceedings of the 2nd International Meeting in Statistical Climatology. Institute national de meteorologiae Geofisca: Lisbon, Portugal 8.8.1-8.5.8.

38. Nicholson SE (1985) African rainfall fluctuations from 1850 to present: spatial coherence, periodic behavior and long-term trends. In 3rd Conference on Climate Variations and Symposium on Contemporary Climate 1850-2100, Los Angeles, American Metrological Society Extended Abstracts, 62-63.

39. Nicholson SE, Entekhabi D (1986) The quasi-periodic behaviour of rainfall variability in Africa and its relationship to the Southern Oscillation. Journal of Climate and Applied Meteorology 34: 311-348.

40. Adejuwon JO, Balogun EE, Adejuwon SA (1990) On the annual and seasonal patterns of rainfall fluctuations in Sub-Sahara West Africa. International Journal of Climatology 10: 839-848.

41. NMA (2007) National Adaptation Program of Action of Ethiopia (NAPA). Final draft report. National Meteorological Agency, Addis Ababa, Ethiopia.

42. Zangvil A (1979) Temporal fluctuations of seasonal precipitation in Jerusalem. The Institute for Desert Research, Ben-Gurion University, Israel, 31: 413-420.

43. Stringer ET (1972) Techniques of climatology. WH Freeman & Co., San Francisco, p. 539