

ANALYSIS OF RAINFALL PATTERN IN THE WESTERN REGION OF GHANA

ABSTRACT

The primary aim for this paper is to examine the pattern of rainfall in the western region of Ghana. Data was obtained from the Ghana Meteorological Agency. The sample include January to September pattern of the amount of rainfall, for the years 2006 to 2016. That is nominal daily rainfall recorded (1485) aggregated into monthly rainfall value (99 data point). The analysis includes fitting an auto regression moving average model (ARMA) model for the data. The series was found to be non-stationary which resulted from the presence of a unit root in it. The series became stationary after eliminating the unit root by finding the first difference in the amount of rainfall. The time series component found in the model were trend and random variation. ARMA (1, 1) which has all parameters significant was fitted for the data and found to be the most suitable model for the conditional mean. A Ljung Box test statistic of 47.207 with a normalized BIC of 6.420 and a Root Mean Square error of 24.16 supported by a probability value of 0.001 show that the fitted model is appropriate for the data. An $R^2 = 0.532$ indicates that about 53% of the variations seen in the pattern of rainfall recorded for the period is being explained by the fitted model. The 18-month forecast for the mean actual rainfall and mean returns could show that the fitted model is appropriate for the data and an increasing trend of rainfall for the forecasted period.

Keywords: Auto Regression Moving Average, Unit root, ACF, PACF, Forecast, Stationarity, Parameter estimates, ADF test statistic

1. INTRODUCTION

Rainfall variability has serious implications for livelihood and food production in developing regions such as West Africa. In this region irrigation is restricted and inter-annual and multi-decadal variability leads to declining rainfall total. The situation is exacerbated by the fact that more than half of the adult population in the sub-region is directly engaged in essentially rain-fed agriculture. Ghana, like the other parts of the sub-continent, has undergone a period of declining annual rainfall total since the early 1970s and she is only recently showing signs of recovery since 2000 [1]. Increases in annual rainfall totals in many parts of Ghana after the year 2000 are evident in the spilling of the Akosombo dam on the Volta River in November 2010. This was the first time in 20 years that the dam had to be spilled due to increases in rainfall [2]. About 42 % of Ghana's 238,540 km² is suitable for crop cultivation but only about 27 % of this is under cultivation as estimated by the Food and Agricultural Organization (FAO) in 2005. In a pilot study in Wenchi, located on the northern fringe of mid-Ghana, [2] identified, in addition to an overall drying, greater reductions in the mean rainfall totals and the mean number of rainy days during the minor rainy season and a slight increase of rains in the short dry spell. This reduction in rainfall and potential diminution of the minor rainy season, if present throughout humid mid-Ghana, is likely to prevent cultivation of crops and crop varieties that have longer growing seasons, as well as the adoption of a single crop per year, instead of the current two crops, under rain-fed agriculture. Such an occurrence will negatively impact on food security. Government agencies and international organizations are currently encouraging the application of seasonal forecast information and weather index insurance as some of the adaptation measures [3, 4].

32 However, to develop a model for predicting changing rainfall patterns or to utilize available forecasted
33 information, it is important to understand both the spatio-temporal nature of the declining and shifting
34 rainfall pattern in the agriculturally important regions in mid-Ghana. According to FAO in 2008, rainfall
35 variability is an inherent part of the African climate and it is deeply entrenched in West Africa. Thus,
36 there is inadequate rain for irrigation in many African countries and as such countries whose
37 economies rely highly on agriculture are greatly vulnerable to economic instability. According to the
38 International Scientific Research (ISR) Journal [2], "in the event of large deviations from the normal
39 rainfall, people are highly affected as floods and droughts are most often the by-products.
40 Government's scarce resources are directed to humanitarian missions to help people affected by
41 floods and other disasters that come with these extreme weather conditions".
42

43 During the last 10 to 15 years, there have been worldwide perceptions that droughts and floods have
44 intensified [5]. In Ghana, it has been observed that the annual rainfall total has generally declined
45 while the total number of extreme events such as droughts has been on the increase [2, 6]. Similar
46 studies conducted by [7] for West Africa, [8] for east Africa, [9] for south Africa, [10, 11] for various
47 parts of Africa show that some regions on the continent, especially west Africa have suffered drastic
48 changes such as prolonged drought and prolonged flood.
49

50 In Ghana, the situation is no different, the Ministry of Finance, in 2007 indicated that the problem of
51 rainfall variability is paramount and continues to have serious consequence on the Ghanaian
52 agriculture, accounting for about 35% of the country's Gross Domestic Product (GDP). Farmers
53 depend on shared knowledge and experience with the weather as well as observations of natural
54 phenomena to forecast forthcoming cropping season and weather condition [12]. However, in recent
55 times, the frequency of change in climate has increased considerably and local experience and
56 knowledge are no longer sufficient to guide agricultural planning and decision making [13]. Hence the
57 initiation of models as a guide to understanding these drastic changes and future circumstances could
58 therefore be predicted based on the knowledge acquired from these models.
59

60 Climate change in Ghana has become a threat to livelihoods. Drought and over flooding in some parts
61 of Ghana have developed into yearly worry to people and government. In the south particularly, the
62 coastal areas, aquatic life is of great importance because of the fishing activity that goes on there, and
63 farmers in these parts also dwell mainly on the rains for farming since there are no major irrigational
64 facility. As such, changes in rainfall affect the level of water bodies as well as crop farming. This
65 problem influences the economic activities in these areas and the country at large. As a result, the
66 Government of Ghana contracts researchers and engineers to come out with ways to solve these
67 problems every now and then [4]. One of the ways used is time series analysis, thus, studying the
68 past and current pattern of rainfall in a systematic approach would help to fit a suitable model for
69 future predictions.
70

71 The major purpose of this study is to identify rainfall pattern in the Western Region by considering the
72 years 2006 to 2016 and fitting an appropriate time series model for forecasting future rainfall pattern
73 (values) in the Western Region. Findings of this paper will be significant since it will enable farmers to
74 plan their farming activities ahead of time and provide empirical evidence to stakeholders on rainfall
75 trends to help them formulate policies that can benefit the region concerned and the nation at large.
76
77

78 2. METHODS AND DATA

79 2.1. Time Series Analysis

80

81 In time series analysis, the past and present behavior of variables are observed and examining them
82 often suggest the method of analysis as well as statistics that will be of use in summarizing any
83 information in the data, so that values predicted from the data may fit the present situation as well as
84 the future. Time series data are often obtained through monitoring industrial processes or tracking
85 corporate business metrics. Data used in time series can be continuous or discrete in nature, it is said
86 to be continuous when the observations are made over time interval and it is described as discrete
87 when observations are made at specific time periods. Usually these observations in time series are
88 taken at regular intervals such as days, months, quarters and years. There are two mutually exclusive
89 approaches usually applied in time series analysis, these are the time domain approach and the
90 frequency domain approach. Conversely, the time domain approach which is adapted in this study is
91 generally motivated by the assumption that correlation between times is explained best in terms of a

dependence of the current value on the past values. This approach focuses on modeling some future value of a time series as a parametric function of the current and past values. A more current method in the time domain approach well-known to statisticians is the use of the additive model or the multiplicative models [14, 15].

Time series data exhibit at least one of the following features; Secular (Trend), Seasonal variations, Cyclical variations, and Irregular (Random) variations. Secular (Trend) are continuous long-term movement in a variable over an extended period that is, a general increase or decrease in a time series data over several consecutive periods. Trend can be linear or nonlinear. A linear trend tends to increase or decrease at a constant rate, however a nonlinear trend is likely to move steadily upwards, as others decline. Seasonal Variation is a wavelike pattern that is repeated throughout a time series with a recurrent period at most one year but, usually on weekly, monthly, quarterly, or annual basis. These are the short-term regular variations in data, generally caused by factors such as weather, holidays, festivals etc.

Seasonal component is a pattern in time series which indicate change of monthly data that repeats itself within a year. A Cyclical Variation exhibits repetitious pattern with a recurrent period longer than one year. This occurs mostly in businesses which indicate variations in the general level of national economic measures such as unemployment, gross national product, stock market index etc. over a relatively long period of time, thus these points toward a cycle. Irregular (Random) Variation is often referred to as the “noise” in the data that are unpredictable in the times series data and cannot be associated with trend, seasonal, or cyclical component of time series. Events such as industrial strike actions, earth quakes, floods, outbreak of epidemics, wars etc., may lead to odd movements in a time series data [14, 15]. The types of patterns of fluctuations in a time series may be represented as;

- T = trend value of the series
- S = value of the seasonal variation
- C = value of the cyclical variation
- I = value of the irregular variation

Thus let;

$$Y_t = \text{observed values of the time series at time } t$$

Hence the additive and multiplicative models may be represented as

$$Y_t = T + S + C + I \text{ and}$$

$$Y_t = T \times S \times C \times I \text{ respectively.}$$

If the data however, do not contain one of the type of variation (e.g., cycle) the value for that missing component is zero. For instance, there is no cycle for a yearly series since cyclical variation cannot be observed over a one-year period, hence the additive model becomes;

$$Y_t = T + S + I.$$

Likewise, in the multiplicative model if trend, seasonal variation, or cycle is missing, then the value is assumed to be 1. So, for series with a period of one year, where there is no cycle then;

$$Y_t = T \times S \times I.$$

2.2. Trend Analysis and Forecasting Techniques

Time series analysis is aimed at projecting trend by fitting a trend line to a series of historical data points through which a model is fit for prediction of future values over a period. Several trend Equations can be developed based on exponential or quadratic models, however the simplest is a linear trend model (least square method- LSM) that is developed using Regression analysis. Equation for Linear Trend is given by

$$T_t = b_0 + b_1 t$$

Where;

T_t = trend value in period t (predicted value)

152 b_0 = intercept of the trend line
153 b_1 = slope of the trend line
154 t = time

155
156 It should be noted that t is the independent (or predictor) variable and T_t is the dependent (response)
157 variable. Computing the Slope (b_1) and Intercept (b_0) using the Least Square Method (LSM). The
158 slope (b_1) is given by;

$$159$$
$$160 \quad b_1 = \frac{n \sum t Y_t - \sum t \sum Y_t}{n \sum t^2 - (\sum t)^2}$$

161
162 and the intercept (b_0) is also given by;

$$163$$
$$164 \quad b_0 = \frac{\sum Y_t}{n} - b_1 \frac{\sum t}{n} = \bar{Y}_t - b_1 \bar{t}$$

165 Where;

166
167 Y_t = actual value in period t
168 n = number of periods in time series

169
170 Quadratic trend model is a non-linear trend model also known as a second-degree polynomial model.
171 It is the simplest curvilinear model with a general equation given by;

$$172 \quad T_t = b_0 + b_1 t + b_2 t^2$$

173
174 Where;

175
176 b_0 estimates the value of T_t when $t=0$
177 b_1 is the linear effect coefficient
178 b_2 is the curvilinear effect coefficient

179
180 Time series data is deseasonalized when the seasonal effects in a time series data is to be removed
181 before trend is fitted and usually seasonal index are computed for such purpose. Seasonal pattern is
182 the short-term cycle occurs within or at most a year. The seasonal variation can be expressed in
183 terms of deviations from the original data in the case of additive model or as percentage of the trend
184 in the case of multiplicative model. Thus, the deseasonalized value for an additive model is given by;

$$185 \quad \text{Deseasonalized value} = \text{time series observation} - \text{seasonal index} = Y_t - I_s$$

186
187 and that of multiplicative model is also given by;

$$188$$
$$189 \quad \text{Deseasonalized value} = \frac{\text{time series observation}}{\text{corresponding seasonal index}} = \frac{Y_t}{I_s}$$

190
191 Thus, applying the LSM, $T_t = b_0 + b_1 t$ in this case, Y_{st} the deseasonalized time series value at time t is
192 used in-place of the actual value of the time series (Y_t). The resulting line equation is therefore used
193 to make trend projections. Projection of trend into the future is usually known as forecasting, the time
194 series data are plotted so that their trends over time are observed. If there is a long term upward or
195 downward trend in the data the least square forecasting method can be considered especially when
196 dealing with annual data. However, if there is no trend then either the moving average or the
197 exponential smoothing forecasting techniques may be employed. Exponential smoothing is a
198 forecasting tool also used predicts future time series data. In this type of forecast technique, the
199 forecast is based on a weighted average of a historic time series data. The weighted average usually
200 represented by alpha (α) [14, 15, 16]. Thus, the forecast value for a current time series is computed
201 as;

$$202 \quad F_{t+1} = \alpha Y_t + (1 - \alpha) F_t$$

203
204 Where;

205 F_{t+1} is the new forecast for time $t + 1$
206 Y_t is the previous period actual demand

207 F_t is the previous forecast for the time t
208 α is the smoothing constants ($0 \leq \alpha \leq 1$)
209

2.3. Measures of Forecast Error (Forecast Error = $(Y_t - F_t)$)

210
211 The forecast error is the deviation of the forecast values (F_t) from the actual values (Y_t). There are
212 four main errors measured in forecast data. These errors include Bias, Mean Absolute Deviation
213 (MAD), Mean percentage deviation error (MAPE) and the mean square error (MSE) [14, 15, 16]. In
214 time series analysis Bias, MAD, and MAPE are the usual errors employed to assess the amount of
215 errors related to a forecast. Bias is similar to the arithmetic mean, that is, the sum of the forecast
216 errors divide by the number of period, T and it is given by
217
218

$$\text{Bias} = \frac{\sum_{t=1}^T (\text{Forecast error})}{T} = \frac{\sum_{t=1}^T (Y_t - F_t)}{T}$$

219
220 Mean Absolute Deviation (MAD) is the sum of the absolute forecast error divide by the number of
221 period, T . Mathematically,
222

$$\text{MAD} = \frac{\sum_{t=1}^T |\text{Forecast error}|}{T}$$
$$= \frac{\sum_{t=1}^T |Y_t - F_t|}{T}$$

223

224

225

226 Mean square deviation is more sensitive measure of usually large forecast error than Mean Absolute
227 Deviation [14, 15, 16]. Mean Absolute Percentage Error (MAPE) [17] is the division of each
228 percentages of the absolute forecast error by their actual values, then all summed and divide by the
229 number of period, T . Hence.
230

$$\text{MAPE} = 100 \frac{\sum_{t=1}^T \frac{|Y_t - F_t|}{Y_t}}{T}$$

231

232 Mean Square Error (MSE) is similar to simple sample variance [14, 15, 17]. Standard Error is the
233 standard deviation of the sampling distribution (the square root of the MSE) given as
234

$$\text{MSE} = \frac{\sum_{t=1}^T (\text{Forecast error})^2}{T}$$
$$= \frac{\sum_{t=1}^T (Y_t - F_t)^2}{T}$$

235

236

237

238

3. RESULTS AND DISCUSSION

239

240
241 The analysis that follows is focused on the pattern of rainfall in the western region of Ghana. The
242 analysis includes fitting an ARMA model for the observed rainfall data. This article considered a model
243 based on information and real data obtained from the Ghana Meteorological Station, Sekondi. The
244 sample include January to September pattern of the amount of rainfall, for the years 2006 to 2016,
245 comprising 99 data points. Time Series Analysis and the statistical computing package R were used
246 for the modeling.
247

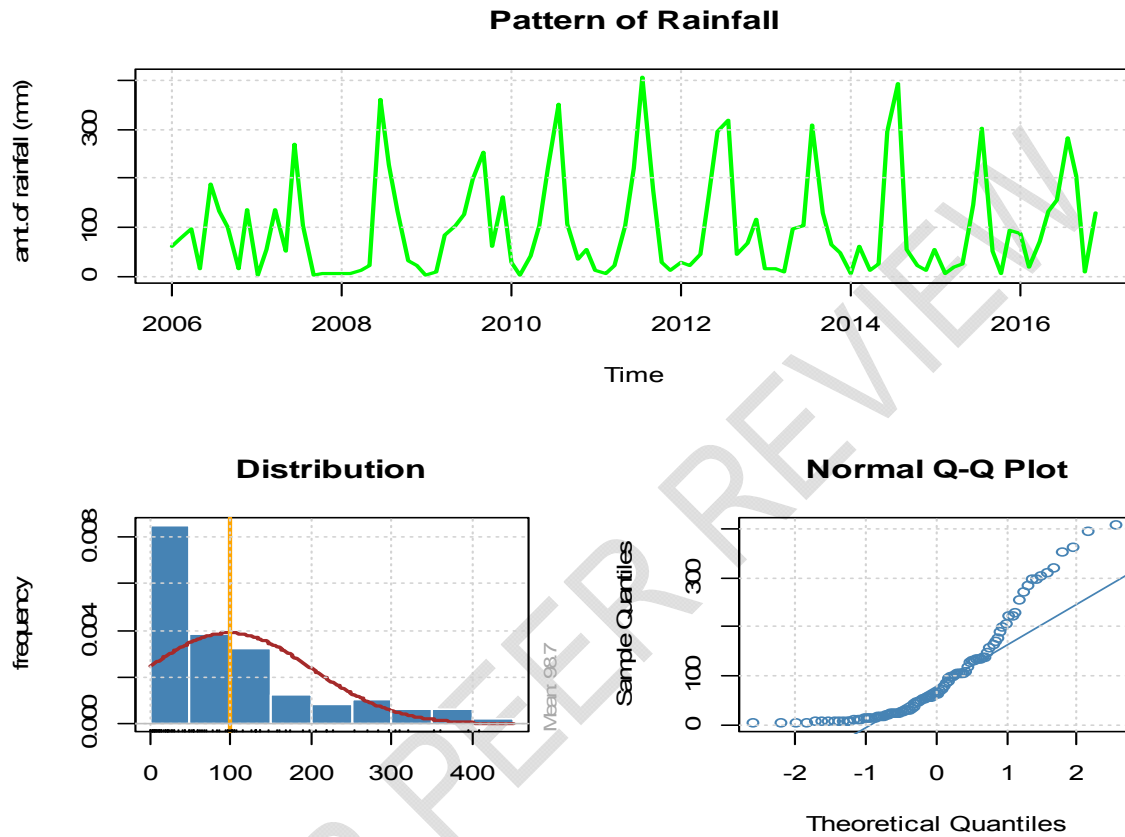
247

3.1. Rainfall Distribution

248

249
250 The time plot of a given series gives a fair idea of the stationarity of the series which is considered as
251 a form of statistical stability. A series with trend or seasonal pattern are considered as non-stationary.
252 That is the mean of the given series change with time. The time plot of the series in Figure 1 shows
253 that the series exhibit a random fluctuation showing a periodic or seasonal variation with maximum
254 value of 408.30 in June. 2011 and minimum value of 1.20 in January. 2009. We also observe that the
255 mean of the amount of rainfall changes over time, which suggest the series is non-stationary. The

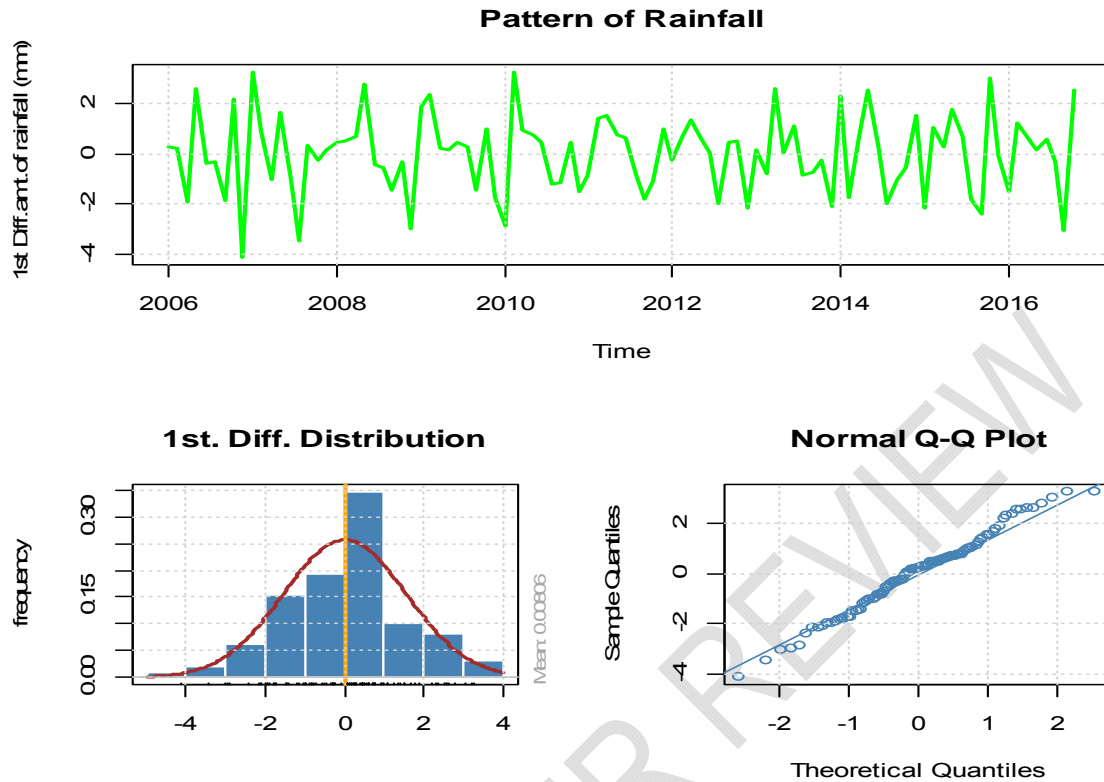
256 histogram with normal curve and normal Q-Q plot indicates that the empirical distribution of the series
 257 is not normally distributed and skewed to the right. By performing the unit root test on the series, we
 258 found that the Augmented Dickey- Fuller (ADF) root test statistic (-1.9453) is higher than the critical
 259 value (-2.86431), at a 5% significance level indicating that we fail to reject the null hypothesis that
 260 there is a unit root in the series which is supported by a p-value of 0.234
 261



262
 263
 264
 265
 266
 267
 268
 269

Figure 1: Time plot, Distribution and Normal Q-Q plot for Monthly Rainfall Series

For us to eliminate the unit root, we found the first difference in the rainfall pattern and conducted the test again. The results of the test show an ADF test statistic for the first difference (-8.2038), with a p-value of 0.01 and critical value (-2.86431) which make us reject the null hypothesis of unit root in the series. Hence, we conclude that the rate return series is stationary.



270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286

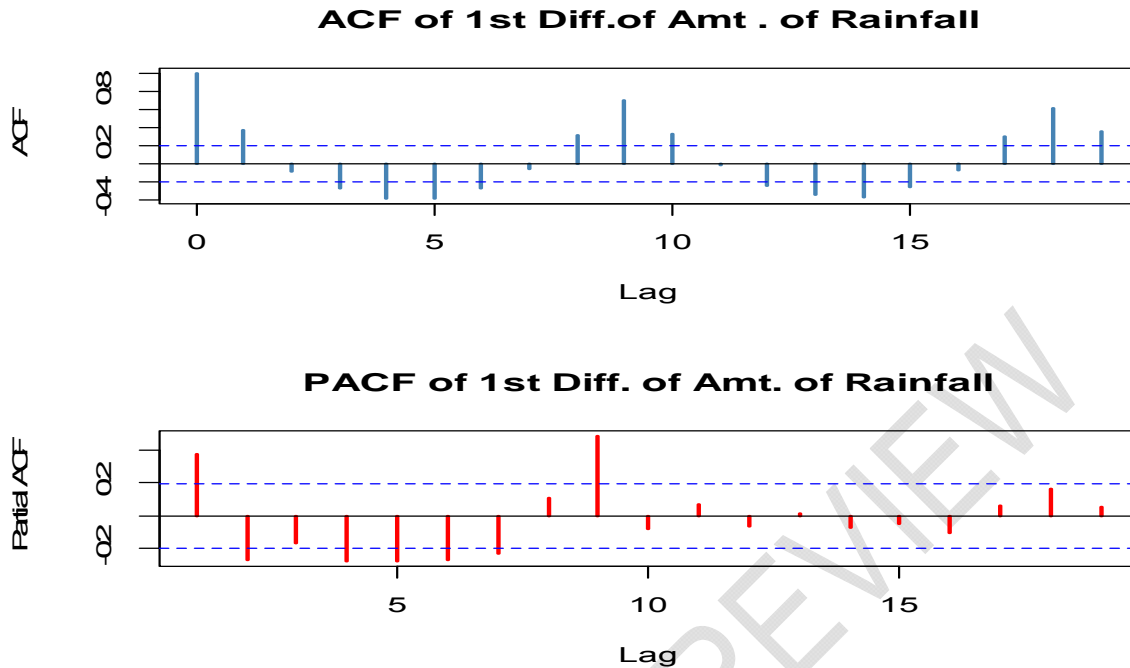
Figure 2.: First Differenced of Monthly Rainfall Series

Figure 2 shows the first difference of amount of rainfall and its distribution. The series appear to be stationary around the mean (top), the histogram look symmetric with heavy tail to the right and the normal Q-Q plot indicates a normal series with few outliers.

3.2. Determining Order of Dependency of 1st Differenced Series

The autocorrelation and partial autocorrelation functions (ACF/PACF) for the first differenced in the amount of rainfall are illustrated in figure 3.

From figure 3, we could observe that both the Autocorrelation and Partial autocorrelation functions showed dependency in the differenced rainfall series. As a result, a correlation structure in conditional mean is required.



287
288
289
290
291
292
293
294
295
296
297

Figure 3: ACF and PACF of 1st differenced monthly rainfall pattern

It can also be observed that the ACF show a significant number of lags of an MA at lag1 and PACF also show a significant number of lags of an AR at lag 1. This indicates that the model for the conditional mean is ARMA (1, 1). This is confirmed by the selection of model using the Akaike Information Criterion shown in Table 1

Table 1: Model Selection by Akaike Information Criterion

| ARMA (p, q) | AIC |
|--------------------|---------------|
| ARMA (1, 0) | 364.91 |
| ARMA (0, 1) | 362.21 |
| ARMA (1, 1) | 339.33 |
| ARMA (1, 2) | 343.17 |
| ARMA (2, 1) | 342.97 |
| ARMA (0, 2) | 342.51 |
| ARMA (2, 0) | 365.42 |
| ARMA (2, 2) | 382.16 |

298
299
300
301
302
303
304
305
306
307

Using the Akaike Information Criterion, we choose the model with the smallest value of AIC. From Table 1, the suitable model for the conditional mean is ARMA (1, 1) with an AIC value of 339.33. The parameter estimates are shown in Table 2

308 **Table 2: ARMA (1, 1) Model's Parameter Estimates and Standard Errors**
 309

| Variable | Coefficient | Standard Error | T-Statistics | Probability |
|----------|-------------|----------------|--------------|-------------|
| Constant | -0.004271 | 0.006400 | -0.667 | 0.505 |
| AR (1) | 0.415772 | 0.096789 | 4.296 | 1.74e-05 |
| MA (1) | -0.996001 | 0.032122 | -0.667 | 2e-16 |

310
 311 $\sigma^2 = 1.757$, conditional sum of squares = 170.2, AIC = 339.33
 312

313 **3.3. Conditional Mean Model the Differenced Rainfall Series**

314 The ARMA(p, q) model states that the current value of some series r_t depends linearly on its own
 315 previous values and a combination of current and previous values of a white noise error term ε_t . In
 316 the general form, the model can be written in the form:
 317

$$y_t = \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{j=1}^q \beta_j \varepsilon_{t-j} + \varepsilon_t$$

318
 319

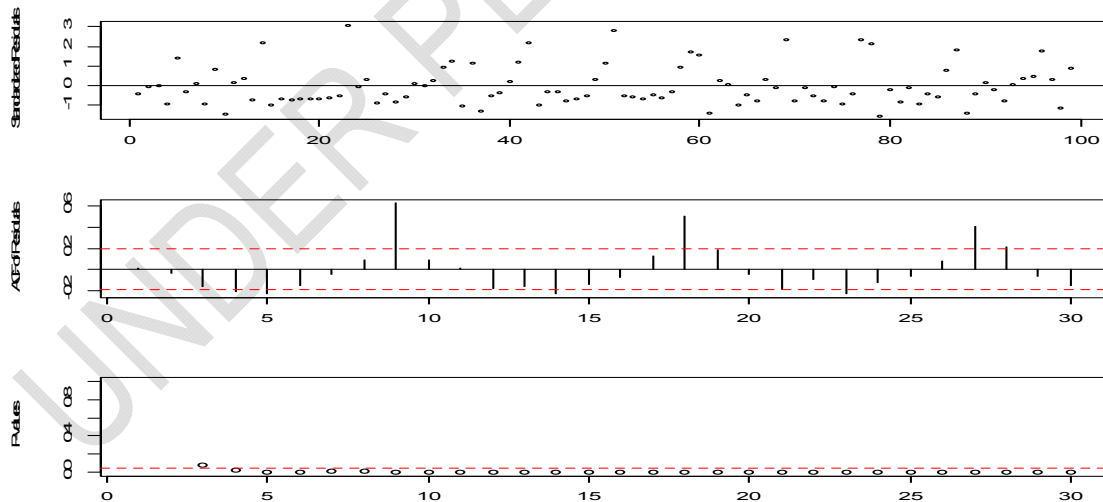
$$E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = \sigma^2, \quad E(\varepsilon_t \varepsilon_s) = 0, \quad t \neq s$$

320
 321

322 Our model for the conditional mean of the differenced rainfall series is ARMA (1, 1) given by

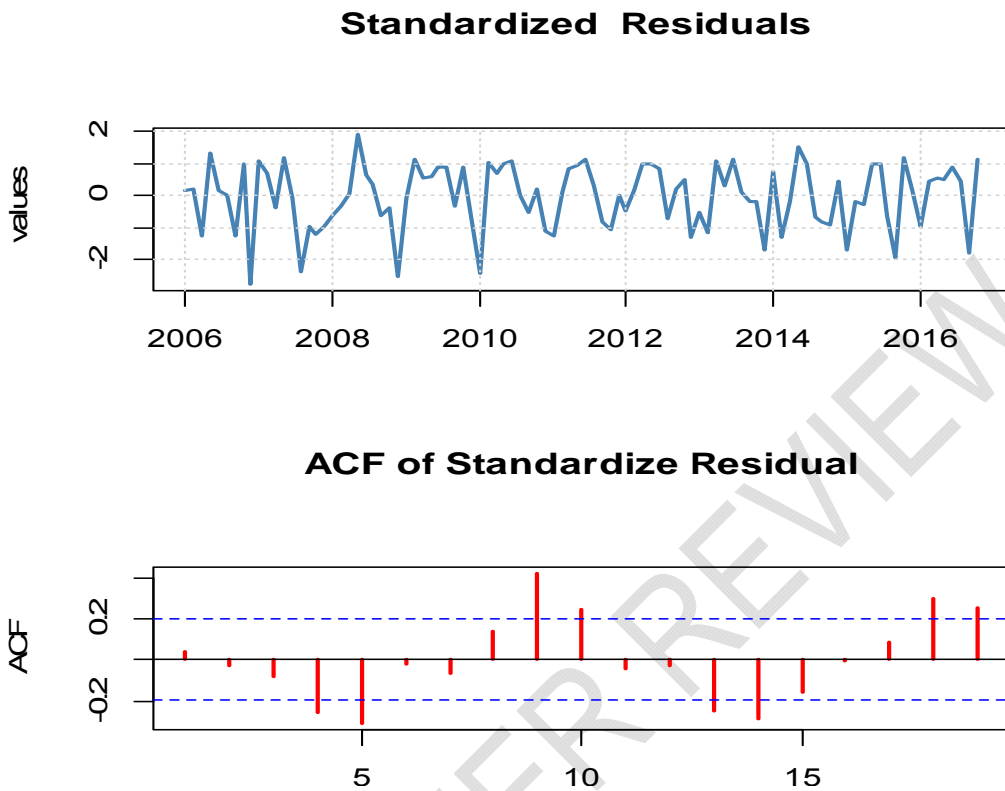
$$y_t = -0.004271 + 0.415772 y_{t-1} - 0.996001 \varepsilon_{t-1} + \varepsilon_t \text{ (see Figure 4).}$$

323
 324



325
 326

327 **Figure 4: Model Diagnosis of ARMA (1, 1)**



328
329

330 **Figure 5: Time plot and ACF of Standardized Residuals**

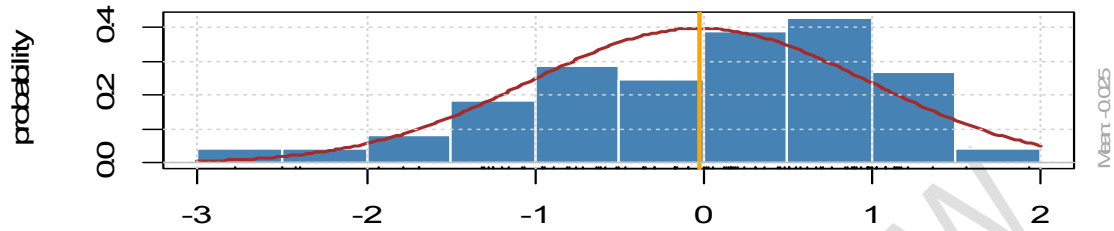
331
332
333
334
335

The time plot of the standardized residuals shows no obvious patterns (does not follow any specific component). The ACF of the standardized residuals and squared standardized residuals show no apparent departure from the model assumptions as shown in Figure 5

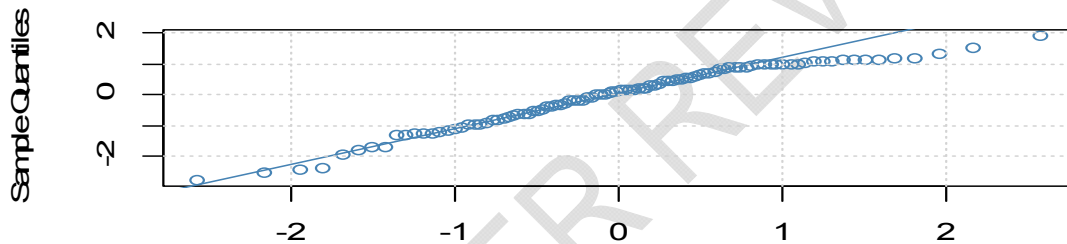
336
337
338

From Figure 6 below the histogram appears to be symmetric and generalized normal q-q plot of the standardized residuals show no departure from model assumptions (i.e. the assumed conditional mean distribution captured the high kurtosis and the heavy tails of the residuals).

Standardized Residuals Distribution



Normal Q-Q Plot



339
340
341
342
343
344
345
346
347
348

Figure 6: Histogram and Normal Q-Q Plot of Standardized Residuals

This suggests the residuals are independent generalized error distribution hence the model seem to be adequate for the data. Consequently, the ARMA (1, 1) is adequate for describing the conditional mean of the differenced rainfall series at 5% significance level.

Table 3: Summary Statistics of Standardized Residuals

| Statistic | Value | Statistics | Value |
|-----------|-----------|------------|----------|
| Mean | 0.001041 | SE mean | 0.101015 |
| Median | -0.303789 | Variance | 1.010203 |
| Minimum | -1.530580 | Std. dev. | 1.005089 |
| Maximum | 3.093621 | Kurtosis | 0.586320 |
| LC L mean | -0.199420 | Skew | 1.089648 |
| UVL mean | 0.201502 | Sum | 0.103062 |
| Nobs | 99.000000 | NAS | 0.000000 |

349
350
351
352
353
354
355
356
357
358
359

The descriptive statistics of standardized residuals in Table 3 shows a standard deviation (1.005) with a general mean (0.001). The empirical distribution of residuals indicates normal kurtosis (0.586) and skewness (1.090). This indicates non-normality of standardized residuals and positively skewed with a lighter tail to the right.

3.4. Model Validation

A model validation test conducted produces a Ljung Box test statistic of 47.207 with a normalized BIC of 6.420 and a Root Mean Square Error of 24.16 supported by a probability value of 0.001. Hence, we fail to reject the null hypothesis that the model is appropriate and suitable for predicting future rainfall

360 figures. An $R^2 = 0.532$ indicates that about 53% of the variations seen in the pattern of rainfall
 361 recorded for the period is being explained by the fitted model i.e. ARMA (1, 1).
 362

363 The fitted model was again used to predict mean actual rainfall for the next two years. That is data up
 364 to 2015 was used to predict the mean actual rainfall for 2016 and from 2016 for 2017 mean rainfall
 365 respectively. It can be observed from the table 4 that the mean rainfall forecasted are very close to
 366 the mean rainfall for the forecasted period suggesting that the fitted model is appropriated for the
 367 data.
 368

369 **Table 4: Mean Forecast of Actual Rainfall for 2016/2017**
 370

| Year (2016) | Actual Rainfall | Forecasted Rainfall | Year (2017) | Actual Rainfall | Forecasted Rainfall |
|-------------|-----------------|---------------------|-------------|-----------------|---------------------|
| Jan. | 86.2 | 1.89 | Jan. | - | 92.21 |
| Feb. | 19.9 | 18.9 | Feb. | - | 33.90 |
| Mar. | 69.8 | 70.1 | Mar. | - | 76.09 |
| Apr. | 131.4 | 129.4 | Apr. | - | 67.23 |
| May. | 156.7 | 158.3 | May. | - | 401.20 |
| Jun. | 283.6 | 290.6 | Jun. | - | 312.76 |
| Jul. | 205.4 | 200.4 | Jul. | - | 138.43 |
| Aug. | 10.0 | 9.8 | Aug. | - | 98.98 |
| Sep. | 130.7 | 128.9 | Sep. | - | 101.90 |

371

372

373 **3.5. Prediction of Next 18 Observations Of Mean Rainfall Returns**
 374

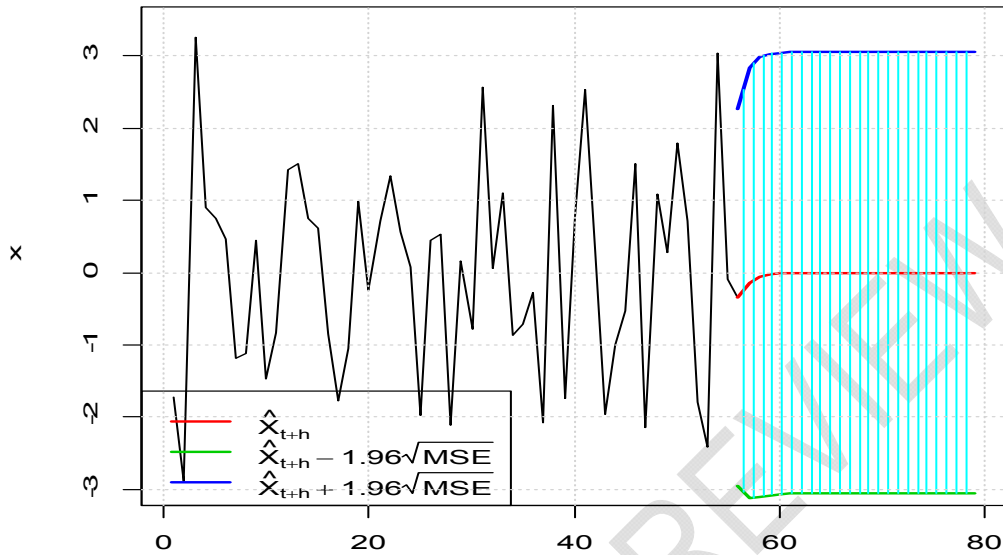
375

376 The fitted model was again employed to predict the mean 1st differenced rainfall for the next two
 377 years. That is data from January, 2006 to December, 2016 was used to forecast 2017/2018 mean
 378 rainfall values. The time plot for the forecasted mean returns is shown in figure 7.

379

380 The up and down movement in black is the actual mean rainfall from January 2006 to December 2016
 381 and the green and blue curve shown is the lower and upper bound of the 95% confidence interval
 382 constructed for the forecasted period. Within the confidence bound is the horizontal broken line which
 383 show the predicted mean rainfall values for the forecasted period. We can observe that the predicted
 384 mean rainfall values for the forecasted period lies within the confidence interval, indicating that the
 model fitted is adequate suitable for the observed rainfall series (see Figure 5).

Prediction with confidence intervals



385
386
387
388
389

Figure 7: Time Plot of 1st Difference Forecasted Rainfall

Table 5: Forecast of 1st Difference in Rainfall for 2017/2018 with Confidence Interval

| Mean Forecast | Mean Error | Standard Deviation | Lower Interval | Upper Interval |
|---------------|------------|--------------------|----------------|----------------|
| 7.804897600 | 101.0965 | 101.0965 | -190.3406 | 205.9504 |
| 5.447387320 | 122.6698 | 101.0976 | - 234.9809 | 245.8757 |
| 3.801975392 | 131.9076 | 101.0985 | -254.7321 | 262.3360 |
| 2.653568772 | 136.1814 | 101.0994 | -264.2571 | 269.5642 |
| 1.852044399 | 138.2161 | 101.1003 | -269.0465 | 272.7506 |
| 1.292624669 | 139.1970 | 101.1011 | -271.5285 | 274.1138 |
| 0.902180604 | 139.6729 | 101.1018 | -272.8517 | 274.6561 |
| 0.629672218 | 139.9046 | 101.1026 | -273.5784 | 274.8377 |
| 0.439476420 | 140.0179 | 101.1032 | -273.9905 | 274.8695 |
| 0.306730261 | 140.0734 | 101.1039 | -274.2321 | 274.8456 |
| 0.214080776 | 140.1009 | 101.1044 | -274.3787 | 274.8068 |
| 0.149416554 | 140.1147 | 101.1050 | -274.4703 | 274.7691 |
| 0.104284500 | 140.1218 | 101.1055 | -274.5293 | 274.7379 |
| 0.072784819 | 140.1256 | 101.1060 | -274.5683 | 274.7138 |
| 0.050799783 | 140.1277 | 101.1065 | -274.5945 | 274.6961 |
| 0.035455442 | 140.1291 | 101.1069 | -274.6125 | 274.6834 |
| 0.024745939 | 140.1300 | 101.1077 | -274.6251 | 274.6746 |
| 0.017271299 | 140.1308 | 101.1077 | -274.6340 | 274.6686 |

390
391
392
393
394
395
396
397
398

Table 5 shows the mean forecasted values of 1st differenced rainfall values for 2017 to 2018. The values obtained indicates that higher rainfall is expected for the period forecasted.

4. CONCLUSION

[The series was found to be non-stationary which resulted from the presence of a unit root in it. The series became stationary after eliminating the unit root by finding the first difference in the amount of rainfall, hence the probability law that governs the behavior of the process does not change over time.

399 The distribution of the 1st differenced series look symmetric with non-constant variance skewed to the
400 right.

401

402 Both the ACF and PACF showed dependency in the 1st differenced series at lag 1, ARMA (1, 1),
403 which has all the parameters to be significant. Thus, the fitted data was found to be the most suitable
404 model for the conditional mean. The model explains the stochastic mechanism of the observed series
405 in ARMA (1, 1). The time series component found in the model were trend and random variation.

406

407 A Ljung Box test statistic of 47.207 with a normalized BIC of 6.420 and a Root Mean Square Error of
408 24.16 supported by a probability value of 0.001 show that the fitted model is appropriate for the data.

409 An $R^2 = 0.532$ indicates that about 53% of the variations seen in the pattern of rainfall recorded for
410 the period is being explained by the fitted model. An 18-month forecast for the mean actual rainfall
411 and mean 1st difference rainfall values made showed that the fitted model is appropriate for the data
412 and an increasing trend of rainfall for forecasted period.

413

414

415 **CONSENT (WHERE EVER APPLICABLE)**

416

417 Not applicable

418

419

420 **ETHICAL APPROVAL (WHERE EVER APPLICABLE)**

421

422 Not applicable

423

424 **REFERENCES**

425

426

427 [1] Owusu K. *Changing Rainfall Climatology of West Africa: Implications for Rainfed Agriculture in*
428 *Ghana and Water Sharing in the Volta Basin*. 2009 Available:

429

430 [2] Boyetey DB, Darkwah KF, Osei-Frimpong E, Acheampomg E and Agyemang, E. Least square
431 determination of spectral time series with trend: Application to rainfall patterns in Ghana.
432 *International Scientific Research Journal*. 2012;1(3):50-61.

433

434 [3] Ukhurebor KE, Abiodun IC. Variation in annual rainfall data of forty years (1978-2017) for south-
435 South Nigeria. 2018; 22(4):511-518.

436

437 [4] Abdul-Aziz AR, Anokye M, Kwame A, Munyakazi L and Nsowah-Nuamah NNN. Modeling and
438 Forecasting Rainfall Pattern in Ghana as a Seasonal Arima Process: The Case of Ashanti
439 Region. *International Journal for Humanities and Social sciences*. 2013; 3(3).

440

441 [5] Boochabun K, Tych W, Chappell NA, Carling PA, Lorsirirat K and Pa-Obsaeng S. *Statistical*
442 *Modeling of Rainfall and Riverflow in Thailand*. 2004. Accessed on March 7, 2013.

443

444 [6] Adiku SGK & Stone RC. Using the Southern Oscillation index for improving rainfall prediction
445 Agricultural water management in Ghana. *Agricultural Water Management*. 1995. 29:85-100.

446

447 [7] Bunting AH, Dennett MD, Elston J and Milford JR. Rainfall trends in the West African Sahel.
448 *Q.J.R. Meteorol. Soc*. 1976; 102:59-64.

449

450 [8] Rodhe H, & Virji H. Trend and periodicities in East African rainfall data. *Mon. Wea Rev*. 1976;104:
451 307-315.

452

453 [9] Tyson PD, Dyer RG and Mametse MN. Sector changes in South African rainfall.1880-972. *Q.J.R.*
454 *Met. Soc*. 1975; 101:817-832

455

456 [10] Ogallo L. Rainfall variability in Africa. *Mon. Wea. Rev*. 1979; 107:1133-1139.

457

- 458 [11] Nicholson SE and Entekhabi D. The quasi-periodic behavior of rainfall variability in Africa and its
459 relationship to the Southern Oscillation. *ArchMeteor Geophys Bioclim Ser.* 1986; 34:31-348.
460
- 461 [12] Mawunya FD. ENSO-Based Rainfall Prediction and Crop Production in Southern Ghana: Interim
462 Report. *International Start Secretariat.* 2007. Accessed on 7/3/2016.
463
- 464 [13] Roncoli C, Ingram P, Kirshen & Jost C. Burkina Faso: Integrating indigenous and scientific rainfall
465 forecasting IK (Indigenous Knowledge). 2001;39.
466
- 467 [14] DeLurgio SA. *Forecasting Principles and Applications.* 1st ed. Irwin/McGraw-Hill, Boston, USA;
468 1998.
469
- 470 [15] Keller G and Warrack B. *Statistics for Management and Economics.* 6th ed. Madrid, Spain:
471 Thompson, Brooks/Cole; 2003.
472
- 473
- 474 [16] Noh J, Engle R and Kane A. *Forecasting Volatility and Option Prices of the S& P 500 Index.*
475 *Journal of Derivatives.* 1994; 2:17-30.
476
- 477 [17] Morley KK. *A Times Series Analysis into the Rainfall Patterns in four selected regions of Ghana.*
478 2011.
479
480
481
482
483
484
485

UNDER PEER REVIEW