THE INTERNATIONAL JOURNAL OF HUMANITIES & SOCIAL STUDIES

Trend Analysis and Forecasting of Precipitation for Southern Nigeria

Dr. Ngozi Isioma Ihimekpen
Student, University of Benin, Nigeria
Sule Joseph
Student, University of Benin, Nigeria

Abstract:
The current study is carried out to determine the potential trend in annual precipitation for Southern Nigeria with a view to assessing its significance, and to model the annual rainfall for future prediction to enhance policy implementation. For this study, ten stations within the southern Nigeria region with continues data series of thirty years and above were subjected to homogeneity analysis using SNHT and Pettitt's test. The widely used Mann-Kendall and Sen's estimator tests were run at 5% significance level on time series data for each of the homogenized stations for the time period, 1971 to 2012. Finally, Holt's exponential smoothing model was developed which allow predictions on future evolutions. These analyses were executed using the R i386 3.4.3 programming language. Three out of the ten stations (Benin, Ikeja and Onitsha) failed both homogeneous tests and were excluded from further analysis. The Mann-Kendall tests indicated that Akure, Enugu, and Port Harcourt stations do not indicate statistically significant results, while the other states show statistically significant increasing trends, Sen's estimator revealed a negative slope for Akure. Holt's Exponential smoothing revealed that the annual rate of increase in rainfall were; -0.514mm/year for Akure, 2.4mm/year for Calabar, 15.232mm/year for Enugu, 40.03mm/year for Ibadan, 27.189mm/year for Owerri, -20.643mm/year for PH and 162.708mm/year for Uyo respectively.

Keywords*: Trends analysis, forecasting, precipitation, southern Nigeria*

1. Introduction

Water resources and hydrologic cycle are largely controlled by climatic factors. The climatic factors are precipitation, temperature, wind velocity, radiation and humidity. Changes in any of these meteorological parameters may affect the quality, quantity and spatial distribution of water. Rainfall is one of the key climatic factor that is very essential to humans e.g. in terms of water availability, agriculture etc., examination of it behavior is important for understanding of climate variability because it is highly variable spatially and temporarily at different local, regional and global scales [1].

Climate variability is not as noticeable as weather variability because it happens over seasons and years. Each "up and down" fluctuation can lead to conditions which are warmer or colder, wetter or drier, stormier or quiescent [2]. Climate fluctuates annually below or above a long-term mean value. Climatic variability can be described as the annual difference in values of specific climatic variables within averaging periods such as a 30-year period [3]. Climate variability may be due to natural internal processes within the climate system (internal variability) or to variations in natural or anthropogenic external forcing (external variability).

On the other hand, climate change is the change within the statistical distribution of the meteorological events over certain periods of time. This can range from centuries, decades or even millions of years [4]. Climate change is slow and gradual, is very difficult to perceive without scientific records. How do scientists detect climate change? They look for long-term continuous changes (trends) in climatological averages and normals and the variety around these averages. One of the commonly used tools for detecting changes in climatic and hydrologic time series is trend analysis. Trend Analysis is the practice of collecting information and attempting to spot a pattern, in the information.

Previous studies have analyzed rainfall trends over entire or part of Nigeria. For example, [5] examines spatial analysis of rainfall in the climatic regions of Nigeria using Insitu data. [6] carried out analysis using long time series rainfall data from Katsina, Zaria and Kano meteorological stations, testing for trend and the results showed a decrease in mean annual rainfall for the three stations. [7] reported significant increase in rainfall over nine stations in northern Nigeria between 1953 and 2002. The results showed a general decline of dry season's contribution to annual rainfall i.e. dry period is getting drier. [8] applied data from the South zone of the country to study seasonality in rainfall distribution. The author considered nine meteorological stations in the Niger Deltan belt and observed significant seasonal pattern from February/March to November and a short dry season from December to January/February. The results further showed a northward increase in rainfall in parts of the eastern side of the Niger Delta. From these and other studies, a range of potential climatic impacts on the hydrologic regime for various geographic areas can be hypothesized.

The rates of change are significantly different among regions. These lead to increasing population pressures, intensive agricultural land-use, overgrazing and bush burning in some regions/cities than others. For this purpose, trend analysis of weather data can be a valuable tool to investigate its variability pattern. Trend analysis of any of the meteorological parameters can be carried out using different methods but, in this study, non-parametric Mann-Kendall method was adopted because it is resilient to skewed distribution, missing values and values that fall outside the detection limit, and to the non-stationary nature of data [9]. Thus, if some values are missing or if an outlier is present, the results would not be affected much because the ranks would not significantly change. Mann-Kendal can be run on different statistical software's like excel, SPSS, R and MATLAB but also in this studying R was chosen because R is free and available for all major platforms. It has excellent built- in helps, it has excellent graphical capacities, it is easy to extend with user defined functions. The aim of this study was to determine the potential trend in annual precipitation for Southern Nigeria with a view to assessing its significance, and to model the annual rainfall for future prediction to enhance policy implementation.

2. Materials and Methods

2.1. Study Areas

This study was carried out in Southern Region of Nigeria. The area lies between lat. 4^0 32' N – lat. 7^0 57' N and long. 2^0 31' E – long. 7^0 48' N. The climate of southern region of Nigeria is tropical in nature and it is characterized by wet and dry seasons. The mean temperature ranges between 19 °C and 30 °C, while the annual rainfall ranges between 924 mm and 2016 mm. Figure 1 shows the location of rain gauge stations used in this study on the map of Nigeria.

Figure 1: Map of Nigeria Showing the Location of the Synoptic Stations (Article.Sapub.Org)

2.2. Source of Data

The precipitation data required for the analysis were sourced from the Nigerian Meteorological Agency (NIMET). A total of ten synoptic stations in the southern region of Nigeria were selected due to their continuous records for thirty years and above namely: Akure, Benin, Calabar, Enugu, Ibadan, Ikeja, Owerri, Onitsha, Port Harcourt and Uyo. The data collected was made up of annual rainfall. Table 1 shows the characteristics of the stations used and the available length of data for the study.

Station Name	**Latitude**	**Longitude**	**Elevation A.M.S.L. (M)**	**Length Of Records (Years)**
Akure	7.25N	5.19	233	31 (1980-2011)
Benin	6.25N	5.30E	138	41 (1971-2012)
Calabar	4.58N	8.21E	99	41 (1971-2012)
Enugu	6.27N	7.29E	183	41 (1971-2012)
Ibadan	7.22N	3.58E	183	41 (1971-2012)
Ikeja	6.40N	3.45E	73	41 (1971-2012)
Onitsha	6.06N	6.42E	127	30 (1975-2005)
Owerri	5.27N	6.59E	127	40 (1972-2012)
PH	4.41N	6.59E	83	41 (1971-2012)
Uyo	5.05N	7.55E	196	32 (1980-2012)

Table 1: Characteristic of Selected Synoptic Stations

2.3. Tests for Homogeneity

In this research, Standard normal homogeneity test and Pettitt's test homogeneity analysis approaches are considered.

2.3.1. The Pettitt's Test

The Pettitt test is a non-parametric test that evaluate the homogeneity of a series based on the ranks of the elements. The statistic is defined as

$$X_k = 2\sum_{i=1}^{k} r_i - k(n+1) \qquad k = 1,2,3,\ldots, n \tag{1}$$

Where $r_1, r_2, r_3, \ldots, r_n$ are the ranks of the tested variables that arranged in ascending order. The value of statistic will reach its maximum or minimum near the year k=E if there is a shift of mean level in year E. The maximum value is given as

$$X_E = \max_{1 \le k \le n} \left| X_k \right| \tag{2}$$

2.3.2. Standard Normal Homogeneity Test (SNHT)

The standard normal homogeneity test assumes that the testing variables are normally distributed and it is able to detect in homogeneity near the beginning and end of the series. The statistic can be expressed in the equation:

$$T_k = K\overline{Z_1^2} + (n-k)\overline{Z_2^2} \qquad k = 1,2,3,\ldots, n \tag{3}$$

where

$$\overline{Z_1} = \frac{1}{k}\sum_{i=1}^{k}(p_i - \overline{p})/s \text{ and } \qquad \overline{Z_2} = \frac{1}{n-k}\sum_{i=k+1}^{n}(Y_i - \overline{Y})/s \tag{4}$$

and P_i is the time series to be tested, $\overline{p}$ is the mean and s is the standard deviation. T_k reaches its peak value near theyear k=K if a break is located at the year K.The test statistic is expressed as:

$$T_0 = \max_{1 \le k \le n} \left| T_{(k)} \right| \tag{5}$$

If T_0 is greater than the critical value, the null hypothesis is rejected and there is a significant break present within the time series.

2.4. Sen's Slope Estimation

The Theil-Sen approach (TSA), a commonly-used method to quantify the significant linear trends in time series, was applied in this study. The TSA is considered more robust than the least-squares method due to its relative insensitivity to extreme values and better performance even for normally distributed data [10]. In general, the slope Q between any two values of a time series x can be estimated from

$$\mathrm{Q} = \frac{\left(x_k - x_j\right)}{\left(k - j\right)}, \mathrm{k} \neq \mathrm{j} \qquad (6)$$

For a time, series x having n observations, there are a possible N = n (n - 1)/2 values of Q that can be calculated. According to Sen's method, the overall estimator of slope is the median of these N values of Q. The overall slope estimator Q* is thus:

$$\mathrm{Q}^* = \begin{matrix} Q_{(N+1)/2}, \left(N \to Odd\right) \\ \dfrac{(Q_{N/2} + Q_{(N+2)/2}}{2}, (N \to Even) \end{matrix} \qquad (7)$$

When significant trends in the data were detected, 95% confidence intervals were calculated using non-parametric techniques as described by [11]. The quantity Cα was first calculated as

$$C\alpha = Z_{1-\alpha/2}\sqrt{V(S)} \qquad (8)$$

where Z is again the standard normal deviate, V(S) is as defined earlier, and α is taken as 0.05. Indices M_1 and M_2 are determined from:

$$\mathrm{M}_1 = \frac{\left(N - C\alpha\right)}{\left(2\right)} \qquad (9)$$

$$\mathrm{M}_2 = \frac{\left(N + C\alpha\right)}{\left(2\right)} \qquad (10)$$

where N is as previously defined. Finally, the confidence limits are defined by the M_1th and (M_2+1)th largest of the ordered estimates of Q, with interpolation as appropriate for non-integer values of M_1 and M_2.

2.5. Analyses of Lag-1 Autocorrelation and the Pre-Whitening Process

Before applying the Mann-Kendall test to the time series of annual precipitation levels, it is necessary to investigate whether the annual precipitation levels are serially correlated. If they are, there would be need to Pre-whitening the series by using the Mann-Kendall test in conjunction with block bootstrapping in order to account for the serial correlation present in the annual precipitation levels. If they aren't, the Mann-Kendall test is applied "as is", without the need to correct the P-value of the test for serial correlation.

2.6. Trend Detection and Characterization

The non-parametric Mann-Kendall test was used to assess the presence of significant trends in the precipitation data. The Mann-Kendall statistic S of the series x is given by:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathrm{sgn}(x_j - x_i) \qquad (11)$$

where sgn is the signum function. The variance associated with S is calculated from:

$$V(S) = \frac{\left(n(n-1)(2n+5) - \sum_{k=1}^{m} t_k(t_k - 1)(2t_k + 5)\right)}{18} \qquad (13)$$

Where m is the number of tied groups and t_k is the number of data points in group k. In cases where the sample size n > 10, the test statistic Z(S) is calculated from:

$$Z(S) = \begin{matrix} \dfrac{S-1}{\sqrt{V(S)}}, for S \triangleright 0 \\ 0, for S = 0 \\ \dfrac{S+1}{\sqrt{V(S)}}, for S \triangleleft 0 \end{matrix} \qquad (14)$$

Positive values of Z(S) indicate increasing trends, while negative Z(S) values reflect decreasing trends. Trends are considered significant if |Z(S)| are greater than the standard normal deviate $Z_{1\alpha/2}$ for the desired value of α (taken as 0.05 in this study).

2.7. The Holt's Exponential Smoothing

If the time series can be described using an additive model with increasing or decreasing trend and no seasonality, Holt's exponential smoothing is used to make short-term forecasts [12].

Holt's exponential smoothing estimates the level and slope at the current time point. Smoothing is controlled by two parameters, alpha, for the estimate of the level at the current time point, and beta for the estimate of the slopeb of the trend component at the current time point. The parameters alpha and beta have values between 0 and 1, and values that are close to 0 mean that little weight is placed on the most recent observations when making forecasts of future values.

This model gives more weight to recent rainfall and the weight slightly decreases for data in the past. For this model, the predicted rainfall was calculated using Equation:

$$F_{t+m} = s_t + b_t m \tag{15}$$

where F_{t+m} is the predicted rainfall in time t + m in which both t and m are in the same unit of time and m is equal to 1. s_t and b_t can be calculated from Equations (16) and (17):

$$s_t = \alpha x_t + (1-\alpha)(s_{t-1} + b_{t-1}) \tag{16}$$

$$b_t = \beta(s_t - s_{t-1}) + (1-\beta) b_{t-1} \tag{17}$$

Where α is the smoothing constant for the actual data and the predicted values, X_t is the actual rainfall in time t and β is the smoothing constant for the actual and estimated trend.

To make forecasts, HoltWinters () function in R is used. To use Holt Winters () forHolt's exponential smoothing, the parameter gamma is set to be equal to FALSE (the gamma parameter is used for Holt-Winters exponential smoothing).

3. Result and Discussion

3.1. Result of Homogeneity Test

As seen from the table 2, SNHT scored four in homogeneities each while Pettitt test scored three in homogeneities. Benin, Ikeja, and Onitsha stations failed the two homogeneity tests while Uyo failed only SNHT. The three stations which failed both SNHT and Pettitt's test were excluded from further analysis.

Station	**Pettitt test**		**SNHT**		**Remark**
	K	P-value	K	P-value	
Akure	62	1.073	1.9	0.872	Accept Ho for both
Benin	194	0.035	10.1	0.017	Reject Ho for both
Calabar	237	0.215	10.3	0.417	Accept Ho for both
Enugu	172	0.192	4.7	0.334	Accept Ho for both
Ibadan	154	0.266	7.5	0.084	Accept Ho for both
Ikeja	266	0.007	9.8	0.021	Reject Ho for both
Onitsha	170	0.007	10.9	0.007	Reject Ho for both
Owerri	302	0.718	16.2	0.452	Accept Ho for both
PH	140	0.424	3.1	0.633	Accept Ho for both
Uyo	101	0.326	11.7	0.004	Accept Ho for PT only

Table 2: Homogeneity Tests for Annual Precipitation

3.2. Analyses of Lag-1 Autocorrelation and the Pre-Whitening Process

The results of lag-1 autocorrelations process of the annual precipitation produced for all the seven homogenized synoptic stations are shown in figure 2(i-v). As shown, the plots suggest that the autocorrelation and partial autocorrelation present in this series do not appear significant and did not require any pre-whitening procedure. (Indeed, most of the vertical spikes in the ACF and Partial ACF plots produced by R fall within the horizontal band defined by the blue dotted lines beyond which autocorrelations and partial autocorrelations would be deemed to be significant.)

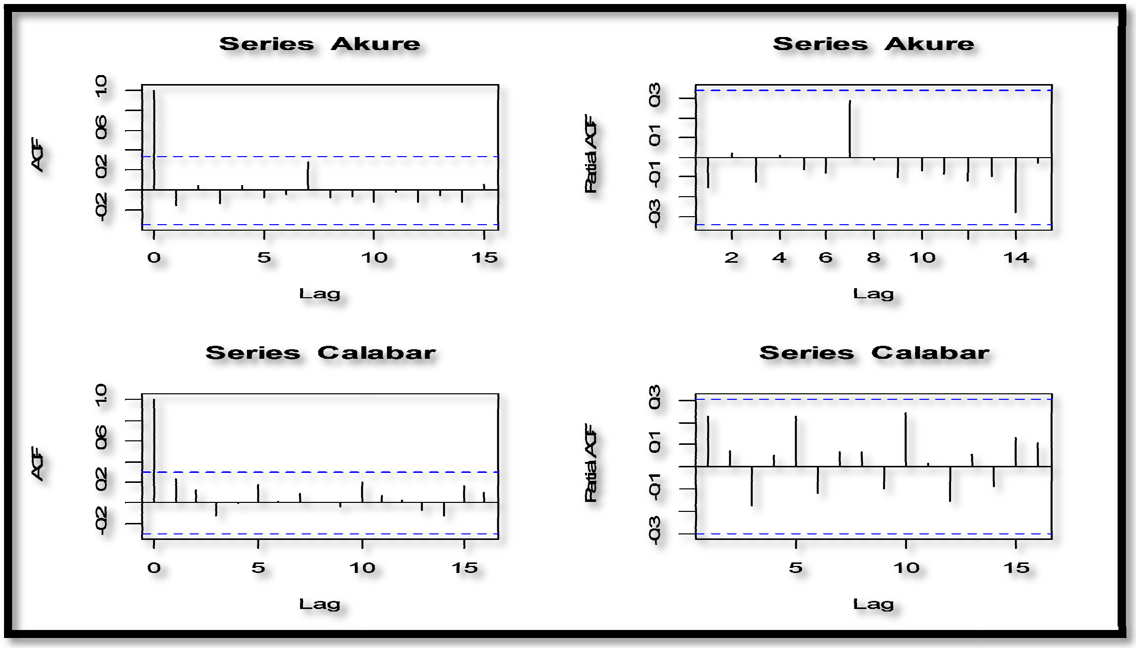

Figure 1: Acf and Pacf Analysis for Akure and Calabar Stations Showing Vertical Spikes Falls within the Horizontal Band Defined by the Blue Dotted Lines

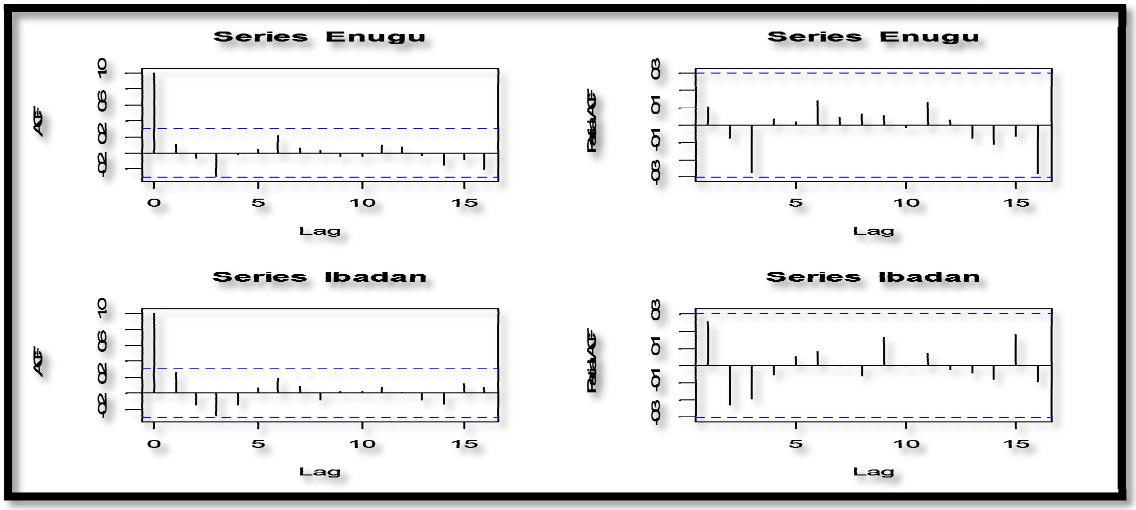

Figure 2: Acf and Pacf Analysis for Enugu and Ibadan Stations Showing Vertical Spikes Falls within the Horizontal Band Defined by the Blue Dotted Lines

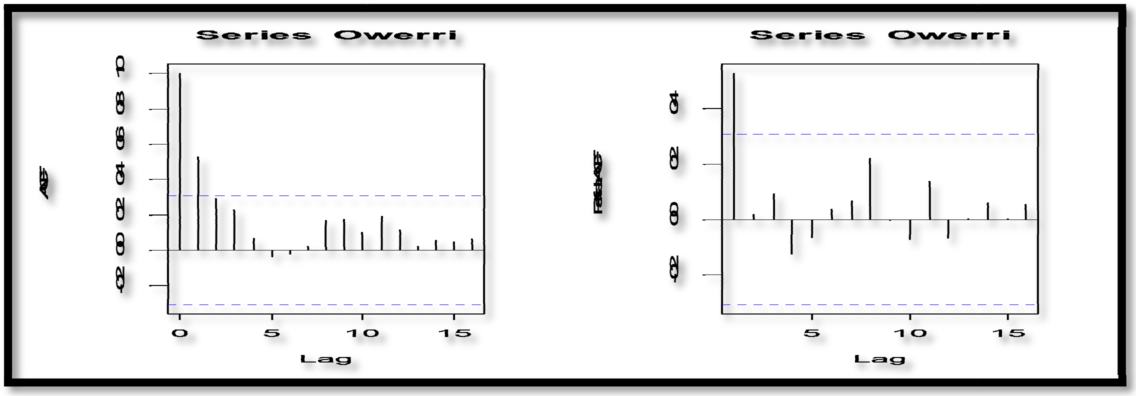

Figure 3: Acf and Pacf Analysis for Owerri Station Showing Vertical Spikes Falls within the Horizontal Band Defined by the Blue Dotted Lines

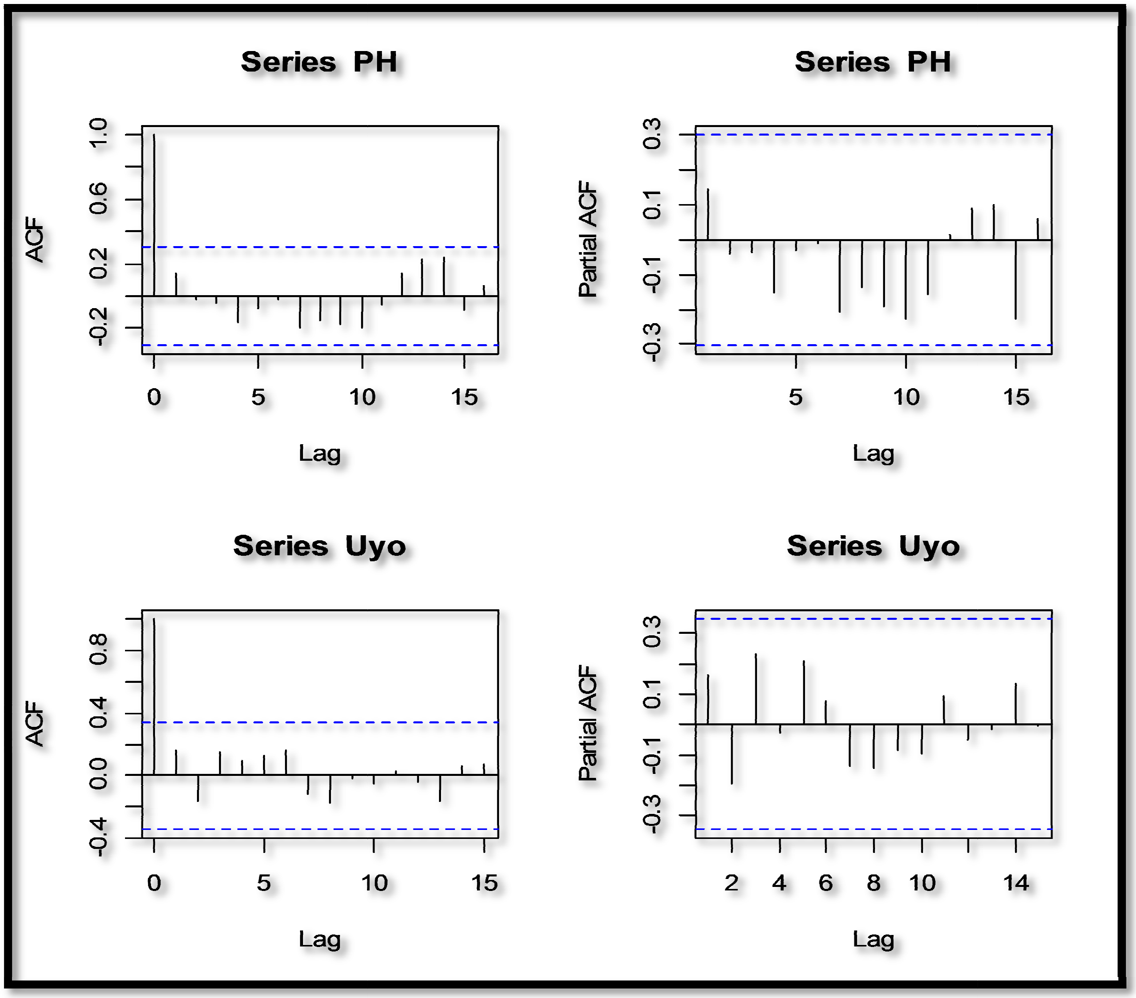

Figure 4: Acf and Pacf Analysis for PH and Uyo Stations Showing Vertical Spikes Falls within the Horizontal Band Defined by the Blue Dotted Lines

3.3. Trend Analysis

On running the Mann-Kendall and Sen's slope tests on the annual precipitation data, the following results in Table 3 were obtained for the seven homogenized stations under the study area. For Mann-Kendall, if the p value is less than the significance level α (alpha) = 0.05, H_0 is rejected. Rejecting H_0 indicates that there is a trend in the time series, while accepting H_0 indicates no trend was detected. On rejecting the null hypothesis, the result is said to be statistically significant. For Sen's slope, positive slope indicates rising trend while negative slope indicates falling trend.

States	Mann-Kendall Statistic (S)	Kendall's Tau	Z-Statistic	Var (S)	P-Value (Two Tailed Test)	Slope	Test Interpretation
Akure	-30	-0.057	-0.4	4165.33	0.631	-0.985	Accept Ho
Calabar	217	0.252	2.3	8514.33	0.019	11.544	Reject Ho
Enugu	123	0.143	1.3	8514.33	0.186	3.700	Accept Ho
Ibadan	184	0.194	2.1	7926.66	0.0398	9.005	Reject Ho
Owerri	360	0.439	4.0	7926.66	5.5e-0.5	21.908	Reject Ho
PH	113	0.131	1.2	8514.30	0.2248	6.150	Accept Ho
Uyo	139	0.246	2.0	3802.66	0.0497	16.260	Reject Ho

Table 3: Results of the Mann-Kendall and Sen's Slope Tests for Precipitation Data

Table 3 indicates that the Null Hypothesis was accepted for only three stations, Akure, Enugu and PH.

3.4. Forecasts Using Holt's Exponential Smoothing

The estimated value of alpha and of beta for each homogenized station are shown in table 4 Uyo and Owerri value is high, which indicate that the estimate of the current value of the level, and of the slope b of the trend component are not based mostly upon very less recent observations in the time series.

Station	Alpha	Beta	Coefficient a	Coefficient b	SSE
Akure	0.0011	1.0000	1390.6	-0.51	971571.4
Calabar	0.0000	0.0000	3050.0	2.40	6576933
Enugu	0.3038	0.1569	1908.22	15.23	3947774
Ibadan	0.1144	0.4605	1529.20	40.0	3246905
Owerri	0.5815	0.0648	2725.97	27.18	6456321
PH	0.3913	0.1816	2414.89	-20.64	11344183
Uyo	0.6753	0.2679	3065.36	162.70	7143981

Table 4: Holt's Exponential Smoothing Parameters, Coefficients and Sum-of-Squared Errors (SSE)

The 'forecast errors' are calculated as the observed values minus predicted values, for each time point. The forecast errors can only be calculated for the time period covered by the original time series for the rainfall data. One measure of the accuracy of the predictive model is the sum-of-squared errors (SSE) for the in-sample forecast errors. The value of the sum-of-squared-errors for each station are also shown on table 4.

The forecast of rainfall for the years 2020, 2025 and 2030 using forecast () function in R are shown below on table 5 while the plot for each station with 80% and 95% prediction level are shown on figure 3.

Station	2020	2025	2030
Akure	1386.56	1383.99	1380.39
Benin	2275.01	2275.01	2275.01
Calabar	3069.20	3081.20	3093.20
Enugu	2030.08	2106.24	2182.40
Ibadan	1889.47	2089.61	2289.76
Ikeja	1787.31	1855.04	1922.77
Onitsha	1959.22	1961.52	1963.81
Owerri	2943.48	3079.43	3215.38
PH	2249.75	2146.54	2043.32
Uyo	4529.73	5343.27	6156.81

Table 5: Projected Annual Rainfall for the Study Area (Mm)

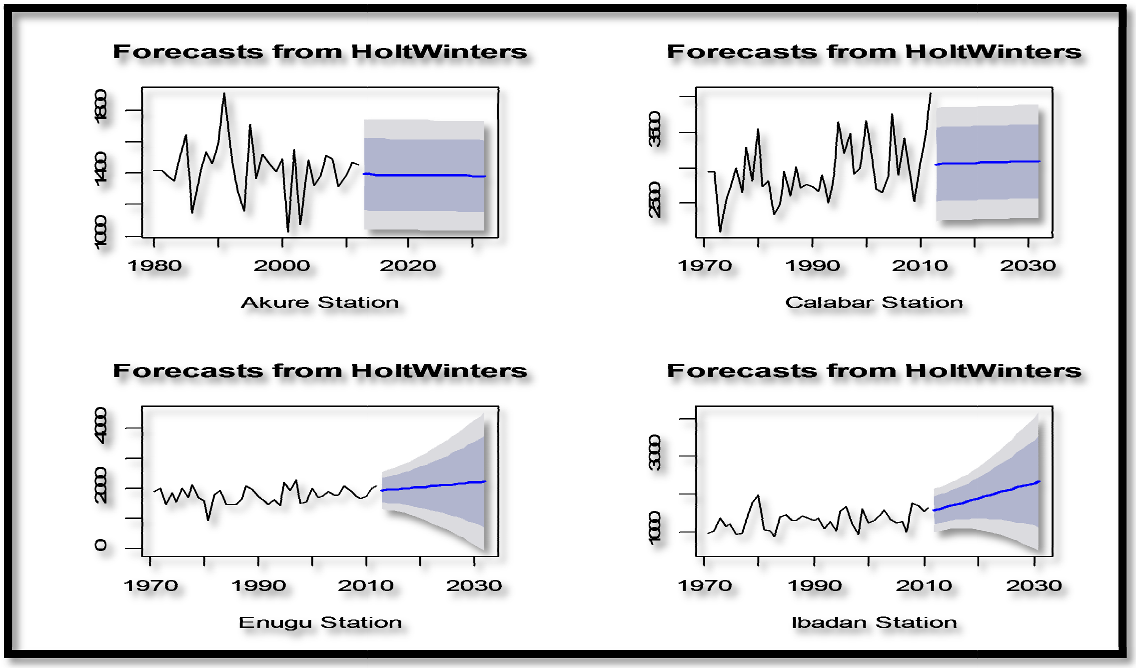

Figure 6

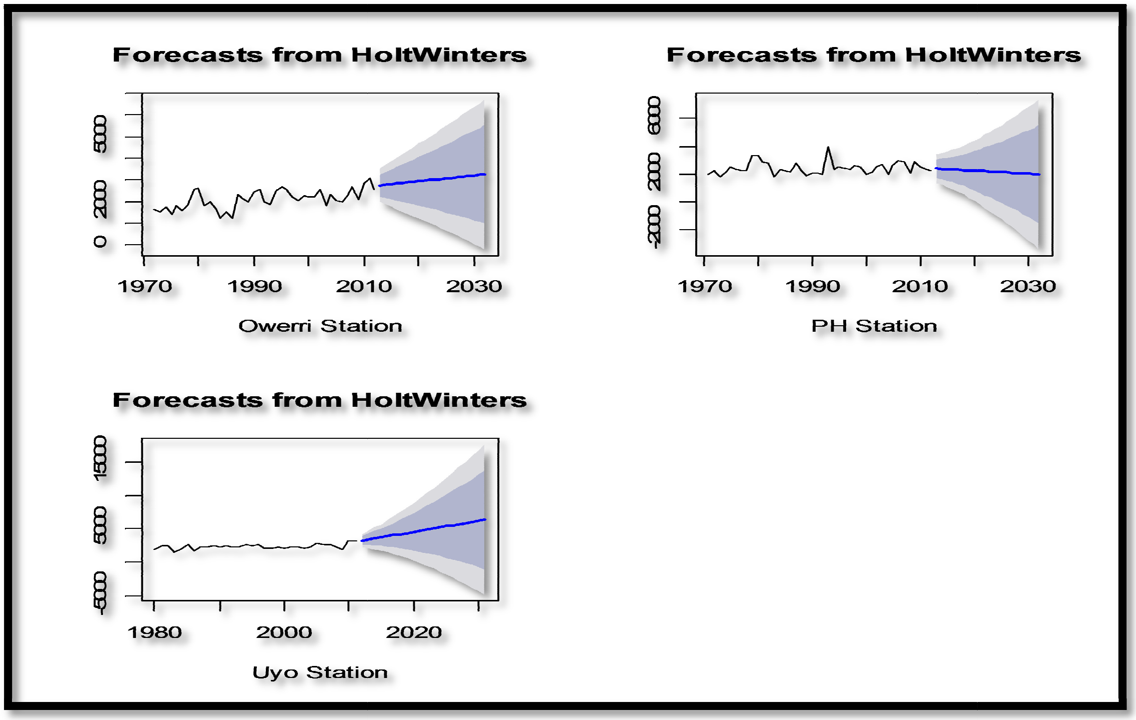

Figure 7: The Graphics of the Annual Precipitation Time Series (Grey); Holt Winters Exponential Smoothing (Red)

Benin, Calabar, Onitsha and Uyo will experience slightly constant annual rainfall. Enugu, Ibadan, Ikeja, Owerri and Uyo will experience Increase in rainfall while Akure and PH will experience slightly decrease in annual rainfall in the future.

3.5. Discussion

Three stations series failed both the SNHT and Pettitt's homogeneous tests and were excluded from further analysis (table 1) as non-homogeneous. Analysis of the serial correlation effect on the seven homogenized stations showed that the data did not revealed a significant lag-1 correlation coefficient therefore pre-whitening was not needed. The results of MK trend analysis test found that statistically significant increasing trend was obtained for Calabar, Ibadan, Owerri and Uyo while Akure, Enugu and PH stations do not indicate statistically significant results. There is consistency in result also for the Sens slope analysis as Akure, Enugu and PH shows a negative and very low slope. The Holt's Winters exponential model revealed that the annual rate of increase in rainfall over the period of investigation were; -0.514mm/year for Akure, 2.4mm/year for Calabar, 15.232mm/year for Enugu, 40.03mm/year for Ibadan, 27.189mm/year for Owerri, -20.643mm/year for PH and 162.708mm/year for Uyo respectively. This indicates potential increase for all the states except for Akure and PH, where they are decrease in the total amount of annual precipitation.

Based on the above results, it is importance to discuss the political, ecological, economic, and social impacts that could result if increasing precipitation trends continue in Enugu, Ibadan, Owerri and Uyo stations in the future. Increased precipitation can influence the water quality and possibly result in the outbreak of waterborne diseases. Excess rainfall could also lead to soil saturation as well as to runoff and soil erosion problems. On the other hand, a decreasing precipitation for Akure and PH will have repercussions in the sustainability of surface water resources and groundwater recharge. It might have serious agricultural implications because most agricultural activities in this area rely on rainfall of this period. Continues decrease in the annual rainfall has an implication in construction of bore holes, wells, and other water resource development projects that depend on water from rivers and ground water sources because of the lowering of the water table.

4. Conclusion

It can be concluded that the stations that exhibit a statistically significant positive trend have a tendency to increase with time and their increment was predictable, while the stations that exhibit statistically significant negative trend have a tendency to decrease and their reduction was also predictable. The results also show that, the climate of the region (southern Nigeria) indicates a tendency towards wetter condition (71.42%) rather than the increasing dryness.

4.1. Recommendation

Enugu, Ibadan Owerri and Uyo should take a proactive measure against flood by constructing more drainages and canals while Akure and PH should take serious measure against drought by introducing water conservation methods.

5. References

i. De Luis, M., Raventós, J., González-Hidalgo, J.C., Sánchez, J.R., Cortina, J. (2000). Spatial Analysis Of Rainfall Trends In The Region Of Valencia (East Spain), Int. J. Climatology. 20 (12), 1451–1469.
ii. NOAA (National Oceanic and Atmospheric Administration). Observing Climate Variability And Change. Retrieved From: Http://Www.Oar.Noaa.Gov/Climate/T_Observing.Html, Accessed On May 10, 2017.
iii. Melillo J.M., Callaghan T.V., Woodward F.I., Salati E., Sinha K. (1990). Climate Change: Effects On Ecosystems. In: Houghton JT, Jenkins GJ, Ephraums JJ (Eds), Climate Change, Cambridge University Press, New York, 283-310.
iv. Ogboi C. O. (2016). Variability and Time Series Analysis Of Rainfall Pattern/Trend In Benin City. M.Eng Thesis Summated To The Civil Engineering Department, University Of Benin, Nigeria.
v. Ojo O., Adedayo K., Emmanuel I. (2015).Spatial Analysis Of Rainfall In The Climatic Regions Of Nigeria Using Insitu Data.Journal Of Environment And Earth Science, Vol.5, No.18.
vi. Ekpoh, I.J. (2007). Climate and Society in Northern Nigeria: Rainfall Variability And Farming. The International Journal Series on Tropical Issues, 8(3), 157-162.
vii. Ati, O. F, Stigter, C. J, Iguisi,E.O, And Afolayan, J. O.(2009). Profile Of Rainfall Changes And Variability In The Northern Nigeria, 1953-2002. Research Journal of Environmental And Earth Sciences, 1(2), 58-63.
viii. Adejuwon, J.O. (2012). Rainfall Seasonality In The Niger Delta Belt, Nigeria. Journal Of Geography And Regional Planning, 5(2): 51-60.
ix. Yue, S., Pilon, P. (2004). A Comparison Of The Power Of The T Test, Mann-Kendall And Bootstrap Tests For Trend Detection. Hydrological Sciences Journal, 49(1): 21-37.
x. Hirsch, R.M.; Slack, J.R.; Smith, R.A. "Techniques of Trend Analysis For Monthly Water Quality Data". Waterresources. Res., 18, 107–121, 1982.
xi. Salmi T, Määttä A, Anttila P, Ruoho-Airola T, Amnell T. "Detecting Trends Of Annual Values Of Atmospheric Pollutants By The Mann-Kendall Test And Sen's Slope Estimates. The Excel Template Application Makesens". Finnish Meteorological Institute Publications On Air Quality No. 31, Helsinki, Finland: 135, 2002.
xii. Avril Coghlan. (2017).A Little Book of R for Time Series. Release 0.2.