Proceedings of the Fourth International Conference on Engineering and Technology Research February 23 - 25, 2016 ISBN: 978-2902-58-6 Volume 4

TREND DETECTION AND COMPARISM OF TWO FORECASTING MODEL FOR HYDRO-METEOROLOGICAL PHENOMENA OF OGUN STATE, NIGERIA

*JIBRIL, Ibrahim [1], MOHAMMED, Jiya Mamman [1], MUSA, John Jiya [2] and ADEOYE, Peter A [2].

1. Niger State College of Agriculture, P M B 109, Mokwa, Nigeria. Tel: +234-8064-485-066. E-mail: abbatee2007@gmail.com
   * Corresponding Author.

2. Department of Agricultural and Bioresources Engineering, Federal University of Technology, P M B 65, Minna, Nigeria. Tel: +234-8094-747-246.
   Email: johnmusa@futminna.edu.ng

Abstract

Extreme events of atmospheric phenomena are often non-deterministic in nature, and this has been a major constraint in achieving agricultural sustainability, which directly affects economic advancement, most especially of developing countries. This call for an urgent look at key climatic phenomena and finding simpler, but most reliable ways of predicting them in order to make proper plan against reoccurrence. To facilitate this research work, 29 years information of the observed relative humidity of Ogun State was obtained from the Federal Ministry of Water Resources, Abeokuta, Nigeria. The data collected covers the periods between 1982 and 2009 and were pre-whitened and aggregated into monthly and annual time series to clear the doubt of outliers. The Mann-Kendal non-parametric test, Long-range dependency test and test for serial dependence were carried out. The Mann-Kendal Z-value obtained was -1.37, which gives no reason to expect the presence of trend in the time series, but the Sen.'s slope trend line indicated slight decreasing trend. The spectral density analysis showed high variance to lower frequency, signifying a positive correlation which was in line with the Durbin-Watson test that gives a d-value of 1.28. No evidence of seasonal effect in the series as clearly depicted by the monthly Periodogram, and the data was therefore treated basically as stochastic. The data was divided into two and the first 20 years was used for model development, while the remaining 9 years was used for validation. Multiple Linear Regression (MLR) and Autoregressive Moving Average (ARMA) models were considered. The results indicated that there may be continues decrease in the amount of air moisture for a while. However, the best predictive model was found to be ARMA, though MLR give better validation. It is therefore recommended that other climatic parameters be looked into for proper planning.

Keywords: ARIMA-model, Climatic Elements, Forecast, MLR-model and Trend

Introduction

Atmospheric water vapor is widely recognized to be a key climate variable. It is the dominant greenhouse gas and provides a key feedback for amplifying the sensitivity of the climate to external forcing (Held and Soden, 2000; Soden and Held, 2006). Water vapor is also an important component of the hydrological cycle. Future increases in water vapor in response to a warming climate are fundamentally linked to the expected changes in moisture convergence, precipitation extremes, meridiunal energy transport and an overall weakening of the atmospheric circulation (Soden and Held, 2006).

Relative humidity represents the amount of water vapor which is in the atmosphere. They primarily come from the evaporation of surface water and superficial layers of soil, from plant and animal respiration and from some technological processes. Millions of such water droplets come together to form clouds. So, if the humidity is more (or relative humidity becomes maximum), it leads to the formation of clouds and subsequent precipitation (Tsoho, 2008). Humidity control is important in many engineering applications, such as space air conditioning, storage warehouses, process industries and many others (Rakesh and Arun, 2014). Information on relative humidity is therefore very important in the life of man and his animals as it is one of the key to changes in atmospheric weather condition.

According to (Mathur, *et al.,* 2001), weather for future is one of the most important attributes to forecast because agriculture sectors as well as many industries are largely dependent on the weather conditions. Weather conditions are required to be predicted not only for future planning in agriculture and industries but also in many other fields like defense, mountaineering, shipping and aerospace navigation etc (Roadknight,*et al.,* 1997).

Regression models are often used for estimating the future events or values using features of a particular time series or other related time series data (Chatfield, 1994). Trend extraction and curve fitting methods are also used to estimate the future behavior of the time series and to fit the future data according to the trend. However, regression models are more of deterministic, which is unlikely of an ideal situation.

Materials and Method

The name 'Ogun basin' is derived from two major rivers that drains within; Rivers Ogun and Osun, though they have smaller tributaries like; Sasa, Ona, Ibu, Ofiki, Yewa rivers etc. The basin under consideration is located in South Western Nigeria (Ewemoje and Ewemooje, 2011). The entire basin is bounded by Oyo state in the north, Osun and Ondo States in the east and Lagos State in the South as shown in (figure 1).

The Ogun basin covers the whole of Ogun State, located in southern Nigeria, bordered geographically by latitudes 6^0 26^I N and 9^0 10^I N and longitudes 2^0 28^I E and 4^0 8^I E. About 2% of the basin area falls outside Nigeria in the Benin Republic. The land area is about 23,000km^2. The relief is generally low, with the gradient in the North-South direction.

The two major vegetation zones that can be identified the area are the high forest vegetation in the north and central parts, and the swamp/mangrove forests that cover the southern coastal and floodplains, next to the lagoon. It has two distinct seasons throughout the year. The monthly rainfall distribution in the study area shows a distinct dry season extending from November through March and a rainy season divided into two periods: April – July and September – October. The mean annual rainfall data for 30 years showed a variation from about 1,150mm in the northern part to around 2,285mm in the southern extremity. The estimates of total annual potential evapotranspiration have been put between 1600 and 1900mm. (Ewomoje and Ewomooje, 2011).

Data Collection and Pre-whitening

The relative humidity data used for this study were obtained from the federal ministry of water resources, Abeokuta, Nigeria. The data collected covered a period of twenty nine years (1982-2009). These values were obtained by the use of GPS (Global Position System) equipment. Data preprocessing is an important task in almost all modeling techniques. The data obtained are the time series types which are collected monthly for the entire period of interest. For the purpose of this study, the mean annual values of the data were first determined and pre-whitened before use.

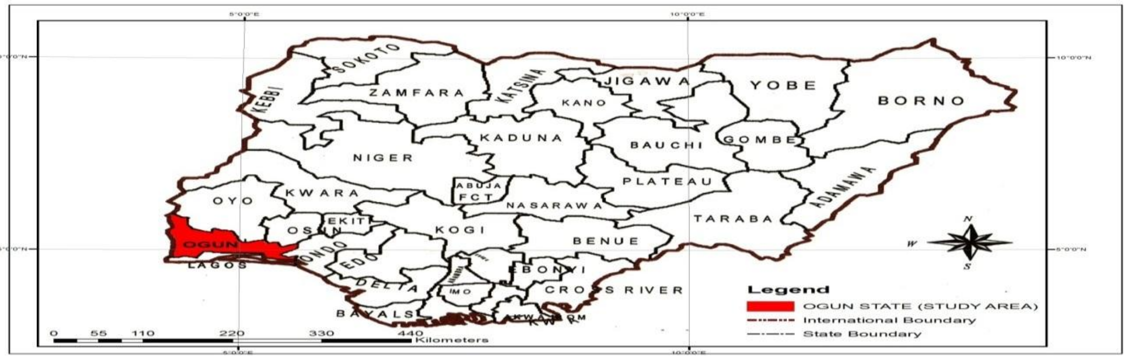

Figure 1: Map of Nigeria showing the study Area.

Test for Trend and Serial Correlation

The Mann-Kendal non-parametric test was considered for trend detection in the annual time series data because of its robustness and unique advantages over other methods. To check for serial correlation, the Durbin-Watson test was considered. The tests were carried out in order to know which modeling techniques will best fit the data to be use.

Time series data are generally represented in the form:

$$T(t) = Tr + P(t) + \varepsilon(t) \qquad 1$$

Where, $T(t) = time\ series, Tr = trend\ component, P(t) = periodic\ component\ and, \varepsilon(t) = stochastic\ component.$

In order to check for the stationarity of the data, the following equations were considered:

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} sgn(X_j - X_k) \; 2$$

Where, X_j and X_k are the annual values in years j and k, j > k, respectively, and

$$sgn(X_j - X_k) = \begin{cases} 1 & if\ X_j - X_K \ > 0 \\ 0 & if\ X_j - X_k \ = 0 \\ -1 & if\ X_j - X_k \ < 0 \end{cases} \; 3$$

$$VAR(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{q} t_p(t_p - 1)(2t_p + 5)\right] \; 4$$

Where, q is the number of tied groups and t_p is the number of data values in the p^{th} group. The values of S and VAR(S) are used to compute the test statistic Z as follows

$$Z = \begin{cases} (S-1)/Var(S)^{1/2} & S > 0 \\ 0 & S = 0 \\ (S+1)/Var(S)^{1/2} & S < 0 \end{cases} \qquad 5$$

The Mann-Kendall test was carried out in accordance with the works of Otache, Ahaneku and Mohammed, (2011); Edwin and Otache, (2014) and Chatfield (2004), with the aid of the excel template of 'MAKESEN's version 1. Lo's modified R/S test was also done to ascertain if the trend persisted. To check for serial correlation, the Durbin-Watson test was considered. The tests were carried out in order to make sure the time series data conforms to the basic criteria for stochastic modeling. No trend and cyclic components were observed from the result obtained, therefore, only the stochastic component was considered.

Autoregressive Moving Average (ARMA) Models for Relative Humidity

The Mann-Kendal test result gives an insignificant Z-value of -1.37, and by this, the null hypothesis of no trend in the time series of relative humidity is accepted. The Box Jenkins (2008) methodology was applied in the model identification, parameter estimation and diagnosis test before the prediction. Based on the ACF and PACF of the training data, in conjunction with the serial iteration, an autoregressive moving average, ARMA – model of order (2, 2) was found suitable for fitting the time series. The results of the serial iteration for model parameters identification was also confirmed by the performance of the Akaike Information Criterion and Bayessian Information Criterion (AIC/BIC) test as presented in Table 1. As shown in the table, the model parameters occupying row with least value of AIC/BIC is considered the best for model building. With the aid of MINITAB software version 16.0, the model equation was built and presented in Table 2. The equation was used to generate values for validating the model and presented graphically as shown in Figure 2. The Lewi's error scaling system (i.e. considering the MAPE), was used in comparing the accuracy of the model, and was found to be accurate (i.e. MAPE values < 10%). Table 3 presents the summary of error determination for the model developed.

Determination of Features and Development of the MLR Model for Relative Humidity

- Moving Average (MA): It is calculated progressively as an average of N number data values over certain period. The term moving is used because every time a new observation becomes available for the time series, it replaces the oldest observation in the equation and a new average is computed. As a result, the average will change, or move, as new observations become available (Anonymous, 2015). Data set is represented by d_t, d_{t-1}, d_{t-2},………………., d_0, where d_t is present and d_0 is the first data value, the moving average with a sliding window of period N is given by:

$$\mathrm{MA} = (d_t + d_{t-1} + d_{t-2} + \cdots d_{t+N})/N \quad 6$$

- Exponential Smoothening (ESM): It also uses a weighted average of past time series values as a forecast; it is a special case of the weighted moving averages method in which we select only one weight, the weight for the most recent observation (NOHC, 2012). The weights for the other data values are computed automatically and become smaller as the observations move farther into the past. The exponential smoothing equation is given as:

$$ESM = F_{t+1} = \alpha Y_t + (1-\alpha)F_t \quad 7$$

Where;

F_{t+1} = forecast of the time series for period t+1

Y_t = actual value of the time series in period t

F_t = forecast of the time series for period t

α = is called the smoothing constant having value ($0 \leq \alpha \leq 1$).

- Oscillator (OSC): Oscillator is used to indicate the rising or falling trend present in the time series when the values are plotted against time. It is defined as difference of moving averages or exponential smoothening of two different periods.

$$OSC = MAN_1 - MAN_2 \quad 8$$

$$\text{Or, } OSC = ESMN_1 - ESMN_2 \quad 9$$

Where, $N_1$and N_2 are different periods and $N_1 > N_2$.

- Rate of Change (ROC): It indicates the rate of change of the variable at present, as compared to the value of the variable at certain period back. Thus percentage ROC at 'a' times back is given by:

$$\mathrm{ROC} = \left(1 - \frac{d_t}{d_{t-a}}\right)100 \quad 10$$

Where, d_t = the value of the time series at present time t

d_{t-a} = the value of the time series at time t – a back.

Using the features or predictors determined, Multiple Linear Regression (MLR) equation is developed using the first part of the features; and the remaining was used for testing the validity. Microsoft Excel 2010 version and Minitab version 16 were used to process the data and present the result in Table 4. Y represent relative humidity while, X_1, X_2, X_3 and X_4 are the moving average, exponential smoothening, oscillator and rate of change respectively. The plot of actual and predicted value of relative humidity is shown in figure 3.

Conclusion

The accessibility to records of hydrological processes is imperative for proper guide and timely preparation against extreme events. Several methods have been used to predict hydrological behaviors, but have shown some weaknesses due to stochastic nature of hydro-meteorological events. The results in this research indicated that there may be decrease in the amount of air moisture of the study area in the nearest future as clearly shown by the Sen's slope. The best predictive model was found to be the ARMA, though MLR give better validation, and has been proved to be the best for local farmers consumption because of its simplicity. It is however recommended that other climatic parameters be looked into for proper planning.

References

Box, G. E. P., Jenkins, G. M., and Reinsel, G. (2008). Time Series Analysis: Forecasting and Control. Wiley, Oxford, 4th Edition.

Chatfield, C., 1994, The Analysis of Time Series- An Introduction, Fourth Edition, London, Chapman and Hall.

Chatfield, C. (2004). The analysis of time series: An Introduction, sixth edition: New York, Chapman & Hall/CRC.

Edwin, A. I. and Otache. M. Y.,(2014). Stochastic Characteristics and Modeling of Monthly

Rainfall Time Series of Ilorin, Nigeria. Open Journal of Modern Hydrology, 4, 67-79. http://dx.doi.org/10.4236/ojmh.2014.43006.

Ewemoje, T. A. and Ewemooje, O. S., (2011). Best Distribution and Fitting Positions of Daily Maximum Flood Estimation at Ona River in Ogun-Osun River Basin, Nigeria. Agricultural Engineering International: CIGR Journal. Vol. 13, No. 3. Manuscript No. 1380.

Held, I. M., and Soden, B. J., (2000). Water Vapor Feedback and Global Warming, Annu. Rev. Energy Environ., 25, 441– 475.

Mathur, S., Shukla, A. K. and Pant, R. P., (2001). International Conference on Optimization Techniques and its Applications in Engineering and Technology. A Comparative Study of Neural Network and Regression Models for Estimating Stock Prices.

Nevada Occupational Health Clinic 'NOHC', (2012). Time Series Analysis and Forecasting: Statistics in Practice. 15(2). Sparks, Nevada.

Otache, M. Y., Ahaneku, I. E. and Mohammed, S. A. (2011). ARMA Modelling of Benue River Flow Dynamics: Comparative Study of PAR Model. Open Journal of Modern Hydrology, Scientific Research, 1, 1-9.

Soden, B. J. and Held, I. M., (2006). Robust Responses of the Hydrological Cycle to Global Warming, J. Clim., 19(21), 5686–5699.

Rakesh, K., and Arun, K. A., (2014). Simplified Mathematical Modeling of Dehumidifier and Regenerator of Liquid Desiccant System. International Journal of Current Engineering and Technology.Vol.4, No.2.E-ISSN 2277 – 4106, P-ISSN 2347 – 5161. Access online on 01, April 2014.

Roadknight, C. M., Balls, G. R., Mills, G.E. and Palmer-Brown, D., (1997). Modeling Complex Environmental Data. IEEE Trans. Neural Networks, vol. 8, No. 4, pp852-861.

Tsoho, U. H., (2008). Growth and History of the Establishment of Makera Assada in Sokoto Metropolis to the year 2007. B. A. Project,History Department, Usmanu Dan-fodiyo University Sokoto.

Table. 1: ARMA - Model order selection for relative humidity

S/NO	Model Order (p,q)	Sum of Sqrs (SS)	AIC/BIC Value	Constant (c)	Mean(µ)
1	1 1	2455.24	230.373	66.201	78.033
2	1 2	2448.26	232.291	35.069	78.012
3	2 1	2522.62	233.159	77.193	78.027
4	2 2	2454.58	136.71	46.726	78.069
4	1 3	1916.03	227.182	15.857	77.542
5	3 1	2345.94	233.052	38.409	77.964
6	2 3	2224.89	233.516	24.986	78.145
7	3 2	1883.31	228.683	61.887	77.875
8	3 3	1933.35	231.443	70.785	78.017
9	1 4	2051.78	231.167	8.589	77.563
10	4 1	2125.42	232.190	56.107	77.773
11	2 4	1690.61	227.552	14.124	79.445
12	4 2	1496.33	224.012	84.376	77.074
13	3 4	1853.25	232.216	11.435	78.130
14	4 3	1487.09	225.832	95.797	77.058
15	4 4	1415.38	226.399	200.55	77.083
16	1 5	1444.03	222.980	31.947	78.243
17	5 1	1571.02	225.425	59.562	79.108
18	2 5	1356.54	223.168	41.395	78.449

Table 2: ARMA Model equation

S/no	Model	Model Order	Mdel Equation
1	ARMA	2, 2	$y_t = 46.726 + 0.3413y_{t-1} + 0.0602y_{t-2} - 0.1023\varepsilon_{t-1} + 0.1416\varepsilon_{t-2} + \varepsilon_t$

Table 3: Summary of Error Values for Relative Humidity Using ARMA Model

Model Type	Model Order	Forecast Error	MSE	RMSE	MAPE (%)
ARMA	2, 2	0.99	84.64	31.64	9.01

Table 4: MLR Model equation

Experiment	Regression equation obtained	R- squared value
Mean annual relative humidity estimation using features of mean annual relative humidity	$Y = 2.37 - 0.422 X_1 + 1.39 X_2 + 0.413 X_3 - 0.793 X_4$	**0.99**

Table 5: Summery of Error Values for Relative Humidity Using MLR Model

Model Type	Model Order	Forecast Error	MSE	RMSE	MAPE (%)
MLR		0.17	2.39	0.71	0.86

Table 6: Observed relative humidity, features and predicted values of relative humidity.

Year	Mean Ann. RH (%)	3 Years MA	6 Years MA	ESM (0.8)	OSC	ROC	Predicted Mean Ann. RH (%)
1	72.02						
2	78.40			72.02		-8.86	
3	87.61			77.12		-11.75	
4	83.78	79.34		85.51		4.38	
5	82.34	83.26		84.12		1.71	
6	80.89	84.57		82.70		1.76	
7	80.92	82.34	80.84	81.25	-1.50	-0.03	80.60
8	82.43	81.38	82.32	80.98	0.94	-1.86	82.06
9	83.55	81.41	82.99	82.14	1.58	-1.36	83.25
10	80.50	82.30	82.32	83.27	0.02	3.65	80.49
11	77.45	82.16	81.77	81.05	-0.39	3.79	77.36
12	76.63	80.50	80.96	78.17	0.46	1.05	76.22
13	72.08	78.19	80.25	76.94	2.05	5.94	71.59
14	81.25	75.39	78.77	73.05	3.38	-12.72	82.16
15	83.50	76.66	78.58	79.61	1.92	-2.77	82.86

16	85.75	78.94	78.57	82.72	-0.38	-2.69	86.18
17	88.88	83.50	79.44	85.14	-4.06	-3.64	88.41
18	85.14	86.04	81.35	88.13	-4.69	4.20	85.27
19	87.30	86.59	82.77	85.74	-3.82	-2.53	87.05
20	73.84	87.11	85.30	86.99	-1.80	15.42	74.32
21	77.84	82.09	84.07	76.47	1.97	-5.42	78.30
22	79.65	79.66	83.13	77.57	3.46	-2.32	78.38
23	81.45	77.11	82.11	79.23	5.00	-2.26	81.71
24	70.86	79.65	80.87	81.01	1.22	13.01	71.03
25	71.82	77.32	78.49	72.89	1.17	-1.36	72.12
26	72.78	74.71	75.91	72.03	1.20	-1.34	72.02
27	66.66	71.82	75.73	72.63	3.91	8.41	66.32
28	62.78	70.42	73.87	67.86	3.45	5.83	62.32
29	90.27	67.41	71.06	63.79	3.65	-43.80	97.29

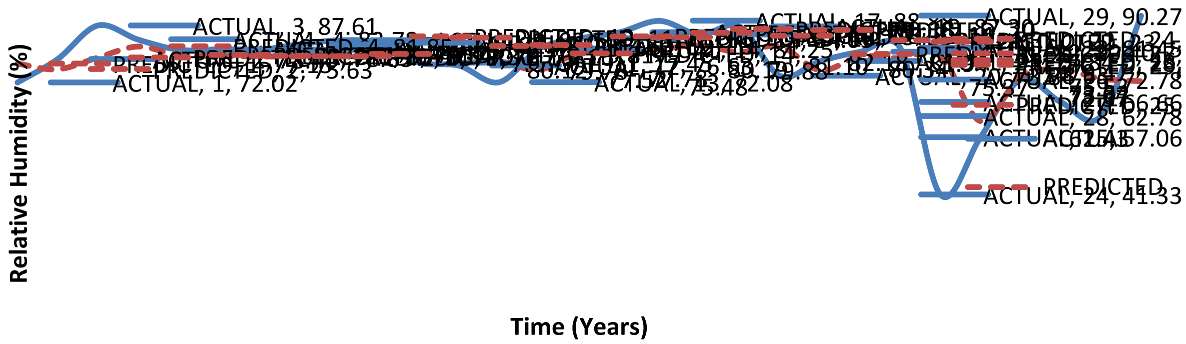

Figure 2: Graph of Observed and Predicted Values for ARMA Model

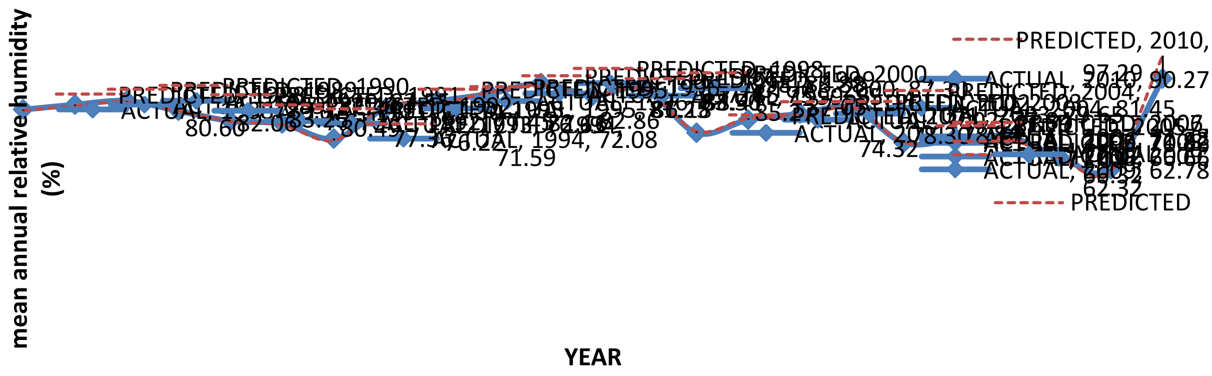

Figure 3: Graph of Observed and Predicted Values for MLR Model

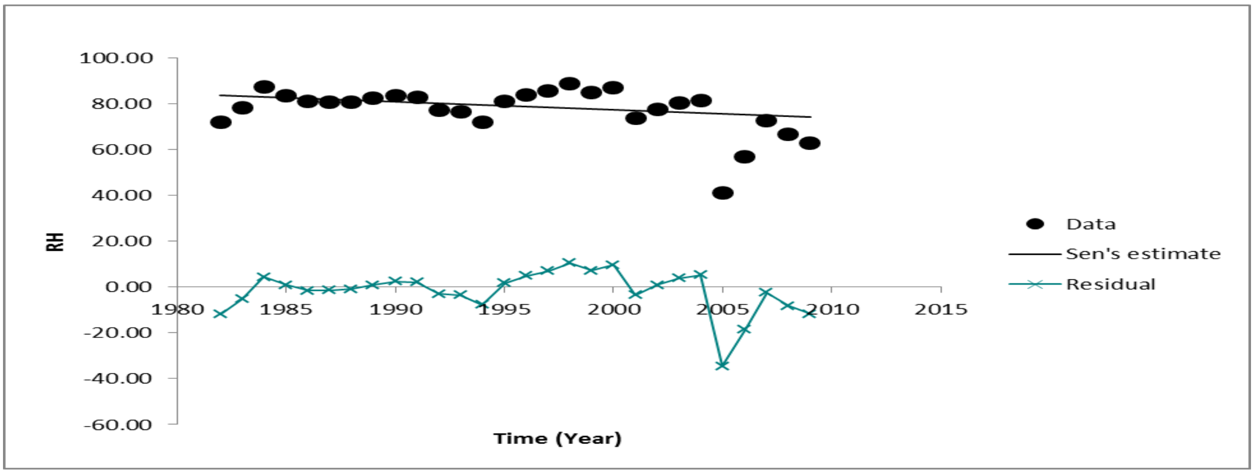

Figure 4: Mann-Kendal Trend Detection Graph

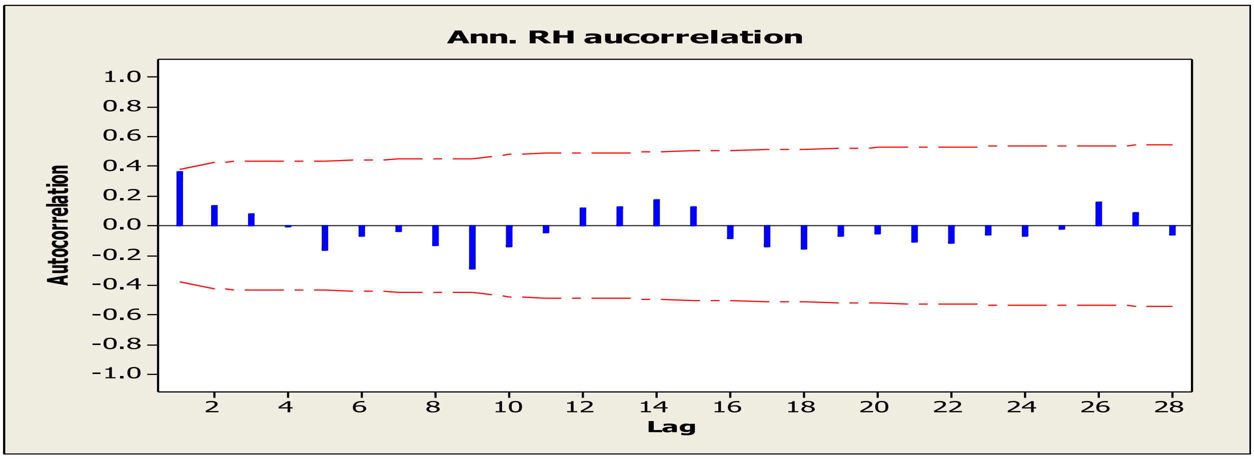

Figure 5: Autocorrelation plot for observed relative humidity

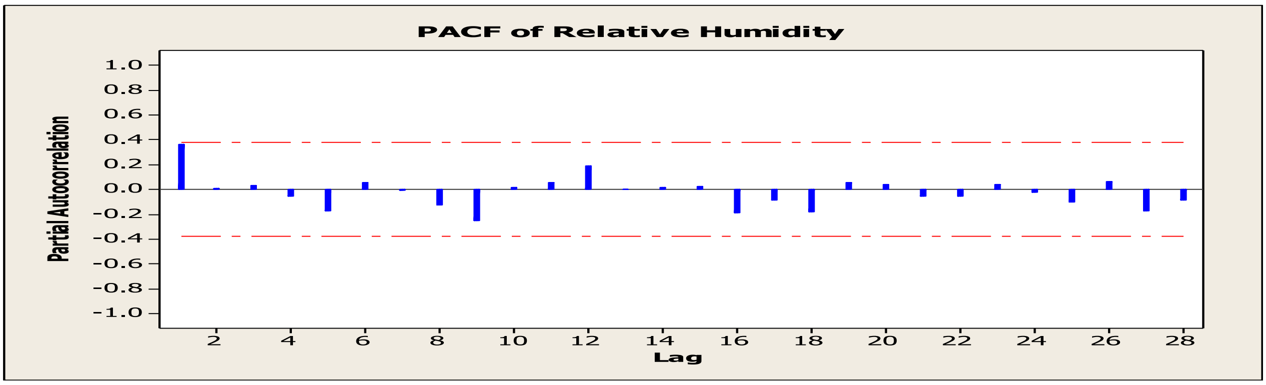

Figure 6: Partial autocorrelation plot for observed relative humidity