FORECASTING RIVER KADUNA DISCHARGE USING HYBRID MOVING AVERAGES AND SMOOTHING EXPONENTIAL METHODS

J. O. Folorunsho[1]

[1]Department of Geography, Ahmadu Bello University, Zaria, Nigeria

Abstract

This paper explores the use of hybrid moving averages and smoothing exponential (HMAs & SE) methods for river flow forecasting. In this study, four major variables for river discharge (rainfall, temperature, relative humidity and stage height), and the discharge values of River Kaduna drainage basin for the period 1993-2004 (Apr-Oct) was used. The variables were used as the input data while the discharge values were used as the output data for the model. The modeling tool was applied to the initial data in order to predict new sets of input values for the period 2005-2015, and subsequently used to generate the output for the same period. The efficiency and accuracy of the HMAs & SE was measured based on the comparison of the initial discharge values (1993-2004) and the forecasted discharge values for the period 2005-2015. From the result using a plot, it was shown that the model provided the best fit and the forecasted discharge trend followed the observed data closely. These result shows that Hybrid Moving Averages and Smoothing Exponential Methods is an effective tool in forecasting river discharge.

Keywords: River Discharge, Hybrid Moving Averages and Smoothing Exponential Methods, River Kaduna

1. Introduction

The discharge forecast model developed for forecasting monthly average of discharge expected in the River Kaduna drainage basin is the Hybrid of Simple Moving Averages and Smoothing Exponential Methods. This is dependent on weather variables namely rainfall, temperature, relative humidity, and stage height and the river discharge data. These variables are basically time series data. According to Abbasov, *et al,*(2003), time series represent a consecutive series of observations of variable(s) taken at regular time intervals over time. One basic requirement for time series is that the data must be displayed in the order in which they occurred, since it is possible that successive observation may probably be dependent (Rojas and Pomares, 2004; Monks, 1996).

The challenge of predicting future values of a time series span a variety of disciplines and as such, time series techniques find applications in such diverse data sets as equity market pricing, disease control, hydrological predictions, meteorological measurements, astronomical observations among others (Rojas *et al,* 2004). The idea of using set of values to certain time, t, to predict the future at another time, t+P, is termed ***Time Series or Stochastic Variables Prediction***.

A time series is a set of a variable at regular intervals over time. A basic requirement for the time series is that the data must be displayed in the order in which they occurred, since it is possible that successive observations may probably be dependent (Rojas, et al., 2004; Monks, 1996). Thus, a time series possess three basic components namely Trend, Seasonal effects, and finally Cyclical factors. The trend is the underlying direction (upward or downward tendency) and rate of change in a time series, when allowance has been made for the other components; seasonal effects describe similar variations occurring during corresponding periods, and cyclical factors describe any regular fluctuations of a time series (Mu'azu, 2006). Time series analysis relies, at least in part, on understanding or exploiting the dependence among the observations. Due to this reliance, the goal of time series prediction can, therefore, be stated thus as: Given a sequence up to time t, the prediction is based on creating a data point, D, scenario "sample every **Δ** units in time **(x(t-(D-1)Δ),..., x(t-Δ), x(t))**, to a predicted future value **x(t+P)"** (Rojas and Pomares, 2004; MATLAB (R2009b). The method to be employed in finding this continuation is a combination of the Simple Moving Average method and Exponential Smoothing, which find applications in stochastic and Erlang distributed data (Kennedy and Davis, 1993).

2. Study Area

River Kaduna drainage basin lies between latitude $9^{0}30'$N and latitude $11^{0}45'$N; longitude $7^{0}03'$E and longitude $8^{0}30'$E with a total basin area of approximately 21,065 km2 (Figure 1). The basin enclosed major rivers such as Kubanni, Galma, Tubo which are tributaries to the main River Kaduna and a greater part of the Kaduna metropolis. The basin lies on the High Plains of Northern Nigeria at altitude of about 670m above sea level and situated within the Northern Guinea Savanna. Typical of the savanna climate, River Kaduna drainage basin experience distinct wet and dry seasons. The wet season (May-Oct) is characterized by conventional rainfall followed by intense lightning and thunderstorms. The annual rainfall can be as high as 2000mm in wet years and as low as

500mm in drought years, but with a long term average of 1000mm (Folorunsho, 2004). The dry season (Nov-Apr) is characterized by a period of low temperature with harmattan season around Dec-Feb; and the hot dry season between Mar-Apr with temperatures as high as 32^{0}C. Relative Humidity is high only during the raining season, but drops during the dry season (Sawa and Buhari, 2011).

3. Materials and Method

Monthly averages of rainfall (Zaria, Kaduna and Jos), temperature, relative humidity for Kaduna metropolis and stage height of River Kaduna (input variables) and discharge data (output) for River Kaduna was used for the forecast. The data covering a period of twelve (12) years (1993 – 2004) and for the rainy months in the basin (Apr-Oct) were sourced from the Hydrology Department, Kaduna State Water Board, Kaduna, and Nigerian Meteorology Agency, Oshodi, Lagos, Nigeria.

Simple Moving Averages (SMA) is the average of the absolute deviations of the series of data above from their mean. Equally, it is a technique that uses a type of average that is adjusted to allow for seasonal, cyclical components of a stochastic variables and help brings either short, mid, long term trends clearer.

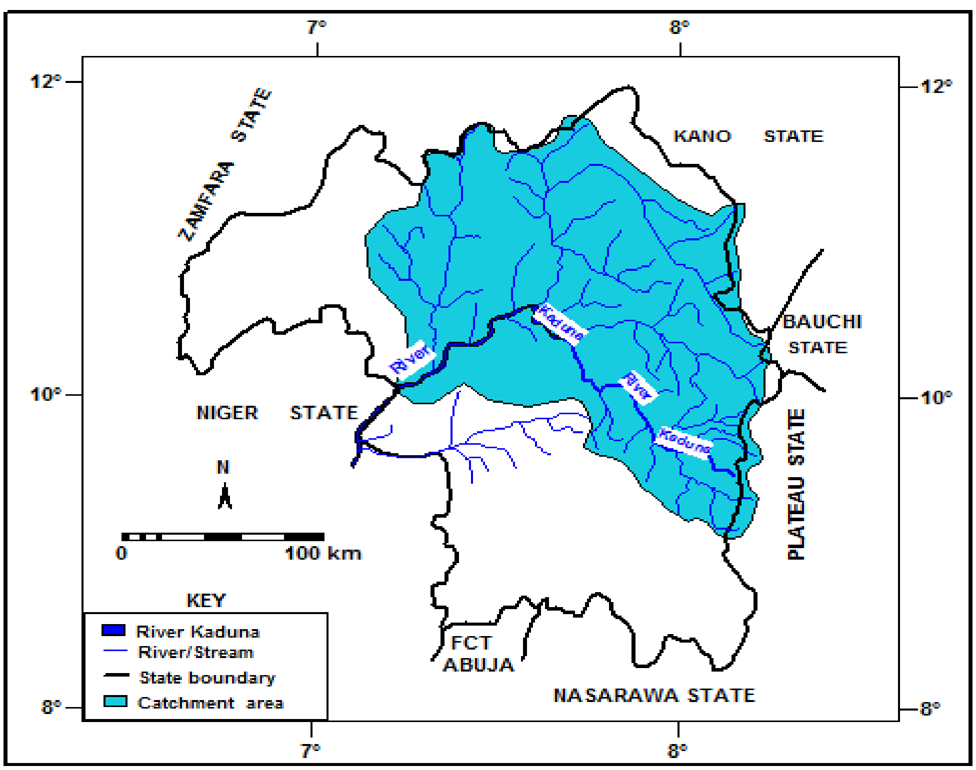

FIG. 3 1 : THE DRAINAGE BASIN OF RIVER KADUNA
Source: Modified from Drainage Map of Kaduna State

In SMA only **n** most recent periods of data points need to be maintained. At the end of each sequence of **n**, the oldest sequence is discarded and the newest data is added to the database. The database is then divided by **n** and used as a forecast for the next period (Monks, 1996; Mu'azu, 2006). The moving average **x(t + 1)** for stochastic data of a period **(t + 1)** is given by equation 1 (Mu'azu (2006):

$$x(t+1) = \frac{[D_1 + D_{t-1} + \cdots + D_{t-n-1}]}{n}, n \leq t \qquad 1$$

where, **n =** number of observations used and **D** = stochastic data

Determination of Number of Observations (n)

The number of observations (n) determined by selecting a value that minimizes the Mean Square Error (MSE) of prediction is given by equation 2 (Mu'azu, 2006):

$$MSE = \sum_{i=1}^{n} \frac{[(D_i - x_i)]^2}{n} \qquad 2$$

where, x_i = mean of the data set

The value x_i was obtained using a moving window of thirty-six (36)-months, fourty-eight (48)-months, sixty (60)-months and seventy-two (72)-months width (**n** = 36, **n** = 48, **n** = 60 and **n = 72**) on the input variables.

Smoothing Exponential method is a forecasting technique that weighs past data in an exponential manner, so that the most recent data carries more weight. The exponentially smoothed moving average is based not on a sequential average of individual stochastic periods, but on the most recent data and the average prior to it adjusted by a smoothing constant, α. The expression for the smoothing constant is in equation 3 (Mu'azu, 2006).

$$\alpha = \frac{1}{\frac{n}{2} - 1}$$

The smoothed series is updated as new observations are recorded using the expression in equation 4:

$$S(t) = \alpha D(t) + (1 - \alpha) S(t-1) \qquad 4$$

where, S(t) is the value of the smoothed stochastic series.

The current smoothed value S(t) is an interpolation between previous smoothed value (S(t-1)) and the current observation, and α controls the closeness of the interpolated value of the recent observation (Mu'azu, 2006).

4. Results and Discussion

Using equations 1 and 2, the simple moving averages operations was performed on the obtained data and the values shown in Tables 1 and Table 2 were obtained.

Subsequently $\boldsymbol{x_i}$ was applied to the input data in order for the choice of the **n** in order to minimize the Mean Square Error (MSE) value to be used. The result is shown in Table 2.

Table 1: Mean ($\boldsymbol{x_i}$) values for River Discharge Input Variables

N	Rainfall (mm)	Temperature (^{O}C)	Rel. Humidity (%)	Stage Height (m)
36	5.5011	25.3914	72.7853	0.2039
48	5.5044	25.3529	73.1463	0.1977
60	5.4325	25.3045	74.0068	0.2013
72	5.5379	25.2576	74.5751	0.2026

Source: Author's analysis, 2011

Table 2: MSE values for number of observations (n's)

Input variables	Mean Square Error		α = 0.03	
	N			
	36	48	60	72
Rainfall	8.35	7.62	7.05	7.53
Temp	3.47	2.97	2.82	2.68
Rel Hum	99.54	87.80	82.96	82.90
Stght	0	0	0	0

Source: Author's analysis, 2011

From Table 2, **n = 72** generate the least Mean Square Error (MSE) value **α**, therefore it was chosen. The selected value of **n was** then used to forecast the input variables values for the drainage basin for the raw data (April- October, 1993 – 2004) Similarly, using equations 3 and 4, **α** value of **0.03** was obtained, and subsequently used to make a complete forecast of input variables values and the discharge values for April to October 2005 – 2015 and the plot of the newly forecasted discharge result is shown in Figure 2.

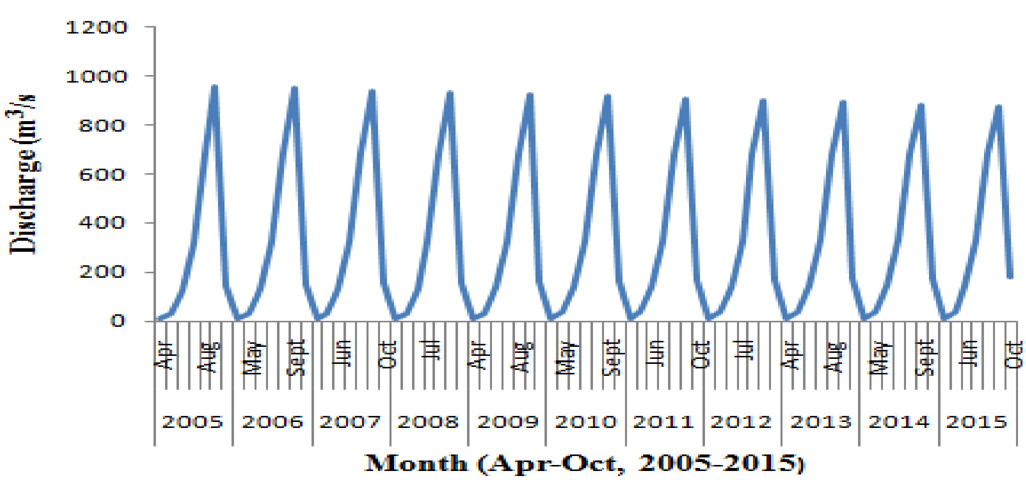

Figure 2: Showing the Plot of River Kaduna Discharge forecasted for

Eight Years (Apr – Oct., 2005 - 2015)

Considering the plot in Figure 2 above (a compressed figure), the monthly forecasted discharge trend between April and October 2005 – 2015 exhibit a similar trend, cycles and seasonality expected in the drainage basin, a definite peak usually observed around August in the study area. These discharge trend correspond with the rainfall seasonality experienced in the drainage basin with the wet season usually covering the months of April – October each year. This condition confirm the research findings by Lafaje (2008); Oloniyo (1995); Musa (1997); and recently by Folorunsho (2004) on River Yobe, River Niger, River Werram and River Kaduna respectively Also, it was discovered from Figure 2 that the discharge values of River Kaduna without doubt has strong correlation to the rainfall trend, cycles and seasonality. . The condition observed in this study and past research findings further confirm the strong seasonality of rainfall and other variables in this drainage basin.

5. Conclusion

The sustainable development of any region is of paramount importance to the planning, development and management of its water resource. Thus, developing model to effectively forecast the river discharge is quite apt. In this research work, the Hybrid Moving Averages and Smoothing Exponential Methods was employed in the forecast of all the input variables and the discharge data for River Kaduna between April – October for the years 2005- 2015 and the results obtained were very promising.

As such, with the trend, cycles and seasonality of the forecasted discharge variable obtained for April – October, 2005 – 2015 is assuring that with the collective and collaborative efforts of the policy makers, researchers and private partners, the water resources development of the drainage basin under the study is near over. From the foregoing, it can be concluded that the Hybrid Moving Averages and Smoothing Exponential Methods is a veritable tool in overcoming the discharge record paucity, inconsistency, unavailability, and unreliability hampering water resources planning and development in the study area.

References

Abbasov, A.M. and Mamedova, M.H. (2003) Application of Fuzzy Time Seriess to Population Forecasting, Vienna University of Technology. www.corp.at/corp/archieve/papers/2003/CORP2003 Abbasov.pdf

Folorunsho, J.O. (2004) An Examination of Some Streamflow Characteristics of River Kaduna State, Unpublished M.Sc. Thesis, Geography Dept., A.B.U., Zaria.

Kennedy G. and Davis, B. (1993) Electronic Communication Systems, New Delhi: Tata, McGraw-Hill Publishing Company Ltd, pp.593-595.

Monks, J.G. (1996) Schaum's Outline of Theory and Problems of Operations Management**,** 2nd ed, McGraw-Hill Inc, pp.39-43.

MATLAB® V7.9.0.529 (R2009b) Neuro-Fuzzy Computing based on Fuzzy Logic, Toolbox, MATLAB Works.

Mu'azu, M.B. (2006) Forecasting and Modeling Statistical Phenomena Using Neuro-fuzzy Logic: A Case Study of Rainfall Forecasting for Zaria. Unpublished Ph.D. Dissertation, Department of Electrical Engineering, A B U, Zaria, pp.95-136.

Musa, S.P. (1997) Stochastic Description of Streamflow Characteristics in the Werram River Catchment Area at Hoss Gauging Station, Plateau State, Unpublished B.Sc.Dissertation, Geography Dept., A.B.U., Zaria.

Oloniyo, G.A. (1995) Some Statistical Characteristics of Some Streamflow in the Niger Basin Area at Lokoja, Unpblished B.Sc. dissertation, Geography Dept., A.B.U., Zaria.

Rojas, I. and Pomares, H. (2004) Soft Computing Techniques for Time Series Forecasting, European Symposium on Artificial Neural Networks, Bruges, Belgium, ESANN'2004 Proceedings, pp.93-102.

Sawa, B.A. and Buhari, B. (2011) Temperature Variability and Outbreak of Meningitis and Measles in Zaria, Northern Nigeria, *Research Journal of Applied Sciences, Engineering and Technology*, 3(5), pp. 399-402.