Time Series Modeling of Water Quality Parameters

Taiwo Abideen Lasisi[1] Olukunmi Olatunji Akanni[2]
1.Department of Mathematics and Statistics, The Polytechnic Ibadan, Nigeria
2.Department of Public Health, Lead City University, Ibadan, Nigeria

Abstract
The inadequate access to safe water can cause people’s health to suffer and resulting to mortality. The study and forecasting of water quality is necessary to improve water quality and reduce mortality rate caused by unsafe water. The study uses time series modeling to monitor the quality of water parameters in Eleyele Dam. ARIMA model was used to generate and forecast the quality of water. Results show that time series modeling is quite capable of water quality forecasting which will help government and decision maker on quality water management system.
Keywords:ARIMA, Mortality, Forecasting, Time Series, Parameters
DOI: 10.7176/JEES/10-9-07
Publication date: September 30th 2020

1.0 Introduction

The world environmental issues has been largely attributed to water quality due to deficiency safe and quality water. Safe water is fundamental to better health, alleviating poverty and community development. Water pollution poses a great threat to sustainable economic development. Water quality could be affected by salinity, overdraw of ground water, urban and domestic wastewater entrance into surface streams as well as agricultural drainage (AbdollahTaheri *et al.*, 2014). Access to safe drinking water and sanitation is a global concern as developing countries, like Nigeria, have suffered from a lack of access to safe drinking water from improved sources and to adequate sanitation services (WHO, 2006). The water quality is associated by its physical, chemical and biological properties; the classification, modeling and interpretations of monitoring data are the most important step in the assessment of water quality (Bayacioglu, 2006). The lack of access to safe water and adequate sanitation is the main symptoms and causes of world poverty, reinforcing the cycle of poverty and incapacity that keeps people trapped and slows the development of societies. The inadequate access to safe water can cause people’s health to suffer, especially children, ranging from reduced growth and life expectancy to critical bouts of diseases, often leading to death (Mathew, 2005). 88% of diarrhoeal disease is attributed to unsafe water supply, inadequate sanitation and poor hygiene, resulting in the deaths of more than two million people every year (WHO, 2006).

Time series analysis is a very useful methods in water quality modeling and forecasting. Many works have been accomplished on hydrological components modeling using time series analysis. The application of this method for water quality forecasting is possible as well. Also evaluation of existing water resources, determining the quality of discharge as well as its quantity, identifying its variation on a watershed scale and forecasting these variables, could be a main step in integrated water resources management (AbdollahTaheri *et al.*, 2014). ARIMA models have been used to examine runoff and river discharge (Papamichail and Georgiou 2001; Yurekli *et al.,* 2005), water levels in lakes (Sheng & Chen, 2011), sediment yield and erosion (Hanh*,* 2010), and water quality (Faruk 2010; Hanh*,* 2010; Voudouris *et al.,* 2010). ARIMA models provides the main statistical characteristics of a hydrologic or environmental time series. The models also provide information about system dynamics and could be used to forecast a time series for the future. A review of Time Series statistical methodology can be found in Khashei and Bijari (2010), Zhang (2003), Panda *et al.,* (2011), Kurunc *et al.,* (2005), Irvine *et al.,* (2011), Halliday *et al.,* (2012), Webb *et al.,* (2003), Antonopoulos *et al.,* (2001), Gonçalves and Alpuim (2011), Gonçalves and Costa (2011), Arya and Zhang (2015) and Hamid *et al.,* (2016).

2.0 Methodology and Study Area

2.1 Study Area and Sampling Methods

The study was conducted on Eleyele dam which is located in north-eastern part of Ibadan, south-western Nigeria within longitude $08^{0}026^{'}00^{''}$N and $08^{'}028^{0}00^{''}$N and Latitude $04^{0}052^{'}00^{''}$E and $004^{0}055^{'}00^{''}$E. The study site is surrounded by Eleyele community toward the south, Apete community toward the east and Awotan community toward the north. The water from the dam serves as a very important source of domestic water supply, transportation, farming, fishing and recreation to the community and people of Oyo State. Water samples were collected daily from the dam using stratified sampling techniques from four different sampling points on the dam for a period of three months. The coordinates of these points are Point A- Latitude $07^{0}26^{'}22.7^{''}$N and Longitude $003^{0}52^{'}26.8^{''}$E; Point B- Latitude $07^{0}26^{'}22.9^{''}$N and Longitude $003^{'}52^{'}29.7^{''}$E; Point C- Latitude $07^{0}26^{'}32.4^{''}$N and Longitude $003^{0}52^{'}32.1^{''}$E; Point D- Latitude $07^{0}26^{'}31.4^{''}$N and Longitude $003^{0}52^{'}38.6^{''}$E. The physical and chemical of the water were measured which include Colour, PH, Total Dissolved Oxygen, Total Hardness, Turbidity, Chloride, Total Alkalinity, Silica, Chlorine

2.2 Methodology

An autoregressive (AR) model has been applied extensively in hydrology for modeling annual and periodic hydrologic time series. Autoregressive (AR) models basically estimate values for the dependent variable y_t as regression function of previous values $y_{t-1}, y_{t-2}, y_{t-3} \ldots y_{t-n}$. The smallest AR model is AR of order 1 i.e. AR (1) is given as:

$$y_t = \varphi_1 y_{t-1} + \varepsilon_t \quad (1)$$

The error term (ε_t) are assumed normally and independently distributed with zero mean and constant variance. Model stationarity requires that the variance of y_t be non-negative and finite (Vandaele, 1983) and for these conditions to be met, $|\varphi_1|$ must be less than 1. Higher order AR can also be obtained just like multiple regressions functions.

Moving average (MA) models incorporate past random fluctuations to represent the time series. The smallest MA model is order 1 which is represented by

$$y_t = \alpha_t - \varphi_1 \alpha_{t-1} \quad (2)$$

The random shocks (α_t) are normally and independently distributed with mean zero and constant variance. The model structure requires the condition of reversibility to be met and therefore $|\ \varphi_1\ |$ must be less than 1. Higher order of MA models is possible, and like the AR model coefficients, the absolute value of each MA coefficient should be less than 1.

A parsimonious model can be achieved using a mixed ARMA model as a combination of a moving average process and an autoregressive process rather than a merely AR or MA model. Therefore, low-order of ARIMA models has been widely used in hydrological practice (Padilla *et al.,* 1996; Montanari *et al.,* 2000). The mixed model structure can provide additional flexibility in describing the result of the interaction between the processes (Salas *et al.,* 1980) as hydrologic time series is the result of several interactive processes that may have both a seasonal and a random fluctuation component. The simplest autoregressive moving average model is ARMA(1,1), and it is given by:

$$y_t - y_{t+1} = \alpha_t - \varphi_1 \alpha_{t-1} \quad (3)$$

3.0 Results and Discussion

The water quality parameters of Eleyele dam were examined. The Colour, Turbidity, PH and Total hardness showed slight negative, Silica showed high negative trend while DO and Alkalinity showed positive trend. Hence, there is need to remove any traces of trend from the series. Chloride and Chlorine showed little or no trend in the series (Figures 1-8). ACF and PACF of the water parameters neither follow AR nor MA (Figures 10-18). A parsimonious model can be achieved using a mixed ARMA model as a combination of a moving average process and an autoregressive process rather than a merely AR or MA model. The mixed model structure can provide additional flexibility in describing the result of the interaction between the processes as hydrologic time series is the result of several interactive processes that may have both a seasonal and a random fluctuation component. The trend of each time series were removed by first order differencing and then the order of model was determined to choose the best model for prediction. The low-order of ARIMA models was adopted which has been proof to be widely used in hydrological practice (Padilla *et al.,* 1996; Montanari *et al.,* 2000). The estimation of the parameters for the physico-chemical properties of the dam is shown in table 1. The forecasts of the parameters were computed for days and the R-squared obtained showed the model is adequate for forecasting (Table 2)

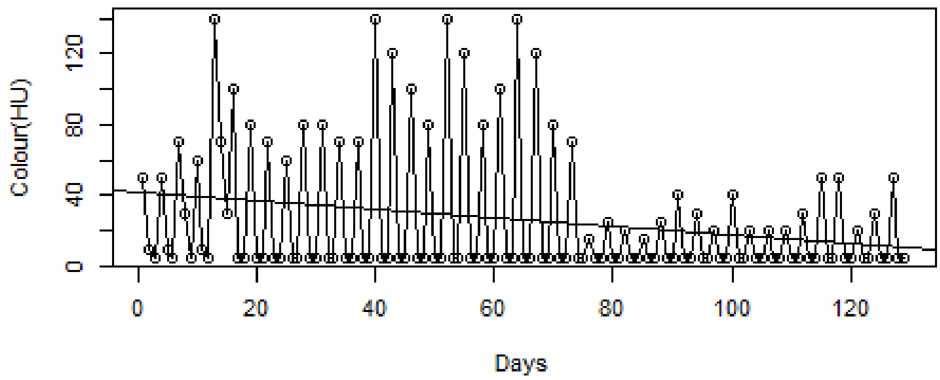

Figure 1: Time plot and Trend of Colour from Eleyele Dam

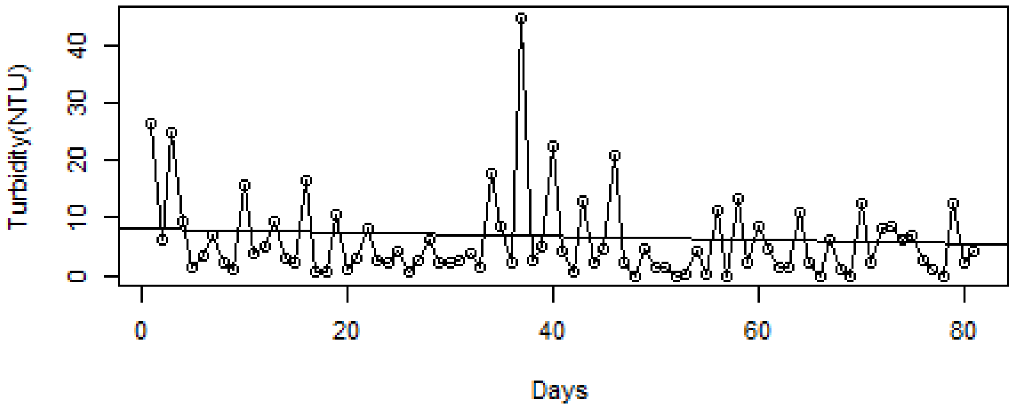

Figure 2: Time plot and Trend of Turbidity from Eleyele Dam

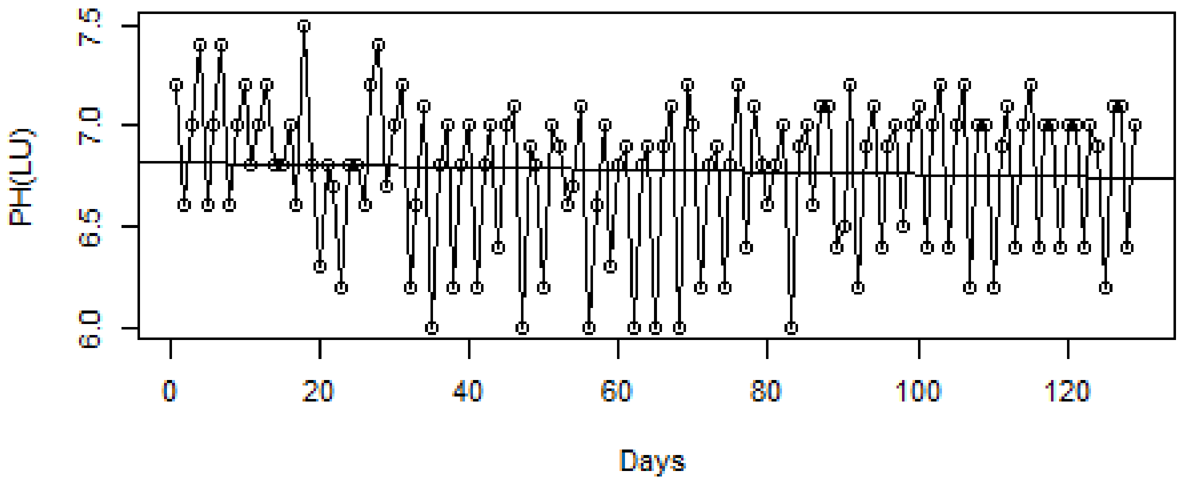

Figure 3: Time plot and Trend of PH from Eleyele Dam

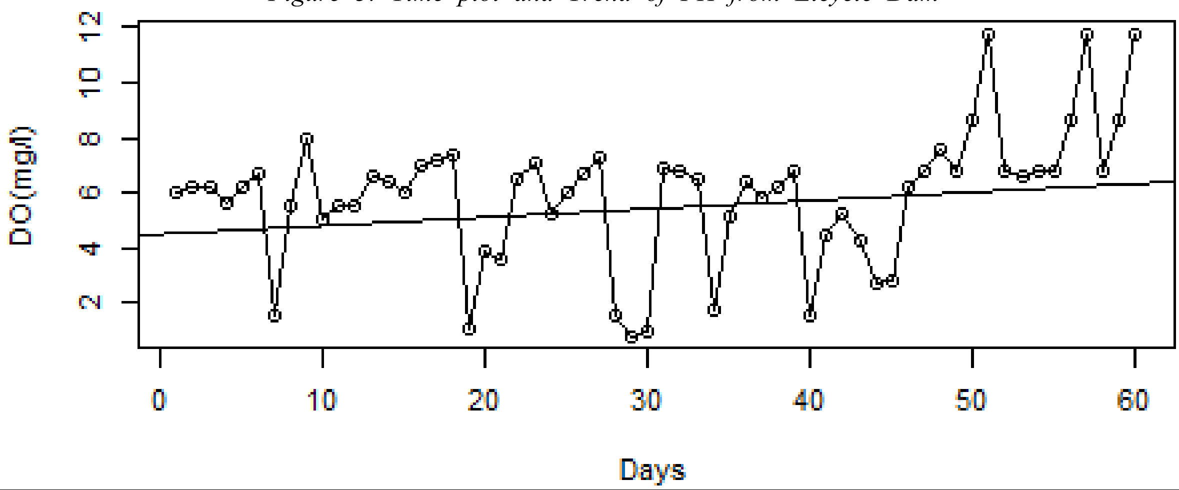

Figure 4: Time plot and Trend of DO from Eleyele Dam

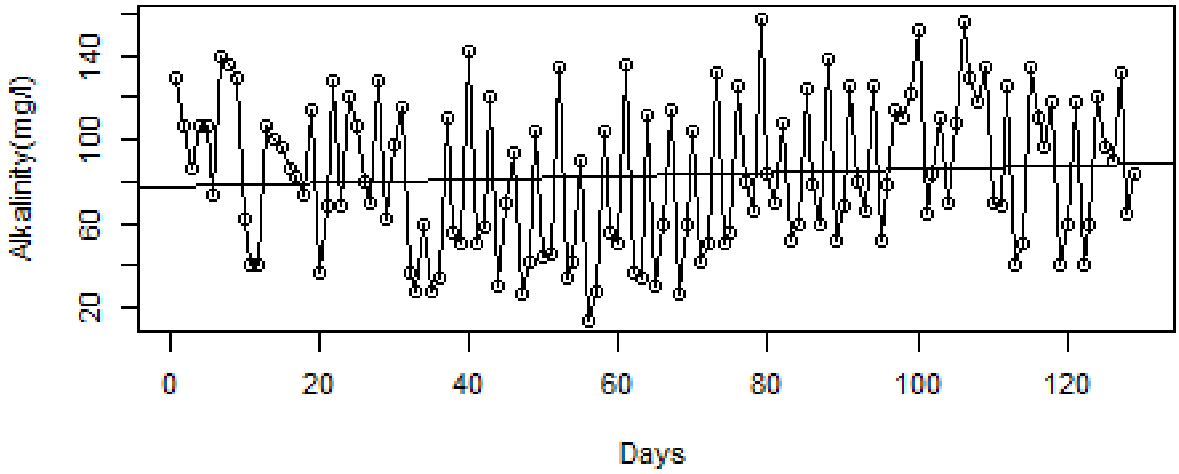

Figure 5: Time plot and Trend of Alkalinity from Eleyele Dam

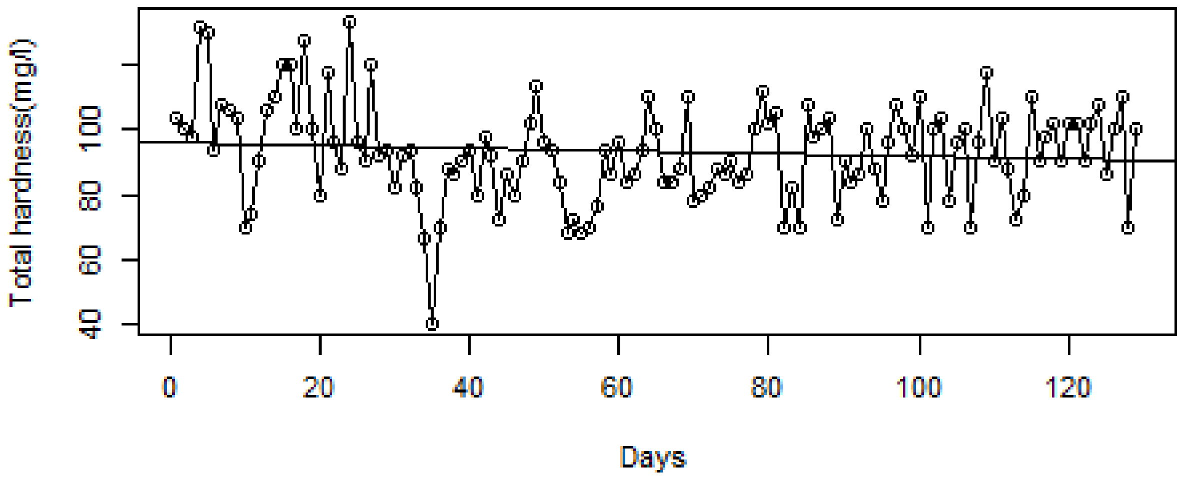

Figure 6: Time plot and Trend of Total Hardness from Eleyele Dam

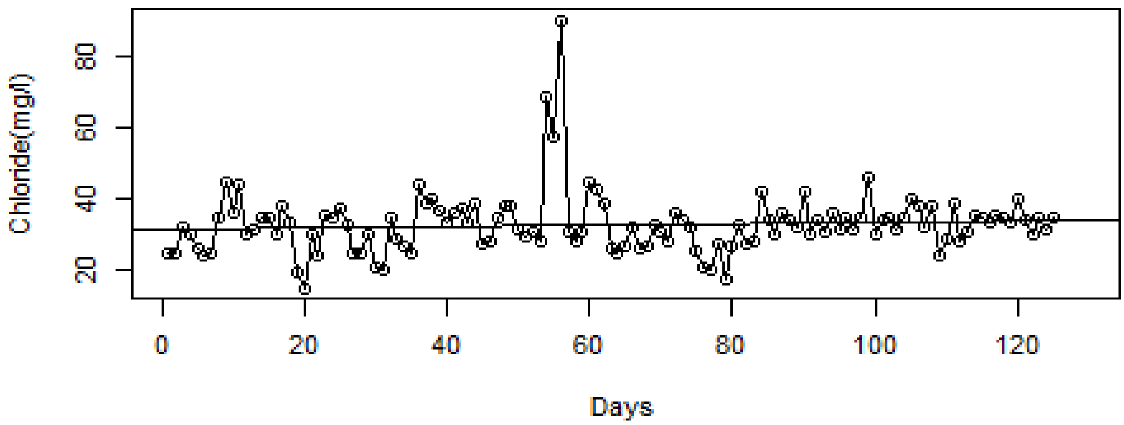

Figure 7: Time plot and Trend of Chloride from Eleyele Dam

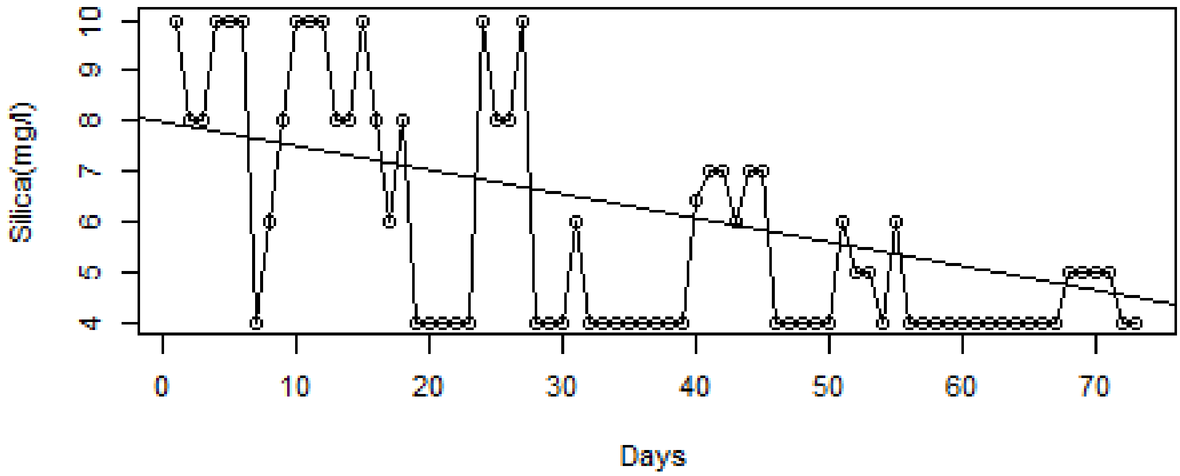

Figure 8: Time plot and Trend of Silica from Eleyele Dam

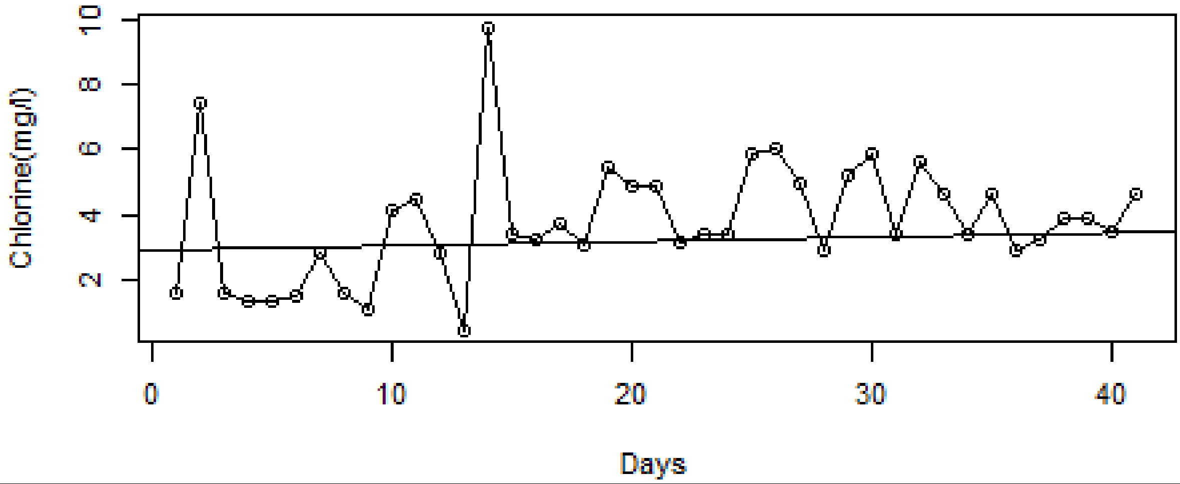

Figure 9: Time plot and Trend of Chlorine from Eleyele Dam

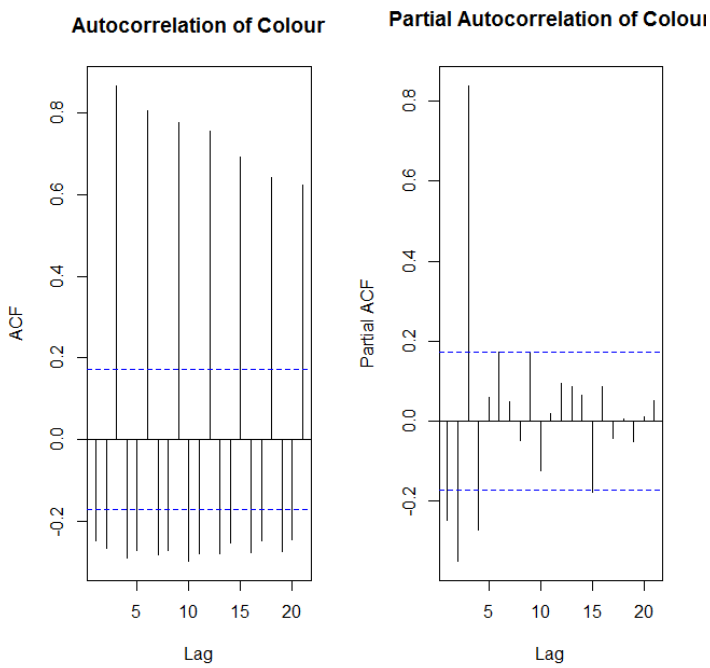

Figure 10: ACF and PACF of Colour from Eleyele Dam

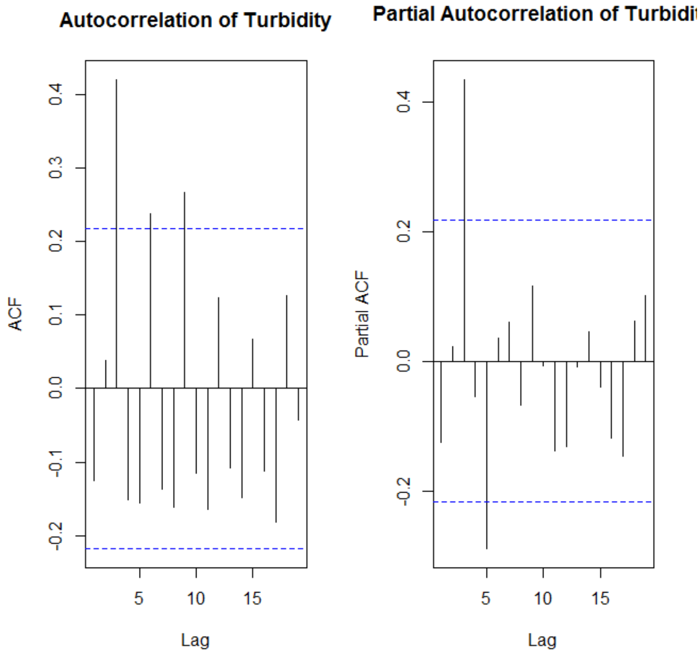

Figure 11: ACF and PACF of Turbidity from Eleyele Dam

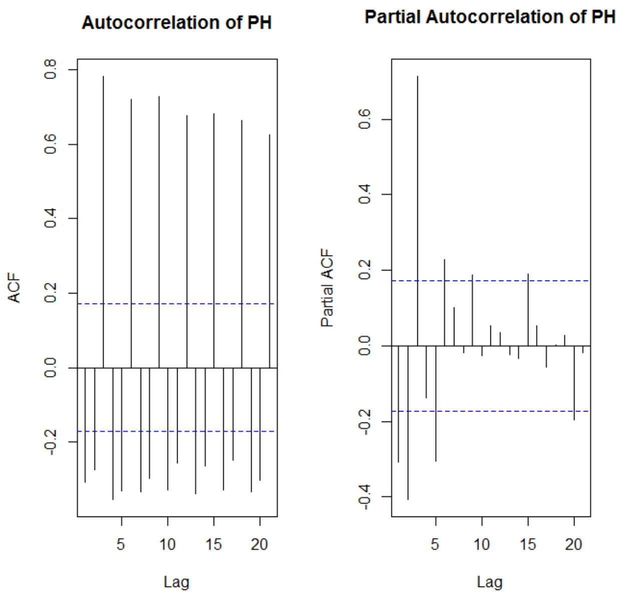

Figure 12: ACF and PACF of PH from Eleyele Dam

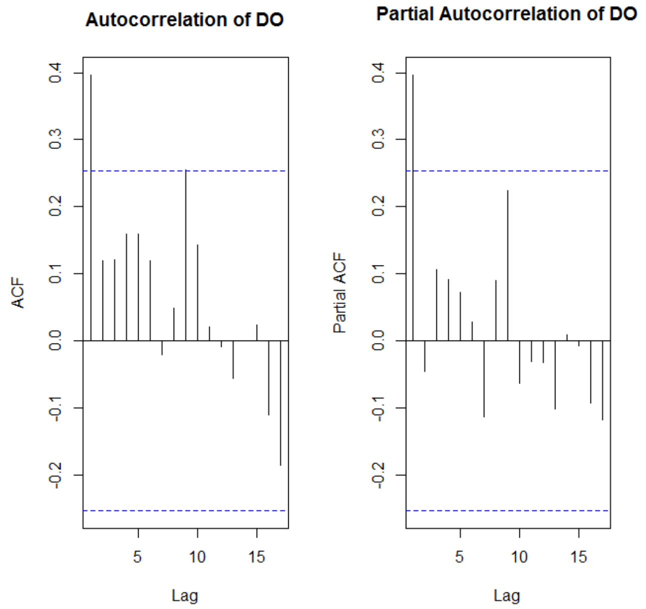

Figure 13: ACF and PACF of DO from Eleyele Dam

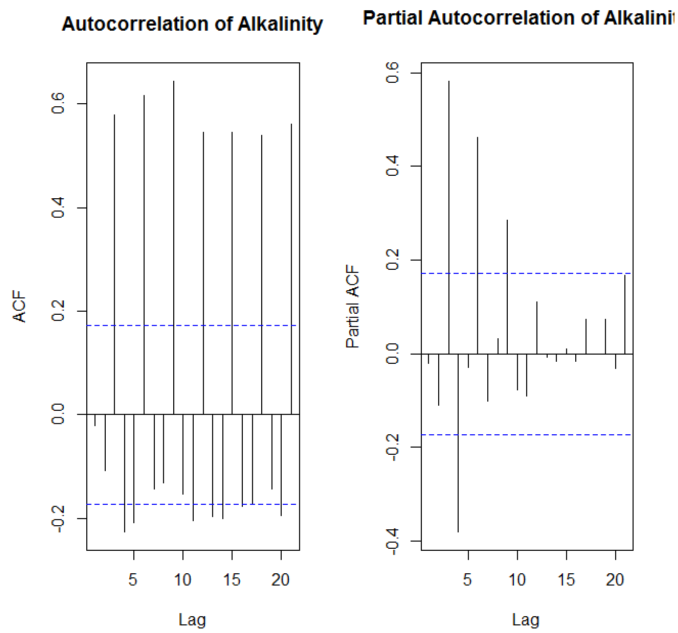

Figure 14: ACF and PACF of Alkalinity from Eleyele Dam

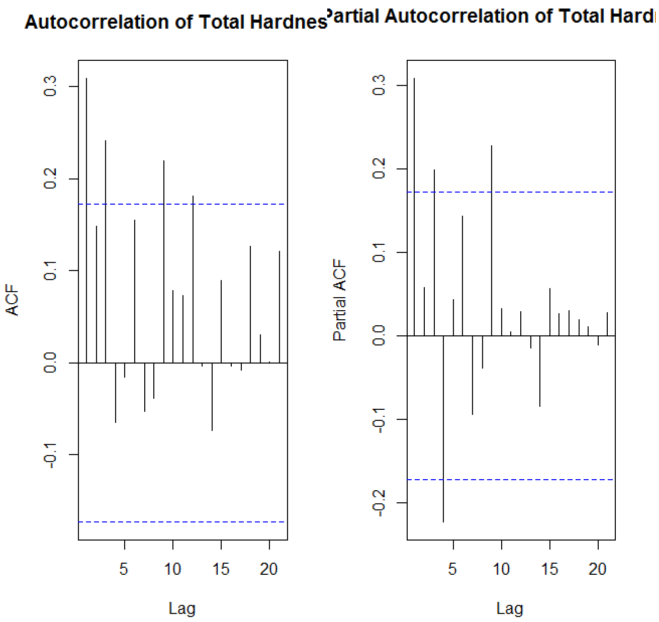

Figure 15: ACF and PACF of Total Hardness from Eleyele Dam

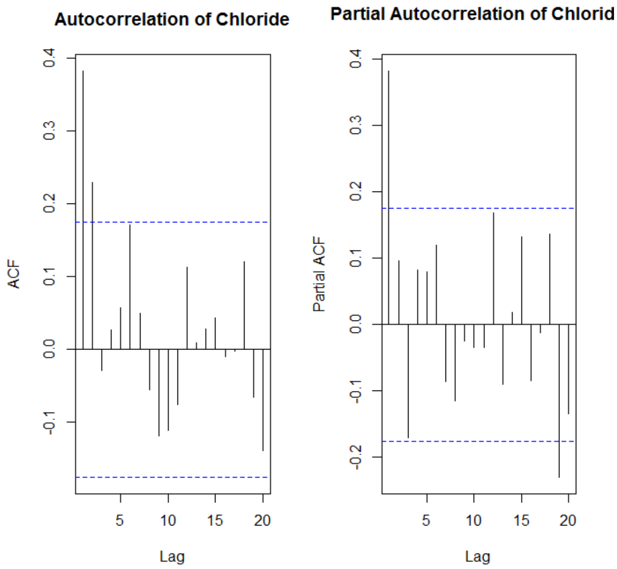

Figure 16: ACF and PACF of Chloride from Eleyele Dam

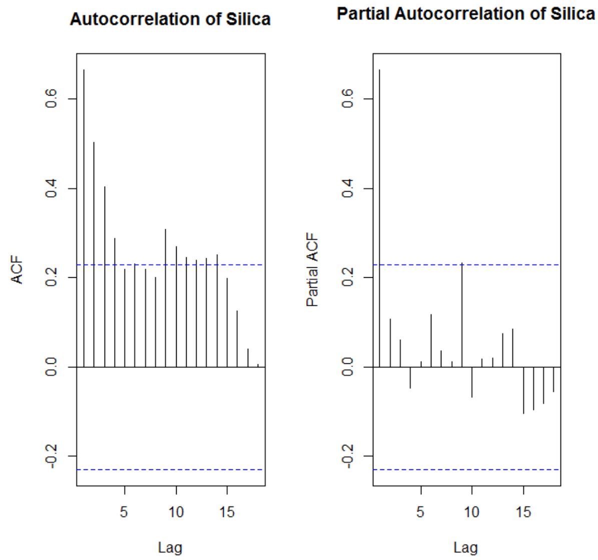

Figure 17: ACF and PACF of Silica from Eleyele Dam

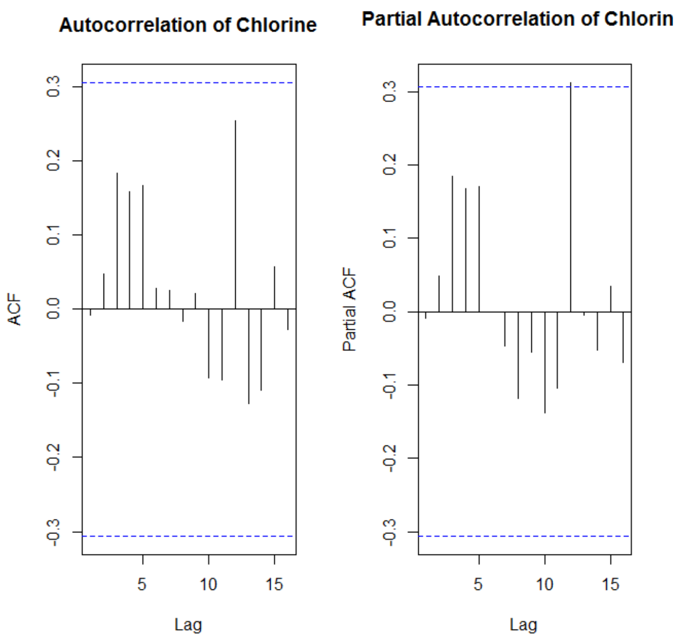

Figure 18: ACF and PACF of Chlorine from Eleyele Dam

PARAMETERS	AR	MA	AIC	LOG LIK
COLOUR	-0.3525	-0.9189	1269.03	-632.51
TURBIDITY	-0.2861	-0.7636	551.78	-273.89
PH	-0.3591	-0.9256	96.05	-46.02
DO	0.2760	-0.8507	263.60	-129.80
ALKALINITY	-0.1430	-0.8978	1279.69	-637.85
TOTAL HARDNESS	0.2249	-0.9238	1052.41	-524.20
CHLORIDE	0.3929	-1.0000	879.40	-437.70
SILICA	0.4779	-0.8797	273.27	-134.64
CHLORINE	-0.1454	-0.8310	166.42	-81.21

Table 1: Parameter Estimation of physico-chemical parameters of Eleyele Dam

PARAMETERS	R SQUARE
COLOUR	0.78
TURBIDITY	0.62
PH	0.57
DO	0.90
ALKALINITY	0.88
TOTAL HARDNESS	0.68
CHLORIDE	0.72
SILICA	0.55
CHLORINE	0.89

Table 2: Model adequacy for 30 days forecast of the parameters

4.0 Conclusions

The increasing trend in the number of patients with reported complaints of water related diseases cannot be overemphasized. In this study, nine water quality parameters of Eleyele dam were studied. Time series trend of the parameters show trend. The series were differenced to eliminate the trend, stationary time series were prepared to work on. The results of modeling show that ARIMA modeling process is suitable in generating and forecasting the water quality parameters.

Acknowledgements

This study was supported by a grant from Tertiary Education Trust Fund (TETFund) through Institution Based Research Committee (IBRC), The Polytechnic Ibadan, Nigeria.

References

AbdollahTaheri T., Maryam G., Pantazis G., Konstantinos V., (2014). Time series analysis of water quality parameters. *Journal of Applied Research in Water &Wastewater*

Antonopoulos, V.Z., Papamichail, D.M., and Mitsiou K.A. Statistical and trend analysis of water quality and quantity data for the Strymon River in Greece. *Hydrology and Earth System Sciences, 5(2001) 679- 691.*

Arya, F.; Zhang, L. (2015). Time series analysis of water quality parameters at Stillaguamish River using order series method. Stoch. Environ. Res. *Risk Assess. 29, 227–239.*

Bayacioglu, H.,2006. Surface water quality assessment using factor analysis, *Water SA, Vol.32, No. 3, July, PP. 389-394.*

Faruk, D.O. (2010). A hybrid neural network and ARIMA model for water quality time series prediction. *Engineering Applications of Artificial Intelligence, 23.586–594.*

Gonçalves, A.M.; Alpuim, T. (2011). Water quality monitoring using cluster analysis and linear models. *Environmetrics 22, 933–945*

Gonçalves, A.M.; Costa, M. (2011). Application of Change-Point Detection to a Structural Component of Water Quality Variables. In AIP Conference Proceedings*; American Institute of Physics: Halkidiki, Greece; pp. 1565–1568.*

Hamid, A.; Bhat, S.; Bhat, S.; Jehangir, J. (2016). Environmetrics techniques in water quality assessment and monitoring: a case study. Environ. Earth Sci. 75, 321.

Halliday, S. J., Wade A. J., Skeffington, R. A., Neal, C., Reynolds, B., Rowland. P., Neal, M. and Norris,D., (2012). An analysis of long-term trends, seasonality and short-term dynamics in water quality data from Plynlimon, *Wales. Science of the Total Environment 434. 186–200.*

Hanh, P.T.M. (2010). Analysis of variation and relation of climate, hydrology and water quality in the lower Mekong River. *Water Sci. Tech 62 (7), 1587–1594.*

Irvine, K. N. Richey, J. E. Holtgrieve, G. W. Sarkkula, J. and Sampson, M.,. (2011): Spatial and temporal variability of turbidity, dissolved oxygen, conductivity, temperature, and fluorescence in the lower Mekong River–Tonle Sap system identified using continuous monitoring, *International Journal of River Basin Management, 9:2, 151-168; http://dx.doi.org/10.1080/15715124.2011.621430*

Khashei, M. and Bijari,M., (2010). An artificial neural network (p,d,q) model for time series forecasting. Expert Systems with Applications 37, 479–489.

Kurunc, A., Yurekli, K. and Cevik,O., (2005). Performance of two stochastic approaches for forecasting water quality and stream flow data from Yesilirmak River, Turkey. *Environmental Modeling & Software, 20. 1195–1200.*

Mathew, B., (2005). iIRC *i*International iWater iand iSanitation iCentre i- iOccasional iPaper iSeries i40.

Montanari, A., Rosso, R., Taqqu, M.S.(2000). A seasonal fractional ARIMA model applied to the Nile River monthly flows at Aswan. *Water Resour. Res. 36, 1249–1259.*

Padilla, A., Pulido-Bosch, A., Calvache, M.L. and Vallejos, A., The ARMA models applied to the flow of karstic springs. *Water Resour. Res. 32 (1996) 917–928.*

Panda, D. K., Kumar, A. and Mohanty, S (2011). Recent trends in sediment load of the tropical (Peninsular) river basins of India. *Global and Planetary Change, 75, 108- 118.*

Papamichail, D.M., and Georgiou, P.E..,(2001). Seasonal ARIMA inflow models for reservoir sizing. *J. Am. Wat. Resourc. Assoc, 37, 877-885.*

Salas, J.D., et al.,(1980) Applied modeling of hydrologic time series. Littleton, CO: *Water Resources Publications.*

Sheng, H. and Y.Q. Chen.,(2011). FARIMA with stable innovations model of Great Salt Lake elevation time series. *Signal Processing, 91, 553–561*.

Vandaele, W. (1983). Applied time series and Box-Jenkins models. *New York: Academic Press, Inc.*

Voudouris, K., Georgiou, P., Stiakakis, E. and Monopolis, D.,(2010). Comparative analysis of stochastic models for simulation of discharge and chloride concentration in Almyroskartsic spring in Greece. *e-Proceedings of the 14th Annual Conference of the International Association of Mathematical Geosciences, IAMG , Budapest, Hungary, 1-15.*

Webb, B.W., Clack, P.D. and Walling, D.E.(2003). Water- Air Temperature Relationships in a Devon River System and the Role of Flow. *Hydrological processes, 17,3069- 3084.*

World Health Organization (2006). In Water, Sanitation and Health World Health Organization.

World Health Organisation (2006) In Water, Sanitation and Health World Health Organisation.

World Health Organization (2004). Guidelines for Drinking-water Quality, World Health Organization, Geneva.

Yurekli, K. and Kurunc, A,(2005). Performance of stochastic approaches in generating low streamflow data for drought analysis. *Journal of Spatial Hydrology, 5, 20–32.*

Zhang, G.P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 50. 159–175.