



## RELIABILITY ASSESSMENT OF THOMAS FIERING'S METHOD OF STREAM FLOW PREDICTION

M. I. Alfa<sup>1,\*</sup>, M. A. Ajibike<sup>2</sup> and D. B. Adie<sup>3</sup>

<sup>1</sup>DEPARTMENT OF CIVIL ENGINEERING, UNIVERSITY OF JOS, JOS, PLATEAU STATE NIGERIA.

<sup>2,3</sup>DEPT. OF WATER RESOURCES & ENVIRONMENTAL ENGR., AHMADU BELLO UNIV. ZARIA, KADUNA STATE, NIGERIA

*E-mail addresses:* <sup>1</sup>[meshilalfa@gmail.com](mailto:meshilalfa@gmail.com), <sup>2</sup>[ajibikefarida2012@gmail.com](mailto:ajibikefarida2012@gmail.com), <sup>3</sup>[donadie2005@yahoo.com](mailto:donadie2005@yahoo.com)

### ABSTRACT

*This study was carried out to assess the reliability of Thomas-Fiering's method of stream flow prediction. The 19 years gauged data of 1955-1973 was extended to 2016 using the model. Model calibration was done by multiple linear regressions of the gauged and synthetic data of 1956-1973. The linear equations developed for January to December were used for adjustment of the three sets of stream flow data generated for 1974-2016. The reliability assessment was done based on the extent to which the unbiased statistics (mean, standard deviation and correlation coefficients) of the 1955-1973 stream flow data were preserved in the synthetic stream flow for 1955-2016. The comparison was done using linear regression and One-Way ANOVA (95% Confidence level) to check for the reliability of the generated data. The coefficients of determination, P-values, F-values and critical F-values were used to estimate the reliability index. Synthetic data was found to be 95.9% reliable.*

*Keywords:* Ofu River, Reliability, Stream flow, Synthetic, Thomas Fiering's model

### 1. INTRODUCTION

Measurement of surface water flow is an important aspect of hydrology related projects such as water quality monitoring, flood studies and forecasting, geomorphology, and aquatic life support amongst others [1]. It is the first and most essential requirement for planning and management of any water resource project [2,3,4]. The availability of a long-term stream flow data covering a minimum of 30 years is required for appropriate hydrological studies [3].

In less developed and developing countries such as Nigeria, the prohibitively high cost of establishing and maintaining a network of stream gauges makes continuous measurement of stream flow nearly impractical [3,5]. As a result of this, several alternative methods of estimating stream flow data for both ungauged and partially gauged catchments such as but not limited to rainfall-runoff modelling and extension of existing stream flow data have been and are being explored [2,3,6]. The respective alternative methods have their unique conditions of application. For instance, estimation of daily runoff peaks from daily rainfall data is only applicable where the rainfall data is available meaning that the unavailability of rainfall data in a particular catchment makes its application

impossible while the extension of existing stream flow data will require the availability of stream flow data for a minimum of 12 years. The researcher or hydrologist is therefore left to choose appropriate alternative method based on the prevailing availability of data.

Ofu River catchment in Kogi State had a gauging station at Oforachi Bridge between 1955 and 1973 which makes about 19 years of monthly stream flow data available. The gauging station was lost due to lack of maintenance leaving engineers and other developers to rely on only 19 years stream flow data which terminated in 1973 for their designs. As a result of these, it becomes imperative to explore viable methods of extending the available stream flow data beyond the 19 years since no rainfall data exist for the site of interest. One of such methods is the application of the Modified Thomas Fiering's autoregressive Markov model to generate synthetic stream flows by extending the existing 19 years stream flow data at Oforachi Bridge hydrometric station since the 12 years minimum data availability criteria is met. The Thomas-Fiering model is an Autoregressive Markov model which has been applied successfully to generate sequentially the monthly, 10-daily or weekly volume of discharge from a serially dependent series. The model

assumes that a monthly or 10-daily variable is dependent only on the just recent one or two variables involving non-stationarity both in mean and standard deviation. Harms and Campbell [7] applied the Thomas-Fierring Model to two representative Pacific Northwest Rivers and concluded that the model provides an authentic representation of stream flow. Stedinger and Taylor [8] used a range of monthly stream flow models including Thomas-Fiering Model to generate synthetic data for the upper Delaware River basin in New York State. They demonstrated that the model was able to reproduce the basic unbiased statistics of the historical data. More so, Lebbe [9] used the lognormal Markov model (a modification of the Thomas-Fierring model) to generate synthetic flow data for five reservoirs in Afghanistan and was able to preserve the statistical properties of the historical data used. Similarly, Maroof *et al.* [10] employed the Thomas-Fierring Model to extend the 12 months discharge data at Ero-Omola falls as a basis to study its hydropower development potential. Celeste *et al.* [11] also utilized Thomas-Fiering stochastic model for synthetic stream flow generation to determine monthly inflow scenarios for the watershed of the reservoir that supplies the city of Matsuyama, Ehime Prefecture. The scenarios were to be used by a stochastic programming model being developed for the optimal operation of the reservoir. The reason for their choice of this model is that it allows for the non-stationarity of seasonal data. They used 20 years of historical data for calibrating the model parameters and generated a new 20-year synthetic series. They concluded that the model can preserve the characteristics of the historical series and effectively incorporate them into the generated data. Furthermore, Ahmad *et al.* [12] used the Thomas-Fiering Model alongside Autoregressive Integrated Moving Average (ARIMA) model and deseasonalised model to forecast water quality of river Ganges in India based on the data collected from 1981 to 1990. Similarly, Kurunc *et al.* [13] evaluated the forecasting performance of two modeling approaches, ARIMA and Thomas-Fiering, for selected water quality constituents and stream flow of the Yeşilirmak River at Durucasu monitoring station using 13-year (1984–1996) monthly time series records to obtain the best model. They used the two approaches to generate 5-year (1997–2001) data for the river. They concluded based on the root mean square error and mean absolute error calculated for the two approaches that Thomas-Fiering model presented more reliable forecasting of water quality constituents and stream flow than ARIMA model. Notwithstanding the wide

application of this model to generate synthetic flow data, the reliability of the synthetic data generated using this model has been a subject of debate among researchers [7-13]. It has been established that the synthetic data generated using the Thomas-Fierring's Model is only reliable to the degree that the unbiased statistics (mean and Standard Deviation and the coefficients of correlation between successive months) of the gauged data is preserved in the synthetic data [8, 9].

The present study was therefore aimed at carrying out reliability assessment of the Thomas-Fierring's Model in the prediction of Ofu River stream flow at Oforachi Bridge hydrometric station. In other words, the study seeks to answer the question: to what extent are the unbiased statistical characteristics (mean, standard deviation and correlation coefficient) of the gauged data preserved in the synthetic data generated using the Thomas-Fierring's Model?

## 2. MATERIALS AND METHODS

### 2.1 Data Collection and Preparation

19 year gauged stream flow data (1955-1973) for Ofu River at Oforachi hydrometric station was obtained from Kogi State Ministry of Water Resources and Rural Development. Identification and treatment of outliers was carried out using Minitab statistical software version 14.12.

### 2.2 Generation of Synthetic Stream flow data

The logarithm transferred flows method described previously by Patra [3] was used to avoid negative flows in the synthetic data. The logarithm to base 10 of the gauged stream flow data (1955 – 1973) was calculated after which the mean and standard deviation for each month were respectively calculated. The correlation coefficient between all the successive months was also calculated. Two sets of normally distributed random numbers with zero mean and unit standard deviation were generated using the analysis tool pack in Microsoft Excel® 2007. The number of variables were set as 12 (12 months) in both cases while the number of random numbers were set as 18 and 43 for the respective sets (18 years for which gauged data exist and 43 years of synthetic data respectively). Consequently, two sets of synthetic stream flow values (logarithm to base 10) were generated respectively for 1956 – 1973 and 1974 – 2016 using Eq. (1).

$$Q_{j+1} = Q_{avj+1} + b_{j,j+1}(Q_j - Q_{avj}) + Z_j S_{j+1} \sqrt{(1 - r_{j+1}^2)} \quad (1)$$

Where,  $j+1$  is the month for which synthetic discharge is generated,  $j$  is the preceding month,  $Q_{j+1}$  is the Discharge during  $(j+1)$ th month,  $Q_j$  is the Discharge during  $j$ th month,  $Q_{avj+1}$  is the mean monthly discharge during  $(j+1)$ th month,  $Q_{avj}$  is the mean monthly discharge during  $j$ th month,  $S_{j+1}$  is the Standard deviation for  $(j+1)$ th month,  $S_j$  is the Standard deviation for  $j$ th month,  $r_{j,j+1}$  is the correlation coefficient between the months  $j$  and  $(j+1)$  and  $Z_j$  is the The random independent variable with zero mean and unit variance while,

$$b_{j,j+1} = r_{j,j+1} \left( \frac{S_{j+1}}{S_j} \right) \quad (2)$$

The synthetic logarithmic discharge values for 1956-1973 were used for the calibration of the model since gauged data exist for these years. Model calibration was done by multiple linear regressions of the gauged ( $y$ ) and synthetic ( $x$ ) data of 1956-1973. 12 linear relationships were obtained for January to December which were of the form:

$$y_i = m_i x_i + c_i \quad (3)$$

Where  $y_i$  is the gauged discharge for month  $i$ ,  $m_i$  is the slope for month  $i$ ,  $x_i$  is the Generated discharge for month  $i$  and  $c_i$  is the intercept on the  $y_i$ 's axis

These linear equations were used to adjust the generated discharge values generated for 1956-1973. Since it is practically impossible to have the same sets of normally distributed random numbers every time synthetic data is required, the process was repeated three times to examine the effect of random numbers. For each set of synthetic data generated for 1956-1973, the gauged data were respectively compared with the adjusted synthetic data using linear regressions in Microsoft Excel® 2007 and Student's t-test using Minitab statistical software version 14.12.0 to assess the goodness of fit. The coefficients of determination ( $r^2$ ) and the  $P$ -value were used as checks for good fit. The developed equations were then used to adequately adjust the synthetic data generated for 1974-2016 in each of the three runs.

### 2.3 Assessment of reliability

As stated earlier, the synthetic data generated using Thomas-Fierring Model is reliable to the degree that the unbiased statistics of the gauged data are preserved in the synthetic data [8, 9]. The mean and standard deviation are principally the determinants of unbiased statistics but since stream flow is a serially dependent phenomenon, the correlation between successive months needs to also be preserved for the generated data to be reliable [3].

The mean, standard deviation and correlation coefficient for the gauged data of 1955-1973 and the synthetic data of 1955-2016 (first, second and third runs) were calculated using the aforementioned statistical software. The reliability assessment was then carried out by comparing these estimates for both datasets using linear regression and One Way Analysis of Variance. The coefficient of determination,  $P$ -value, calculated  $F$  and Critical  $F$  values at 5% level of significance were used to infer whether there is statistically significant difference or not.  $P < 0.05$  was taken to be statistically significant.

In order to estimate the reliability index, the average of the three coefficients of determination ( $R^2$ ) was recorded as the score for  $R^2$ ,  $P$ -value greater than 0.05 was given a score of 1 while, Calculated  $F$  less than Critical  $F$  was given a score of 1. The reliability index was calculated as the ratio of the total score obtained to the total obtainable score.

### 3. RESULTS AND DISCUSSIONS

The results of the reliability assessment of the Thomas-Fierring's Model in the prediction of Ofu River stream flow at Oforachi Bridge hydrometric station are presented and discussed herein. A comparison of the mean, standard deviation and correlation coefficients for the gauged data and the synthetic data generated via the three runs are presented in Fig. 1-3.

A minimal observable difference between the mean of the gauged data and the synthetic data can be observed from Fig. 1 which gives an indication of the possibility of significant reliability. A similar observation can be seen in Figs. 2 and 3 for the standard deviations and coefficients of correlation. Although these descriptive presentations are not sufficient basis for deciding the reliability of this method, it serves as pointers in the right direction.

Furthermore, the Coefficients of Determination,  $P$ -value,  $F$ -value, Critical  $F$ -value and Reliability Index are presented in Table 1. The coefficients of determination for the comparison of mean and standard deviation were very close to unity which is indicative of a very good fit which agrees with the findings of Dashora *et al.* [14] who concluded that Thomas-Fierring's model performed very well in the prediction of high stream flows for the Narmada River Basin given by the high coefficient of determination and the Root Mean Square Error.

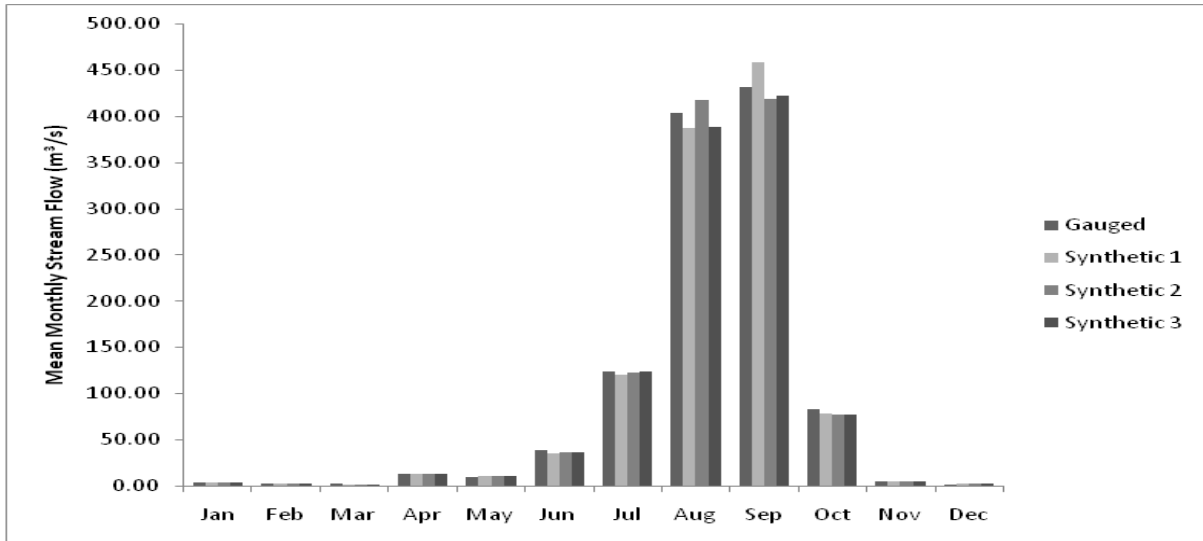


Fig. 1: Gauged and Synthetic Mean Stream Flow Data for Ofu River at Oforachi Bridge Station

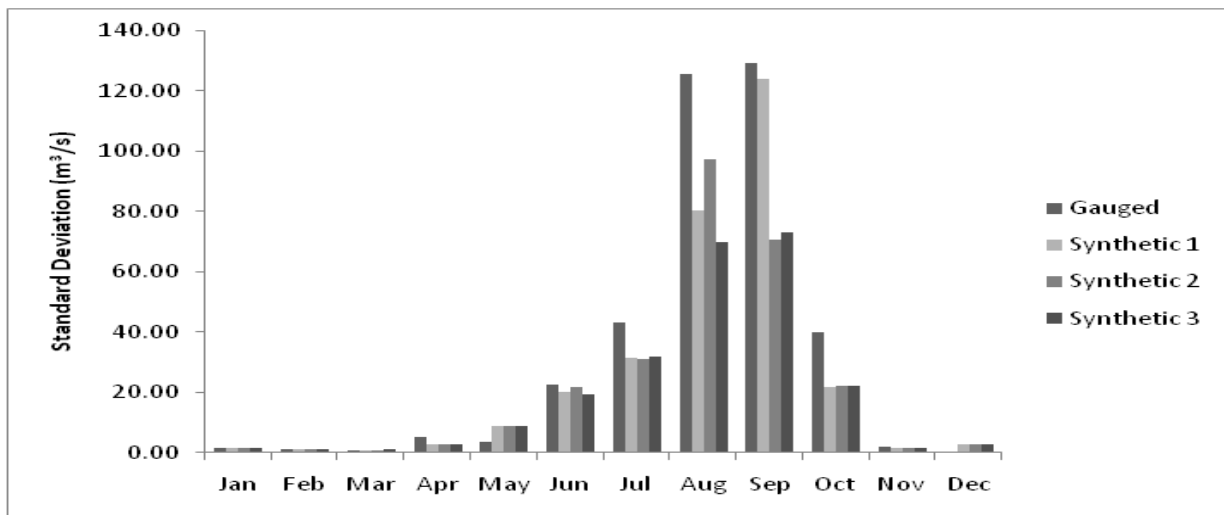


Fig. 2: Standard deviation of Gauged and Synthetic Stream Flow Data for Ofu River at Oforachi Bridge Station

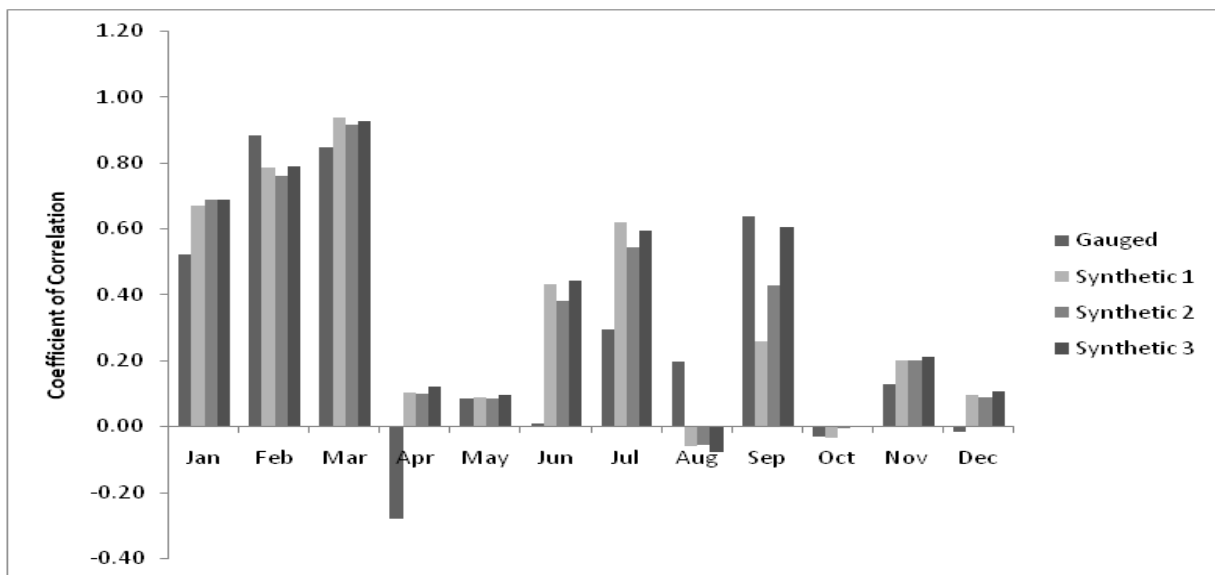


Fig. 3: Coefficients of Correlation of Gauged and Synthetic Stream Flow Data for Ofu River at Oforachi Bridge Station

Table 1: Coefficients of Determination, P-value, F-value, Critical F-value and Reliability Index

Parameter	Mean (m <sup>3</sup> /s)		Standard Deviation (m <sup>3</sup> /s)		Correlation Coefficient	
	Value	Score	Value	Score	Value	Score
R <sup>2</sup>	0.998	0.998	0.956	0.956	0.677	0.677
P-Value	1.000	1	0.879	1	0.905	1
F	0.001	1	0.224	1	0.187	1
F-Critical	2.816		2.816		2.817	
Total Score		2.998		2.956		2.677
Reliability Index	0.9590					

The coefficient of determination for the comparison of correlation coefficient on the other hand is very low (0.677) which shows that only about 67% of the correlation coefficients of the gauged data were preserved in the synthetic data. The p-values obtained for the mean, standard deviation and correlation coefficient were all greater than 0.05 which shows that the difference between these parameters were not statistically significant. This was further buttressed by the calculated F-value which was less than the critical F-value for the three parameters compared. The above results are indicative of the fact that the unbiased statistics of the gauged data especially mean and standard deviation were sufficiently preserved in the synthetic data [8,9]. Related studies by other researchers [14 – 17] agree with the findings in this present study in the aspect of the unbiased statistics of the gauged data being preserved in the synthetic data but most concluded that in comparison with other models such as the Autoregressive Integrated Moving Average (ARIMA) model, the Thomas-Fiering's model has the tendency of over-estimating the stream flow especially for months of low flow. Kurunc *et al.* [13] however was of a contrary opinion that the Thomas-Fiering's Model is more reliable in stream flow prediction than the ARIMA model.

The overall reliability index obtained was 0.959. This implies that the synthetic stream flow data generated using the modified Thomas Fiering's model is 95.9% reliable. While this results gives a significant level of credibility to the Thomas-Fiering's model, a comparison with other models such as the ARIMA models is recommended for further studies.

#### 4. CONCLUSIONS

This study succeeded in generating 62 years synthetic stream flow data (1955-2016) for Ofu River at Oforachi hydrometric station. The study has demonstrated that the generated synthetic data is about 95.9% reliable implying that it could be used for hydrological studies and projects with respect to Ofu River.

#### 5. REFERENCES

- [1] Meals D. W. and Dressing S. A. "Surface water flow measurement for water quality monitoring projects, Tech Notes 3" developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA, 16 p, March 2008. Available online at [www.bae.ncsu.edu/programs/extension/wqg/31\\_9monitoring/tech\\_notes.htm](http://www.bae.ncsu.edu/programs/extension/wqg/31_9monitoring/tech_notes.htm). Accessed February 10, 2016.
- [2] Mustafa S. and Yusuf M. I. *A Textbook of Hydrology and Water Resources, Revised Edition*, Topsmerit Page Publishing Company, Abuja, 2012.
- [3] Patra, K. C. *Hydrology and Water Resources Engineering, Second Edition*, Narosa Publishing House, New Delhi, 2008.
- [4] Daffi R. E. "Flood Hazard Assessment for River Dep Catchment Using Remote Sensing and Geographic Information System Techniques", Unpublished PhD Thesis in The Department of Water Resources and Environmental Engineering Ahmadu Bello University, Zaria, 2013.
- [5] World Meteorological Organization. Manual on Stream Gauging. [wmo.int/pages/prog/hwrp/publications/stream\\_gauging/1044\\_vol\\_II\\_en.pdf](http://wmo.int/pages/prog/hwrp/publications/stream_gauging/1044_vol_II_en.pdf). Accessed August 22, 2010.
- [6] Chow, V. T. "Hydrologic studies of floods in the United States". *International Association of Hydrological Sciences*, Vol. 42, , Pp 134-170,1956.
- [7] Harms, A. A., & Campbell, T. H. "An extension to the Thomas-Fiering Model for the sequential generation of streamflow." *Water Resources Research*, Vol. 3 Number 3, Pp 653-661, 1967.
- [8] Stedinger, J. R., & Taylor, M. R. "Synthetic streamflow generation: 1. Model verification and validation." *Water Resources Research*, Vol. 18, Number 4, 909-918, 1982.
- [9] Lebbe, J. M. E. "Reliability Studies of Multi-Reservoir System Using Stochastic Approach." An M.Sc. Civil Engineering at King Fahd University of

- Petroleum & Minerals, Dhahran, Saudi Arabia, 1991.
- [10] Maroof, L. K., Sule, B. F., & Ogunlela, O. A. "Economic Sustainability of Integrated Hydropower Development of Ero-Omola Falls, Kwara State, Nigeria". In *Decision Making and Knowledge Decision Support Systems*, pp. 143-164. Springer International Publishing, 2015.
- [11] Celeste, A. B., Suzuki, K., Kadota, A., & de Farias, C. A. "Stochastic Generation of Inflow Scenarios to be used by Optimal Reservoir Operation Models". *Annual Journal of Hydraulic Engineering*, Vol. 48, Pp 451-456, 2004.
- [12] Ahmad, S., Khan, I. H., & Parida, B. P. "Performance of stochastic approaches for forecasting river water quality." *Water research*, Vol. 35 Number 18, 4261-4266, 2001.
- [13] Kurunç, A., Yürekli, K., & Çevik, O. "Performance of two stochastic approaches for forecasting water quality and streamflow data from Yeşilirmak River, Turkey." *Environmental Modelling & Software*, Vol. 20 Number 9, 2005, 1195-1200.
- [14] Dashora, I., K Singal, S., & K Srivastav, D. "Software Application for Data Driven Prediction Models for Intermittent Streamflow for Narmada River Basin", *International Journal of Computer Applications*, Vol. 113, Number 10, 9-17, 2015.
- [15] Teymouri, M., & Fathzadeh, A. "Stochastic modeling of monthly river flow forecasting (Case study: Atrak River Basin, Iran)". *Journal of Selçuk University Natural and Applied Science*, Vol. 4, Number 2, 38-48, 2015.
- [16] Cui, Q., Wang, X., Li, C., Cai, Y., and Liang, P. "Improved Thomas-Fiering and wavelet neural network models for cumulative errors reduction in reservoir inflow forecast", *Journal of hydro-environment research*, Vol. 13, 134-143, 2016.
- [17] Sathish, S., & Babu, S. K. "Stochastic time series analysis of hydrology data for water resources". In *IOP Conference Series: Materials Science and Engineering*, Vol. 263, Number. 4, pp. 042140, 2017.