

# Evaluation of Best-Fit Probability Distribution Models for the Prediction of Inflows of Kainji Reservoir, Niger State, Nigeria

Authors: Mamman, Mohammed J, Martins, Otache Y, Ibrahim, Jibril, and Shaba, Mohammed I

Source: Air, Soil and Water Research, 10(1)

Published By: SAGE Publishing

URL: https://doi.org/10.1177/1178622117691034

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at <u>www.bioone.org/terms-of-use</u>.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

## Evaluation of Best-Fit Probability Distribution Models for the Prediction of Inflows of Kainji Reservoir, Niger State, Nigeria

### Mohammed J Mamman<sup>1</sup>, Otache Y Martins<sup>2</sup>, Jibril Ibrahim<sup>1</sup> and Mohammed I Shaba<sup>2</sup>

<sup>1</sup>Department of Agricultural Technology, Niger State College of Agriculture, Mokwa, Nigeria. <sup>2</sup>Department of Agricultural and Bioresources Engineering, Federal University of Technology, Minna, Minna, Nigeria.

Air, Soil and Water Research Volume 10: 1-7 © The Author(s) 2017 Reprints and permissions: sagepub.co.uk/journalsPermissions.nav DOI: 10.1177/1178622117691034



ABSTRACT: The analysis of time series is essential for building mathematical models to generate synthetic hydrologic records, to forecast hydrologic events, to detect intrinsic stochastic characteristics of hydrologic variables, as well as to fill missing and extend records. To this end, various probability distribution models were fitted to river inflows of Kainji Reservoir in New Bussa, Niger State, Nigeria. This is to evaluate the probability function that is best suitable for the prediction of their values and subsequently using the best model to predict for both the expected maximum and minimum monthly inflows at some specific return periods. Three models, ie, Gumbel extreme value type I (EVI), lognormal (LN), and normal (N), were evaluated for the inflows and the best model was selected based on the statistical goodness-of-fit test. The values of goodness-of-fit test for Kainji hydropower dam are as follows: r = 0.96,  $R^2 = 0.99$ , SEE = 0.0087,  $\chi^2 = 0.0054$ , for Gumbel (EVI); r = 0.96,  $R^2 = 0.99$ , SEE = 0.0087,  $\chi^2 = 0.0054$ , for Gumbel (EVI); r = 0.96,  $R^2 = 0.99$ , SEE = 0.0087,  $\chi^2 = 0.0054$ , for Gumbel (EVI); r = 0.96,  $R^2 = 0.99$ , SEE = 0.0087,  $\chi^2 = 0.0054$ , for Gumbel (EVI); r = 0.96,  $R^2 = 0.99$ , SEE = 0.0087,  $\chi^2 = 0.0054$ , for Gumbel (EVI); r = 0.96,  $R^2 = 0.99$ , SEE = 0.0087,  $\chi^2 = 0.0054$ , for Gumbel (EVI); r = 0.96,  $R^2 = 0.99$ , SEE = 0.0087,  $\chi^2 = 0.0054$ , for Gumbel (EVI); r = 0.96,  $R^2 = 0.99$ , SEE = 0.0087,  $\chi^2 = 0.0054$ , for Gumbel (EVI); r = 0.96,  $R^2 = 0.99$ , SEE = 0.0087,  $\chi^2 = 0.0054$ , for Gumbel (EVI); r = 0.0054,  $R^2 = 0.0054$ , for Gumbel (EVI); r = 0.0054,  $R^2 = 0.0054$ ,  $R^2 = 0.005$ 0.79,  $R^2 = 0.85$ , SEE = 0.02,  $\chi^2 = 0.31$  for LN; and r = 0.0.75,  $R^2 = 0.0.68$ , SEE = 0.056,  $\chi^2 = 1376.39$  for N. For the Kainji hydropower dams, the Gumbel (EVI) model gave the best fit. These probability distribution models can be used to predict the near-future reservoir inflow at the Kainji hydropower dams.

KEYWORDS: Probability distribution models, predicted inflow, Kainji dam, goodness-of-fit tests

#### RECEIVED: August 31, 2016. ACCEPTED: January 6, 2017.

PEER REVIEW: Four peer reviewers contributed to the peer review report. Reviewers' reports totaled 705 words, excluding any confidential comments to the academic editor.

**TYPE:** Original Research

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article

#### Introduction

The assessment of the dynamics and regime of a particular hydrologic phenomenon is imperative, especially the timebased characteristics. Time-based characteristics of hydrologic data are of great significance in the planning, designing, and operation of water systems. This significance is informed more largely due to the variability and oscillatory behavior of hydrologic sequences. Therefore, against this backdrop, according to Ahaneku and Otache,<sup>1</sup> as noted by Kottegoda,<sup>2</sup> the lack of complete understanding of the physical processes involved and the consequent uncertainties in the magnitudes and frequencies of future events highlight the importance of time series analysis. Thus, the main objective of any time series analysis is to understand the mechanism that generates the data and also, but not necessarily, to produce likely future sequences over a short period of time.<sup>1</sup> This is usually not without taking cognizance of the appurtenant uncertainty resulting from spatiotemporal variability of hydrologic processes. The analysis of time series is essential for building mathematical models to generate synthetic hydrologic records, to forecast hydrologic events, to detect intrinsic stochastic characteristics of hydrologic variables, as well as to fill missing and extend records (Ahaneku and Otache, 2014).<sup>1</sup>

River is a natural stream of water flowing in a channel to the sea or to the lake or joining another river, and the area drained DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Mohammed J Mamman, Department of Agricultural Technology, Niger State College of Agriculture, Mokwa, PMB 109, Mokwa Niger State, Nigeria. Postal code 913104, Zip code +234. Email: ikwan1565@gmail.com

by a river and its tributaries is called river basin. There are 2 major rivers in Nigeria: River Niger and River Benue. River Niger and its tributaries have great potential for the socioeconomic transformation of the West African subregion, including Nigeria. Kainji hydropower dam is situated along the Niger River; a dam is a barrier built across a water body to hold back the flow of water, thereby creating a large body or pool of water called reservoir. Dams can be classified according to the purpose for which they are meant and also according to the materials used in constructing them, such as concrete and earth fill. The functions of dam include regulation of the river flow and controlling the water release in accordance with the electricity production or irrigation requirement. Despite the usefulness of dams, they usually have significant impact on economy, geology, environment, hydrology, and meteorological variables.

The probability distribution is a hydrologic tool most widely used in flood estimation and prediction. The importance of reservoir inflow analysis at any hydropower dam to our daily life makes it imperative that the appropriate probability distribution model be established to determine the discharge into the reservoir. According to Olukanni and Salami,<sup>3</sup> Larry and Murray<sup>4</sup> stated that the choice of the probability distribution model is almost arbitrary as no physical basis is available to rationalize the use of any particular function. In



Creative Commons Non Commercial CC-BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 3.0 License (http://www.creativecommons.org/licenses/by-nc/3.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage). Downloaded From: https://bioone.org/journals/Air,-Soil-and-Water-Research on 08 Oct 2024 Terms of Use: https://bioone.org/terms-of-use

general, the search for a proper distribution function has been the subject of several studies. Salami<sup>5</sup> studied the flow along the Asa River and established probability distribution models for the prediction of the annual flow regime. For minimum and maximum flows, log-Pearson type III (LP3) and Gumbel extreme value type I (EVI), respectively, were recommended. Salami<sup>6</sup> considered flood levels at 4 gauging stations along the River Niger, below the Jebba hydropower dam. The maximum and minimum flood level data were fitted with 4 probability models and compared graphically with the observed data. The EVI distribution fits the data best, and it was used to predict flood levels with return periods of 10, 50, and 100 years. Olukanni and Salami reported that Onoz and Bayazit<sup>7</sup> dealt with the probability distribution of largest available flood sample with the aim of determining the distribution that best fits the observed flood. According to Olukanni and Salami,<sup>3</sup> the Water Resources Council of the Unite States conducted a study with the objective of developing a uniform technique of determining flood frequency. The work applied the available methods to flood records at 10 stations in various parts of the United States. Record length varied and 5 methods were used, namely, Gamma, EVI, log-Gumbel, log-normal (LN), and LP3 distributions. However, no statistical test was applied to determine the goodness of fit; instead, flood discharge for various return periods (2-50 years) was obtained from the probability plot and was compared with the corresponding values from the 5 hypothesized distributions. Among these methods, the LP3 distribution was preferred for common use and for being capable of fitting skewed data. Salaudeen and Yusuff<sup>8</sup> reported that LN, LP3, EVI distributions for the flood data from 108 stations in Italy. Statistical tests such as  $\chi^2$ , Kolmogrov-Smirnov, and probability plot correlation coefficient were applied, and the best fitting distribution was found to be LN by the  $\chi^2$  test, whereas EVI and LP3 were found to be the best by the other test. According to Olukanni and Salami,<sup>3</sup> estimated 1000-year floods at 300 stations in the United States with 4 different models (LN, Gamma, log-Gumbel, and LP3). Log-normal and LP3 came close to reproducing the expected exceedances and were concluded to be the best. Vogel et al9 explored the suitability of various models applied to the flood flow data at 38 sites in the Southwest United States. The probability distribution models adopted include N, LN, EVI, and LP3, which were compared graphically with the observed data. Ajayi et al<sup>10</sup> estimated the occurrence of flood events and its frequency at the lower Niger basin, Nigeria, using hydrologic data, including river discharges, runoff records, and meteorologic data from different gauging stations within the basin. The data collected were subjected to various statistical analyses and plotting position, and probability distributions were determined. The results showed that various plotting positions and probability distributions could be used to fit the available discharge records of the River Niger. The EVI distribution was the best of the

applied models for peak average reservoir inflow and peak discharge at the River Kaduna (Wuya gauging station). The LN distribution best predicted the peak runoff discharge of River Niger (Lokoja gauging station) and peak discharge at Baro gauging station. The predicted models that compared favorably with the observed values are considered the best distribution models. Busari et al (2013)<sup>11</sup> evaluated best-fit probability distribution models for the prediction of rainfall and runoff volume in Tagwai Dam, Minna, Nigeria. The N distribution model was found most appropriate for the prediction of yearly maximum daily rainfall of 131.21 mm, and the Log-Gumbel distribution model was the most appropriate for the prediction of yearly maximum daily runoff of 1124.73 m<sup>3</sup>/s. According to Chowdhury and Stedinger,12 various probability distribution models were fitted to the peak reservoir inflows, and the suitable model was selected based on the goodness-of-fit tests. The Gumbel (EVI) probability distribution model was found to be appropriate for Kainji. The objective of this article, similar to any modeling research, is to obtain synthetic sequences of stream flow with the same statistical properties as the historic ones.

To this end, stochastic characteristics of the inflows were analyzed. Three methods of probability distribution analysis for the prediction of mean reservoir inflow at Kainji hydropower dam in Nigeria were applied, ie, LN, N, and Gumbel. This study could serve as a guide to the responsible institutions and dam managers in determining available flow that will generate maximum discharge for hydropower dams and prevent flood waters overtopping the dam, thereby causing subsequent release of a flood wave and averting loss of life and properties.<sup>12</sup> The information can also be a valuable tool for preventive flood forecasting.

#### Materials and Methods

#### The study area

Geographically, Kainji Hydroelectric Dam is located in New Bussa town, now headquarter of Borgu local government area of Niger State, Nigeria (Figure 1). The lake is created behind the dam and span between latitude 9° 8′ to 10° 7′ N and between longitude 4° 5′ to 4° 7′ E with reference point 9.54 N and 4.38 E northwest of the Federal Capital Territory (Abuja).<sup>13</sup>

#### Hydrology of the Niger River system

The average rainfall at the headwaters of Niandan and Milo rivers at the source of the Niger at the Fouta Djallon Mountains in Guinea and its exit to the sea in Nigeria is 2200 mm. The river flow regime is characterized by 2 distinct flood periods occurring annually, namely, the white and black floods. The black flood derives its flow from the tributaries of the Niger outside Nigeria (flow lag October to May) and arrives at Kainji



Figure 1. Location of Kainji Hydroelectric Dam. Source: Salami.<sup>15</sup>

Reservoir (Nigeria) in November and lasts until March at Jebba after attaining a peak rate of about 2000 m<sup>3</sup>/s in February.<sup>14</sup> The white flood is a consequence of flows from local tributaries, especially the Sokoto-Rima and Malendo River systems. The white flood is heavily laden with silts and other suspended particles (flow lag June to September) and arrives Kainji in August in the pre-Kainji Dam River Niger having attained a peak rate of 4000 to 6000 m<sup>3</sup>/s in September to October in Jebba. The critical low flow period into the Kainji reservoir is March and July each year. The maximum capacity is  $19 \times 10^9$  m<sup>3</sup>, minimum capacity is  $3.5 \times 10^9$  m<sup>3</sup>, surface area is 1270 km<sup>2</sup>, length is 135 km, maximum width is 30 km, and maximum elevation (m a.s.l.) is 141.9 m.

#### Data collection

The reservoir inflow data for a total of 25 years (1990-2014) were collected from the hydrologic unit of Kainji hydropower station in Nigeria.

# Data analysis and evaluation of probability distribution models

The mean inflow data were evaluated with 3 methods of probability distribution to determine the best probability distribution function for the inflows. The methods adopted were as follows: Gumbel EV1, LN, and N, respectively.

*Gumbel (EV1).* According to Salami<sup>5</sup> and Wilson,<sup>16</sup> the Gumbel (EV1) distribution model is based on the probability that any of the events would equal or exceed a particular value having return period ( $T_r$ ) as given in equation (6).

The Gumbel distribution on the basis of equations (1) and (2) is given according to equation (3):

$$P = 1 - e^{-e^{-T}}$$
(1)

$$Y_T = -\ln(-\ln(1-P))$$
 (2)

where *P* is the probability of occurrence of event,  $Y_T$  is the reduced variate, and ln is the natural log (Yusuf and Salami, 2009)<sup>18</sup>:

$$Q_{T_{e}} = Q_{av} + \sigma \left( 0.78Y_{T} - 0.45 \right) \tag{3}$$

where  $Q_{av}$  is the average of all values of inflows,  $\sigma$  is the standard deviation of the series,  $Q_{T_r}$  is the inflow with return period  $T_r$ , and  $Y_T$  is the reduced variate. The mean values and standard deviation of the inflows were determined using equations (4) and (5), respectively, as stated below:

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{N} x_i \tag{4}$$

where  $\overline{x}$  is the measure of central tendency, n is the size of the sample, x is the observed parameter,  $\Sigma$  is the summation symbol, and N is the number of observations:

$$\sigma_{p} = \frac{\sqrt{\sum (x-\mu)^{2}}}{n-1} \tag{5}$$

where  $\sigma_{p}$  is the standard deviation, x is the observed parameter,  $\mu$  is the mean, and n is the size of the sample.

The average monthly inflows were ranked based on the magnitude of the inflow and the return period computed using equation (6) as in the works of Yusuf and Salami (2009)<sup>18</sup>:

$$Tr = \frac{N+1}{m} \tag{6}$$

where  $T_r$  is the return period, *m* is the series of events ranking, and N is the number of observations in the series. The probability of occurrence of the inflow was computed using equation (7):

$$P = \frac{1}{Tr} \tag{7}$$

The reduced variate  $(Y_T)$  was computed as in equation (8); this is in accordance with the works of Yusuf and Salami (2009)<sup>18</sup>:

$$Y_T = -\ln(-\ln(1-P))$$
(8)

where  $Y_T$  is the reduced variate and P is the probability of occurrence of an event.

The values of the mean and standard deviation of the inflows obtained were substituted into the general Gumbel equation (9) to obtain a new Gumbel distribution (EV1) model (equation (9)) for the reservoir inflows. This equation is used to simulate the inflows of the reservoir and to determine its best fit:

$$Q_T = 817.14 + 675.28 \, y_T \tag{9}$$

The computed reduced variates  $Y_T$  obtained were then substituted into equation (9) to obtain the simulated values of inflow  $Q_T$ . This is to ascertain the fitness of the model and whether the data are significantly correct.<sup>17</sup> The  $R^2$  and r values were used to determine the degree of correlation and linearity between the model prediction and the observed.

*LN distribution.* The LN probability distribution was applied by first finding the seasonal mean of the inflow and then determining the log of the mean values and using equation (10) to evaluate the seasonal standard deviation of the inflows. The standard variable z was determined using equation (11), and the z values obtained were used to determine the K values from the probability distribution table:

$$\sigma = \sqrt{\sum \left(\frac{x - \overline{x}}{n - 1}\right)^2} \tag{10}$$

The standard variable, ie, z values, was estimated using the relationship, as shown in equation (11):

$$z = \frac{\log x - \log \mu}{\sigma} \tag{11}$$

where  $\log x$  is the logarithm of the inflows and  $\log \mu$  is the seasonal mean of the inflows (Zhou, 2000)<sup>20</sup>. The general relationship for the LN probability distribution (equation (12)) was then used as in the works of Yusuf and Salami (2009)<sup>18</sup>:

$$\operatorname{Log} Q_T = \overline{\log Q} + K \sigma_{\log Q} \tag{12}$$

where  $\log Q$  depicts the mean of the log of inflow.

The values of average mean of the log of the inflows and the standard deviation obtained were substituted into equation (12), and a new LN probability distribution model for the reservoir inflows was obtained, as shown in equation (13):

$$Log Q_T = 2.8764 + 0.5384 K_T \tag{13}$$

The obtained values of  $K_T$  were then substituted into equation (13) to have the simulated values of the  $\log Q_T$ . This was done to test the fitness of the LN distribution model.

Normal or Gaussian distribution. The general relationship for the normal probability distribution is given in equation (14) (Busari et al, 2013)<sup>11</sup>. The normal probability distribution was determined by first finding the averages of the reservoir inflows and then determining the standard variable (*z*) using equation (15):

$$Q_t = Q_{AV} + K\sigma \tag{14}$$

$$Z = \frac{x - \mu}{\sigma} \tag{15}$$

The z values computed were used to obtain K values from normal probability distribution table. The values of the average inflow and standard deviation were substituted into equation (14) to obtain a new normal probability distribution model (equation (16)) for the reservoir inflows:

$$Q_t = 1206.72 + 865.7371K \tag{16}$$

The obtained K values were then substituted in equation (16) to obtain the simulated values of the inflow after which the observed and the simulated values of the inflows were plotted to test the fitness of the developed normal probability distribution model.

#### **Testing of the Probability Distribution Models**

The acceptability and reliability of the developed probability distribution models were tested using statistical tests (goodness-of-fit test), such as  $\chi^2$ , probability plot coefficient of correlation (*r*), coefficient of determination (*R*<sup>2</sup>), and standard error of estimate (SEE). The equations, respectively, are presented below:

#### $\chi^2$ test

The expression for the analysis of  $\chi^2$  is as given in equation (17):

$$\chi^{2} = \sum_{j=1}^{N} \frac{\left(O_{j} - e_{j}\right)^{2}}{e_{j}}$$
(17)

where *O* is the observed flow, *e* is the predicted flow, N is the total frequency, and the level of confidence is 95%.

### Probability plot coefficient of correlation (r)

The following equation was adopted for the estimation of the probability plot coefficient of correlation (r):

$$r = \pm \sqrt{\frac{\sum (Q_{est} - Q_{mean})^2}{\sum (Q_{obs} - Q_{mean})^2}}$$
(18)

where  $Q_{est}$  is the value of inflow estimated with the probability function,  $Q_{mean}$  is the mean value of the observed inflow, and  $Q_{obs}$  is the value of the observed inflow.

#### Coefficient of determination $(R^2)$

According to Dibike and Solomatine  $(1999)^{21}$ , the coefficient of determination  $(R^2)$  is given as in equation (19). This is to determine the strength between the observed inflow and the predicted inflows:

$$R^2 = \frac{E_o - E}{E_O} \tag{19}$$

Table 1. Model equations for the probability distributions.

S. NO	HYDROPOWER DAM	PROBABILITY DISTRIBUTION	DEVELOPED EQUATION
1	Kainji	Gumbel (EVI)	$Q_T = 817.118 + 675.28 y_T$
2		Log-normal	$Q_T = 2.8764 + 0.5384 K_T$
3		Normal	$Q_t = 1206.718 + K_T 865.7371$

Abbreviation: EVI, extreme value type I.

 $Q_{\tau}$  is the expected discharge associated with a particular probability of occurrence;  $Y_{\tau}$  is the reduced variate;  $K_{\tau}$  is selected from the normal distribution table.

Table 2.	Results of	goodness-of-fit	tests and th	ne selected	best-fit model	for the	inflow
----------	------------	-----------------	--------------	-------------	----------------	---------	--------

S. NO.	INFLOW	BEST-FIT DISTRIBUTION MODEL	X <sup>2</sup>	R <sup>2</sup>	R	SEE
1		Gumbel	0.0054	0.99	0.96	0.00876
2		Log-normal	0.31	0.85	0.79	0.02
3		Normal	1376.39	0.68	0.75	0.056

SEE: standard error of estimate.



Figure 2. Relationship between average observed inflows and Gumbel (EVI) simulated inflow.

where  $E_o = \sum_{i=1}^{N} (Q_{i(obs)} - Q_{i(mean)})^2$  and  $E = \sum_{i=1}^{N} (Q_{i(obs)} - Q_{i(est)})^2$  $Q_{i(est)}$  is the model output in the *i*th time period,  $Q_{i(obs)}$  is the

 $Q_{i(est)}$  is the model output in the *i*th time period,  $Q_{i(obs)}$  is the observed data in the same period, and  $Q_{i(mean)}$  is the mean over the observed periods.

#### Standard error of estimate

The relationship in equation (20) was adopted in the estimation of SEE:

$$SEE = \sqrt{\frac{1 - r^2}{N - 2}}$$
(20)

#### **Results and Discussion**

The results of the fitted probability functions, ie, Gumbel (EVI), LN, and N for the monthly inflow, are presented below.

The reservoir inflow data were evaluated using various probability distribution functions to determine the best fitting model; the mathematical representations of the evaluated probability functions are presented in Table 1. Also, for the purpose of theoretical determination of best-fit probability function, statistical tools (goodness-of-fit test) were adopted. The results of goodness-of-fit tests and best-fit models are presented in Table 2.

Figures 2 to 4 compare the average monthly inflow of Kainji Reservoir with the model-predicted inflow of the reservoir. The distribution models are the EVI, LN, and N distribution. The inflow at Kainji Reservoir station has values of  $\chi^2$ ,  $R^2$ , r, and SEE as 0.0054, 0.99998, 0.95518, and 0.00876, respectively, for Gumbel distribution. From this result, the value of the ratio of calculated  $\chi^2$  to the  $\chi^2$  table is less than 1, and the model gives the correlation coefficient (r) value of 0.955518,  $R^2$  value of 0.99998, and SEE value of 0.00876







Figure 4. Relationship between observed inflows and normal simulated inflow.

Table 3.	Inflow	values	(Gumbel	distribution)	) with	return	periods.
Table 0.	mmow	values	Gumber	uistinbution	/ ••••	roturn	perious.

MONTHS	AVERAGE INFLOW (M <sup>3</sup> /S)	RETURN PERIOD ( <i>T<sub>R</sub></i> )	PROBABILITY OF OCCURRENCE	EXTREME VALUE TYPE I (EVI) SIMULATED INFLOW (M <sup>3</sup> /S)
April	131.9457	2.60	.384615	558.6428
Мау	81.45882	1.85	.538462	181.0456
June	122.2636	1.44	.692308	393.7838
July	655.9912	1.30	.769231	847.8401
August	1862.457	1.08	.923077	1720.657
September	2743.942	1.18	.846154	2522.331
October	2030.267	1.63	.615385	2025.489
November	1548.73	4.33	.230769	1140.973
December	1708.727	13.0	.076923	1492.682
January	1617.771	6.50	.153846	1305.048
February	1334.489	2.16	.461538	990.8207
March	642.5718	3.25	.307692	706.1196

which shows that the model is strong and there is a strong linearity between the observed and the predicted reservoir inflow. The same applies to the LN distribution except for N distribution where the  $\chi^2$  value is greater than 1 (Table 2). Also, based on the graphical comparison (Figures 2-4), the EVI distribution model is a better fit than the other probability distribution models. Hence, EVI is the most appropriate model for the reservoir inflow at Kainji Reservoir. The flow in the month of May is low as a result of incipient rainfall within the month. The margin between the observed and simulated data could be as a result of errors in course data taken (Table 3).

#### Conclusions

Owing to the purpose of identifying a more realistic modeling scheme for the inflow series, assessment of the stochastic characteristics was done to be able to understand the dynamics of monthly series. Sequel to this, various probability distribution models were fitted to the reservoir inflow records to evaluate the model that is most appropriate for prediction at Kainji hydropower station in Nigeria. Three models were established for the hydropower station, and the most suitable model was selected based on the goodness-offit tests. The EVI model was found to be appropriate for Kainji Reservoir.

#### **Author Contributions**

Conceived and designed the experiments: MJM. Analyzed the data: MJM. Wrote the first draft of the manuscript: MJM Contributed to the writing of the manuscript: MJM. Agree with manuscript results and conclusions: MJM. Jointly developed the structure and arguments for the paper: MJM. Made critical revisions and approved final version: MJM. All authors reviewed and approved of the final manuscript.

#### REFERENCES

- Ahaneku IE, Otache YM. Stochastic characteristics and modelling of monthly rainfall time series of Ilorin, Nigeria. Open J Mod Hydrol. 2014;4:67–79.
- Kottegoda NT. Stochastic Water Resources Technology. London, England: Macmillan Publishers; 1980:2-3,21,112–113.
- Olukanni DO, Salami AW. Fitting probability distribution functions to reservoir inflow at hydropower dams in Nigeria. *J Environ Hydrol*. 2008;16:16–18.
- Larry JS, Murray RS. Theory and Problems of Statistics. 3rd ed. New Delhi, India: McGraw-Hill; 2000:314–316.
- Salami AW. Prediction of the annual flow regime along Asa River using probability distribution models. *AMSE Periodical*. 2004;65:41–56.
- Salami AW. Fitting probability curves for flood levels at Jebba, Belle, Oshin and Yelwa gauging station, along River Niger. Technical Paper presented: Department of Civil Engineering, University of Ilorin. 28 February 2002; Ilorin, Nigeria.
- Onoz B, Bayazit M. Best distributions of largest Available flood samples. J Hydrol. 1995;167:195–208.
- Salaudeen FA, Yusuff OK. Comparison of probability distribution models for the prediction of annual flow regime along River Niger and Benue in Nigeria. *J Inform Knowl Res Civ Eng.* 2008;5(1):68–83.
- Vogel RM, Thomas WO, McMahon TA. Flood flow frequency model selection in South Western United States. J Water Res Pl: ASCE. 1993;119:353–366.
- Ajayi AE, Olufayo AA, Affinnih TA. Historical study of flood events in the lower Niger River Basin, Nigeria. *Geophys Res Abstr.* 2007;9:10883.
- Busari AO, Mohammed S, Ajibola O. Evaluation of best fit probability models for the prediction of rainfall and runoff volume (Case Study of Tagwai Dam, Minna-Nigeria). Int J Eng Technol. 2013;3(2):4–6.
- 12. Chowdhury JU, Stedinger JR. Goodness of fit tests for regional generalized extreme value flood distributions. *Water Resour Res.* 1991;27:1765–1776.
- 13. Binnie GM. Early Victorian Water Engineers. London, England: Thomas Telford; 1981.
- Dukiya JJ. Spatial analysis of the impacts of Kainji hydropower dam on the down stream communities an overview. J Geoinformatics Geostat. 2013;9:1–5, doi:http:// dx.doi.org/10.4172/2327-4581. S1-009.
- Oyebande L, Balogun I. Water resources management in the semiarid areas of Northern Nigeria. Can J Dev Stud. 1992;13:209–226.
- 16. Salami AW. An overview on reservoir operational Impact of Kainji, Jebba and Shiroro dams on the environment. Paper presented at: One day seminar on Reservoir Operation, organized by Federal Ministry of Water Resources, Dams & Reservoir Operation department; October 29, 2010; Abuja, Nigeria.
- Wilson EM. Engineering Hydrology. 3rd ed. Houndmills, UK; London, England: Macmillan Publisher; 1990.
- Yusuf K.O, Salami AW. Influence of Jebba hydropower dam on statistical distribution of hydro-meteorological variables in Niger River Basin, Nigeria. In: 1st Annual Civil Engineering Conference; University of Ilorin, Nigeria, 26–28 August 2009.
- Zhang J, Hall MJ. Regional flood frequency analysis for the Gan-Ming River basin in China. J Hydrol. 2004;296:96–115.
- Zhou Y. Comparison of various floods frequency distributions using annual flood peaks data of river in Anatolia. J Hydrol. 2000;136:1–31.
- Dibike BY, Solomatine DP. River flow forecasting using artificial neural network. Paper presented at: European Geophysical Society (EGS) XXIV General Assembly. The Haugue, The Netherlands; 1999:1–11.