VOLUME 10, NO. 1 - 2023
ISSN 2465-7425

NIGERIA JOURNAL OF ENGINEERING AND APPLIED SCIENCES (NJEAS)

NIGERIA JOURNAL OF ENGINEERING AND APPLIED SCIENCES (NJEAS)

Nigeria Journal of Engineering and Applied Sciences – NJEAS (ISSN: 2465-7425) is a publication of the two schools of Engineering: School of Infrastructure, Process Engineering and Technology (SIPET) and School of Electrical Engineering and Technology (SEET) of Federal University of Technology, Minna, Nigeria.

The journal is committed to promoting and publishing original high quality research papers and stresses on academic excellence, knowledge distribution as well as joint scholarly efforts in order to support theoretical, experimental and practical research. Papers for publication are selected through peer review to ensure originality, relevance, and readability.
The subject areas related to this journal include but are not limited to:
1. Chemical, Petroleum and Gas Engineering 2. Civil Engineering 3. Computer Science and Information Technology. 4. Electrical Engineering 5. Electronics & Telecommunication Engineering 6. Mechanical & Mechatronics Engineering and 7. Allied Applied Sciences.

EDITORIAL BOARD MEMBERS

Engr. Prof. A. A. Amadi	Editor-in-Chief
Engr. Prof. O. M. Olaniyi	Managing Editor
Engr. Prof. (Mrs) Z. D. Osunde	Member
Engr. Prof. A. S. Kovo	Member
Engr. Prof. J. O. Okafor	Member
Engr. Prof. A. M. Aibinu	Member
Engr. Prof. M. Abdullahi	Member
Engr. Prof. A. S. Abdulrahman	Member
Engr. Prof. S. A. Lawal	Member
Engr. Dr. P. Adeoye	Member
Engr. Prof. M. Nwohu	Member
Engr. Prof. J. Tsado	Member

Published by:

School of Infrastructure, Process Engineering & Technology and School of Electrical Engineering & Technology, Federal University of Technology, Minna, Nigeria.

All Correspondences to:
Editor-In-Chief,
Nigeria Journal of Engineering and Applied Sciences (NJEAS),
Federal University of Technology, P. M. B. 65, Minna, Nigeria.
E-Mail: njeas@futminna.edu.ng, Website: njeas.futminna.edu.ng

TABLE OF CONTENTS

PREDICTING THE FLOW CHARACTERISTICS OF RIVER NIGER USING ARTIFICIAL INTELLIGENCE MODELS

[1]Gbadebo, O.A., [1]Busari, A. O., [1,2]Sadiku, S. and [1]Saidu M.

[1]Civil Engineering Department, Federal University of Technology Minna, Niger State, Nigeria.

[1,2]Confluence University of Science and Technology, Osara, Kogi State, Nigeria.

*Corresponding author email: g.olukemi@futminna.edu.ng

ABSTRACT

Artificial intelligence (AI), as a branch of computer science, is capable of analysing long-series and large-scale hydrological data. In recent years, AI technology has been applied to the hydrological forecasting modelling. It is essential to determine the hydrological system of River Niger, which is the major water sources of the annual flood in Lokoja, Kogi State, Nigeria. This paper investigates and compares the forecasting capability of three algorithms namely Artificial Neural Network (ANN), Support Vector Machine Regression (SVM Reg.) and Random Forest (RF) to determine the optimal model for forecasting downstream river flow. Daily discharges data from 2001 to 2019 were obtained from National Inland Waterways Authority at Lokoja, Kogi State, Nigeria and applied in the forecasting analysis. Discharge data were divided into 65:35 percent for training and testing respectively. The results of evaluation criteria based on Root Mean Square Error (RMSE), Nash-Sutcliffe Efficiency Coefficient (NSEC), Coefficient of correlation (CC) and Accuracy (ACC) showed that all the models applied gave perfect results except the value obtained for uncertainty analysis in ANN model which was 1.4445 and 0.6219, was slightly high when compare with the values of RF 0.1634 and 0.0134 and SVM Regression models 0.1634 and 0.1210 in testing and training phases respectively. This is caused by the failure of ANN model to carry out pre-processing of discharge data, to remove all the error present in the data unlike the SVM Regression and RF models. Therefore, the RF and SVM Regression algorithms are considerably more adaptive in optimizing the forecasting problem for the river flow prediction.

Keywords: Artificial Neural Network, Machine Learning Models, Random Forest, River Prediction, Support Vector Machine regression.

INTRODUCTION

River flow prediction is a requirement for various uses of water resources like design of reservoir and flood warning systems. The process of hydrology of river flow is very complicated in which a simple data driven model cannot expound its characteristics. It is very essential to investigate the suitable models to a high degree for estimating uncertainties in stream-flow, the nonlinearity and seasonal flow of river. In this study, Artificial Intelligence models which consists of Artificial Neural Network - ANN, Support Vector Machine Regression – SVM Regression and Random Forest – RF were used to predict the streamflow of River Niger, Lokoja, Kogi State, Nigeria.

The SVM technique depends on the principle of statistical learning (Vapnik 1998). The SVM is one type of neural networks that has gain increasing attention in the classification of pattern and in the estimation of nonlinear regression because of its generalization performance (Cao and Tay Francis, 2003). SVM is a type of supervised machine learning algorithm which belongs to kernel-based learning techniques and it uses a linear high dimensional hypothesis space called feature space and thus, the SVM has gained a wide popularity. The basic principle of the SVM is that it uses kernel functions implicitly, mapping the data to a higher dimensional space (Bhagwat and Maity, 2012). Random Forests (RF) are supervised machine learning algorithms that have of recent gained popularity in water resource applications. It has been used in a various water resource research domain, which include simulation of discharge and water level. Random forest is an alternate approach to physical and conceptual hydrological models for large-scale hazard assessment in various catchments due to its inexpensive setup and operation costs (Jibril *et al*., 2022). Ighile *et al*. (2022) applied Machine learning and GIS to predict flood prone areas in Nigeria from 1985 – 2020. They used Receiver operating characteristic curve and Area under Curve to evaluate the ANN and Logistic Regression models and found that both models can predict flood prone areas well.

Miller *et al*. (2018) applied RF to quantify monthly flow of river from 1950 to 2015 and obtained very high coefficient of NSEC of 0.85. The actual/predicted ratio of 94% implied a better consistency between predicted and actual river flow at almost 2000 gaging station. Sha *et al*. (2017) compared daily discharge using five various algorithms, given as: Basic extreme learning machine, extreme learning machine with kernels, random forest, back-propagation neural network, and support vector machine. The results indicated that the extreme learning machine with kernels algorithm have the best performance when compared with the other four algorithms, and the basic extreme learning machine algorithm has the least performance. The RF algorithm has good performance in peak flow prediction, while the extreme learning machine with kernels algorithm performed best in low flow prediction. Tongal *et al*. (2018) Predict and Simulate discharge data with the use of Support Vector Machine Regression (SVM Reg.), Artificial Neural Networks (ANNs), and Random Forest (RF) using precipitation (P), temperature (T), and Potential EvapoTranspiration (PET) as its function. Mohammad *et al*., (2016), Modelled River discharge time series using support vector machine and artificial neural networks and compared the performance with conventional method of rating curve and Multi linear regression. The result obtained indicated that the SVM and ANN has a better performance than that of Rating Curve and MLR, which are convectional method. Qiu *et al*. (1998) combined the use of fuzzy pattern recognition activation function with an ANN model for prediction of runoff. **Activation function** is a mathematical equation which determines the output of a neural network. This function grouped the runoff into monsoon and non-monsoon periods, which pointed nonlinear and periodic behaviour of the river system (Chen *et al*., 2015). Liong and Chandrasekaran, (2007) applied a machine learning algorithm of SVM and ANN for flood prediction at Dhaka, Bangladesh, and discovered that the predictive ability of SVM is better than that of ANN.

River Niger in Lokoja, Kogi State is always flooded during the peak of the raining season of every year. Thereby leading to the destruction of lives, infrastructures and properties in the area (Jimoh & Salami, 2020). Abdulkadir *et al*., (2012) applied ANN model to the management of hydropower Reservoir along River Niger in Nigeria. They used ANN to predict reservoir storage capacity along River Niger in Nigeria. The results obtained yield a better prediction both in training and testing phases for Jebba (0.95 and 0.97) and Kanji Reservoir (0.69 and 0.75). They failed to study the predictive power of flow of River Niger as a result, little work has been done on River Niger flow prediction with the use of combination of ANN and Machine learning models. Hence the need to study the flow prediction of the River Niger using the combination of three AI models (ANN, SVM Regression and RF). In view of this, the aim of this paper is to compare the forecasting performance of the three (3) models used for river Niger flow prediction and to find the uncertainties associated with each of the model.

METHODOLOGY

Collection of Data

All the relevant discharges (daily discharges data) from Lokoja gauging station at National Inland Waterways Authority which were available from the past were collected from 2001 to 2019. The collected information formed the sample space for the random variable under consideration.

Artificial Intelligence and Machine Learning Models

Artificial Neural Network (ANN) model

Artificial neural network (ANN) is a computer programs that is biologically inspired and designed to simulate the way in which the human brain processes information. ANNs gather their knowledge by detecting the patterns and relationships in data and learn through experience, not from programming. The use of ANN models has contributed to an increase of interest within hydrology and hydraulic community. A number of ANN models have been used in hydrological modelling (Renaud & Robert, 2022; Kumar *et al.*, 2016; Tanty, & Desmukh, 2015 and Gunathilake *et al.*, 2021). The Artificial Neural Network (ANN)-based machine learning methods have made great progress than ever before, such as the deep learning and reinforcement learning (Kan *et al.*, 2020). It consists of a three-layer feed forward ANN, the input layer consists of nodes, which are joined or linked with an activation function to the hidden layer and it also consists of nodes in the output layer. An objective function is obtained by comparing the differences in the actual and predicted output. ANN model has three layers called input, hidden and output layers. A concept of ANN is introduced from input to hidden layer and is defined as follows:

$$Q_i = \sum_{i=1}^{n} w_{ij} q_j + b \tag{1}$$

Where Q_i are nodes in the hidden layer (i= 1, 2…, n) and q_j^{in} connotes nodes in the input layer (j= 1, 2..., k). The connecting factor w_{ij} denotes the weight parameters from the input to the output layers and b is the bias. Fig. 1 shows a basic overview of ANN topology.

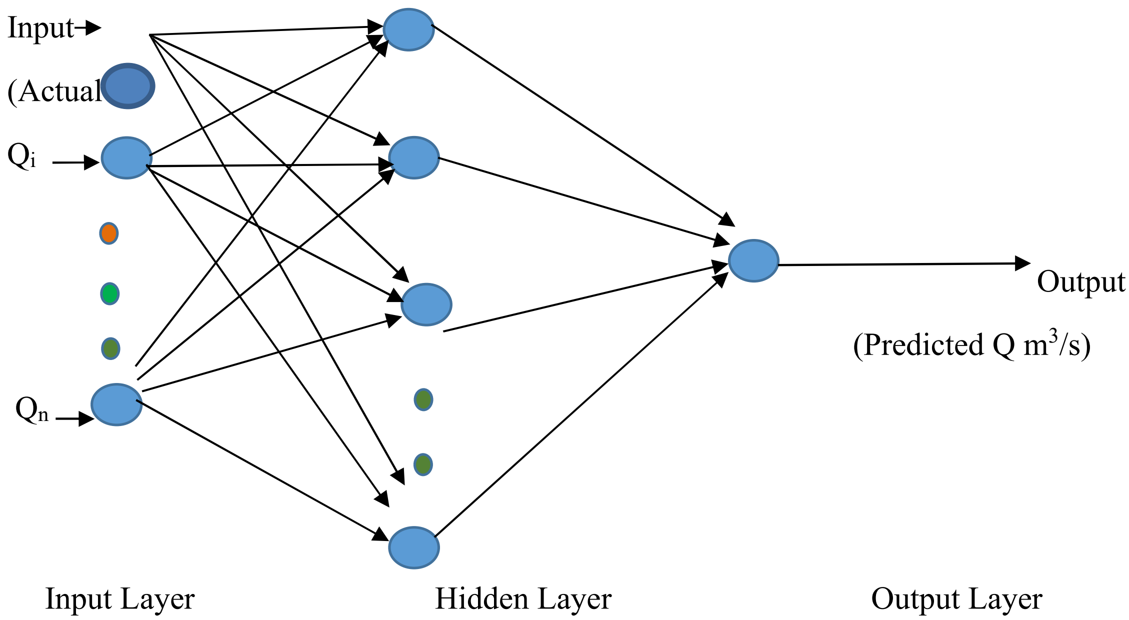


Fig. 1 shows a basic overview of ANN topology.

Support Vector Machine Regression (SVM Regression) Model

SVM is a powerful supervised machine learning algorithm which employs various classifications and regression problems. It is based on Structural Risk Minimization as opposing the principle of Empirical Risk Minimization selected by conventional regression methods. It is one of the robust techniques for flood prediction. It (SVM Reg.) is a probabilistic approach whereas Support Vector Machine is based on statistical approaches (Anshul, 2021). It uses a hyper-plane that divides and segments the data and classes. SVM find the maximum margin between the hyper-planes that means maximum distances between the two classes. The equation of hyper-plane for a linearly separable data is obtained from straight line equation and given by:

$$y = ax + b \quad (2)$$

By replacing x with x_1 and y with x_2, and substitute into equation (ii) the equation becomes:

$$ax_1 - x_2 + b = 0 \quad (3)$$

If x is defined by (x_1, x_2) and ω is given by (a, -1), which is a vector normal to hyperplane and b is the offset. Then the equation of the hyper-plane is given as:

$$\omega x + b = 0 \quad (4)$$

The result of the hyper-plane was used to make the river prediction for the stream-flow with the following hypothesis function:

$$Q(x_i) = \begin{Bmatrix} +1 if \omega.x + b \geq 0 \\ -1 if \omega + b \angle 0 \end{Bmatrix} \quad (5)$$

The point above the hyper-plane is classified as +1 and below is -1. The aim of this algorithm is to find the hyper-plane that could separate the dataset successfully.

Random Forest (RF)Model

Random Forest is a tree-based computer algorithms, and a supervised machine learning methods which employs the principle of ensemble learning methods, that is widely used in classification and regression problems (Sruthi, 2021). It is a predictive model with high accuracy, stability and ease of interpretation (Li et al, 2016). RF performs better in regression and classification work. Hyper-parameters are employed in random forest to improve the performance and predictive power of models. The ensemble learning technique is a combination of multiple models. It uses bagging and boosting method. Bagging creates many trainings subset from the training data set with replacement and the output depends on majority voting. Boosting makes weaker learner to be stronger ones by making a sequential model in a way that the final model will have the highest accuracy. It is calculated as the decrease in node impurity weighted by the probability of reaching that node. The node probability can be calculated by the number of samples that reach the node, divided by the total number of samples.

For classification Gini impurity formula was applied. For regression, variance reduction using mean square error was adopted. Variance reduction can also be estimated using mean absolute error as well in Scikit learn. The equation used for classification is given by Gini impurity as:

$$\sum_{i=1}^{n} f_i(1 - f_i) \quad (6)$$

Regression equation for variance or mean square error is given as:

$$\frac{1}{n}\sum_{i=1}^{n}(Q_i - \overline{Q}_i)^2 \quad (7)$$

Regression equation for variance or mean absolute error is given as:

$$\frac{1}{n}\sum_{i=1}^{n}(Q_i - \overline{Q}_i) \quad (8)$$

Where f_i is the frequency of label i at a node and n is the number of instances or observation, Q_i is the actual discharge and $\overline{Q}_i$ is the mean of actual discharge.

Evaluation criteria for Model Performance

The evaluation and comparison of three models is carried out to determine the model with the best performance. The model employed are ANN, Random Forest (RF) and Support Vector Machine regression (SVM Regression) models. Four (4) evaluation criteria were used to determine the best model performance, which are the root mean square error (RMSE) in m^3/s, Nash Sutcliffe efficiency coefficient, correlation coefficient (CC), and the accuracy (ACC). The following evaluation criteria equations were applied in the study as shown:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Q_i - \hat{Q}_i\right)} \qquad (9)$$

$$\text{NSEC} = 1 - \sum_{i=1}^{n}\frac{\left(Q_i - \hat{Q}_i\right)^2}{\left(Q_i - \overline{Q}\right)^2} \qquad (10)$$

$$\text{CC} = \frac{\sum_{i=1}^{n}\left(Q_i - \overline{Q}\right)\left(\hat{Q}_i - \overline{\hat{Q}}\right)}{\sqrt{\sum_{i=1}^{n}\left(Q_i - \overline{Q}\right)^2\left(\hat{Q}_i - \overline{\hat{Q}}\right)^2}} \qquad (11)$$

$$\text{ACC} = 1 - \sum_{i=1}^{n}\frac{\left(Q_i - \hat{Q}_i\right)}{\left(Q_i\right)} \qquad (12)$$

where Q_i, $\hat{Q}_i$, $\overline{Q}$ and n represented the actual discharge, predicted discharge, mean discharge and number of instances respectively. RMSE which is the evaluation criteria is a measure of an absolute error. The smaller value of mean square error indicate that the model performance is better. RMSE values ranges from 0 – infinity, while NSEC, CC and ACC values ranges from zero to one (0 - 1), which implies no fit for a value of zero and a perfect or flawless fit for 1.

RESULTS AND DISCUSSION

Evaluation for Model Performance.

The evaluation criteria were done for the ANN, SVM Regression and RF in training and testing but due to the bulkiness of the data, smaller portion of evaluation measures for ANN were presented in Table 1.

Table 1: Evaluation Criteria for ANN Model in Training phase.

ANN Model			RMSE		NSEC	ACC
Time	Actual (Q)	Predicted	(Qi - Q^i)^2		(Qi-Q-)^2	(Qi-Q/Qi)
01-Feb-01	2902	2894.912	50.239	-7.088	11075337.930	0.002
02-Feb-01	2941	2929.425	133.980	-11.575	10817277.810	0.003
03-Feb-01	2949	2963.988	224.640	14.988	10764718.410	-0.005
04-Feb-01	2949	2971.083	487.658	22.083	10764718.410	-0.007
05-Feb-01	2902	2971.083	4772.460	69.083	11075337.930	-0.023
06-Feb-01	2889	2929.425	1634.180	40.425	11162033.970	-0.013
07-Feb-01	2811	2917.915	11430.817	106.915	11689308.200	-0.038
08-Feb-01	2760	2848.972	7916.016	88.972	12040643.430	-0.032
09-Feb-01	2760	2804.001	1936.088	44.001	12040643.430	-0.015
10-Feb-01	2760	2804.001	1936.088	44.001	12040643.430	-0.015
11-Feb-01	2760	2804.001	1936.088	44.001	12040643.430	-0.015
12-Feb-01	2915	2804.001	12320.778	-110.999	10988979.890	0.038
13-Feb-01	2712	2940.94	52413.523	228.940	12376063.880	-0.084
14-Feb-01	2688	2761.753	5439.505	73.753	12545502.110	-0.027
15-Feb-01	2640	2740.657	10131.831	100.657	12887834.560	-0.038
16-Feb-01	2628	2698.523	4973.493	70.523	12974137.670	-0.026
17-Feb-01	2640	2688.001	2304.096	48.001	12887834.560	-0.018
18-Feb-01	2664	2698.523	1191.837	34.523	12716092.330	-0.012
19-Feb-01	2664	2719.580	3089.136	55.580	12716092.330	-0.020
20-Feb-01	2664	2719.580	3089.136	55.580	12716092.330	-0.020
21-Feb-01	2671	2719.580	2360.016	48.580	12666217.850	-0.018

22-Feb-01	2676	2725.726	2472.675	49.726	12630653.220	-0.018
23-Feb-01	2664	2730.116	4371.325	66.116	12716092.330	-0.024
24-Feb-01	2628	2719.580	8386.896	91.580	12974137.670	-0.034
25-Feb-01	2640	2688.001	2304.096	48.001	12887834.560	-0.018
26-Feb-01	2346	2698.523	124272.465	352.523	15085168.820	-0.150
27-Feb-01	2301	2442.110	19912.032	141.110	15436750.490	-0.061
28-Feb-01	2301	2403.117	10427.881	102.117	15436750.490	-0.044
29/02/2001	2251	2403.117	23139.581	152.117	15832146.800	-0.067
30/02/2001	2168	2359.870	36814.096	191.870	16499543.660	-0.088
01-Mar-01	2616	2288.265	107410.230	-327.735	13060728.780	0.125
02-Mar-01	2616	2677.484	3780.282	61.484	13060728.780	-0.023
03-Mar-01	2616	2677.484	3780.282	61.484	13060728.780	-0.023
04-Mar-01	2616	2677.484	3780.282	61.484	13060728.780	-0.023
05-Mar-01	2616	2677.484	3780.282	61.484	13060728.780	-0.023
06-Mar-01	2616	2677.484	3780.282	61.484	13060728.780	-0.023
07-Mar-01	2616	2677.484	3780.282	61.484	13060728.780	-0.023
08-Mar-01	2616	2677.484	3780.282	61.484	13060728.780	-0.023
09-Mar-01	2616	2677.484	3780.282	61.484	13060728.780	-0.023
SUM	30838317		3480735048	445465.217	1.56784E+11	
RMSE			838.5575696	0.1202		
MEAN	6229.963					
NSEC					0.9777	
Uncertainty	0.0144					
Analysis	1.4445					

Comparison of AI Model Performance.

Table 2: Summary of the Model Performances for ANN, RF and SVM Reg.

Algorithm	Training	Testing						
	RMSE	NSEC	CC	ACC	RMSE	NSEC	CC	ACC
ANN	0.1202	0.9778	0.9879	1.0033	**6.4284**	0.9820	0.9910	1.0020
Random F.	**0.0116**	**0.9961**	**0.9983**	1.0010	3.2955	**0.9922**	**0.9956**	**1.0016**
SVM reg.	0.0348	0.9887	0.9982	1.0010	3.2955	**0.9922**	0.9914	**1.0016**

Fig. 2 shows the scattered plot of actual and predicted discharges by the three algorithms applied (SVM Regression RF and ANN) in testing phases.

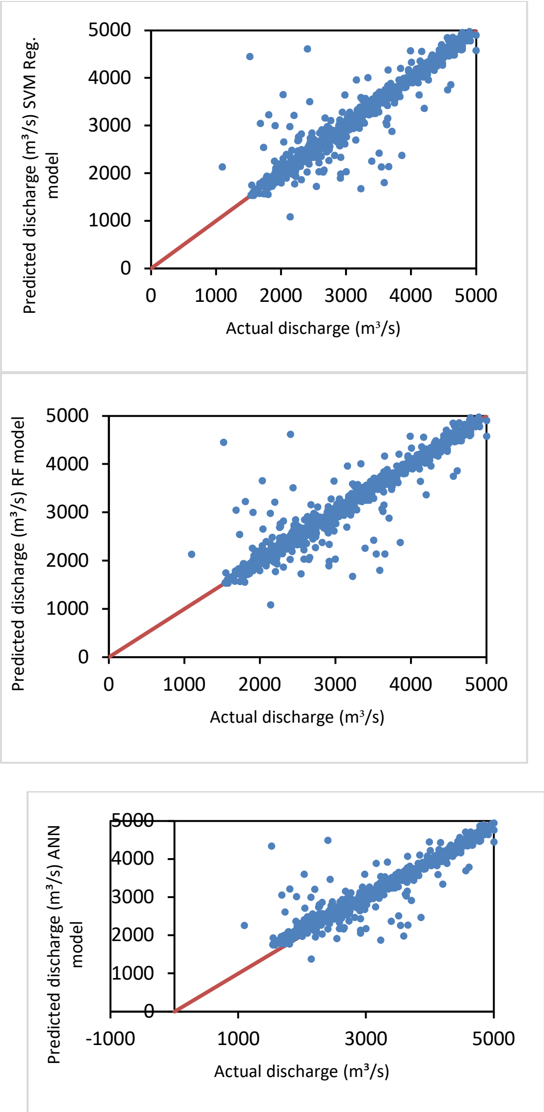
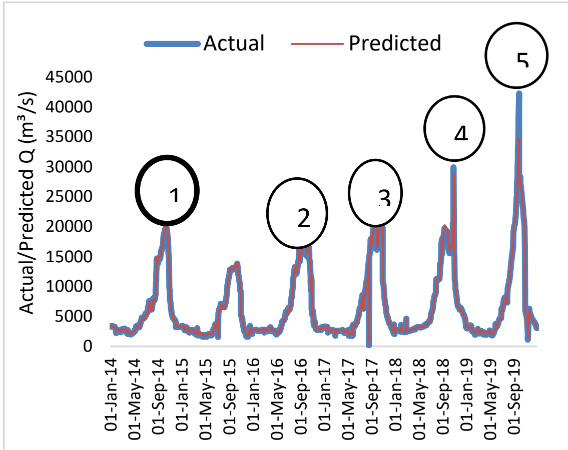
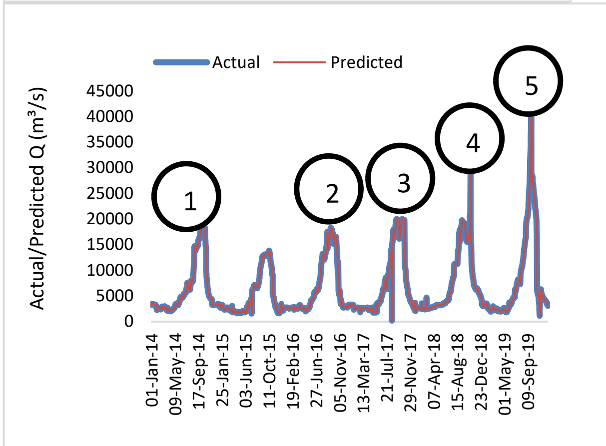


Fig. 2: Actual and Predicted discharges of different algorithms in the testing stages.

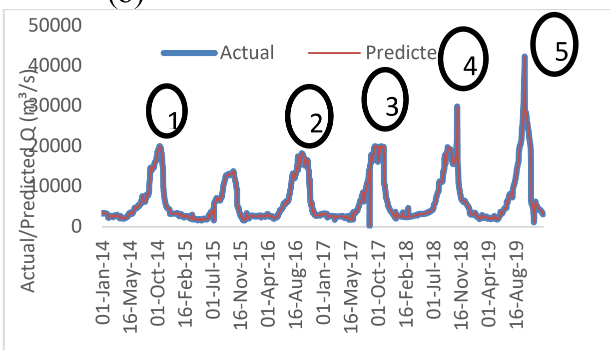
The SVM Regression and RF models portray good agreement between the actual and predicted discharges in testing stages. The plot of the intensively distributed dots along the ideal line from 1500 - 2000m^3/s implies that the low river flow are mostly well predicted. The reason been that the regular occurrence of low values allows an improved or good generalization of the trained model. The performance of ANN model is not good as compared with that of RF and SVM Reg. in the low flow during the testing phase from 1500 – 2000m^3/s. Fig. 3 (a – c) depicts the time series of actual and predicted discharges by the ANN, SVM and RF algorithms and starred five actual extreme values. The values are provided in Table 3 together with the algorithms employed.





(a)

(b)



(c)

Fig. 3: Testing from 2014-2019 (a) ANN; (b) Support Vector Machine Regression (c) Random Forest

It was observed that all the three algorithms employed captured all the extreme values. It starred the extreme values for the actual discharges. The peak value was obtained at extreme point 5 for all the models. The peak value was under predicted by the ANN model and have a good prediction for both RF and SVM Regression due to the relative closeness algorithms result. Extreme values 2, 3 and 4 over-predicted the actual discharges, while SVM Regression predicts well the actual discharge in extreme value 1. RF slightly under predict and ANN slightly over predict the actual extreme value at point 1. The relative mean error between the actual and predicted discharges are the same for both RF and SVM Regression which was 0.0090, while ANN has a slightly higher relative mean error of 0.0100 when compare with that of RF and SVM Regression respectively. In addition, the values of uncertainty analysis 1.4445 and 0.6219 for ANN is very high in training and testing respectively, while RF has the uncertainty values of 0.0134 and 0.1634 in training and testing respectively. SVM Reg. has the uncertainty analysis values of 0.1210 and 0.1634 in training and testing.

Table 3 presented the summary of the values of the extreme points and relative mean error for the three models employed in this study. The values for the extreme points were displayed and the least relative mean error was obtained by SVM Reg. and RF, while the ANN has a slightly high relative mean error when compare with that of SVM Reg. and RF.

Table 3: Extreme and Relative Mean Error values obtained for the algorithm

Algorithm	Extreme Value 1 (m^3/s)	Extreme Value 2 (m^3/s)	Extreme Value 3 (m^3/s)	Extreme Value 4 (m^3/s)	Extreme Value 5 (m^3/s)	Relative Mean Error
Actual Q	19695	17280	10855	19640	42248	
ANN	19965.22	17914.08	20299.73	28822.85	33963.24	0.0100
SVM Reg.	19633.11	17495.93	19660.19	29937.88	42361.77	0.0090
RF	19408.46	17495.93	19854.75	29937.88	42361.77	0.0090

Confidence Interval (CI)

Confidence Interval is the range in which the true mean value will lie with a high probability. In order to calculate CI, the distribution function of the mean values or variables in the observation is essential. Assume that the distribution is normally distributed, the CI for the mean value is given in equation (13) as:

$$CI = \overline{Q} \pm z * \frac{\sigma}{\sqrt{n}} \tag{13}$$

where CI is the confidence interval, $\overline{Q}$ is the mean discharge, z is the value for the confidence interval for 95% and 99% CI which is given by 1.96 and 2.57 respectively, σ is the standard deviation and n is the number of instances or observation (DATAtab, 2023). The mean, variance, 95% and 99% CI of the three algorithms employed in this study are provided in Table 4.

The statistical analysis is essential in order to examine the significance or importance of the differences. The mean value obtained from SVM Regression and RF are closer to the actual mean value as compare with that of ANN that is a slightly above the actual mean. The three models obtained variances that is slightly above the actual variance, this implies that the predicted results are widely distributed. The closest value to the actual is achieved by RF algorithm for the mean with 95% CI, but there is a slight difference in the actual value when compared with the other algorithm employed. As a result, the predicting performances by these three algorithms are comparable.

Table 4: Statistical Analysis of different algorithm used for testing period.

Algorithm	Mean (m^3/s)	Variance (m^6/s^2)	Mean of 95% CI (m^3/s)	Mean of 99% CI (
Actual Q	6645.01	36398877	6.8994	6.9786
ANN	6686.34	36585845	6.9414	7.0208
SVM Reg.	6655.87	36608461	6.9110	6.9905
RF	6655.87	36608461	6.7125	6.9905

CONCLUSION

In conclusion, the performances of the three models employed are compared. The purpose is basically to consider a relatively reliable model for river prediction at the downstream part of the river. The input variables employed were the discharges of River Niger from 2001 to 2019. The capability to capture extreme values and four statistical evaluations (RMSE, NSEC, CC and ACC) were employed to estimate the predicting performances. The results indicated that the three algorithms perform well on generalization and forecasting of daily discharge data used. In addition, Random Forest performs excellently with high efficiency due to its ability to handle binary, continuous and categorical data. It also has the ability to reduce the risk of over-fitting, reduction in training time, fixing of the missing data and the stability is very high. ANN displayed high rate of uncertainty analysis 1.4445 in this study when compared with the RF and SVM Regression. Therefore, the best model with the least error in the stream flow data set is RF followed by SVM regression. The models adopted in this study can be used to solve the nonlinear and non-differential problems in multidimensional space.

REFERENCES

Abdulkadir, T. S., Sule B. F. and Salami A. W. (2012). APPLICATION of ANN Model to the Management of Hydropower Reservoirs along River Niger, Nigeria. Faculty of Engineering Hunedoara. Pp 419 – 424. e-ISSN 26012332. 1077766189.

Anshu, S. (2021) Support Vector Machine a complete guide for beginners. *Analytics Vidhya.* Analyticsvidhya.com.

Cao L. J. and Tay Francis E. H. (2003). Support vector machine with adaptive parameters in Financial time series forecasting. IEEE Trans Neural Netw 14:1506–1518. https://doi:10.1109/TNN.2003.820556

Chen, X. Y. Chau, K. W. and Busari, A. O. (2015). A comparative study of population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model. Engineering application of Artificial intelligence 46, 258 – 268.

DATAtab Team (2023) DATAtab: Online statistical calculator. DATAtab e.U.Graz, Australia. URL https://datatab.net

Gunathilake, M. B., Karunanayake, C., Gunathilake, A. S., Marasingha, N., Samarasinghe, J. T., Bandara, I. M. and Rathnayake U. (2021) Hydrological Models and Artificial Neural Networks (ANNs) to Simulate Stream-flow in a Tropical Catchment of Sri Lanka. Applied Computational Intelligence and Soft Computing. Volume 2021 | Article

ID 6683389 | https://doi.org/10.1155/2021/6683389

Jimoh U. U. & Salami H. (2020). Spatio-Temporal Analysis of Flooding in Lokoja Kogi State Nigeria. International Journal of Scientific Research in Research Paper Multidisciplinary Studies, E- ISSN: 2454-9312 Vol.6, Issue.2, pp58-72, P-ISSN: 2454-6143

Jibril, M.M., Bello A., Aminu I. I., Ibrahim A. S., Bashir A., Malami S. I., Habibu M. A., and Magaji M. M., (2022). An overview of streamflow prediction using random forest Algorithm. GSC Advanced Research and Reviews, 2022, 13(01), 050–057. https://doi.org/10.30574/gscarr.2022.13.1.0112.

Ighile, E.H., Shirakawa, H. and Tanikawa, H. (2022). Application of GIS and Machine Learning to Predict Flood Areas in Nigeria. Sustainability, 14, 5039. https://doi.org/10.3390/ su14095039 Hiroaki Shirakawa and Hiroki Tanikaw

Kan, G., Liang, K., Yu, H., Sun, B., Ding, L., Li, J., He, X. and Shen, C. (2020), Hybrid Machine learning hydrological model for flood forecast purpose. *Open Geosciences*, *12*(1), 813-820. https://doi.org/10.1515/geo-2020-0166

Kumar, P. S., Praveen T.V. and Prasad M. A. (2016) Artificial Neural Network Model for Rainfall-Runoff -A Case Study *International Journal of Hybrid Information Technology* Vol.9, No.3, pp. 263-272 http://dx.doi.org/10.14257/ijhit.2016.9.3.24

Li, B., Yang, G., Wan, R., Dai, X., & Zhang, Y. (2016). Comparison of random forests and other statistical methods for the prediction of Lake water level: a case study of the Poyang Lake in China. *Hydrology Research* (2016) 47 (S1): 69–83. https://doi.org/10.2166/nh.2016.264

Liong S. & Chandrasekaran S. (2007) Flood Stage Forecasting With Support Vector Machines JAWRA Journal of the American Water Resources Association 38(1):173 – 186.
DOI: 10.1111/j.17521688.2002.tb01544.x

Maity, R., & Bhagwat, P. (2012). Multistep ahead River Flow Prediction using Support Vector Regression for non-linear function at Daily scale. *Journal of Water Resources and Protection*. 4 528-539. http://dx.coi.org/10.4236/jwarp.2012.47062

Miller, M. P., Carlisle, D. M., Wolock, D. M. and Wieczorek, M. (2018) A Database of Natural Monthly Streamflow Estimates from 1950 to 2015 for the Conterminous United States," J. Am. Water Resour. Assoc., vol. 54(6),1258–1269, doi: 10.1111/1752-1688.12685.

Mohammad A. G., Rahman K., Arun G., Mohammad H. F., Atefeh A. (2016). Modelling River discharge time series using support vector machine and artificial neural Networks. Environ Earth Sci. 75:685 DOI 10.1007/s12665-016-5435-6

Qiu I., Chen, S. Y., Nie, X. T., (1998) A forecast model of Fuzzy recognition neural network and its application. Advances in water science 9 (3), 258 – 264.

Renaud J., & Robert L., (2022). Short-Term Hydrological Forecast Using Artificial Neural Network Models with Different Combinations and Spatial Representations of

Hydrometeorological Inputs, *Water*, *14*(4), 552.

https://doi.org/10.3390/w14040552

Sha, J., Li, X. and Wang, Z. L. (2019). Comparison of daily streamflow forecasts using extreme Learning machines and the random forest method. Hydrol. Sci. J., 64(15), 1857–1866,

Doi 10.1080/02626667.2019.1680846.

Sruthi E. R. (2021) Random Forest Algorithms comprehensive guide with examples. Analythic Vidhya. www.analyticsvidhya.com

Tanty, R., & Desmukh T. S. (2015) Application of Artificial Neural Network in Hydrology- A Review International Journal of Engineering Research & Technology (IJERT), 4(6). ISSN:2278-0181

Tongal, H. and Booij, M. J. (2018) Simulation and forecasting of stream flows using machine Learning models coupled with base flow separation, J. Hydrol., 564, 266–282.

doi: 10.1016/j.jhydrol.2018.07.004.

Vapnik VN (1998) Statistical learning theory. Wiley, New York.