Techno-computing Journal 1(2) 1-17 (2025)
https://doi.org/10.71170/tecoj.2025.1.2.1-17

Research Paper

Ensemble Machine Learning Technique Based on Gaussian Algorithm for Stream Flow Modelling

Ismail I. Aminu [1*], U. J Muhammad [2], Ismail A. Mahmoud [3], Usman U. Aliyu [4], A. Abdullahi [5], Mahmud M. Jibril [6], Salim I. Malami[1,6]

[1] *School of Energy, Geoscience, Infrastructure and Society, Institute of Sustainable Built Environment, Heriot-Watt University, Edinburgh, United Kingdom*
[2] *Department of Civil Engineering, Bayero University Kano, Kano, Nigeria*
[3] *Faculty of Science, Department of Physics, Northwest University, Kano, Nigeria*
[4] *Department of Civil Engineering, Federal University Dutsin-ma, Katsina, Nigeria*
[5] *Department of Planning and Engineering Services, Federal Roads Maintenance Agency, Abuja, Nigeria*
[6] *Department of Civil Engineering, Kano University of Science and Technology, Kano, Nigeria*
**Corresponding Author Email:* ii2009@hw.ac.uk

Abstract

Streamflow modelling is regarded as a crucial part of managing and planning water resources. Water resources engineers face a variety of challenges when predicting streamflow. These difficulties are caused by complex natural processes that involve non-linearity, non-stationarity, and randomness. This research investigates the application of machine learning (ML)-based models for forecasting streamflow discharge (Q) using input variables, including temperature and rainfall. The study attains the essential stationarity required for precise modelling using Augmented Dickey-Fuller tests, data normalization, and transformation. In contrast, unit root tests identify initial-level non-stationarity and call for first-differencing. Correlation matrix analysis identifies relevant input combinations. The result findings are supported by statistical metrics including mean squared error (MSE), mean absolute error (MAE), and Pearson correlation coefficient (PCC). Notably, in terms of prediction accuracy, the Gaussian Process Regression (GPR) GPR-M3 model stands out as a notable performer, with a low MAE value of 0.034 in the calibration phase and 0.027 in the verification phase. The success of these techniques is further supported by first and second-order ensemble algorithms, with some models reaching a perfect PCC score during both the calibration and verification phases. The study emphasizes the significance of preprocessing, model selection, and ensemble procedures in improving the accuracy of streamflow prediction models. In addressing complex nonlinear interactions, artificial intelligence (AI)- based models are valuable tools for both technical applications and practical understanding.

Keywords: Streamflow Prediction, Artificial Intelligence, Machine Learning, Gaussian Model, Kano State

1. Introduction

Reliable streamflow forecast plays a vital role in water resource planning and management. It provides critical information that can be used to avoid natural disasters, such as floods and droughts, and offers valuable insights for proper water allocation in a regime. Research has shown that the evolution of the hydrological cycle system is primarily influenced by the combined effects of weather, climate, ocean, and the underlying surface [1]. The complex nature of the water cycle results in streamflow with natural

(Received 16 Nov 2024; revised 04 Dec 2024; accepted 30 Dec 2024; first published online 23 July 2025); Journal homepage https://technocomputing.org/index.php/tecoj

features, including spatial and temporal distribution, as well as alternating periods of dryness and wetness. Moreover, scientific studies have demonstrated the complex nature of streamflow, which results from the interplay between natural variables such as non-linearity, randomness, and non-stationarity [2].

Despite all the challenges, researchers in the field of hydrology continue to conduct their studies to achieve a certain level of accuracy in streamflow prediction under various environmental conditions (Figure 1). The dependence on streamflow prediction by water resources planners and decision-makers has continued to inform the design of sustainable hydraulic structures, improve effective monitoring systems, and optimize operations for the flood control system [3]. However, the literature shows that streamflow is modeled using two approaches: physical models, which employ partial differential equations, and data-driven models, which utilize artificial intelligence. It involves physical data collection, which tells information about the behavior of the river, while a data-driven model uses a regression method. The streamflow prediction by the conventional regression cannot be accurately achieved due to the ingrained non-linear interrelationship between input and output variables, which makes it a continuous scientific problem [3]. Concerning this, it has led to the development of AI models, as they are more reliable and non-linear tools for hydrological analysis.

1.1 Research Background

Studies on streamflow simulation have shown that classical regression tools are commonly used, despite being associated with a low degree of accuracy. This has led to the development of AI models that are accurately considered, as well as non-linear hydrologic tools. [2]. Studies have revealed that early AI models primarily focused on regression models, such as autoregressive moving average (ARMA), and linear regression models [1]. However, due to non-linearity behavior of streamflow, the mentioned regression models do not provide reliable forecasting accuracy, because of this reason numerous AI models were developed such as adaptive neuro-fuzzy inference system (ANFIS), artificial neural network (ANN), complementary wavelet-AI model, hybrid evolutionary computing model, and support vector machine (SVM) model [2]. Random forest (RF) and extreme learning machine (ELM) models have indicated the potential of streamflow prediction [1]. Moreover, the models have their shortcomings, which include time-consuming and non-automated modeling processes. Due to these reasons, it is necessary to conduct research that will enhance streamflow prediction accuracy. This could be achieved by augmenting and integrating two or more models.

The objective of this study is to employ four different machine learning models; ANN, GPR, SVM, and stepwise regression (SWR) to predict streamflow. The results of the models will be compared through three performance evaluations. While approaches such as ANN, SVM, GPR, and SWR have revolutionized streamflow forecasting, significant gaps remain, including a lack of comparative research on combining different methods into ensemble models to utilize their complementary capabilities and quantify uncertainty bounds. Furthermore, current studies are limited to specific regions, necessitating research into the transferability and generalizability of ensemble approaches across diverse hydrological conditions. Although promising, ensemble approaches often lack interpretability; therefore, enhancing clarity could increase confidence and acceptance. Finally, despite their potential, the operational deployment of ensemble forecasting into real-time decision support systems is restricted, emphasizing the need for frameworks that enable smooth integration and updating. In conclusion, deficiencies in comparative ensemble model assessment, uncertainty estimates, model generalization, interpretability, and real-time monitoring necessitate further study to fully realize the benefits of ensemble methodologies for streamflow forecasting and water management.

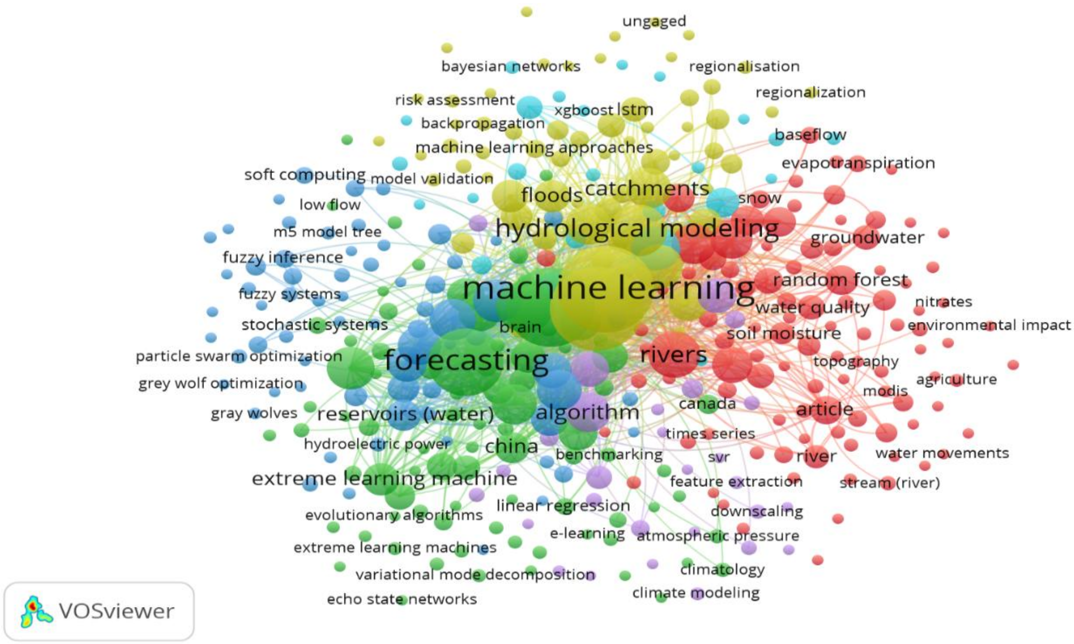

Figure 1: *The terminology most frequently used in the literature on streamflow prediction employing machine learning models.*

2. Literature Review

An essential component of ensuring sustainable management and use of water resources is the accurate runoff forecasts. In situations where it is impossible to get the underlying physical link directly, artificial intelligence techniques can open new possibilities for runoff prediction. The estimation of streamflow has been improved in numerous ways due to the advancement of AI, making the process easier and more accurate. The AI has improved the analysis of streamflow by automating data processing, developing accurate predictive models, integrating diverse data sources, and enabling real-time monitoring and decision support. These advancements have made streamflow estimation more efficient, reliable, and accessible, thereby assisting in water resource management, flood forecasting, and ecosystem monitoring. Nevertheless, there are limited reports that have assessed the performance of different AI techniques in predicting daily time series of streamflow for sustainable water resource management. Below is a glance at the study that forecasts various aspects of water resources.

To forecast the daily and monthly streamflow, Cheng et al. [4] employed two AI models, namely Long Short-Term Memory (LSTM) & ANN, to predict the streamflow at Nan River Basin, Thailand, within the 1974 to 2014 period. The models were trained and validated using rainfall-runoff datasets collected at the study area. In a China case study, Niu and Feng, [5] employed five AI-based models (ANN, ELM, GPR, SVM, and ANFIS) to predict the daily streamflow. Four statistical indices were used to anticipate performance accuracy during the study for the performance evaluation metric. The research indicates that the AI-based model achieves satisfactory forecasting outcomes in the study area. In conclusion, the outcomes depict that three of the AI-based models outperform the other two models in terms of accuracy. In another case study of Algeria, Aichouri et al. [6] employed an ANN to predict the relationship between rainfall-runoff in semiarid and Mediterranean climates. The model proves to be a promising method for predicting flow to the study region. The research revealed that the model is a sufficient tool for anticipating rainfall-runoff relations.

In another study, downstream of Agra City. Abba et al. [7] employed three AI models (MLR, ANN, and ANFIS) to anticipate water resource problems, which enabled the calculation of oxygen concentration dissolution upstream, in the middle stream, and downstream. Afterwards, the study indicates that the ANN model is marginally better than ANFIS in use and has shown a significant advantage over the MLR model, in a case study of the Palla station of the Yamuna River, India. Gaya et al. [8] employed several AI models to estimate the water quality index, presenting the MLR model as a suitable approach for determining water quality. This study found that other classical models are the most reliable for forecasting, based on the research results, which indicates a high improvement in the accuracy of ANN and ANFIS models over MLR when evaluating the accuracy and goodness of fit (GOF) of the models using MSE, RMSE, and the Coefficient of Determination (R-squared). The study shows that the precision between the two models is negligible, indicating that both models are reliable for the case study estimation, with minimal impact on accuracy performance.

However, Ali and Shahbaz [9] employed an ANN to determine the actual prediction of streamflow, which aids in water resources planning and management, including irrigation systems, hydropower plants, flood hazards, and dam control. The accuracy of the model was evaluated using four different performance metrics: RMSE, R-squared, correlation coefficient (R), and Nash–Sutcliffe efficiency (NSE). However, the study proves that ANN is an effective tool for solving hydrological problems. In Albert River, Queensland, Australia, Yaseen et al. [10] utilized an ENN to detect hourly river flows for flood forecasting and risk management of adverse events. The ENN model displayed outstanding performance in terms of accuracy, unlike the other model. The results clearly describe the application of the ENN model as a promising AI technique for accurate performance in real-time. In the United States study. Parisouj et al. [11] Employed three AI-Models (Artificial neural network with backpropagation (ANN-BP), support vector regression (SVR), and ELM) to estimate the significant role of streamflow in water resources. The feature selection method of recursive feature elimination (RFE) for support vector machines (SVM) was employed to select the most suitable predictor variable. The performance of the developed model was measured using selected statistics. The findings reveal that the SVR model yielded better results than the ANN-BP and ELM at the monthly and daily scales for streamflow simulation.

3. Materials and methods

3.1 Case Study and Data Description

Kano State has a large, continuous area of land, i.e., a total of 20,131 km^2, positioned in the northern region of Nigeria. It is the most populous city in Nigeria, with a population density of 470 people per square kilometer, a total population exceeding 9,383,682, and a growth rate of 2.9% per annum. The state borders to the Northeast, Northwest, Southwest, and Southeast with Jigawa, Katsina, Kaduna, and Bauchi states, respectively [2]. The state is positioned at the GPS coordinates of 120°0’0.0000” N and 80°31’0.0012” E, equivalent to a latitude of 12.000000 and a longitude of 80.516667. The study map area is shown on (Figure 2). Kano state's temperature mainly varies between a low of 15.80 °C and a high of 330 °C. The state has a long dry season and an average rainy season of 4 to 5 months. The metropolis of the state has a mean annual rainfall of almost 800mm. The primary reservoirs in the state are in the watershed of the Challawa River, Jakara River, and Watari River. The runoff from the city is studied and determined for these three reservoirs.

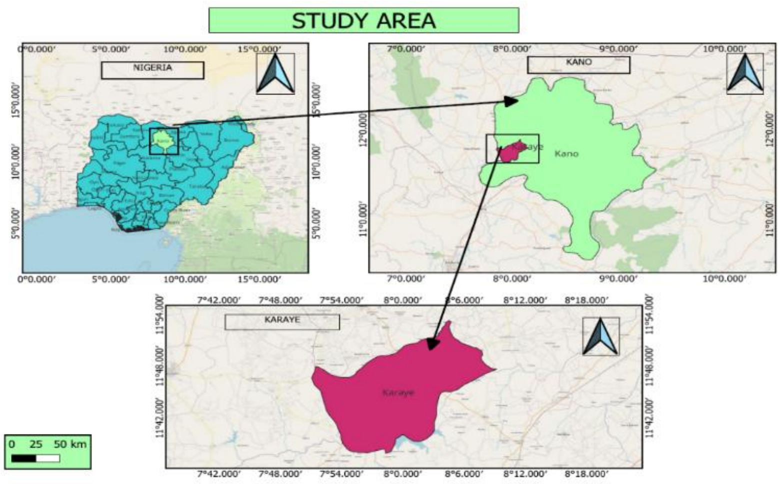

Figure 2: *Study Map and Locations*

3.2 Methodologies

Streamflow forecasting plays a critical role in hydrology and water resources management, as it supports decision-making processes related to flood control, reservoir operations, irrigation planning, and sustainable water allocation. Accurate and timely predictions are crucial for mitigating the adverse impacts of hydrological extremes and ensuring the optimal utilization of water resources. In this study, four machine learning algorithms; ANN, SVM, GPR, and SWR, viz: were employed to model and forecast streamflow discharge based on meteorological input variables, including rainfall and temperature (Figure 3). These models were selected due to their ability to capture nonlinear and complex relationships in hydrological systems. Despite the potential of these models, several methodological challenges must be addressed to ensure prediction accuracy. One major issue is the quality and availability of historical data, which directly influences model performance; missing values, noise, and inconsistent records can significantly distort model training and outputs. Another critical challenge is the non-stationarity of streamflow data, which often exhibits temporal fluctuations due to changing climatic and land-use conditions. This variability complicates the model's ability to generalize over time. Proper data preprocessing, including normalization, outlier removal, and feature engineering, is crucial for enhancing data quality and model robustness. Additionally, model selection plays a critical role; choosing the most suitable algorithm requires a balance between accuracy, interpretability, and computational efficiency. Equally important is model calibration, which involves tuning hyperparameters to optimize performance and prevent overfitting. To address these challenges, this study implemented a systematic modeling framework involving training, validation, and testing phases, using statistical evaluation metrics such as the PCC, MAE, and MSE to assess model performance. Cross-validation techniques and visual tools, such as response curves and scatter plots, were also employed to ensure reliability and interpretability. Furthermore, ensemble learning techniques were explored to assess their potential in enhancing prediction performance by combining the strengths of individual models. The adopted methodology provides a comprehensive foundation for developing data-driven streamflow forecasting tools with enhanced accuracy and reliability under varying hydrological conditions.

3.3 Data Processing

The data is scaled during normalization to a point where it is dimensionless, meaning all characteristics or variables are on the same scale. This is significant because some algorithms may experience numerical instability when features have various units or orders of magnitude. You prevent any characteristic from having an undue influence on the model by normalizing the data. Before implementing the AI-based models, normalization is a crucial step. This normalization aims to reduce data redundancy and enhance its integrity. Proceeding with the pre-processing procedure has two main reasons [12]. According to [13], this pre-processing step is vital, as it ensures that the variables receive equal attention during training, thereby enhancing the algorithm's efficiency. For this research, the min-max scaling normalization method was employed.

3.4 Model Development

Researchers are typically interested in determining predicted output from inputs based on historical data in hydrological forecasting models. The objective in predicting streamflow using predecessor values is to generalize a relationship of the following form [14]. The selection of a suitable model input vector plays a crucial role in the practical application of AI techniques, as it is typical of any data-driven forecasting model, providing the fundamental details about the system being modeled. The number of flow values identified as input variables is used to determine the runoff lags that significantly affect the expected flow. The variables used in the research include discharge (Q) as the output variable, temperature (Temp) as the input variable, and rainfall (R) as the input variable. The correct model input must be chosen when using an AI model for forecasting streamflow to achieve a positive result. Moreover, this employed four models for streamflow prediction, namely ANN, SVM, SWR, and GPR. A feed-forward ANN model of a typical three-layer ANN was built for the determination of predicted monthly discharge time series. The conventional BP training algorithm is a supervised training mechanism commonly used in most engineering applications [14]. The training epoch in this paper is set to 500, and the training algorithm is a scaled conjugate gradient algorithm. The neurons used a Tan-sigmoid transfer function in the hidden layer, while the linear transfer function determines the output layer. The determination of the ideal number of neurons involves a trial-and-error method in the hidden layer, where the number of hidden neurons is altered from 2 to 13. The training and testing sets are further divided to form training data. The hidden neurons were decided upon using the cross-validation approach and the RMSE. This shows that the performance of the feed-forward model is not significantly impacted by the number assigned to the hidden layer. When three hidden neurons are present, the training error is the closest to the testing error.

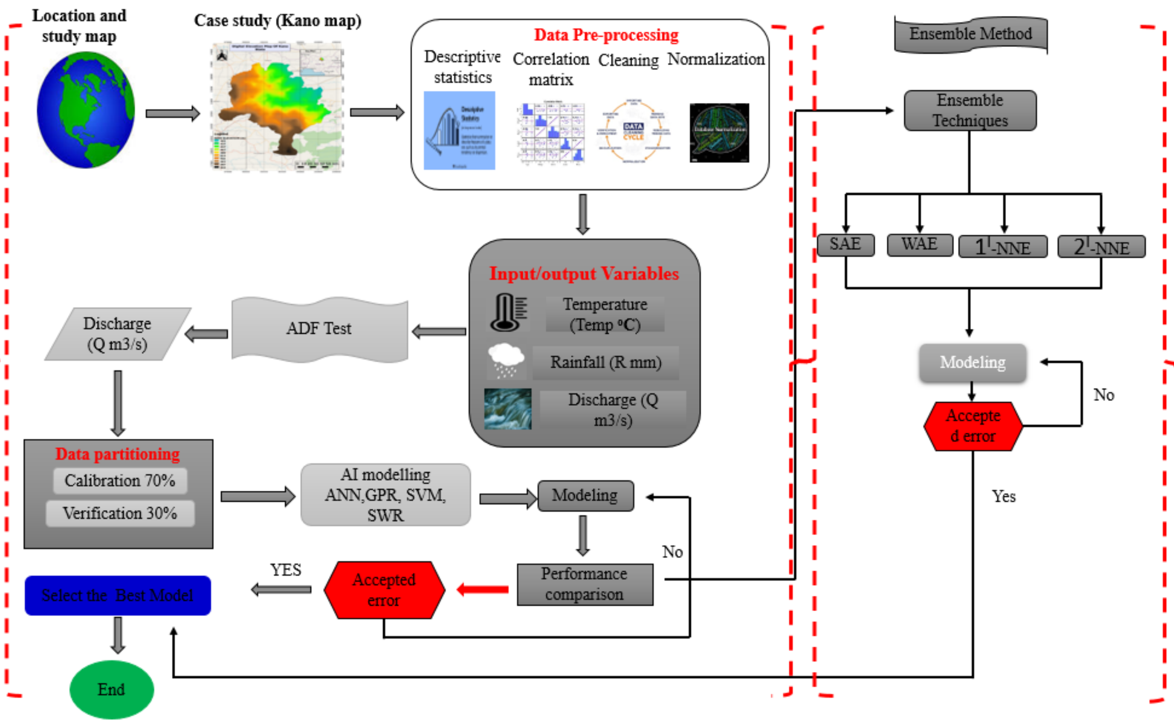

***Figure 3:** The proposed methodology flowchart*

The link between input and output, as explained by the most excellent computer program, can be created using GPR. The correct GPR evolution parameters must be used in this research to develop the best monthly streamflow time series forecasting model. Although the algorithm's fine-tuning is the focus of this research, several startup and run approaches were identified, and the chosen GP parameters are reported. the evolutionary processes, which include setting the genetic operators, fitness function, including crossover, reproduction, and mutation, as well as termination, are comparable to GAs. When utilizing SVM, a Kernel function must be chosen from the qualifying functions. It was stated by [15] that various Kernels in SVR for modeling rainfall-runoff, showing that the radial basis function (RBF) performs better than other Kernel functions. The selection of model parameters has a significant impact on SVM models. There aren't any organized techniques for choosing parameters, though. As a result, a calibration of the model parameters was performed for this work.

3.5 Theory of Models

3.5.1 Artificial Neural Network (ANN)

ANNs are admired by biological neural networks; however, ANNs come along with neurons, linking the process of information to search for a relationship between input and output variables. ANN is among the most efficient tools for data-driven methods to model a nonlinear system [16]. ANN model as a nonlinear mapping for literature and observation to estimate the fourth coming values [17].

3.5.2 Gaussian Process Regression (GPR)

Gaussian process regression (GPR) is a non-parametric machine learning method that can be applied for modeling and multivariate regression problems. Compared to other techniques, such as SVR and random forests, GPR often demonstrates superior performance. Its Gaussian process framework enables the approximation of personalized probability distributions required for robust and flexible regression models. This makes GPR a valuable technique for solving many complex engineering challenges [18].

3.5.3 Support Vector Machine (SVM)

SVMs became paramount in the most widely used statistical learning performance familiarized in 1995 [19]. The SVM's ability to form accurate generalizations makes it a valuable tool for applications in various scientific fields. For those less disposed to over-fitting of data. While SVMs allow for concurrent error minimization, the use of a kernel function makes the original inputs separable in a plotted high-dimensional feature dimensional [19]

3.5.4 Stepwise Regression (SWR)

SWR serves as the appropriate method for selecting both the combination of backward and forward techniques for problem-solving. SWR became one of the most popular models at a unique stretch; it became an alteration of both regressive and headlong ranges. So that after a variable is added to each contender, the candidate variables in the model must be checked to see if their significance is reduced below the actual tolerance level. having to find mutable, which is non-significant, it must be removed from the model [20].

3.5.5 Ensemble Learning Technique (ELT)

Ensemble models are a set of models that are used to provide a prediction that is often more accurate and robust than any individual model by combining the predictions of numerous base models (also known as weak learners). Ensemble models are a set of learning classifiers that combine their decisions to obtain more accurate and reliable predictions in supervised and unsupervised learning problems [21]. It is known that combining two or more predictors can enhance prediction performance for a time series [22]. The literature indicates that ensembling the outputs of two or more models is a more effective method, which can enhance the prediction efficiency of time series (Figure 4).

Technique 1: Simple averaging ensemble (SAE)

The principle behind SAE is to take the average of the forecasts from these base models to arrive at a final prediction. SAE often relies on a few base models, which can be various algorithms or variations of a single method with different hyperparameters. Decision trees, support vector machines, random forests, and neural networks are examples of base models. SAE can be applied to problems involving both regression and classification. In this paper, all four models — ANN, GPR, SVM, and SWR — were tested and trained independently, and their average outputs were compared and evaluated against the observed test values. See the formula as shown below (Eq.1).

$$f(t) = \frac{1}{N}\sum f(ti) \tag{1}$$

N (number of models), ti = Output of single models (ANN, GPR, SVR, and SWR), f(t)= time

Technique 2: Weighted average ensemble (WAE).

This prediction is achieved by assigning different weights to the various outputs based on the level of significance of each production. When creating the final prediction in a weighted averaging ensemble model, the prediction from each base model is multiplied by its weight first. The aim is to assign higher weights to more reliable or accurate models and lower weights to those that are less reliable. WAE is used in improving predictive accuracy and robustness. WAE is applicable in classification, regression, and anomaly detection. The weighted average ensemble is expressed as in (Eq.2).

$$f(t) = \sum_{I=1}^{N} Uifi(t) \quad (2)$$

Where Ui is obtained from a formula below, Ui= Weight given on the output of the i_{th} model, DZi= Performance efficiency of the i_{th} single model (Eq. 3).

$$Ui = Dzi / \sum_{i=1}^{n} Dzi \quad (3)$$

Technique 3: Non-linear neural ensemble (NNE)

NNE model is another techniques using neural network models that are coupled or integrated non-linearly to boost a machine learning system's overall performance. This ensemble strategy is frequently employed in deep learning and machine learning to enhance the robustness and predictive capabilities of models. For the sake of a non-neural ensemble technique, non-linear averaging is done by training a separate neural network. The selected model output is used to feed the input of the neural ensemble model; one neuron in the input layer is assigned to all the selected models. In the case of the FFNN ensemble model, the activation function is tangent sigmoid for both hidden and output layers. The network can easily be trained through the application of the BP algorithm. Moreover, using a trial-and-error procedure, the best structure and epoch number of the ensemble network can be obtained. For the sake of this study, the non-linear ensemble used is FFNN because it's a standard AI method.

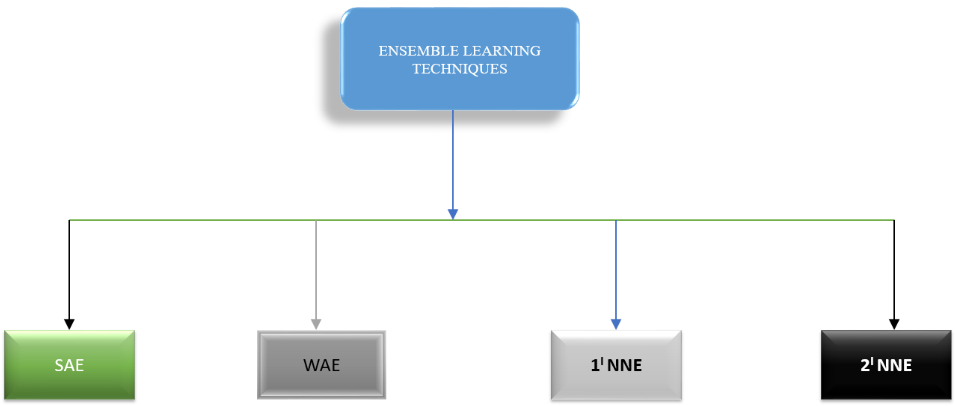

***Figure 4:** Flowchart for ELT*

3.6 Evaluation Performance

To evaluate the accuracy and reliability of the developed streamflow prediction models, three statistical performance metrics were employed. These metrics are widely recognized for assessing regression models and provide comprehensive insight into both the magnitude of prediction errors and the strength of the relationship between predicted and observed values. MSE measures the average of the squared differences between predicted and observed values, giving more weight to larger errors. This makes it especially useful in detecting models that produce occasional large deviations. A lower MSE value indicates that the model predictions are closer to the actual values, signifying higher accuracy. MAE calculates the average absolute difference between predicted and observed values. Unlike MSE, it treats all errors equally, offering a straightforward interpretation of the average size of the prediction errors. A lower MAE value reflects better model performance with fewer deviations from the actual observations. PCC assesses the strength and direction of the linear relationship between predicted and observed values. A PCC value closer to 1 indicates a strong positive correlation, meaning the model successfully captures the overall pattern and trend in the data. A high PCC confirms the model's ability to mirror the variability in streamflow behavior [23].

4.0 Results and Discussion

4.1. Result Application and Analysis

Using the discharge (Q) data collected at Kano, ANN, GPR, SVM, and SWR models were constructed for comparison purposes. The following sections present the results and applications of Artificial Intelligence and regression models in streamflow prediction. The study's variables include temperature (TEMP) and rainfall (R) as input variables, while discharge (Q) is the output variable. Before modeling, both the input and target were normalized. As part of the preprocessing stage, the Unit Root Test-based Augmented Dickey-Fuller (ADF) test in E-Views was used to transform the data from non-stationary to stationary. The descriptive statistics of the datasets and critical data used in model development are shown in Table 1 [24], [25]. The most frequent and efficient input combinations with the target variable were examined using a correlation matrix in a conservative sensitivity analysis, as illustrated in Fig. 5. The matrix identifies the specific type of linear relationship between the variables. It indicates the primary indication for probable correlation between variable sets. Positive correlation values demonstrate direct associations between two variables, whereas a probability of less than 0.05 indicates a significant and stationary relationship.

Table 1: *Statistical relationship between the input and output parameters.*

Parameters	Temp (°C)	R (mm)	Q (m³/s)
Mean	0.00	-2E-16	0.01
Median	0.00	0.00	0.00
Standard Deviation	1.78	67.18	19.27
Kurtosis	0.62	9.00	13.79
Skewness	0.31	0.88	-0.35
Range	11.60	744.20	247.80
Minimum	-5.00	-321.2	-140.8
Maximum	6.60	423.00	107.00

4.2. Preliminary Results

It is crucial to guarantee the stability and consistency of the dataset to precisely execute the stochastic process and time series analysis when creating a model. Testing for unit roots and ensuring stationarity of all HMs variables was done using the Augmented Dickey-Fuller (ADF) analysis (see Table 2). This was done to get more trustworthy and accurate findings. [26] Both computational intelligence and numerical analysis depend heavily on stability analysis. Recently, feature selection methods have used stability analysis, and the current study used linear feature selection methods.

Table 2: *Input-output variable ADF tests*

Variables	t-Statistic	5% Critical Value	Prob	Decision
TEMP (OC)	-2.00	-2.87	0.28	I (0)
	-17.56	-2.87	0.00	I (1)
R (mm)	-2.37	-2.87	0.15	I (0)
	-14.79	-2.87	0.00	I (1)
Q (m³/s)	-2.35	-2.87	0.1562	I (0)

	-16.37	-2.87	0.00	I (1)

However, the results of the unit root test, presented in Table 2 (ADF), demonstrate that all variables are non-stationary at level I (0). The procedure uses a least squares exogenous approach to create the Schwarz information criteria. Due to its link with Q (m3/s), the findings of all variables reached the stationarity level at the first difference, which was unexpected. The fact that the probability and t-statistics values are above the critical values for all degrees of freedom lends credibility to the findings. According to [27], when Cronbach's alpha values surpass 0.7, a stationary test is deemed valid. It is clear from Table 2 that the means and variances of none of the variables are constant; to ensure stationarity, we must create the first difference. It is significant to highlight that ADF empowers analysts to select suitable modeling approaches and get more precise forecasts of the series' future values. Although various nonlinear viable alternatives may be utilized based on the data's nature, the combination of input variables for the ADF test was selected using standard correlation analysis. The outcomes of this combination are shown in Equation 8 as presented in Figure 5.

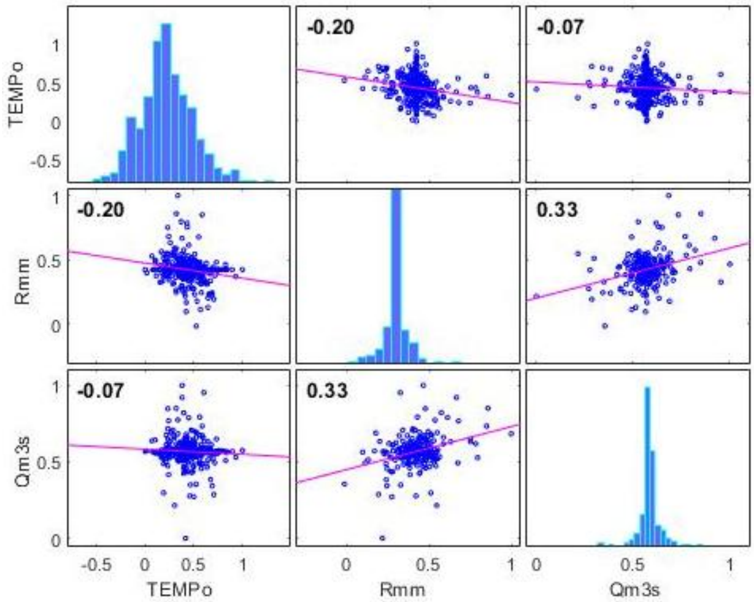

Figure 5: *Matrix of correlations for the parameters used to model discharge (Q m3/s)*

4.3 Analysis of Predictive Models

The practical assessment analysis results for prediction models are summarized in Table 3. While considering errors, the statistical indices (PCC, MSE, and MAE) are used to assess the models' capacity to predict outcomes and their effectiveness in making estimates. According to Table 3, almost all the combinations satisfy the model's accuracy level in terms of the statistical requirements (M1, M2, and M3). Based on the results, it is acknowledged that these approaches can handle models with several unrestrained parameters, minimize the error function, and resolve data fitting issues. For highly complex nonlinear problems, they have proven to be an average solution. More than 50 per cent of the models met the statistical benchmarks for accuracy (MAE value less than 0.05). The obtained SVM, GPR, and SWR model combinations meet the criteria of MAE values less than 0.05. The ANN model's ANN-M1 and ANN-M2 both met the necessary standards. The GPR-M3, SVM-M2, ANN-M2, and SWR-M1 models are the most effective models for predicting streamflow discharge (Q, m^3/s), with MAE values of 0.034, 0.05, 0.037, and 0.037, respectively, in the training phase. The results indicate that the GPR-M3 model outperforms the other models, with a minimum MAE Value of 0.034 & 0.027 in both the calibration and verification phases. The results suggest that these approaches help minimize the error function by managing models with multiple uncontrolled parameters and addressing issues with data fitting. They have developed into a common strategy for extremely complex nonlinear situations.

Table 3: *Result of the performance criteria in the first scenario*

Models	**Calibration phase**			**Verification phase**		
	PCC	MSE	MAE	PCC	MSE	MAE
ANN-M1	0.16	0.01	0.04	0.20	0.01	0.03
ANN-M2	**0.34**	**0.01**	**0.04**	**0.63**	**0.01**	**0.03**
ANN-M3	0.24	0.01	0.05	0.40	0.01	0.04
GPR-M1	0.01	0.01	0.04	0.56	0.01	0.03
GPR-M2	0.40	0.00	0.04	0.65	0.01	0.03
GPR-M3	**0.53**	**0.00**	**0.03**	**0.65**	**0.00**	**0.03**
SVM-M1	0.04	0.01	0.04	0.12	0.01	0.03
SVM-M2	**0.30**	**0.01**	**0.04**	**0.58**	**0.01**	**0.03**
SVM-M3	0.30	0.01	0.04	0.59	0.01	0.03
SWR-M1	0.01	0.01	0.04	0.59	0.01	0.03
SWR-M2	**0.30**	**0.01**	**0.04**	**0.58**	**0.01**	**0.03**
SWR-M3	0.30	0.01	0.04	0.58	0.01	0.03

Furthermore, as shown in Figure 6, the error plot can be used to depict the maximum error calculated by each model, providing a comparison of all four models employed in this study. The bar chart illustrates the MAE for various machine learning models during both the calibration (red bars) and verification (blue bars) phases of streamflow prediction. Most models show lower MAE values in the verification phase, indicating good generalization and minimal overfitting. However, one model (second from the left) exhibits significantly higher MAE in both phases, indicating poor performance and potential data sensitivity issues. The models toward the right demonstrate consistent and lower MAE values across both phases, reflecting higher accuracy and robustness in streamflow forecasting.

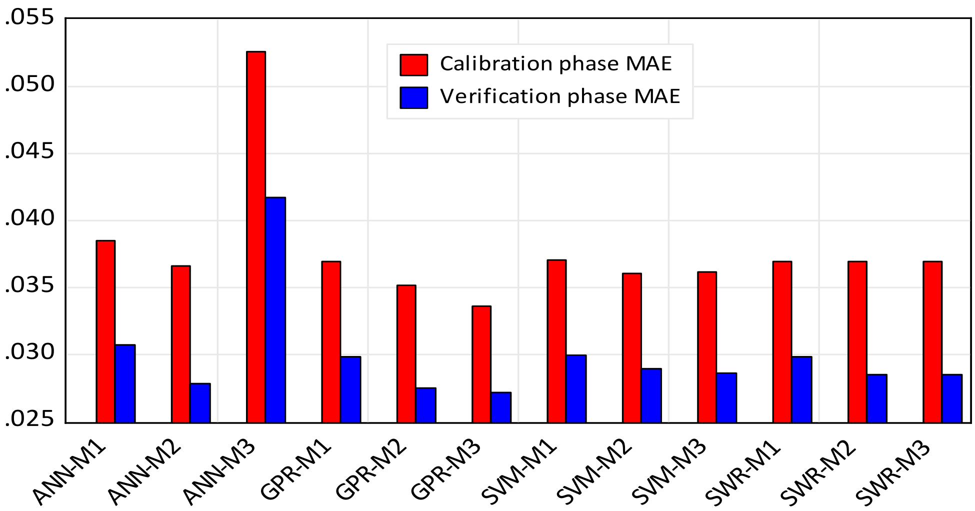

Figure 6*: Error plot for all four best models*

4.4 Result of the ELT model

The accuracy of standalone models was improved in the second scenario of this study by using an ensemble, which offered more accurate predictions by combining the strengths of various models. The SA and the NNE model strengths were associated with WAE. The maximum level of accuracy, as shown on Table 4 for MAE (100%), was attained for SA-GPR, WA-GPR, and 1^{I}–NN–GPR. This wasn't unanticipated given that 1^{I} – NN-GPR is adaptable to changes in environmental conditions or input data and can manage a wide range of input data types and formats. Figure 7 presents the performance of all the models, comparing the results of the four models used in this study. An error plot can be used to depict the maximum error calculated by each model, which is summed up in the first-order ensemble techniques. Thus, ensemble techniques are more adaptable and flexible in providing solutions to challenging issues related to discharge (Q).

Table 4: *First-order ensemble algorithm result*

Model	**Calibration Phase**			**Verification Phase**		
	PCC	MSE	MAE	PCC	MSE	MAE
SA-ANN	0.31	0.01	0.04	0.54	0.01	0.03
SA-GPR	**0.48**	**0.00**	**0.03**	**0.67**	**0.01**	**0.03**
SA-SVM	0.30	0.01	0.04	0.59	0.01	0.03
SA-SWR	0.30	0.01	0.04	0.58	0.01	0.03
WA-ANN	0.31	0.17	0.39	0.54	0.17	0.38
WA-GPR	**0.48**	**0.01**	**0.05**	**0.67**	**0.01**	**0.05**
WA-SVM	0.30	0.04	0.19	0.59	0.04	0.19
WA-SWR	0.30	0.05	0.21	0.58	0.05	0.21
1^{I} - NN-ANN	0.31	0.01	0.04	0.58	0.01	0.03
1^{I} - **NN-GPR**	**0.53**	**0.00**	**0.03**	**0.65**	**0.00**	**0.03**
1^{I} - NN-SVM	0.33	0.01	0.04	0.48	0.01	0.03
1^{I} - NN-SWR	0.31	0.01	0.04	0.58	0.01	0.03

As can be seen from Table 4, in SA techniques the SA-GPR outperforms all other SA techniques with minimum MAE of 0.034 & 0.028 in both calibration and verification phase, in addition looking how the WA performance is effective, it was also depicted that WAA-GPR outperforms all other WA techniques with minimum MAE of 0.052 & 0.049 in both calibration and verification phase. In first-order NNE, it was shown that the 1^{I} - NN-GPR surpasses other first-order NNE with MAE of 0.034 & 0.027 in both calibration and verification phases. In conclusion, due to the lower performance of PCC in both SA, WA, and NN ensemble techniques, Table 5 presents the second-order result of NN ensemble techniques As shown in Figure 6, which compares all four models used in this study, the error plot can be used to depict the maximum error calculated by each model in the first-order ensemble techniques.

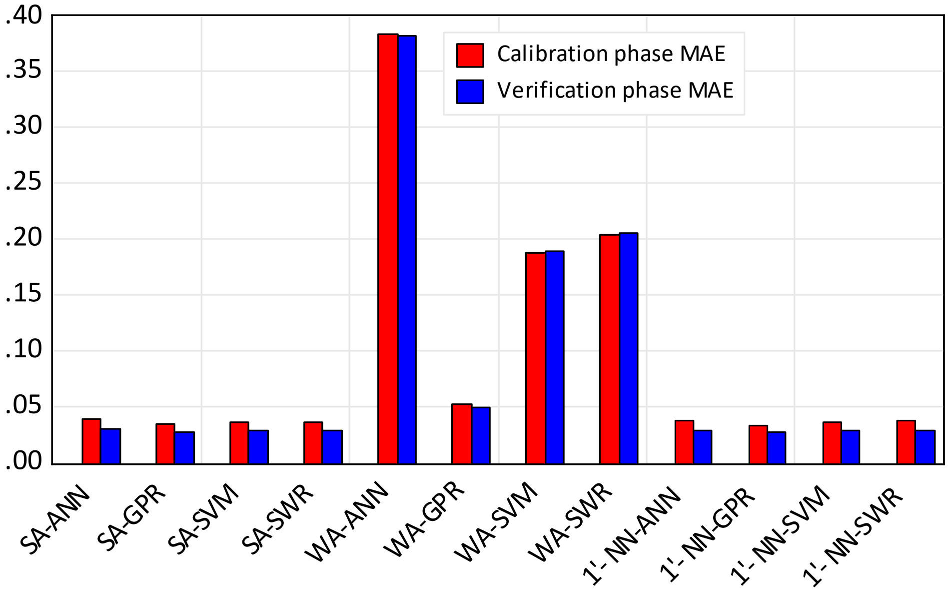

Figure 7: *Error plot for the first-order ensemble algorithm*

Table 5: *Result of the second other ensemble algorithm of NN*

Model	Calibration Phase			Verification Phase		
	PCC	MSE	MAE	PCC	MSE	MAE
2^{I} - NN-ANN	1.00	8.51E-20	2.48E-10	1.00	7.57E-20	2.23E-10
2^{I} - NN-GPR	1.00	8.51E-20	2.48E-10	1.00	7.57E-20	2.23E-10
2^{I} - NN-SVM	1.00	8.51E-20	2.48E-10	1.00	7.57E-20	2.23E-10
2^{I} - NN-SWR	1.00	8.51E-20	2.48E-10	1.00	7.57E-20	2.23E-10

As shown in Table 5, all models achieve the same level of accuracy as the second-order ensemble algorithm with a PCC of 1 in the calibration phase, nevertheless, except for 2^{I}-NN-GPR, which outperforms all other second-order ensembles with the highest value of PCC equal to 1 in both the calibration and verification phases. Scatter plots can identify trends, linkages, and patterns in data and are frequently used in data analysis and scientific research. The response curve shows a robust correlation between the experimental and predicted values (R = 1), as seen in Figure 8. The response plot can also demonstrate the contemporaneous agreement between the two variables (measured and expected), as shown in Figures 9.

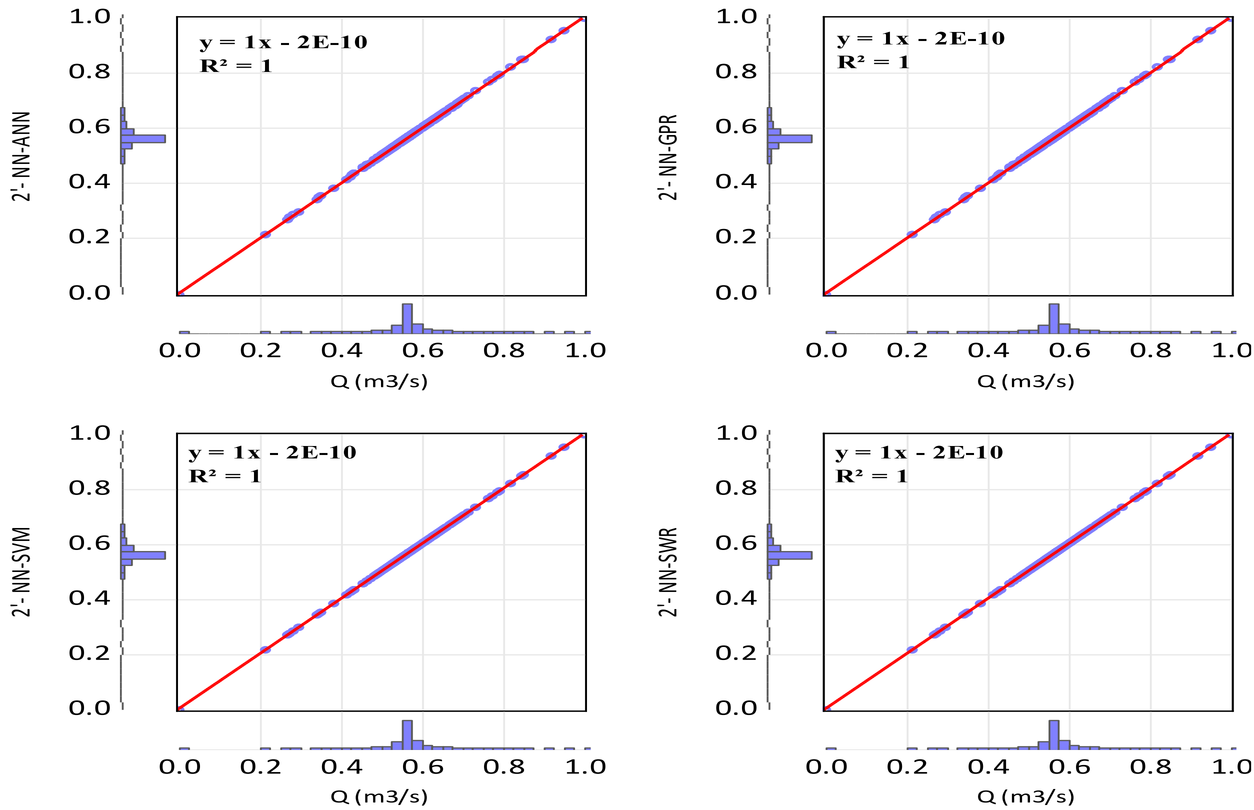

Figure 8: *Scatter plot of the second-order ensemble algorithm*

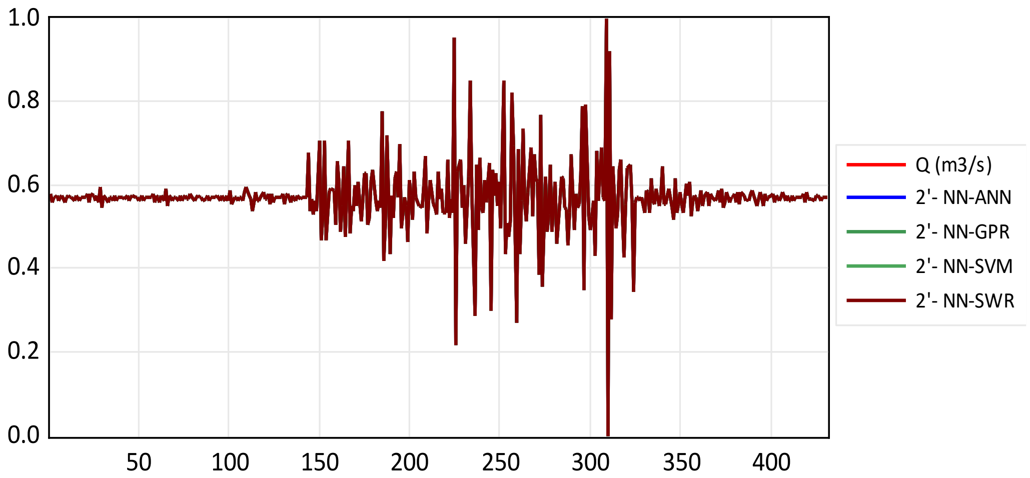

Figure 9: *Time series of the second-order ensemble algorithm*

Further examination, as shown in Figure 10, can be conducted to assess the expected accuracy of the models during the testing phase. The plot has shown that ANN-M2, SVM-M2, GPR-M3, and SWR-M2 demonstrated contemporaneous agreement with a similar pattern after accounting for the observed strength values. In the literature on civil and material engineering, time series are frequently utilized, for instance. However, [28]–[31] highlighted the necessity of understanding time series to grasp the precision of a data set. Using the radar diagram, all of the models were compared during the modeling phase using the PCC performance evaluation criterion. In Figure 10, the second-order ensemble model outperforms other models in terms of prediction accuracy. This article explains why modeling with AI is suitable for engineering and academic research.

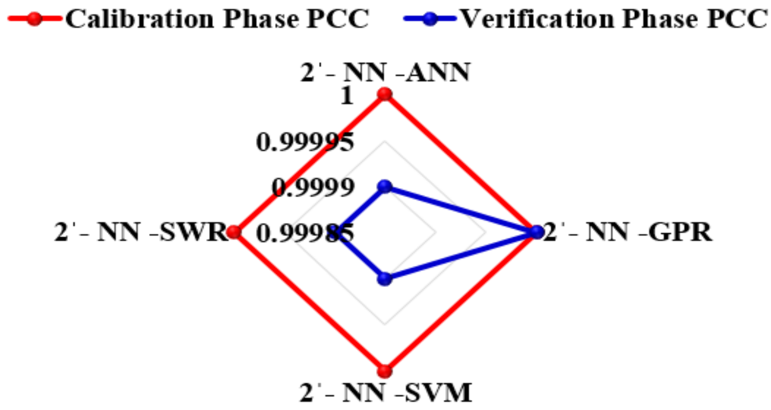

Figure 10: *Radar plots for second-order ensemble techniques.*

5.0 Conclusion

In this research, the primary focus was on the application and comparative analysis of multiple machine learning models, namely ANN, SWR, SVR, and GPR, for predicting streamflow discharge based on meteorological input variables such as temperature and rainfall. The results revealed that all selected models demonstrated acceptable performance, as supported by statistical evaluation metrics, including the PCC, MAE, and MSE, which underscored their predictive capability. Notably, the GPR-M3 model outperformed others, achieving the lowest MAE values of 0.034 in the calibration phase and 0.027 during verification, indicating its strong ability to capture the underlying non-linear patterns in streamflow data. Furthermore, the study showed that employing ensemble modeling techniques significantly enhanced the performance of individual models. Second-order ensemble approaches demonstrated remarkable accuracy, with some models achieving a perfect PCC of 1 in both the calibration and verification stages. Visual validation tools such as scatter plots, response curves, and radar charts further affirmed the robustness, precision, and consistency of the models. However, this study is limited by the scale and scope of the dataset used, which was derived from a specific geographic and climatic region, potentially affecting the generalizability of the models to other areas with different hydrological behaviors. Also, the models did not account for human interventions, such as dam operations or land-use changes, which could influence streamflow dynamics. Future studies should explore hybrid deep learning frameworks, integrate remote sensing data and catchment-specific physical parameters, and investigate model transferability across multiple catchments. It is also recommended to enhance real-time forecasting capabilities through data assimilation and to adopt uncertainty quantification techniques for improved decision-making in water resources management.

Competing Interests: The authors declare that they have no competing interests.

Data Availability Statement: The supported data associated with this researcher is available upon request from the corresponding author.

References

[1] N. Sun, S. Zhang, T. Peng, N. Zhang, J. Zhou, and H. Zhang, "Multi-Variables-Driven Model Based on Random Forest and Gaussian Process Regression for Monthly Streamflow Forecasting," Water (Switzerland), vol. 14, no. 11, 2022, doi: 10.3390/w14111828.

[2] S. J. Hadi, S. I. Abba, S. S. H. Sammen, S. Q. Salih, N. Al-ansari, and Z. M. Yaseen, "Non-Linear Input Variable Selection Approach Integrated With Non-Tuned Data Intelligence Model for Streamflow Pattern Simulation," IEEE Access, vol. 7, pp. 141533–141548, 2019, doi: 10.1109/ACCESS.2019.2943515.

[3] A. N. Ahmed, T. Van Lam, N. D. Hung, N. Van Thieu, O. Kisi, and A. El-Shafie, "A comprehensive comparison of recent developed meta-heuristic algorithms for streamflow time series forecasting problem," Appl. Soft Comput., vol. 105, p. 107282, 2021, doi: 10.1016/j.asoc.2021.107282.

[4] M. Cheng et al. Pre-proofs, "Long lead-time daily and monthly streamflow forecasting using machine learning methods," J. Hydrol., p. 125376, 2020, doi: 10.1016/j.jhydrol.2020.125376.

[5] W. Niu and Z. Feng, "Evaluating the performances of several artificial intelligence methods in forecasting daily streamflow time series for sustainable water resources management," Sustain. Cities Soc., vol. 64, no. September 2020, p. 102562, 2021, doi: 10.1016/j.scs.2020.102562.

[6] I. Aichouri, A. Hani, N. Bougherira, and L. Djabri, "River Flow Model Using Artificial Neural Networks," Energy Procedia, vol. 74, pp. 1007–1014, 2015, doi: 10.1016/j.egypro.2015.07.832.

[7] S. I. Abba, S. Jasim, and J. Abdullahi, "ScienceDirect ScienceDirect River water modelling prediction using multi-linear regression , artificial neural network , and adaptive neuro-fuzzy inference system techniques," Procedia Comput. Sci., vol. 120, pp. 75–82, 2018, doi: 10.1016/j.procs.2017.11.212.

[8] M. S. Gaya, S. I. Abba, A. M. Abdu, and A. I. Tukur, "Estimation of water quality index using artificial intelligence approaches and multi-linear regression," vol. 9, no. 1, pp. 126–134, 2020, doi: 10.11591/ijai.v9.i1.pp126-134.

[9] S. Ali and M. Shahbaz, "Streamflow forecasting by modeling the rainfall – streamflow relationship using artificial neural networks," Model. Earth Syst. Environ., no. 0123456789, 2020, doi: 10.1007/s40808-020-00780-3.

[10] Z. M. Yaseen, S. R. Naganna, and Z. Sa, "Hourly River Flow Forecasting : Application of Emotional Neural Network Versus Multiple Machine Learning Paradigms," 2020.

[11] P. Parisouj, H. Mohebzadeh, and T. Lee, "Employing Machine Learning Algorithms for Streamflow Prediction : A Case Study of Four River Basins with Different Climatic Zones in the United States," 2020.

[12] S. G. Meshram, C. Meshram, C. A. G. Santos, B. Benzougagh, and K. M. Khedher, "Streamflow Prediction Based on Artificial Intelligence Techniques," Iran. J. Sci. Technol. - Trans. Civ. Eng., vol. 46, no. 3, pp. 2393–2403, 2022, doi: 10.1007/s40996-021-00696-7.

[13] Ismail I Aminu, "A novel approach to predict Water Quality Index using machine learning models: A review of the methods employed and future possibilities," Glob. J. Eng. Technol. Adv., vol. 13, no. 2, pp. 026–037, 2022, doi: 10.30574/gjeta.2022.13.2.0184.

[14] W. C. Wang, K. W. Chau, C. T. Cheng, and L. Qiu, "A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series," J. Hydrol., vol. 374, no. 3–4, pp. 294–306, 2009, doi: 10.1016/j.jhydrol.2009.06.019.

[15] N. With, P. By, Y. B. Dibike, S. Velickov, D. Solomatine, and M. B. Abbott, "M i s v m : i a," no. July, pp. 208–216, 2001.

[16] V. K. Vidyarthi, A. Jain, and S. Chourasiya, "Modeling rainfall-runoff process using artificial neural network with emphasis on parameter sensitivity," Model. Earth Syst. Environ., vol. 6, no. 4, pp. 2177–2188, 2020, doi: 10.1007/s40808-020-00833-7.

[17] M. T. Vu, T. Aribarg, S. Supratid, S. V. Raghavan, and S. Y. Liong, "Statistical downscaling rainfall using artificial neural network: significantly wetter Bangkok?," Theor. Appl. Climatol., vol. 126, no. 3–4, pp. 453–467, 2016, doi: 10.1007/s00704-015-1580-1.

[18] O. P. Okeke et al., "Performance analysis and control of wastewater treatment plant using Adaptive Neuro-Fuzzy Inference System (ANFIS) and Multi-Linear Regression (MLR) techniques," pp. 1–16, 2022.

[19] S. Shabani, P. Yousefi, and G. Naser, "Support vector machines in urban water demand forecasting using phase space reconstruction," Procedia Eng., vol. 186, pp. 537–543, 2017, doi: 10.1016/j.proeng.2017.03.267.

[20] S. Kumar, A. Pandey, and B. Yadav, "Assessing the applicability of TMPA-3B42V7 precipitation dataset in wavelet-support vector machine approach for suspended sediment load prediction," J. Hydrol., vol. 550, pp. 103–117, 2017, doi: 10.1016/j.jhydrol.2017.04.051.

[21] I. W. A. Publishing, "Using ensemble systems to study natural processes Stanislav Yamashkin , Milan Radovanovic , Anatoliy Yamashkin," pp. 753–765, 2018, doi: 10.2166/hydro.2018.076.

[22] E. Sharghi, V. Nourani, and N. Behfar, "Earthfill dam seepage analysis using ensemble artificial intelligence based modeling," J. Hydroinformatics, vol. 20, no. 5, pp. 1071–1084, 2018, doi: 10.2166/hydro.2018.151.

[23] S. I. Malami et al., "Implementation of soft-computing models for prediction of flexural strength of pervious concrete hybridized with rice husk ash and calcium carbide waste," Model. Earth Syst. Environ., 2021, doi: 10.1007/s40808-021-01195-4.

[24] V. Nourani, G. Elkiran, and S. I. Abba, "Wastewater treatment plant performance analysis using artificial intelligence – an ensemble approach," Water Sci. Technol., 2018, doi: 10.2166/wst.2018.477.

[25] S. I. Abba, S. J. Hadi, and J. Abdullahi, "River water modelling prediction using multi-linear regression, artificial neural network, and adaptive neuro-fuzzy inference system techniques," in Procedia Computer Science, 2017, vol. 120, pp. 75–82. doi: 10.1016/j.procs.2017.11.212.

[26] T. Zhou, F. Wang, and Z. Yang, "Comparative analysis of ANN and SVM models combined with wavelet preprocess for groundwater depth prediction," Water (Switzerland), vol. 9, no. 10, 2017, doi: 10.3390/w9100781.

[27] J. . Hair, W. . Black, B. . Babin, and R. Anderson, Multivariate Data Analysis, 7th ed. Upper Saddle River, NJ, USA.: Prentice-Hall, Inc., 2010.

[28] A. Tahsin, J. Abdullahi, A. Rotimi, F. H. Anwar, S. I. Malami, and S. I. Abba, "Multi-state comparison of machine learning techniques in modelling reference evapotranspiration: A case study of Northeastern Nigeria," 2021 1st Int. Conf. Multidiscip. Eng. Appl. Sci. ICMEAS 2021, pp. 1–6, 2021, doi: 10.1109/ICMEAS52683.2021.9692355.

[29] R. Adhikari, M. Bijari, and G. P. Zhang, "Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model Related papers A new class of hybrid models for t ime series forecast ing Mehdi Bijari A

Homogeneous Ensemble of Art ificial Neural Net works for T ime Series Forecast ing Rat nadip Adh," Neurocomputing, vol. 50, pp. 159–175, 2003.
[30] S. S. Sammen, M. Ehteram, S. I. Abba, R. A. Abdulkadir, A. N. Ahmed, and A. El-Shafie, "A new soft computing model for daily streamflow forecasting," Stoch. Environ. Res. Risk Assess., vol. 35, no. 12, pp. 2479–2491, 2021, doi: 10.1007/s00477-021-02012-1.
[31] B. Mohammadi et al., "Adaptive neuro-fuzzy inference system coupled with shuffled frog leaping algorithm for predicting river streamflow time series," Hydrol. Sci. J., vol. 00, no. 00, pp. 1–14, 2020, doi: 10.1080/02626667.2020.1758703.