Hybrid Machine Learning Ensemble Techniques for Modeling Dissolved Oxygen Concentration

SANI ISAH ABBA[1], NGUYEN THI THUY LINH[2,3], JAZULI ABDULLAHI[4], SHABAN ISMAEL ALBRKA ALI[4,5], QUOC BAO PHAM [6,7], RABIU ALIYU ABDULKADIR[8], ROMULUS COSTACHE[9,10], VAN THAI NAM[11], AND DUONG TRAN ANH[11]

[1]Department of Physical Planning Development, Yusuf Maitama Sule University, Kano 700221, Nigeria
[2]Institute of Research and Development, Duy Tan University, Danang 550000, Vietnam
[3]Faculty of Environmental and Chemical Engineering, Duy Tan University, Danang 550000, Vietnam
[4]Department of Civil Engineering, Faculty of Engineering, Near East University, 99138 Nicosia, Cyprus
[5]Sustainable Urban Transport Research Centre (SUTRA), Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia
[6]Environmental Quality, Atmospheric Science and Climate Change Research Group, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam
[7]Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam
[8]Department of Electrical Engineering, Kano University of Science and Technology, Wudil 713101, Nigeria
[9]Research Institute of the University of Bucharest, 050663 Bucharest, Romania
[10]National Institute of Hydrology and Water Management, 013686 Bucharest, Romania
[11]Ho Chi Minh City University of Technology (HUTECH), Ho Chi Minh City 700000, Vietnam

Corresponding author: Quoc Bao Pham (phambaoquoc@tdtu.edu.vn)

ABSTRACT The reliable prediction of dissolved oxygen concentration (DO) is significantly crucial for protecting the health of the aquatic ecosystem. The current research employed four different single AI-based models, namely long short-term memory neural network (LSTM), extreme learning machine (ELM), Hammerstein-Weiner (HW) and general regression neural network (GRNN) for modeling the DO concentration of Kinta River, Malaysia using available water quality (WQ) parameters. Afterwards, the first scenario used four different ensemble techniques (ET). Two linear, i.e. simple averaging ensemble (SAE) and weighted averaging ensemble (WAE) and two nonlinear namely; backpropagation neural network ensemble (BPNN-E) and HW ensemble (HW-E). The second scenario employed a hybrid random forest (RF) ensemble in order to enhance the prediction accuracy of the single models. The WQ parameters were subjected to a different pre-analysis test to ascertain their stability. The four-model combinations are generated using the nonlinear sensitivity input selection approach. The modeling performance was assessed using the statistical measures of Nash-Sutcliffe coefficient efficiency (NSE), Willmott's index of agreement (WI), root mean square error (RMSE), mean absolute error (MAE) and mean square error (MSE) and correlation coefficient (CC). The results of the single AI-based models demonstrated that HW (M3) served as the best model for predicting DO concentration. For ensemble results, BPNN-E (WI = 0.9764) was superior to the other three ET with average decreased of more than 2% with regards to MAE. Investigation on the hybrid RF ensemble demonstrated the reliable accuracy for all the hybrid models with better predictive skill shown by the HW-RF (CC = 0.981) ensemble. The overall results verified the promising impact of HW-M3, ET and hybrid RF ensemble for the prediction of the DO concentration in the Kinta River, Malaysia.

INDEX TERMS Dissolved oxygen, artificial intelligence, ensemble techniques, hybrid random forest ensemble, Kinta River.

I. INTRODUCTION

The health of a river solely depends on the dynamic and uncertain behaviour of water quality (WQ) parameters, which can be described by their physico-chemical and biological characteristics [1]–[3]. Determination of WQ is indispensable

The associate editor coordinating the review of this manuscript and approving it for publication was Yuan Zhuang.

to protect the ecosystem and attain sustainable development. The dynamic nature in terms of concentration and permissible fluctuation of River WQ parameters may have complex consequences throughout the aquatic environment [4]–[6]. Among the WQ variables, dissolved oxygen (DO) is widely known as one of the most important parameters indicating the health and the state of aquatic ecosystems, and therefore its modeling is essential for River WQ, watersheds, wetland

ponds, and other hydro-environmental analysis [7], [8]. DO also serves as a significant index for the evaluation of pollution concentration; in terms of its effect, low or high concentrations of DO threatens the aquatic life and endanger the environmental balance of the entire aquatic environment. As such, it is crucial to provide an enhanced approach for modeling uncertain DO. Based on previous studies on Artificial Intelligence (AI), different AI models have been widely used to simulate and predict DO concentration owing to their accuracy, fast learning speed and non-linear nature [9].

Based on the developed literature on the prediction and simulation of DO concentration, the classical linear regression tools have been widely established, which have generally been reported as being associated with low prediction accuracy. This necessitates the development of a strong, accurate and non-linear hydro-environmental method known as the AI approach. In parallel with this, different types of AI based models have been explored for DO prediction such as artificial neural network (ANN) [11]–[15], support vector machine (SVM) [9], [16]–[20], adaptive neuro-fuzzy inference system (ANFIS) [13], [15], [21]–[24], complementary wavelet-AI model, and hybrid evolutionary algorithms [20], [25], [26], [29], [30] extreme learning machine (ELM) [30], [31], deep learning neural network [32], [33] and fuzzy logic models [34]–[36] for the modeling and prediction of DO concentration.

Furthermore, Heddam [37] applied general regression neural network (GRNN) for the modeling and simulation of DO concentration, and the results were compared with the traditional multi-linear regression analysis (MLR) model. The predictive performance skill showed the superiority of the GRNN model over the linear MLR model. Emamgholizadeh *et al.* [22] developed the application of multilayer perceptron (MLP), radial basic function neural network (RBNN) and ANFIS models for the prediction of DO, biological oxygen demand (BOD) and chemical oxygen demand (COD) in the Karoon River, Iran. For this purpose, nine input variables were used for the models. The results showed that the performance accuracy of RBNN and ANFIS were in close agreement; on the other hand, the MLP model exhibited better performance accuracy over the RBNN and ANFIS models. Nemati *et al.* [38] investigated the use of MLR, ANN and ANFIS for the estimation of DO in the Tai Po River, Hong Kong. The data measured from the river included chloride (Cl), pH, electrical conductivity, temperature (Temp), nitrate-nitrogen (NO_2N), nitrate-nitrogen (NO_3-N), ammonia nitrogen (NH_4-N) and total phosphorus (T-P). The prediction accuracy of the models based on different performance efficiencies indicated that the ANN model had better accuracy than MLR and ANFIS models in modeling the DO concentration Kisi *et al.* [26] studied the performance of three AI-based models, namely MLP, ANFIS and genetic programming (GP), for the estimation of DO concentration in the South Platte River at Englewood, Colorado. For this purpose, various input combinations were

used, and the results demonstrated the capability of GP over the MLP and ANFIS models. Elkiran *et al.* [10] employed three different data-driven models including FFNN, ANFIS and MLR for multi-station prediction of the DO concentration in the Yamuna River, India. The results demonstrated the capability of the AI- based models over the MLR model. Comparison based on the performance accuracy revealed that ANFIS was more effective than the FFNN model.

Besides the above studies, Yaseen *et al.* [9] employed different data-intelligence algorithms including least square support vector machines integrated with a bat algorithm (LSSVM-BA), M5-Tree and multivariate adaptive regression splines (MARS) models for the prediction of DO concentration in a river using different input combinations. The outcomes revealed that LSSVM-BA outperformed the other models with considerable accuracy. The obtained results also indicated that both the employed models were capable of predicting DO concentration Alizadeh and Kavianpour [39] applied ANN and WNN for the estimation of WQ parameters (DO) in Hilo Bay, Pacific Ocean using various combinations of WQ input parameters. The results depicted that the performance of the WNN models was better than the ANN models Zhu and Heddam [31] studied the application of the ELM model for the estimation of DO concentration in four different Rivers in China. The study invoves different set of WQ paameters. The predictive results were compared with the traditional MLP, and the outcomes showed that both ELM and MLP were capable of modeling the DO concentration. The performance of the models also indicated that MLP slightly outperformed the ELM model Yalin *et al.* [34] developed a study based on fuzzy neural network for the prediction of DO concentration in a crab pond, and the prediction results indicated the suitability of the fuzzy neural network in DO prediction over a grey neural network. Similarly, Zounemat-Kermani and Scholz [35] reported the capability of the fuzzy model in modeling and predicting DO concentration. Besides the application of MLP, SVM, ANFIS, the deep learning (DL) neural network has also been reported in some recent studies such as Ta and Wei [32] who proposed the application of a convolutional neural network (CNN) and compared it with the traditional BPNN for modeling the DO concentration using data obtained from Mingbo Experimental Base in Shandong Province. The results indicated that the superiority of the CNN in terms of performance accuracy over the traditional BPNN Liu *et al.* [33] employed Long Short-term Memory (LSTM) deep neural networks for the prediction of WQ parameters using data measured from Guangzhou water source of the Yangtze River in Yangzhou, Malaysia. The results showed the feasibility of the LSTM model for estimating WQ parameters.

In fact, all the previous researches justified the reliability of the black box models (e.g. MLP, ANFIS, SVM) for modeling and predicting DO concentrations over different geographical location around the globe. Despite the satisfactory records of such AI-based models, it is evident that no single model has been verified as the most effective for all kinds of data set,

as certain issues with different models may lead to different outcomes. The data characteristics, like size, linearity, normality, size, and so on, have an effect on the model's predictive performance [40]. According to Sharghi *et al.* [40] and Raj Kiran and Ravi [41], the combination of several models could enhance the forecasting performance. The overall concept of combining the models (ensemble modeling) is to take advantage of the unique features for the constituent models to bring about different patterns presented in the dataset. Khan and Chai [42] employed an ensemble of ANN and ANFIS models for the prediction of water quality index in the Hog Island Channel Monitoring Station. The obtained results proved that the ensemble techniques produced better results than the single model. Nourani *et al.* [43] proposed and applied different types of ensemble approaches for the prediction of a wastewater treatment plant in Nicosia, Cyprus, using the output of FFNN, ANFIS, SVM, and MLR models. The outcomes indicated that the ensemble model has better prediction accuracy. This technique has been also applied in various fields of hydro-environmental engineering, such as precipitation [44], earth-fill dam seepage analysis [40], evapotranspiration [45], and River WQ [46].

One of the factors affecting the accuracy of the models is the model input determination, which depends on the identification of the AI-based models, and others include model configuration, prediction horizon, etc. According to Hadi *et al.* [47] different approaches have been reported for determining the most suitable input variables including principle component analysis (PCA), autocorrelation function (ACF), partial autocorrelation function (PACF) and Pearson correlation analysis. Conversely, those techniques are associated to the input-output linear relationship [48]. To overcome this drawback as part of the current study, nonlinear sensitivity analysis will be used to extract the dominant input variables. However, the proposed techniques serve as the first study depicting the application of enhanced hybrid AI at Kinta River and up-to-date to the best knowledge of the authors here is no conducted technical research using this approach. The objectives of this study are: (i) To develop a neuro-sensitivity approach using MLP to determine the influence of each WQ parameter on DO concentration; (ii) To develop and compare the potentials of different data-intelligence models including ELM, LSTM, HW, and GRNN models; (iii) To improve the prediction accuracy using four different ensemble techniques including two nonlinear models, i.e. backpropagation neural network ensemble (BPNN-E) and HW ensemble (HW-E), and two linear models, namely weighted averaging ensemble (WAE) and simple averaging ensemble (SAE), for scenario 1 and a hybrid random forest (RF) ensemble for the black box model in scenario 2.

The main motivation of this study in the realm of environmental research is the inspection of potential nonlinear sensitivity analysis for selecting the most dominant attribute to the target value (DO). Some robust modeling techniques are considered in this study, which is relatively new to the application of DO concentration. Moreover, for complex and chaotic systems such as river dynamics, single models often provide an unsatisfactory forecast. For this purpose, the hybrid model is established to improve the prediction performance of the single models. The use of hybrid ensemble RF with the AI-based models (RF-LSTM, RF-ELM, RF-HW, and RF-GRNN) for DO modeling has not received much attention in the literature. With the aforementioned references in the literature section, various studies were explored using several compactional intelligent models indicating the real environmental simulation. Yet, for utilizing the appropriate decision making, emerging algorithms need to be incorporated for hydro-metrological and environmental modeling. Similarly, most AI-based models, are intricate and thus their calibration involve high computing costs. Recently, AI models such as ELM, HW and LSTM have gradually become popular in various water management application, owing to their simplicity, robustnesss and high computational efficiency in handling large data compared to several other AI methods [33], [48], [51]. It is evident with the above reviews that the application of the relatively simple LSTM, ELM, GRNN, SAE, WAE, BPNN-E, HW-E and hybrid RF ensemble has not been evaluated before at Kinta River, Malaysia.

II. MATERIAL AND METHODS

A. LONG SHORT-TERM MEMORY (LSTM)

Long short-term memory neural network (LSTM) is a special type of recurrent neural network (RNN) that can solve the issues of gradient explosion and gradient disappearance effectively during RNN training as well as increase the accuracy of RNN [33]. This model was designed to minimise the weakness of the classical RNN, which generally does not have the ability to remember sequences with a length of 10 or more. All RNNs take the form of chained repeating modules of the neural network. LSTMs, which involves the usage of special memory cells used in storing information, also have this chain with an almost identical structure [33], [51], [52].

B. EXTREME LEARNING MACHINE (ELM)

Normally, the traditional ANN model requires that the parameters of the hidden neuron are tuned. ELM has recently been developed as a novel technique that map the internal features without the requirement of a traditional ANN [53]. From several pre-assigned neurons in the ELM, weightings of the input and hidden neurons are calculated at random. These values do not have to pass through all the neurons. Also, the ELM's generalization capacity requires less computing time [50]. The ELM was first suggested in [53] as a new developed data-driven black-box model consisting of a single hidden layer feed-forward network (SLFN). The ELM somewhat distinct from the conventional feed-forward NN (FFNN) by handling the issues of slow learning, overfitting, and local minima [54]. In particular, the potentials of the ELM may be ascribed to the generalization capacity

and high learning speed. Hence, ELM has been widely implemented in hydro-environmental studies [50], [55].

C. HAMMERSTEIN-WEINER MODEL (HW)

The Hammerstein-Weiner (HW) model is another form of a black-box model, developed for nonlinear systems identification [56]. In HW model a linear dynamic system is sandwiched between two nonlinear blocks. In general, the configuration of the HW model is contained three blocks: a static input nonlinear block, followed by a linear dynamic block, and then another static output nonlinear block [57]. The HW model converts a given nonlinear inputs into a piecewise linear function blocks, and then transform it into a nonlinear output function.

D. GENERAL REGRESSION NEURAL NETWORK (GRNN)

GRNN is a type of ANN that has an important and attractive characteristic of self-learning ability and can handle complex non-linear problems. The modeling using GRNN can be performed accurately without employing large data sets [16]. Also, GRNN has the capability to resolve problems concerning smooth functions, approximation, and can also generate consistent prediction accuracies [37]. Due to this, the algorithm exhibits fast learning speed, which demonstrates excellent results in the field of environmental modeling.

E. PROPOSED SINGLE MODELING SCHEMA

Based on a review of the literature, it is apparent that numerous studies using data-intelligence algorithms have been conducted and have shown promising performance for modeling complex systems. For the current study, the DL neural network (LSTM), emerging self-adaptive predictive model (ELM), non-linear system identification model (HW) and recently developed traditional neural network (GRNN) were proposed separately for modeling the DO concentration of Kinta River, Malaysia. Afterward, two different scenarios were applied using a novel ensemble approach and hybrid random forest ensemble, as highlighted in section 1 above. Although the AI models proved robust, the determination of appropriate input variables is the major problem in most of the techniques. According to Hadi et al. [47] and Yaseen et al. [50], excessive inputs deteriorate the model performance while too little inputs may not reveal all of the hidden information in the time series. Therefore, in this study, nonlinear sensitivity analysis was carried out using MLP owing to its promising capability for modeling DO concentration and other WQ parameters. Other feasible alternatives to sensitivity analysis and input variables selection may also be used. After that, the first scenario used four different ensemble techniques (ET) (two linear (i.e. simple averaging ensemble (SAE) and weighted averaging ensemble (WAE)) and two nonlinear techniques, i.e. BPNN-E and HW ensemble (HW-E)) and the second scenario employed a hybrid random forest (RF) ensemble in order to enhance the prediction accuracy of the single models. The general proposed flowchart is presented in Fig. 1.

1) ENSEMBLE LEARNING TECHNIQUE (ELT)

The ability to combine models (ensemble technique) to improve the final prediction has been successful in various fields including classification, hydro-environmental, water resources and traffic engineering [58]. ELT is a discipline in the field of machine learning used to combine the process of obtaining multiple predictors by single models to enhance the final prediction performance. The main target of the ensemble is to produce higher accuracy and reliable estimates than could be achieved through a single model [59]. As reported by Khan and Chai [42], Elkiran et al. [46], there are two ensemble techniques: (1) linear ensemble method, which includes linear ensembles by simple averaging, weighted averaging and weighted median; and (2) nonlinear ensemble method, which involve the use of black-box model as nonlinear kernels to obtain an ensemble output. Other researchers have categorized the ELT into two, namely homogeneous and heterogeneous ensembles; when ELT comprised of the same learning algorithm (e.g. neural network), it is called homogeneous, but if it consists of different learning algorithms, it defined as heterogeneous. As suggested by [40], [43], the heterogeneous ensemble is recommended for overcoming the model diversity and for attaining prediction accuracy.

Therefore, two linear (i.e. simple averaging ensemble (SAE) and weighted averaging ensemble (WAE)) and two nonlinear techniques, namely BPNN-E and HW ensemble (HW-E) were employed in this study to simulate the DO concentration of the Kinta River.

a: LINEAR ENSEMBLE APPROACH

The proposed SAE approach is carried out by considering the arithmetic average of the predicted model's outputs (ELM, LSTM, GRNN, and HW) as:

$$DO_{(t)} = \frac{1}{N} \sum_{i=1}^{n} DO_{(t)i} \qquad (1)$$

Similarly, the WAE approach can be obtained by assigning a unique weight to each of the individual outputs, and the final predicted outcomes are obtained by averaging the models. The WAE provides more reliable predictive skills than the SAE owing to the nature of assigning weights, and it can be expressed as:

$$DO_{(t)} = \sum_{i=1}^{N} w_i DO_{(t)i} \qquad (2)$$

where w_i is the assigned weight on the output of the ith model, $DO_{(t)}$ is the output of the ensemble model (SAE or WAE), $DO_{i(t)}$ is the ith single model output (here outputs of ELM, LSTM, GRNN, and HW) and N is the total number of the single models (here, N = 4). The term w_i can be expressed as:

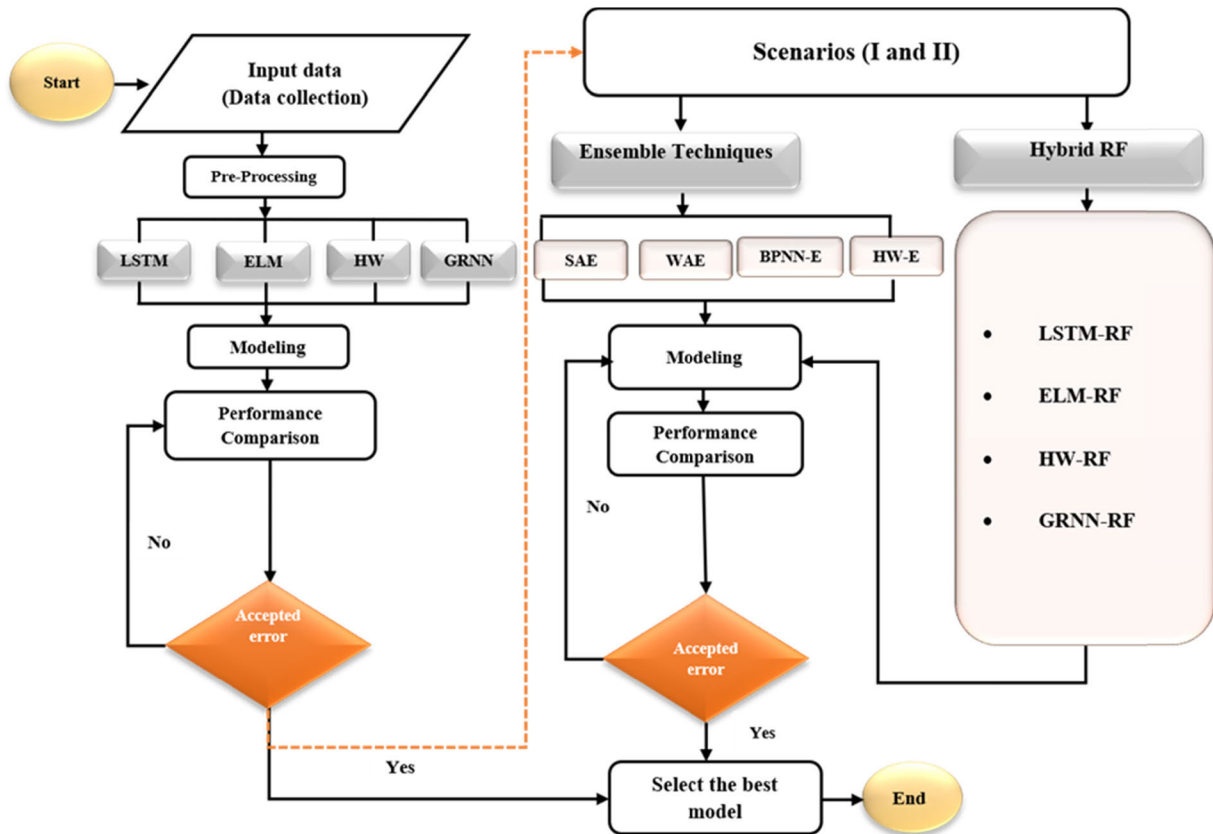$$w_i = \left[\frac{DC_i}{\sum_i^N DC_i} \right] \qquad (3)$$

FIGURE 1. General proposed methodology.

b: NONLINEAR LINEAR ENSEMBLE (NLE) APPROACH

The approach of NLE is similar to the traditional BPNN model where the outputs of the single models, i.e. LSTM, ELM, GRNN, and HW models are imposed and trained using a new neural network (NN). The procedure follows the same trend of the traditional NN in terms of best architecture selection. Recently, the application of the NN ensemble has received attention in different fields of hydro-environmental engineering, including earth-fill dam seepage analysis [40], vehicular traffic noise [60] etc, and all have reported the superiority of the NLE over the single model. Moreover, the above studies suggest the use of other nonlinear kernel functions as an alternative for such nonlinear ensembling. Hence, this study proposes an ensemble using the HW model as the additional nonlinear kernel function due to its advantages for single modeling prediction. Even though the HW modelling technique has not been implemented for DO prediction, the technique has demonstrated enormous potentials in water resource researches [56]. Fig. 2 shows the schematic of the proposed NLE techniques using the BPNN model.

2) SCENARIO II: HYBRID RANDOM FOREST (RF) ENSEMBLE

The Traditional ANN was previously considered to have the strong capability to handle complex nonlinear relationships, but recent studies have reported various defects and difficulties in modeling WQ parameters with the traditional ANN and other AI-based models. Consequently, researchers are no longer solely reliant on single AI-based models to capture the nonlinear nature of hydro-environmental systems. Hence, hybrid models have been proposed to enhance the evaluation accuracy [61]. In this section, the hybridization of an RF ensemble with four data intelligence algorithms (LSTM, ELM, HW and GRNN) was employed.

RF is one of the powerful ensemble machine learning algorithms proposed by Breiman (2001) by adding an additional randomness layer to the bagging method [62], [63]. RF performs its function by generating multiple decision trees using a randomization process. These processes produce a large ensemble of trees, and the overall predictions are achieved from the averaged outcomes. The decisions are generated using bootstrapping or a random selection of inputs that are used to create the various base trees. Very recently, there has been increasing interest in RF, and it has been applied in different areas [64], [65]. This is of course, due it's an advantage of overcoming overfitting difficulties, which are reported as one of the most severe problems of Decision Trees (DT). For the implementation of RF, understanding the values of the two parameters are essential, which are *ntree* and *mtry*. The *ntree* stands for the number of trees in the forest and is used for finding the optimal value, while *mtry* indicates the number of parameters in the random subset at each node [66].
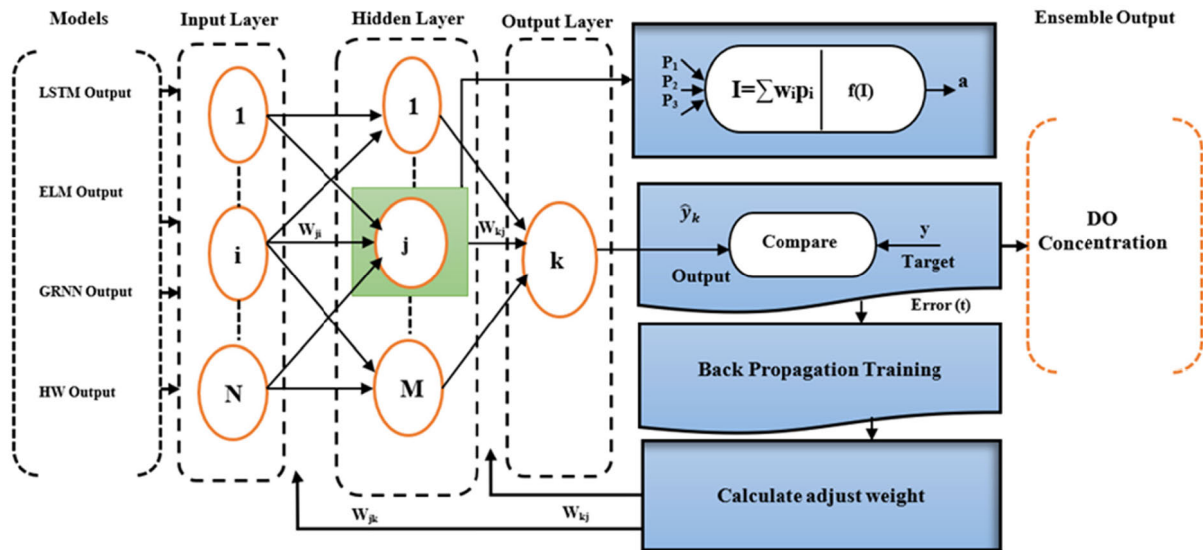
FIGURE 2. Procedure of the developed method for BPNN-E modeling.

As an ensemble learning technique on its own, RF predictions offer a suitable relationship between dependent and independent variables. Therefore, the proposed hybrid RF with AI-based models employed in this study combines the best fitted single models in the form of a regression tree to each bootstrap sample taken from the original predicted values. In RF, each of the different decision trees contributes to the overall RF result, which serves as the weighted average of all the results. Although several hybrid models have been proposed using optimizations algorithms to improve prediction and evaluation of accuracy, to the very best knowledge of the author, the hybrid combining the ensemble (i.e., RF) method with the AI-based models (RF-LSTM, RF-ELM, RF-HW and RF-GRNN) has not been considered.

F. MODELS' EVALUATION CRITERIA

For validating the performance efficiency of the models, various measures were considered as a multi-criteria approach for the models evaluation. The predictive performance of the models is evaluated using three statistical error viz; Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Square Error (MSE) and three goodness-of-fit measures, including Nash-Sutcliffe coefficient efficiency (NSE), Willmott's Index of agreement (WI), and Correlation Coefficient (CC). MAE describe the actual error difference by disregarding the influence of negative values. Low values of MAE suggest that the accuracy of the prediction model is high. MSE is the sum of the square of the differences between the predicted and actual values, although residual show proportionality effect at an individual level in MAE, the consequence of square raise MSE always higher than the MAE and the effect of outliers can easily get recognize. NSE usually take values in the range of $-\infty$ and 1. The match between the predicted and observed data is considered to be perfect for $NSE = 1$. The closer the NSE is to 1,

the more accurate the prediction is. However, with regards to regression techniques, the NSE is the same as the coefficient of determination (R^2), hence ranging between 0 and 1. Similarly, WI account for the degree of the forecast error and ranges from 0 to 1. Where WI value of 1 signifies a perfect match between the observed and predicted data, whereas WI equal to 0 shows that there is no agreement between the predicted and the observed data. However, WI is very sensitive to extreme values because of the squared differences. For more information about performance criteria referred to [56], [68]

$$NSE = 1 - \frac{\sum_{i=1}^{N} (DO_{obs_i} - DO_{pred_i})^2}{\sum_{i=1}^{N} (DO_{obs_i} - \bar{DO}_{obs})^2},$$
$$-\infty < NSE \leq 1 \tag{4}$$

$$WI = 1 - \left[\frac{\sum_{i=1}^{N} (DO_{obsi} - DO_{prdi})^2}{\sum_{i=1}^{N} (|DO_{predi} - \bar{DO}_{obs}| + (|DO_{obsi} - \bar{DO}_{obs}|))}\right],$$
$$0 \leq WI \leq 1 \tag{5}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (DO_{obs_i} - DO_{prd_i})^2}{N}} \tag{6}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |DO_{obsi} - DO_{predi}| \tag{7}$$

$$CC = \frac{\sum_{i=1}^{N} (DO_{obsi} - \bar{DO}_{obsi})(DO - \bar{DO}_{predi})}{\sqrt{\sum_{i=1}^{N} (DO_{obsi} - \bar{DO}_{obsi})^2 \sum_{i=1}^{N} (DO_{predi} - \bar{DO}_{predi})^2}} \tag{8}$$

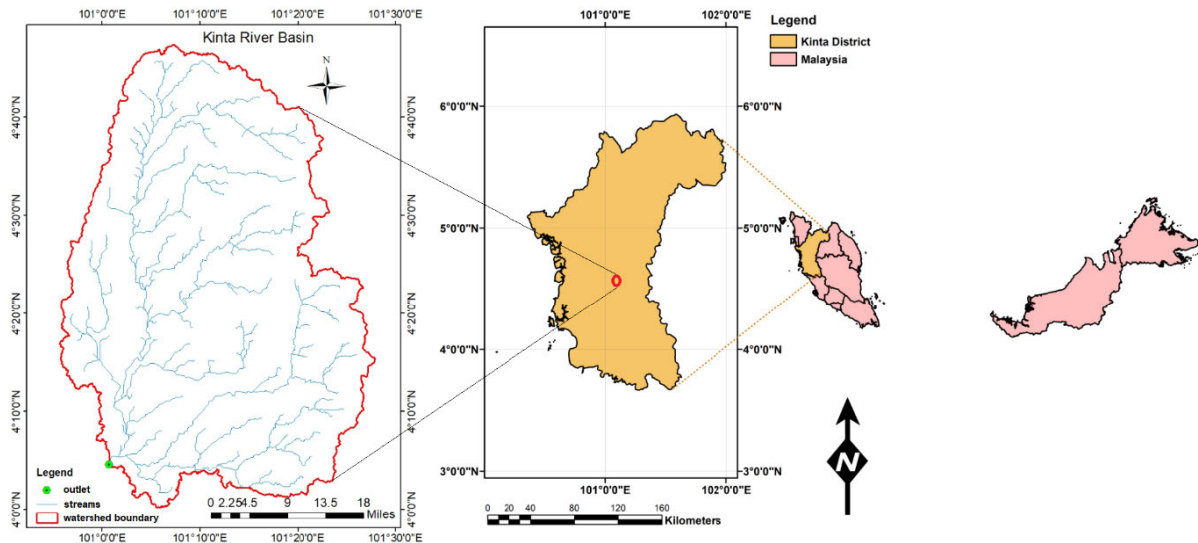$$MSE = \frac{1}{N} \left[DO_{obsi} - DO_{predi}\right] \tag{9}$$

FIGURE 3. Location of (a) Kinta district and (b) drainage Rive Basin, Malaysia.

where N, DO, $\overline{DO}_{obsi}$ and DO_{predi} are data number, observed DO, average value of the observed DO and predicted DO value, respectively.

III. STUDY LOCATION AND DATA DESCRIPTION

The Kinta catchment area is found to lie along latitude N 040 07' 102' and N 040 40' 115', longitude S1010 01' 284' and S 1010 09' 400'. It covers an area of about 2,500 km² with a length of almost 100km. Generally, it is utilized for industries, residential areas and farming. For example, the River Kinta is subdivided into three main divisions: downstream, upstream and undulating. This is one of the largest rivers that cross through Ipoh, where the other two Rivers are the River Pinji and the River Pari. It also flows at about 200m above sea level from Gunung to Perak River [68]. This river, together with its tributaries, drains a basin that covers an area of almost 2,500 km². The river also flows via heterogeneous as well mixed-use land where, besides extensive forest cover, the main uses of the land in the basin are palm oil planting, mining, rubber planting, logging and urban development. The Kinta River is the primary source of water for irrigation and drinking in Ipoh as well the main tributary of the River Perak, which serves as the main source of water for drinking and irrigation in Perak [69]. Currently, there is only one dam on the Kinta River, in which was constructed in 2000 to increase the water supply of Perak to 25%. This dam can provide almost 650,000 m³ of water daily and satisfies the water demands of the people. The location district and drainage map of the study area are presented in Fig. 3a and b, respectively.

The WQ data is collected from Department of Environment (DoE) (Malaysia), consisting of 301 instances for the period of 12-year (2002-2013) with 301 instances. The dataset is composed of six monitoring stations' records which are located at the Kinta River, Malaysia. The WQ parameters includes DO, BOD, COD, temperature (Temp), ammonia (NH3), total solids (TS), chlorides (Cl), calcium (Ca),

PH and sodium (Na). The data were portioned into 75% for calibration and 25% for verification of the model. However, for any AI-based model, the primary target is to fit the model to the given dataset with the aim of attaining a reliable estimation on the unknown datasets. In this regard, k-fold cross-validation method is employed for the dataset validation. The k-folding method has the ability to optimize the model by using separate dataset for the training and validation phases. In this study similar to many other hydro-enviromental studies, 10 k-folding cross-validation procedure is used, however, other alternatives may also be used. The descriptive statistical analysis of the datset under the study is given Table 1, and normalization of the data was done using Eq. 10.

$$y = 0.05 + \left(0.95 \left(\frac{x - x_{min}}{x_{max} - x_{min}} \right) \right) \qquad (10)$$

where y is the normalized data, x is the measured data, x_{max} and x_{min} are the maximum and minimum value of the measured data, respectively.

IV. RESULTS AND DISCUSSION

As stated above, the current study has three main objectives, namely to apply four different intelligence algorithms for modeling DO, to enhance the prediction performance using two linear and two nonlinear ensemble techniques and finally, to propose a hybrid RF ensemble using the best outputs of the single models. Hence, this section presents the obtained results of the two scenarios (ensemble techniques and hybrid RF ensemble).

A. PRELIMINARY AND RELIABILITY ANALYSIS

For any time-series data, pre-analysis of the individual data, i.e. the individual inputs, is paramount because their accuracy can significantly contribute to determining the efficiency of the individual models. As such, reliability and stationary
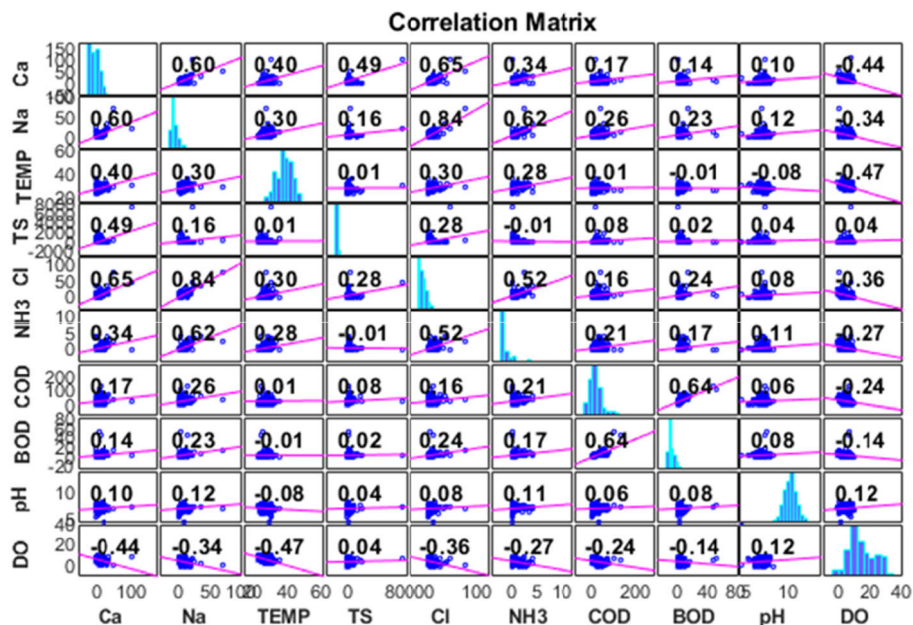
FIGURE 4. Pearson's correlations coefficients among the observed WQ parameters.

TABLE 1. Statistic of WQ parameters.

Parameters	X_{mean}	S_x	C_{sx}	X_{min}	X_{max}
DO	4.82	2.12	0.29	0.28	9.93
BOD	4.09	4.7	6.97	1	54
COD	26.34	15.84	2.37	3	138
pH	6.99	0.42	-0.87	4.54	8.29
NH3	0.43	0.71	2.68	0.01	4.22
TEMP	27.94	1.99	0.05	23.06	35.49
TS	216.67	481.62	12.59	17	7613
Cl	6.38	7.19	4.55	0.5	77
Ca	12.48	9.81	3.13	0.05	105
Na	6.56	5.72	5.74	0.05	71

tests were conducted using the Cronbach's alpha method and unit root test (i.e. using Augmented Dickey-Fuller (ADF)) analysis to ascertain the stability of the data. According to Hair *et al.* [70] variables of a dimension are internally consistent if their Cronbach's alpha values exceed the threshold of 0.7. Dickey *et al.* [71] reported that the ADF test is carried out to have more reliable and valid outcomes and to ensure the stationarity of all the variables. The data used in this study certified all the stationarity requirement with regard to the unit root test.

Besides the pre-analysis, determination of the correlation coefficient is also crucial for the development of data-driven models since the directional sign (negative or positive) indicates the relationship between the dimensions and also shows the proportionality of the independent and dependent variables (see, Fig.4). The selection of 4 input combinations is predominantly based on the nonlinear sensitivity analysis. The advantage of using the nonlinear sensitivity input variables selection approach to carefully determine the most relevant factors has been reported in various studies [47], [72], [73]. Therefore, a sensitivity analysis using MLP between the observed WQ parameters was evaluated, and the average values of RMSE and NSE were considered and ranked in hierarchical order of the WQ variables as presented in Table 2.

From Fig. 4, despite the linearity function of the Spearman Pearson Correlation but still depicted good relationship between the parameters. However, neither the direct nor sign signifies the strength and weakness of the correlation bonding. In any relationship matrix there exist a positive and negative correlation in which the former indicates the corresponding increases between the WQ paper and later shows an inverse pattern [74].

Except for TS, all the WQ variables show excellent inverse relationship with the DO. Even though studies such as [43], [46], [47], [57], [75] have criticized the classical linear input variable selection and recommended the use of nonlinear approaches; they are applicable for input selection and the determination of linear patterns between the variables.

The observed WQ parameters were analysed, and the statistical overview of the data was obtained as presented in Table 1. The terms S_x, and C_{sx}, in the table represent the standard deviation and skewness while X_{min}, X_{max} and X_{mean} indicate the minimum, maximum, and mean values respectively. According to Table 2, the minimum and maximum values of the WQ parameters demonstrate variability trend. The DO concentration varies from 0.28 to 9.93 mg/L.
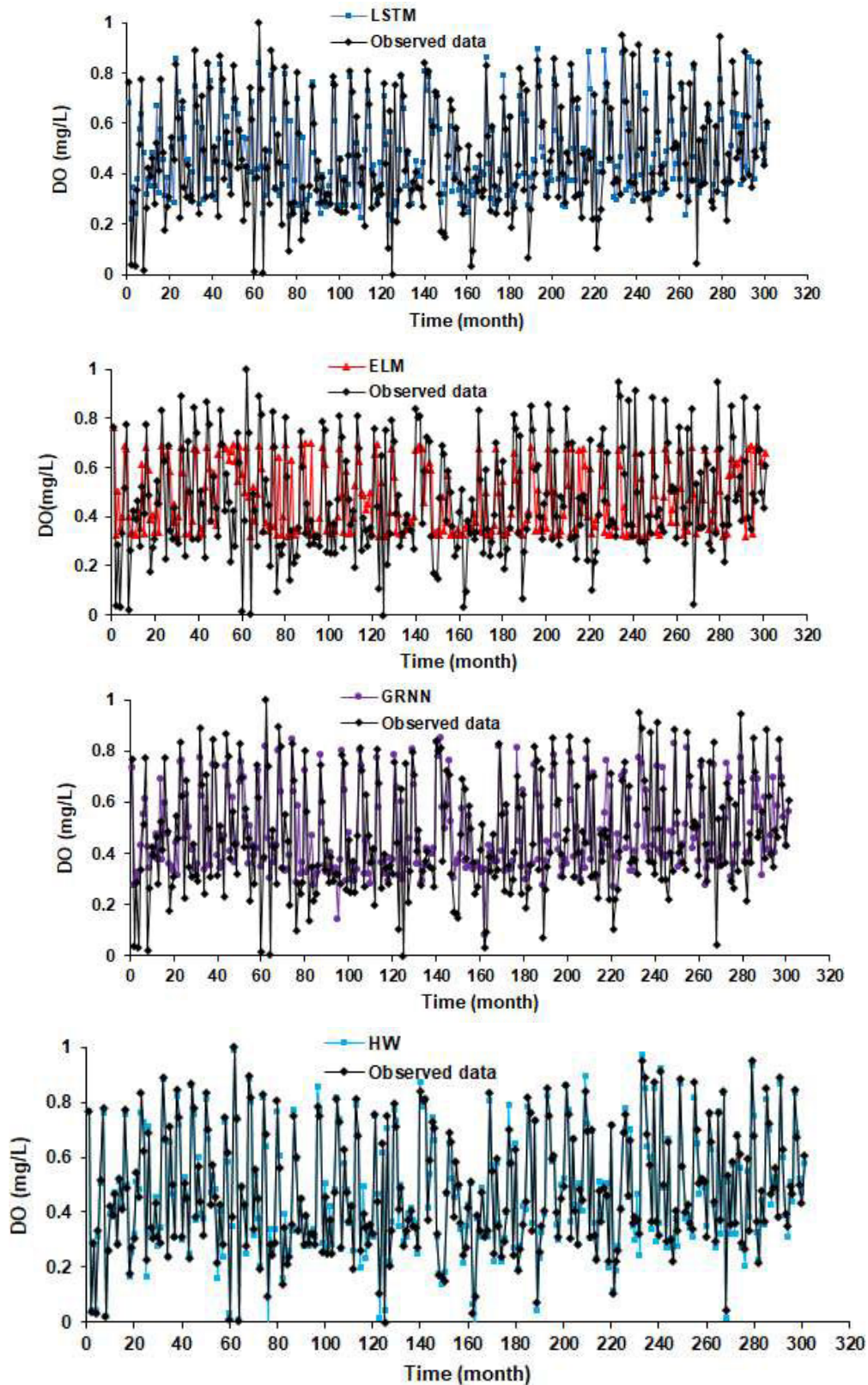
FIGURE 5. Time series graph for LSTM-M3, ELM-M1, GRNN-M4 and HW-M3 between the observed and predicted values.

TABLE 2. Sensitivity analysis results.

Parameters	Units	RMSE	NSE	Ranking
Ca	(mg/L)	0.0098	0.3950	1
Na	(mg/L)	0.0109	0.2441	2
TEMP	°C	0.0110	0.2364	3
TS	(mg/L)	0.0112	0.1995	4
Cl	(mg/L)	0.0113	0.1872	5
NH3	(mg/L)	0.0121	0.0800	6
COD	(mg/L)	0.0123	0.0507	7
BOD	(mg/L)	0.0125	0.0073	8
pH	-	0.0143	-0.2893	9

In recent years it was recorded that impurity related to WQ and pollution in Kinta River are associated to the extreme anthropogenic activities owing to the significant amount of non-biodegradable. The WQ of Huaxi and Yipin River was defined to be within the pollution level with range of DO: 2.9-10 and DO 5.2-10.5, respectively [31]. According to Olyaie et al. [17], the predictive performance of the ANN model may be substantially affected by too much skewness of the parameters, while a low skewness is more suitable to model. The pattern variation of the WQ parameters affirms the complexity and non-linearity of the DO concentration modelling.

B. RESULTS OF SINGLE DATA-INTELLIGENCE ALGORITHMS

For modeling DO concentration, four different models were derived based on the sensitivity analysis as M1, M2, M3 and M4. These models were separately imposed into four data-intelligence algorithms (LSTM, ELM, HW and GRNN), and the performance results were evaluated using NSE, MAE, WI and RMSE. A performance comparison of the four data intelligence models was carried out, and the obtained results are reported in Table 3. It should be noted that the best hyperparameters structure were attained using trial and error for all four models. Among the AI-based models, HW (M3) served as the best model for predicting DO concentration followed by the GRNN (M4), ELM (M1) and LSTM (M3) models. Fig. 5 presents the time series variation graph between the observed and predicted DO concentrations by the single models. Further analysis of the time-series results shows that HW-M3 and GRNN-M4 performed well, while moderate accuracy was observed for both LSTM-M3 and ELM-M1. The promising capability of the HW model is certainly not surprising, because it is an evolving non-linear system identification technique and has shown better predictive ability in various studies. The quantitative examination of HW-M3 with regard to NSE = 0.9702 and MAE = 0.0013 indicates that the model outperformed the other three models. Also, these comparisons revealed that

TABLE 3. Performance results of different single models in modeling do concentration.

	Models		NSE	MAE	WI	RMSE
Calibration	(a) LSTM	M1	0.5509	0.0136	0.7235	0.0077
		M2	0.5869	0.0201	0.7424	0.0094
		M3	0.599	0.0195	0.7519	0.0093
		M4	0.5909	0.0199	0.758	0.0094
	(b) ELM	M1	0.644	0.0273	0.7602	0.011
		M2	0.6407	0.0289	0.7602	0.0113
		M3	0.6297	0.0343	0.7514	0.0123
		M4	0.6362	0.0311	0.7527	0.0117
	(c) HWM	M1	0.6358	0.0313	0.7509	0.0118
		M2	0.6362	0.0311	0.7459	0.0117
		M3	0.9733	0.0013	0.9867	0.0024
		M4	0.7228	0.0135	0.801	0.0077
	(d) GRNN	M1	0.6014	0.0194	0.7573	0.0093
		M2	0.5673	0.0211	0.7258	0.0097
		M3	0.6505	0.017	0.7789	0.0087
		M4	0.7656	0.0167	0.878	0.0086
Verification	(a) LSTM	M1	0.524	0.0248	0.6201	0.1576
		M2	0.5169	0.0351	0.6584	0.1873
		M3	0.529	0.03	0.645	0.1731
		M4	0.519	0.0342	0.6595	0.185
	(b) ELM	M1	0.6189	0.0313	0.7594	0.1751
		M2	0.6133	0.0324	0.7617	0.1799
		M3	0.6167	0.0309	0.7634	0.1759
		M4	0.6073	0.0307	0.7636	0.1753
	(c) HWM	M1	0.6489	0.0216	0.7444	0.1469
		M2	0.6341	0.0278	0.767	0.1668
		M3	0.9702	0.0013	0.9851	0.0355
		M4	0.7325	0.0113	0.8662	0.1063
	(d) GRNN	M1	0.6307	0.0293	0.6536	0.171
		M2	0.6325	0.0285	0.6627	0.1688
		M3	0.6411	0.0249	0.7056	0.1577
		M4	0.7429	0.0241	0.8714	0.1553

even for the same input combinations, the effect of each model on independent WQ parameters behaves differently, for example, HW-M3 and LSTM3.

Another reason for the poor performance of the other model combinations could be associated with the inverse relations, which was identified by the negative correlation between the observed DO concentration and the WQ parameters except for the pH and TS values. This observation was similar to the findings reported by Zhu and Heddam [31].
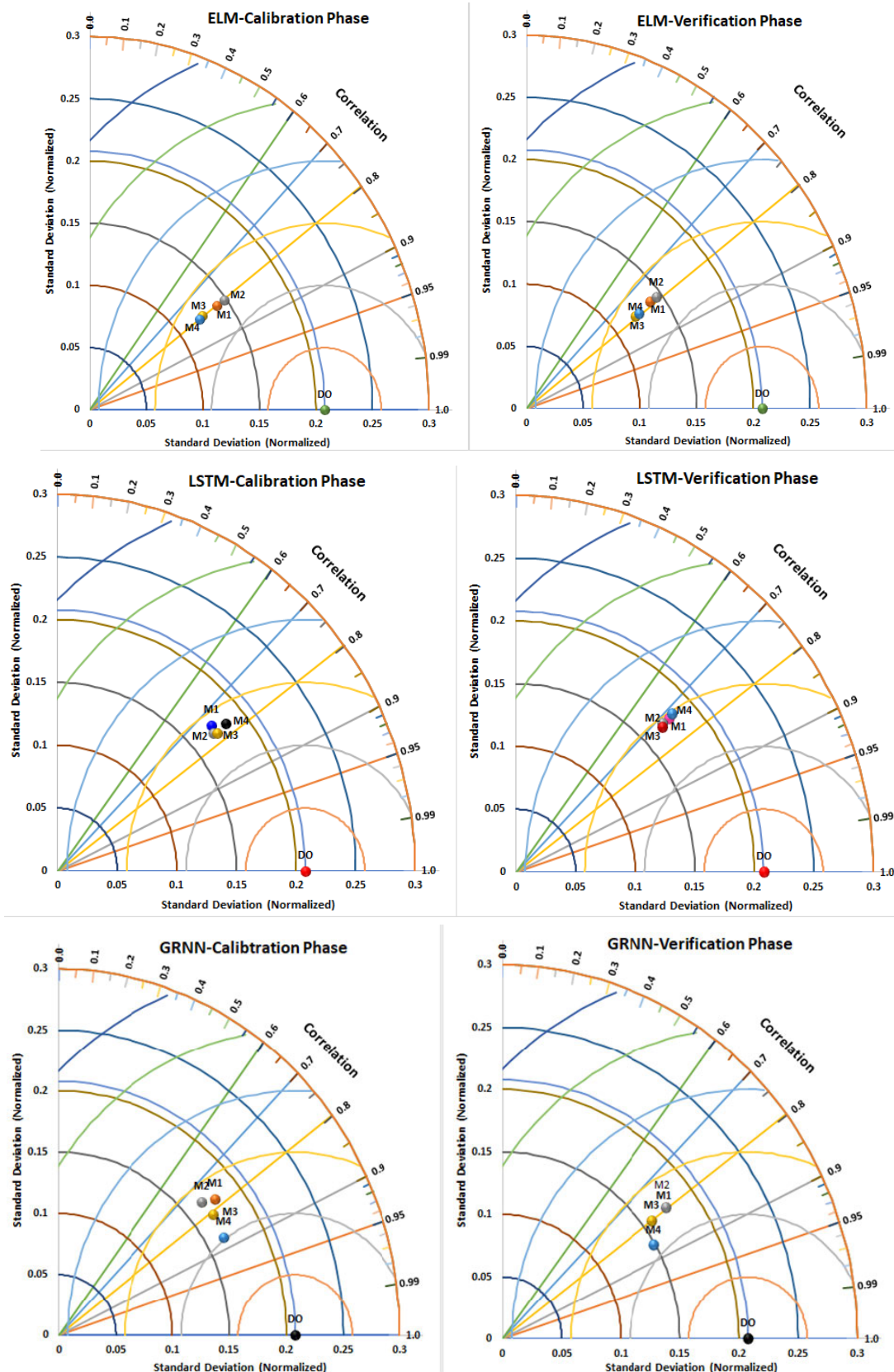
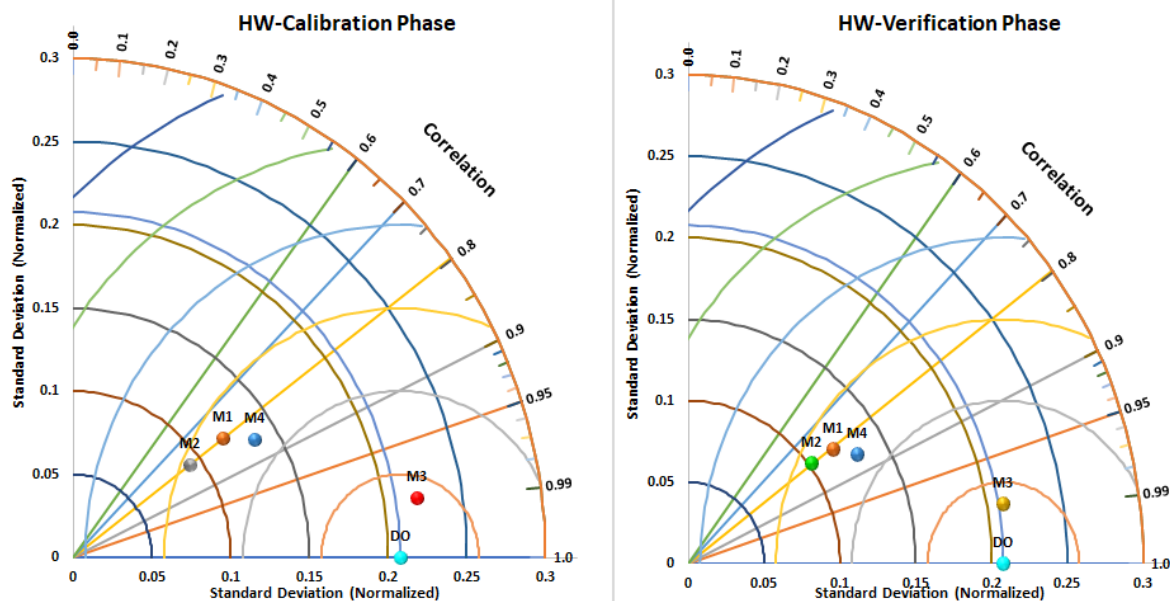**FIGURE 6.** Two-dimensional Taylor diagram for LSTM, ELM, GRNN and HW for both calibration and verification.

FIGURE 6. *(Continued.)* Two-dimensional Taylor diagram for LSTM, ELM, GRNN and HW for both calibration and verification.

TABLE 4. Results of four ensemble techniques.

Ensemble Techniques	Calibration				Verification			
	NSE	MAE	WI	RMSE	NSE	MAE	WI	RMSE
SAE	0.9179	0.0025	0.9562	0.0033	0.8069	0.0643	0.9035	0.0794
WAE	0.9198	0.0024	0.9569	0.0033	0.8128	0.0611	0.9064	0.0782
BPNN-E	0.9714	0.0018	0.9877	0.0019	0.9670	0.0331	0.9764	0.0182
HW-E	0.9633	0.0019	0.9775	0.0029	0.9555	0.0341	0.9757	0.0176

The comparison evaluation for the best single models is provided using two-dimensional Taylor diagram, as presented in Fig. 6. The Taylor diagram highlights and summarizes several statistical indices such as correlation (R), RMSE, and standard deviation between the observed and computed values [57], [76]. From Fig. 6, it can be observed that the DO concentration achieved better goodness-of-fit using HW-M3 with the value of R = 0.9849, LSTM (R = 0.7273), ELM (R = 0.7867) and GRNN(R = 0.8619) in the verification phase.

The results lead to the conclusion that for both calibration and verification, HW-M3 is capable of capturing the complex nonlinear patterns between the WQ variables. Furthermore, among all the four models, M3 with eight inputs combination (Ca, Na, Temp, TS, Cl, NH3, pH, COD) proved merit and hence served as the most satisfactory and reliable combination for the simulation of DO concentration. The promising capability of this combination is not surprising owing to the fact that it comprises important factors affecting the performance of DO in a River (pH, Temp) as indicated in the studies by [10], [16], [17], [27]. Generally, the average

performance of the LSTM was recorded as unsatisfactory in both calibration and verification phases. Despite the predictive skills demonstrated by LSTM as one the state-of-the-art and DL models, this was not revealed in our study. This is also in line with the investigations of Vijai and Bagavathi Sivakumar [77]. Also, Zhang *et al.* [52] employed the LSTM with SVR and FFNN, and the obtained results disclosed that LSTM slightly outperformed FFNN and SVR with 0.7% and 0.17%, respectively, in terms of R2. It should be noted that the Kinta River received unprocessed sewage from various industries and agricultural activities which heavily contributed to the deterioration of WQ.

C. RESULTS OF THE ENSEMBLE MODELING
As mentioned earlier, the ensemble techniques introduced in this study are aimed at improving the accuracy of the individual models (i.e. LSTM, ELM, HW and GRNN). For this purpose, the advantages of single models are combined, and the outputs are considered as the subsequent input parameters. The two nonlinear ensemble models (BPNN-E and HW-E) were modelled using a similar method to
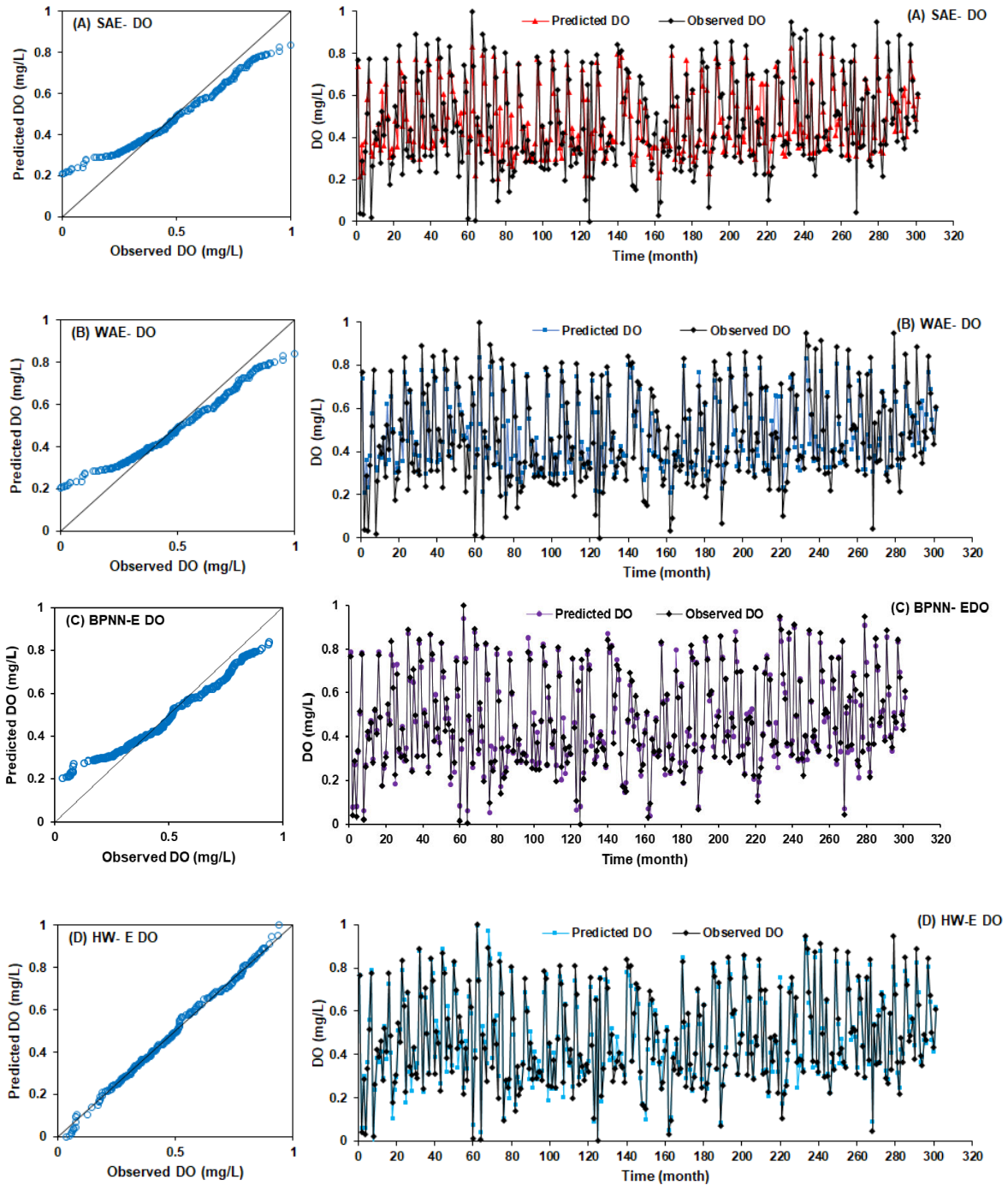
FIGURE 7. Scatter and time series graph for (A) SAE (B) WAE (C) BPNN-E (D) HW-E between the observed and predicted values.

that of the respective single models. For the two linear ensembles (SAE and WAE), the modeling was carried out using Equations (1) and (2). The performance results of the ensemble techniques are presented in Table 4. According to the results, it is evident that all four ensemble techniques showed higher performance than the single models except for HW-M3. This statement leads to the conclusion that

for the prediction of DO concentration in Kinta River, the ensemble approaches served as the most reliable method. The factors that cause the superiority of HW-M3 over the ensemble techniques could be assigned to the drawback of other single models. In other words, the predictive skills of the ensemble techniques depend on the efficiency and accuracy of each of the single models. For example, in SAE,
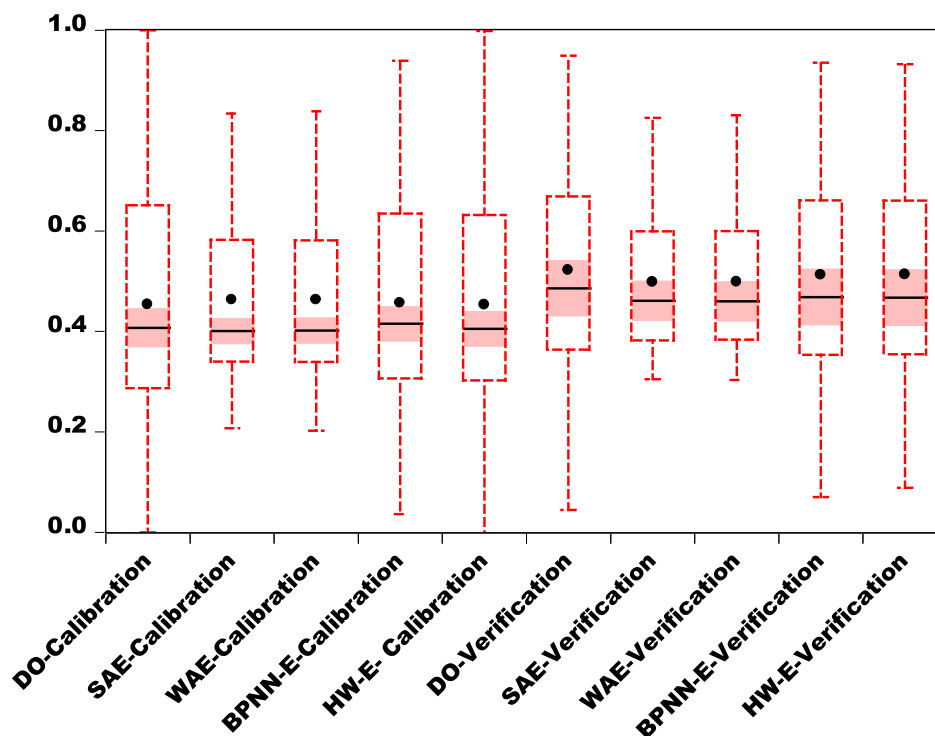
FIGURE 8. The comparison box-plot for the four ensemble models during calibration and verification phase.

averaging of all the single models is generated, while for WAE, the weights are assigned based on relative importance to enhance the prediction accuracy. These phenomena could behave as the weakness for improving the prediction accuracy of the ensemble method.

According to [74] few cases have shown that a single model can outperform the ensemble techniques. From Table 4, the direct comparison results reveal a slight increase of WAE over SAE with regard to the accuracy and BPNN-E over HW-E. It can be clearly seen that BPNN-E was more effective than other three ensembles in both calibration and verification. This robustness of the neural network ensemble was proved in different studies [40]. The visual investigation, scatter, and time series plots of the four different ensemble techniques are depicted in Fig. 7(A-D). The best predictive model emerged as BPNN-E in which the prediction pattern was closer to being in agreement with the observed DO concentration. With regard to the numerical comparison between the best single model (HW-M3), the outcomes indicated that HW-M3 yielded a high-performance accuracy up to 8% over SAE, WAE and 0.9% over HW-E, BPNN-E in terms of WI criteria in the verification phase.

Similarly, both the calibration and verification results showed negligible increases among the performance criteria of the models. However, in general, the model that gives high NSE, WI and low RMSE, MAE values should be considered as the best. With regard to the quantitative assessment of the four ensemble techniques, the predictive accuracy in terms of MAE indicates that BPNN-E decreases by 3%, 3%,

and 1% compared with SAE, WAE and HW-E, respectively, in the verification phase. The exploratory analysis for the ensemble models is better visualized through the boxplots in Fig. 8. Boxplots are powerful graphical representation of data that gives an overview and a numerical summary of a data set. According to Fig. 8, the closest of all the models to the observed values is selected to be the best model based on the mean value, the plot contained (box and whisker median, mean and staples). The extent of the spread values between the observed and predicted models indicates that BPNN-E ranked as the best model among all the models.

D. RESULT OF HYBRID RF ENSEMBLE
In this article, a hybrid of the RF ensemble (ELM-RF, LSTM-RF, GRNN-RF and HW-RF) was developed to compare the ensemble techniques discussed in section 2.6. The predictive performance of the models was evaluated using NSE, RMSE, MSE and CC. The calibration and verification results of the hybrid RF ensemble are presented in Table 5. As seen from Table 5, the HW-RF has the highest values of NSE = 0.9566 and CC = 0.981 and the lowest values of RMSE = 0.0428 and MSE = 0.018 in the verification phase. By considering the other performance criteria, the results proved the superiority of all the hybrid models in spite of the better predictive skill shown by HW-RF. Fig 9. shows the scatter plots between the observed and predicted values of the four-hybrid ensemble. According to Fig. 9, it is clear that the closeness agreement between the observed and predicted
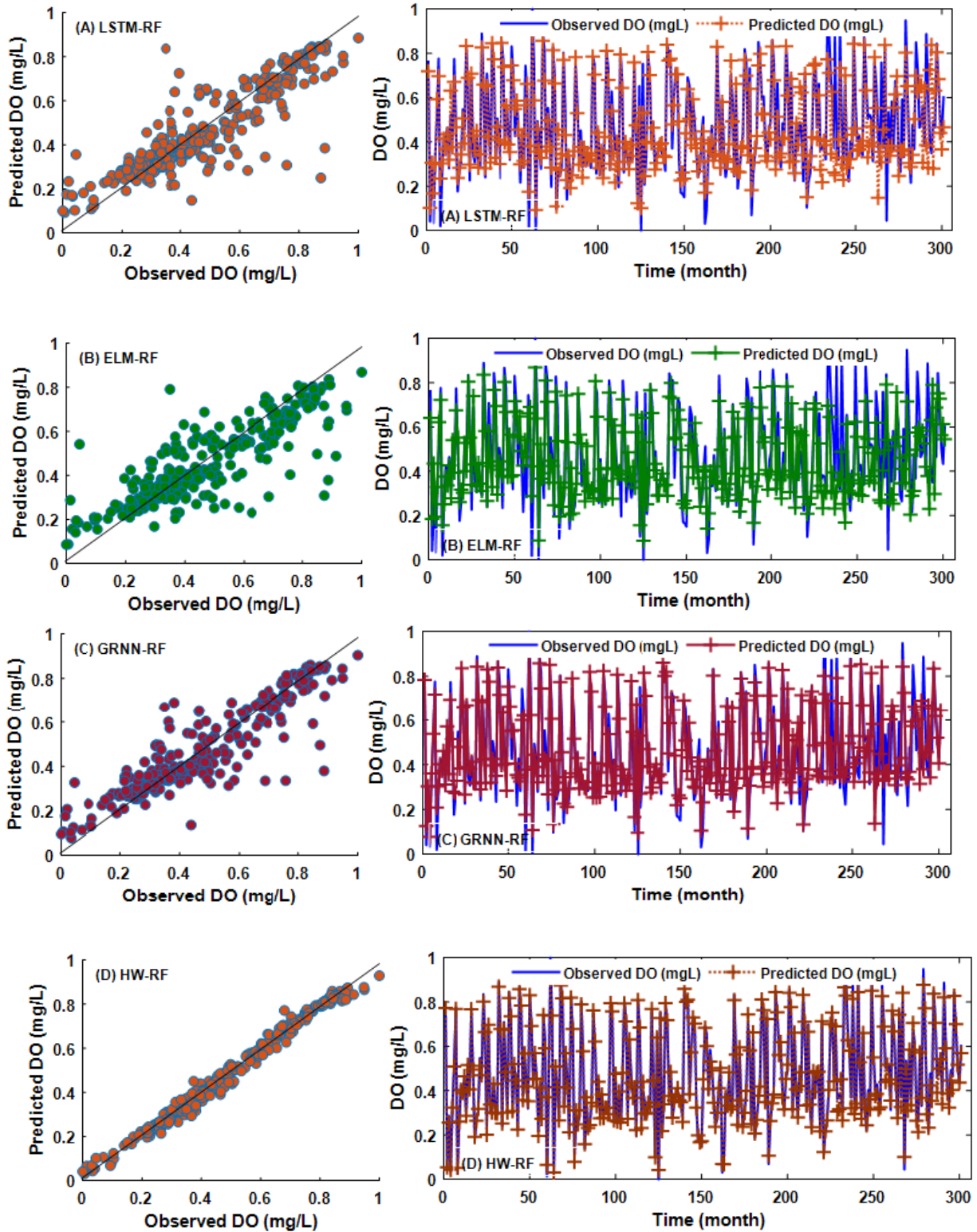
FIGURE 9. Scatter and time series graph for (A) LSTM-RF (B) ELM-RF(C) GRNN-RF (D) HW-RF between the observed and predicted values.

values were attained in the following order: HW-RF>GRNN-RF>LSTM>ELM. Furthermore, the CC value of all the models were found to be greater than 0.7, which conforms

to the conclusion reached by [31] that CC values higher than 0.70 are considered acceptable; thus the results of all four models are acceptable (see, Table 5).
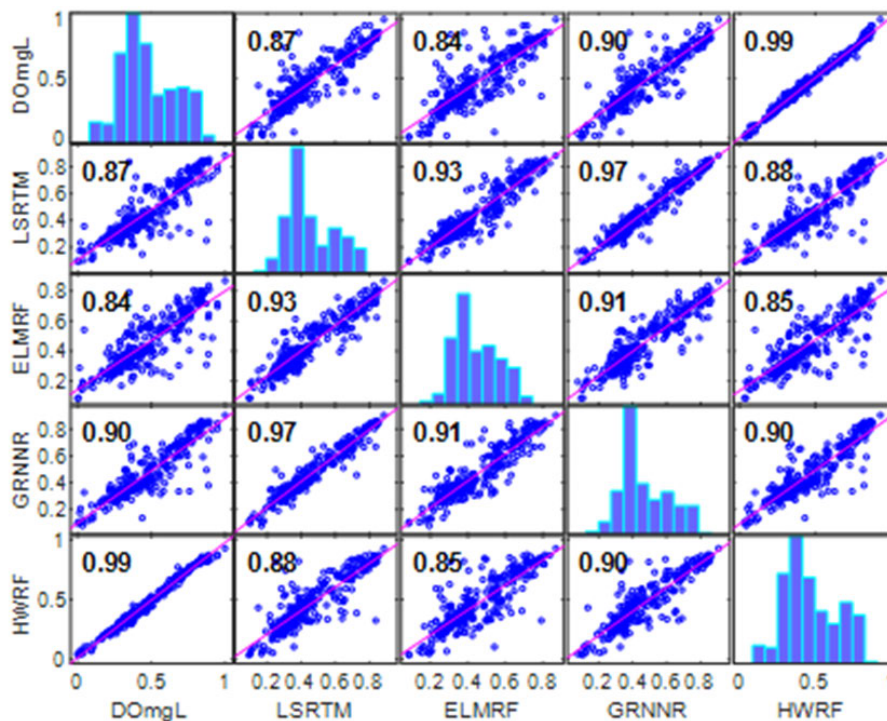
FIGURE 10. Distribution, scatterplot and CC of each predicted hybrid RF against the observed DO value.

TABLE 5. Results of hybrid RF ensemble models.

	Hybrid Models	NSE	RMSE	MSE	CC
Calibration	ELM-RF	0.8787	0.0051	0.0059	0.9374
	LSTM-RF	0.9104	0.0044	0.0043	0.9541
	GRNN-RF	0.9269	0.0040	0.0035	0.9627
	HW-RF	0.9931	0.0012	0.0003	0.9966
Verification	ELM-RF	0.6233	0.0205	0.0420	0.7895
	LSTM-RF	0.8201	0.0184	0.0338	0.9056
	GRNN-RF	0.8390	0.0161	0.0258	0.9160
	HW-RF	0.9566	0.0428	0.0018	0.9781

Generally, the comparison of the predictive performance between the two ensemble approaches (i.e. sections 2.6 and 2.7) demonstrated that the best model of hybrid RF ensemble (i.e. HW-RF) outperformed all the four-ensemble techniques (SAE, WAE, BPNN-E and HW-E). This is due to the robust nature of the ensemble RF on its own besides integrating it with highly promising nonlinear robust models (LSTM, ELM, GRNN and HW). Another factor is that RF performs a significant ensemble function by generating multiple decision trees using the randomization process; these processes produced a large ensemble of trees, and the general predictions are accomplished from the averaged results. In the same way, the comparison results showed that BPNN-E and HW-E are superior to other three hybrid RFs (i.e. LSTM-RF, ELM-RF and GRNN-RF). A closer examination of the observed DO concentration and predicted values using both the single, ensemble techniques and hybrid RF ensemble indicated the importance of employing both the ensemble techniques and hybrid RF ensemble in improving the prediction accuracy of the individual models. It is evident in Fig.10 that the distribution, scatter plot and CC and 1:1 lines of the correspondence with predicted hybrid RF were very close to the observed DO concentration. More reliable and accurate prediction of the DO concentration in a river can enable better management of the aquatic environment; as such, the ensemble techniques and hybrid RF used in this study are suitable for implementation in management practice and the other decision-making processes. However, considering the single model's results and the discrepancy in the model's performance, it should be recommended that more studies are needed using both conventional and AI integrated with optimization algorithms to bridge the variations between the measured and computed WQ variables.

Despite AI models has massive potential advantages but still suffered from certain limitations ranges from various degrees of inaccuracy and insufficiency, mostly when an extremely non-stationary hydroenvironmental process involved. Hence, the AI models may not meet the desired outcomes if there is no prior preprocessing of input-output data. On the other hands, ensemble and hybrid learning techniques is proposed in computational festimation to improve forecasting skill with their performance efficiency of capturing highly nonlinear patterns of the data.

V. CONCLUSION

The present study proposed the application of four different AI-based models, namely the LSTM, ELM, GRNN and HW models for the prediction of DO concentration in the Kinta River, Malaysia. To enhance the prediction accuracy of the single models, four different ensemble techniques were subsequently employed including two linear (SAE and WAE) and two nonlinear ensembles (BPNN-E and HW-E) and hybrid random forest (RF) ensemble were separately used for the same prediction purpose. The performance efficiency of the models was evaluated using various efficiency criteria (NSE, WI, RMSE, MAE, CC and MSE). For the pre-analysis of the data, reliability and stationary test were conducted using the Cronbach's alpha method and unit root test (i.e. using Augmented Dickey-Fuller (ADF)) to ascertain the stability of the data. Sensitivity analysis was conducted between the WQ variables using nonlinear input variables selection approaches and four different models were considered as M1, M2, M3 and M4.

The results of the single AI-based models demonstrated that HW (M3) served as the best model for predicting DO concentration followed by the GRNN (M4), ELM (M1) and LSTM (M3) models Furthermore, according to the numerical comparison between the best single model (HW-M3), the outcomes indicated that HW-M3 yielded high-performance accuracy up to 8% over SAE, WAE and 0.9% over HW-E, BPNN-E in terms of WI criteria. For the four ensemble results, BPNN-E proved superior to the other three ensembles in both calibration and verification. With regard to the quantitative assessment of the four ensemble techniques, the predictive accuracy in terms of MAE indicated that BPNN-E decreases by 3%, 3%, and 1% compared with SAE, WAE and HW-E, respectively in the verification phase. The hybrid results were better for all the hybrid models (LSTM-RF, ELM-RF, GRNN-RF and HW-RF) with the best predictive skill shown by the HW-RF ensemble.

The overall outcomes of the current study demonstrated the promising impact of the ensemble techniques and hybrid RF ensemble for the prediction of DO concentration in the Kinta River, Malaysia. Hence, the study also suggests the application of other possible alternatives of emerging optimization algorithms, deep learning models and other black box models coupled with the promising ensemble approaches to enhance the prediction accuracy. However, other hydro-environmental phenomena could also be modelled using the proposed hybrid ensemble techniques.

ACKNOWLEDGMENT

CONFLICT OF INTEREST

No

REFERENCES

[1] A. K. Verma and T. N. Singh, "Prediction of water quality from simple field parameters," *Environ. Earth Sci.*, vol. 69, no. 3, pp. 821–829, 2013.

[2] W. Lin, J. Dai, R. Liu, Y. Zhai, D. Yue, and Q. Hu, "Integrated assessment of health risk and climate effects of black carbon in the Pearl River Delta region, China," *Environ. Res.*, vol. 176, May 2019, Art. no. 108522.

[3] Y. Wu, G. Li, Y. Yang, and T. An, "Pollution evaluation and health risk assessment of airborne toxic metals in both indoors and outdoors of the Pearl River Delta, China," *Environ. Res.*, vol. 179, Jul. 2019, Art. no. 108793.

[4] S. I. Abba, M. S. Gaya, M. L. Yakubu, M. U. Zango, R. A. Abdulkadir, M. A. Saleh, A. N. Hamza, U. Abubakar, A. I. Tukur, and N. A. Wahab, "Modelling of uncertain system: A comparison study of linear and nonlinear approaches," in *Proc. IEEE Int. Conf. Autom. Control Intell. Syst. (I2CACIS)*, Jun. 2019, pp. 1–6.

[5] G. Pan, Y. Xu, Z. Yu, S. Song, and Y. Zhang, "Analysis of river health variation under the background of urbanization based on entropy weight and matter-element model: A case study in Huzhou city in the Yangtze River Delta, China," *Environ. Res.*, vol. 139, pp. 31–35, May 2015.

[6] D. Wang, D. Liu, H. Ding, V. P. Singh, Y. Wang, X. Zeng, J. Wu, and L. Wang, "A cloud model-based approach for water quality assessment," *Environ. Res.*, vol. 148, pp. 24–35, Jul. 2016.

[7] O. Kisi, N. Akbari, M. Sanatipour, A. Hashemi, K. Teimourzadeh, and J. Shiri, "Modeling of dissolved oxygen in river water using artificial intelligence techniques," *J. Environ. Inform.*, vol. 22, no. 2, 2013.

[8] O. Kisi and A. Murat, "Comparison of Ann and Anfis techniques in modeling dissolved oxygen," in *Proc. 6th Int. Water Technol. Conf.*, 2012, pp. 1–10.

[9] Z. Yaseen, M. Ehteram, A. Sharafati, S. Shahid, N. Al-Ansari, and A. El-Shafie, "The integration of nature-inspired algorithms with least square support vector regression models: Application to modeling river dissolved oxygen concentration," *Water*, vol. 10, no. 9, p. 1124, Aug. 2018.

[10] V. N. G. Elkiran, S. I. Abba, and J. Abdullahi, "Artificial intelligence-based approaches for multi-station modelling of dissolve oxygen in river," *Global J. Environ. Sci. Manage.*, vol. 4, no. 4, pp. 439–450, 2018.

[11] X. Wen, J. Fang, M. Diao, and C. Zhang, "Artificial neural network modeling of dissolved oxygen in the Heihe river, Northwestern China," *Environ. Monitor. Assessment*, vol. 185, no. 5, pp. 4361–4371, May 2013.

[12] S. Heddam, "Simultaneous modelling and forecasting of hourly dissolved oxygen concentration (DO) using radial basis function neural network (RBFNN) based approach: A case study from the Klamath river, Oregon, USA," *Model. Earth Syst. Environ.*, vol. 2, no. 3, p. 135, Sep. 2016.

[13] W.-B. Chen and W.-C. Liu, "Artificial neural network modeling of dissolved oxygen in reservoir," *Environ. Monitor. Assessment*, vol. 186, no. 2, pp. 1203–1217, Feb. 2014.

[14] K. M. A. Bayram, "Artificial neural network modeling of dissolved oxygen concentration in a Turkish Watershed," *Polish J. Environ. Stud.*, vol. 24, no. 4, pp. 1507–1515, 2015.

[15] B. Keshtegar and S. Heddam, "Modeling daily dissolved oxygen concentration using modified response surface method and artificial neural network: A comparative study," *Neural Comput. Appl.*, vol. 30, no. 10, pp. 2995–3006, Nov. 2018.

[16] X. Ji, X. Shang, R. A. Dahlgren, and M. Zhang, "Prediction of dissolved oxygen concentration in hypoxic river systems using support vector machine: A case study of Wen-Rui Tang river, China," *Environ. Sci. Pollut. Res.*, vol. 24, no. 19, pp. 16062–16076, Jul. 2017.

[17] E. Olyaie, H. Z. Abyaneh, and A. D. Mehr, "A comparative analysis among computational intelligence techniques for dissolved oxygen prediction in Delaware river," *Geosci. Frontiers*, vol. 8, no. 3, pp. 517–527, May 2017.

[18] S. Malek, M. Mosleh, and S. M. Syed, "Dissolved oxygen prediction using support vector machine," *Int. J. Comput. Inf., Syst. Control Eng.*, vol. 8, no. 1, pp. 46–50, 2014.

[19] A. S. A. Yahya, A. N. Ahmed, F. B. Othman, R. K. Ibrahim, H. A. Afan, A. El-Shafie, C. M. Fai, M. S. Hossain, M. Ehteram, and A. Elshafie, "Water quality prediction model based support vector machine model for Ungauged river catchment under dual scenarios," *Water*, vol. 11, no. 1231, pp. 1–16, 2019.

[20] C. Li, Z. Li, J. Wu, L. Zhu, and J. Yue, "A hybrid model for dissolved oxygen prediction in aquaculture based on multi-scale features," *Inf. Process. Agricult.*, vol. 5, no. 1, pp. 11–20, Mar. 2018.

[21] V. Ranković, J. Radulović, I. Radojević, A. Ostojić, and L. Čomić, "Prediction of dissolved oxygen in reservoirs using adaptive network-based fuzzy inference system," *J. Hydroinf.*, vol. 14, no. 1, pp. 167–179, Jan. 2012.

[22] S. Emamgholizadeh, H. Kashi, I. Marofpoor, and E. Zalaghi, "Prediction of water quality parameters of Karoon river (Iran) by artificial intelligence-based models," *Int. J. Environ. Sci. Technol.*, vol. 11, no. 3, pp. 645–656, Apr. 2014.

[23] A. Najah, A. El-Shafie, O. A. Karim, and A. H. El-Shafie, "Performance of ANFIS versus MLP-NN dissolved oxygen prediction models in water quality monitoring," *Environ. Sci. Pollut. Res.*, vol. 21, no. 3, pp. 1658–1670, Feb. 2014.

[24] 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," *Procedia Comput. Sci.*, vol. 120, pp. 75–82, 2017.

[25] S. Liu, H. Tai, Q. Ding, D. Li, L. Xu, and Y. Wei, "A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction," *Math. Comput. Model.*, vol. 58, nos. 3–4, pp. 458–465, Aug. 2013.

[26] O. Kisi, N. Akbari, M. Sanatipour, A. Hashemi, K. Teimourzadeh, and J. Shiri, "Modeling of dissolved oxygen in river water using artificial intelligence techniques," *J. Environ. Informat.*, vol. 22, no. 2, pp. 92–101, Dec. 2013.

[27] P. Shi, G. Li, Y. Yuan, G. Huang, and L. Kuang, "Prediction of dissolved oxygen content in aquaculture using clustering-based softplus extreme learning machine," *Comput. Electron. Agricult.*, vol. 157, pp. 329–338, Dec. 2019.

[28] B. Shi, P. Wang, J. Jiang, and R. Liu, "Applying high-frequency surrogate measurements and a wavelet-ANN model to provide early warnings of rapid surface water quality anomalies," *Sci. Total Environ.*, vols. 610–611, pp. 1390–1399, Jan. 2018.

[29] D. Ö. Faruk, "A hybrid neural network and ARIMA model for water quality time series prediction," *Eng. Appl. Artif. Intell.*, vol. 23, no. 4, pp. 586–594, Jun. 2010.

[30] Z. Chengyun, W. Rong, and T. Qiaoying, "Prediction model of dissolved oxygen based on SADE-RVM," in *Proc. 3rd Int. Conf. Robot., Control Autom. (ICRCA)*, 2018, pp. 131–134.

[31] S. Zhu and S. Heddam, "Prediction of dissolved oxygen in urban rivers at the Three Gorges Reservoir, China: Extreme learning machines (ELM) versus artificial neural network (ANN)," *Water Qual. Res. J.*, vol. 55, no. 1, pp. 1–13, 2019.

[32] X. Ta and Y. Wei, "Research on a dissolved oxygen prediction method for recirculating aquaculture systems based on a convolution neural network," *Comput. Electron. Agricult.*, vol. 145, pp. 302–310, Dec. 2018.

[33] P. Liu, J. Wang, A. K. Sangaiah, Y. Xie, and X. Yin, "Analysis and prediction of water quality using LSTM deep neural networks in IoT environment," *Sustainability*, vol. 11, no. 7, pp. 1–14, 2019.

[34] L. Yalin, W. Yaoguang, and C. Yingy, "Dissolved oxygen prediction model which based on fuzzy neural network," in *Proc. IFIP Adv. Inf. Commun. Technol.*, vol. 420, 2014, pp. 544–551.

[35] M. Zounemat-Kermani and M. Scholz, "Modeling of dissolved oxygen applying stepwise regression and a template-based fuzzy logic system," *J. Environ. Eng.*, vol. 140, no. 1, pp. 69–76, Jan. 2014.

[36] E. Giusti and S. Marsili-Libelli, "Spatio-temporal dissolved oxygen dynamics in the orbetello lagoon by fuzzy pattern recognition," *Ecol. Model.*, vol. 220, no. 19, pp. 2415–2426, Oct. 2009.

[37] S. Heddam, "Generalized regression neural network (GRNN)-based approach for colored dissolved organic matter (CDOM) retrieval: Case study of Connecticut river at Middle Haddam Station, USA," *Environ. Monitor. Assessment*, vol. 186, no. 11, pp. 7837–7848, Nov. 2014.

[38] S. Nemati, M. H. Fazelifard, Ö. Terzi, and M. A. Ghorbani, "Estimation of dissolved oxygen using data-driven techniques in the Tai Po River, Hong Kong," *Environ. Earth Sci.*, vol. 74, no. 5, pp. 4065–4073, Sep. 2015.

[39] M. J. Alizadeh and M. R. Kavianpour, "Development of wavelet-ANN models to predict water quality parameters in Hilo Bay, Pacific Ocean," *Mar. Pollut. Bull.*, vol. 98, nos. 1–2, pp. 171–178, Sep. 2015.

[40] E. Sharghi, V. Nourani, and N. Behfar, "Earthfill dam seepage analysis using ensemble artificial intelligence based modeling," *J. Hydroinf.*, vol. 20, no. 5, pp. 1071–1084, Sep. 2018.

[41] N. Raj Kiran and V. Ravi, "Software reliability prediction by soft computing techniques," *J. Syst. Softw.*, vol. 81, no. 4, pp. 576–583, Apr. 2008.

[42] Y. Khan and S. S. Chai, "Ensemble of ANN and ANFIS for water quality prediction and analysis—A data driven approach," *J. Telecommun., Electron. Comput. Eng.*, vol. 9, no. 2, pp. 117–122, 2016.

[43] V. Nourani, G. Elkiran, and S. I. Abba, "Wastewater treatment plant performance analysis using artificial intelligence—An ensemble approach," *Water Sci. Technol.*, vol. 78, no. 10, pp. 2064–2076, Dec. 2018.

[44] V. Nourani, S. Uzelaltinbulat, F. Sadikoglu, and N. Behfar, "Artificial intelligence based ensemble modeling for multi-station prediction of precipitation," *Atmosphere*, vol. 10, no. 2, pp. 1–28, 2019.

[45] V. Nourani, G. Elkiran, and J. Abdullahi, "Multi-station artificial intelligence based ensemble modeling of reference evapotranspiration using pan evaporation measurements," *J. Hydrol.*, vol. 577, Mar. 2019, Art. no. 123958.

[46] G. Elkiran, V. Nourani, and S. I. Abba, "Multi-step ahead modelling of river water quality parameters using ensemble artificial intelligence-based approach," *J. Hydrol.*, vol. 577, Jul. 2019, Art. no. 123962.

[47] 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.

[48] M. Liu and J. Lu, "Support vector machine—An alternative to artificial neuron network for water quality forecasting in an agricultural nonpoint source polluted river?" *Environ. Sci. Pollut. Res.*, vol. 21, no. 18, pp. 11036–11053, Sep. 2014.

[49] Q. B. Pham, S. I. Abba, A. G. Usman, N. T. T. Linh, V. Gupta, A. Malik, R. Costache, N. D. Vo, and D. Q. Tri, "Potential of hybrid data-intelligence algorithms for multi-station modelling of rainfall," *Water Resour. Manage.*, vol. 33, no. 15, pp. 5067–5087, Dec. 2019.

[50] Z. M. Yaseen, S. O. Sulaiman, R. C. Deo, and K.-W. Chau, "An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction," *J. Hydrol.*, vol. 569, pp. 387–408, Oct. 2019.

[51] L. Yunpeng, H. Di, B. Junpeng, and Q. Yong, "Multi-step ahead time series forecasting for different data patterns based on LSTM recurrent neural network," in *Proc. 14th Web Inf. Syst. Appl. Conf. (WISA)*, Jan. 2017, pp. 305–310.

[52] D. Zhang, E. S. Hølland, G. Lindholm, and H. Ratnaweera, "Hydraulic modeling and deep learning based flow forecasting for optimizing inter catchment wastewater transfer," *J. Hydrol.*, vol. 567, pp. 792–802, Nov. 2018.

[53] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.

[54] Z. M. Yaseen, M. F. Allawi, A. A. Yousif, O. Jaafar, F. M. Hamzah, and A. El-Shafie, "Non-tuned machine learning approach for hydrological time series forecasting," *Neural Comput. Appl.*, vol. 30, no. 5, pp. 1479–1491, 2016.

[55] S. I. Abba, Q. B. Pham, A. G. Usman, N. T. T. Linh, D. S. Aliyu, Q. Nguyen, and Q.-V. Bach, "Emerging evolutionary algorithm integrated with kernel principal component analysis for modeling the performance of a water treatment plant," *J. Water Process Eng.*, vol. 33, Oct. 2020, Art. no. 101081.

[56] M. S. Gaya, M. U. Zango, L. A. Yusuf, M. Mustapha, B. Muhammad, A. Sani, A. Tijjani, N. A. Wahab, and M. T. M. Khairi, "Estimation of turbidity in water treatment plant using Hammerstein–Wiener and neural network technique," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 5, no. 3, pp. 666–672, 2017.

[57] S. I. Abba, V. Nourani, and G. Elkiran, "Multi-parametric modeling of water treatment plant using AI-based non-linear ensemble," *J. Water Supply, Res. Technol.-Aqua*, vol. 68, no. 7, pp. 547–561, Nov. 2019.

[58] B. B. Gulyani and A. Fathima, "Introducing ensemble methods to predict the performance of waste water treatment plants (WWTP)," *Int. J. Environ. Sci. Develop.*, vol. 8, no. 7, pp. 501–506, 2017.

[59] N. M. Baba, M. Makhtar, S. Abdullah, and M. K. Awang, "Current issues in ensemble methods and its applications," *J. Theor. Appl. Inf. Technol.*, vol. 81, no. 2, pp. 266–276, 2015.

[60] V. Nourani, H. Gökçekuş, I. K. Umar, and H. Najafi, "An emotional artificial neural network for prediction of vehicular traffic noise," *Sci. Total Environ.*, vol. 707, Mar. 2020, Art. no. 136134.

[61] Z.-Y. Chen, T.-H. Zhang, R. Zhang, Z.-M. Zhu, J. Yang, P.-Y. Chen, C.-Q. Ou, and Y. Guo, "Extreme gradient boosting model to estimate PM2.5 concentrations with missing-filled satellite data in China," *Atmos. Environ.*, vol. 202, pp. 180–189, Apr. 2019.

[62] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[63] J. C.-W. Chan and D. Paelinckx, "Evaluation of random forest and AdaBoost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery," *Remote Sens. Environ.*, vol. 112, no. 6, pp. 2999–3011, Jun. 2008.

[64] A. Liaw and M. Wiener, "Classification and Regression by random forest," *R News*, vol. 2, no. 3, pp. 18–22, Dec. 2002.

[65] M. Al-Mukhtar, "Random forest, support vector machine, and neural networks to modelling suspended sediment in tigris river-baghdad," *Environ. Monitor. Assessment*, vol. 191, no. 11, p. 673, Nov. 2019.

[66] H. Mohammed, I. A. Hameed, and R. Seidu, "Random forest tree for predicting fecal indicator organisms in drinking water supply," in *Proc. Int. Conf. Behav., Econ., Socio-cultural Comput. (BESC)*, Jan. 2017, pp. 1–6.

[67] D. R. Legates and G. J. McCabe, "Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation," *Water Resour. Res.*, vol. 35, no. 1, pp. 233–241, Jan. 1999.

[68] N. M. Gazzaz, M. K. Yusoff, A. Z. Aris, H. Juahir, and M. F. Ramli, "Artificial neural network modeling of the water quality index for Kinta river (Malaysia) using water quality variables as predictors," *Mar. Pollut. Bull.*, vol. 64, no. 11, pp. 2409–2420, Nov. 2012.

[69] S. Y. Muhammad, M. Makhtar, A. Rozaimee, A. Abdul, and A. A. Jamal, "Classification model for water quality using machine learning techniques," *Int. J. Softw. Eng. Appl.*, vol. 9, no. 6, pp. 45–52, Jun. 2015.

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

[71] D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *J. Amer. Stat. Assoc.*, vol. 74, no. 366a, pp. 427–431, 1979.

[72] V. Nourani and M. Sayyah, "Advances in engineering software sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes," *Adv. Eng. Softw.*, vol. 47, no. 1, pp. 127–129, 2012.

[73] V. Nourani, G. Elkiran, J. Abdullahi, and A. Tahsin, "Multi-region modeling of daily global solar radiation with artificial intelligence ensemble," *Natural Resour. Res.*, vol. 28, no. 4, pp. 1217–1238, Oct. 2019.

[74] R. Eisinga, M. T. Grotenhuis, and B. Pelzer, "The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown?" *Int. J. Public Health*, vol. 58, no. 4, pp. 637–642, Aug. 2013.

[75] V. Nourani, "An emotional ANN (EANN) approach to modeling rainfall-runoff process," *J. Hydrol.*, vol. 544, pp. 267–277, Jan. 2017.

[76] M. Zounemat-kermani, Y. Seo, and S. Kim, "Can decomposition approaches always enhance soft computing models? Predicting the dissolved oxygen concentration in the St. Johns River, Florida," *Appl. Water*, vol. 9, no. 2534, pp. 2–24, 2019.

[77] P. Vijai and P. Bagavathi Sivakumar, "Performance comparison of techniques for water demand forecasting," *Procedia Comput. Sci.*, vol. 143, pp. 258–266, Jan. 2018.

NGUYEN THI THUY LINH received the master's degree in the International Master Program on Natural Hazards Mitigation and Management (iNHMM) and the Ph.D. degree from National Cheng Kung University. Her Ph.D. thesis was related to water resources management. Her research interests include developing novel technologies for water management and hazard mitigation.

JAZULI ABDULLAHI received the B.Sc. degree from KUST, Wudil, Nigeria, in 2012, and the M.S. degree from NEU, Cyprus, in 2016, where he is currently pursuing the Ph.D. degree with the Civil and Environmental Engineering Faculty. His research interests include hydro climatological studies, water and waste water management, and artificial intelligence.

SHABAN ISMAEL ALBRKA ALI received the M.Sc. and Ph.D. degrees from the National University of Malaysia (UKM), Bangi, Malaysia. He is currently a Lecturer with the Department of Civil Engineering, Faculty of Civil Environmental Engineering, Near East University. His principal research interests include transportation and highway engineering and artificial intelligence applications.

QUOC BAO PHAM received the Ph.D. degree in machine learning and water resources management from National Cheng Kung University, Taiwan. He is currently a Researcher with the Environmental Quality, Atmospheric Science and Climate Change Research Group, Ton Duc Thang University, Ho Chi Minh City, Vietnam. He has collaborated actively with researchers in several other disciplines of engineering.

SANI ISAH ABBA received the B.Sc. degree from Bayero University Kano (BUK), the M.Tech. degree from Sharda University, India, and the Ph.D. degree from Near East University (NEU), Cyprus. He is currently a Researcher with Yusuf Maitama Sule University, Kano, Nigeria, and the DESAM Institute Bioinformatics and Computational Biology Research, NEU. His research interests include water resources, hydrology engineering, hydro-climatology, hydro-environmental modeling, computational engineering, soft computing, and artificial intelligence. In addition, he has an excellent expertise in machine learning, GIS, and advance data analysis and classification.

RABIU ALIYU ABDULKADIR received the bachelor's degree in electrical engineering from the Kano University of Science and Technology, Wudil, Nigeria, and the master's degree in instrumentation and control from Sharda University, India. He is currently a Lecturer with the Department of Electrical Engineering, Kano University of Science and Technology. His research interests include control systems design, robotics, computer vision, image processing, and artificial intelligence.
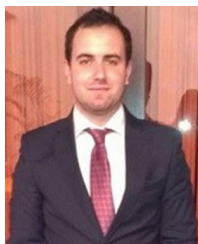
ROMULUS COSTACHE received the B.Sc., M.Sc., and Ph.D. degrees in geography from the University of Bucharest, Romania. He is currently a Researcher with the National Institute of Hydrology and Water Management, Romania, and also a Postdoctoral Researcher with the Research Institute of University of Bucharest. His research interests include hydrology, natural hazards, geographic information science, bivariate statistics, machine learning, and artificial intelligence applied in the natural hazards susceptibility assessment.

VAN THAI NAM received the master's degree from Vietnam National University Ho Chi Minh City (VNUHCM), Vietnam, in 2003, and the Ph.D. degree from Osaka University (Handai), Japan, in 2011. He is currently an Associate Professor with the Institute of Applied Sciences (HIAS), Ho Chi Minh City University of Technology (HUTECH). He is also working as the Deputy Director of HIAS and a Senior Lecturer in the field of sustainable energy, environmental engineering, and management. His research interests include environmental management systems, environmental engineering, water and soil resources engineering, environmental, and health risk assessment.

DUONG TRAN ANH received the master's degree from the Water Engineering and Management, Asian Institute of Technology, and the Ph.D. degree from the Technical University of Munich, Germany. His Ph.D. thesis was related to water resources management, climate change, and artificial intelligence in Mekong Delta. He is currently a Postdoctoral Researcher with the Institute of Applied Sciences (HIAS), Ho Chi Minh City University of Technology (HUTECH). He has collaborated actively with researchers in several other disciplines of computer science, climatology, and computational data. He is in charge for the Research Group of Artificial Intelligence and Water Resources Engineering, as a Coordinator. He also collaborates with many colleagues from U.K., The Netherlands, Singapore, and USA. His research interests include hydrological, hydrodynamic modeling, downscaling and climate change, and artificial intelligence.

● ● ●