

Multi-parametric modeling of water treatment plant using AI-based non-linear ensemble

S. I. Abba, Vahid Nourani and Gozen Elkiran

ABSTRACT

In this research, general regression neural network (GRNN), Hammerstein-wiener (HW) and non-linear autoregressive with exogenous (NARX) neural network, least square support vector machine (LSSVM) models were employed for multi-parametric (Hardness (mg/L), turbidity (Turb) ($\mu\text{s}/\text{cm}$), pH and suspended solid (SS) (mg/L)) modeling of Tamburawa water treatment plant (TWTP) at Kano, Nigeria. The weekly data from the TWTP were used and the predictive models were evaluated based on several numerical indicators. Four different non-linear ensemble techniques (GRNN-E, HW-E, NARX-E, and LSSVM-E) were subsequently employed. The comparison of the results of modeling showed that HW served as the best model for the simulation of Hardness, Turb, and SS while the NARX model demonstrated high capability in predicting pH. Yet, the HW and NARX as system identification techniques attained best overall predictive performance among the four modeling approaches. The HW model offers, therefore, a reliable and intelligent approach for predicting the treated Hardness, Turb, and SS and should be explored as a predictive tool in TWTP. Among the non-linear ensemble models, GRNN-E proved of high merit and increased the accuracy of the best single models significantly up to 30% for Hardness and Turb, 34% for pH, and 37% for SS with regards to the single models.

Key words | artificial intelligence, non-linear ensemble, water treatment plant

S. I. Abba

Department of Physical Planning Development and Maintenance,
Yusuf Maitama Sule University,
Kano,
Nigeria

Vahid Nourani

Gozen Elkiran (corresponding author)
Faculty of Civil and Environmental Engineering,
Near East University,
Near East Boulevard, Nicosia 99138, North Cyprus
via Mersin 10,
Turkey
E-mail: gozen.elkiran@neu.edu.tr

Vahid Nourani

Center of Excellence in Hydroinformatics,
Faculty of Civil Engineering,
University of Tabriz,
Tabriz,
Iran

INTRODUCTION

Water is essential to sustain life, and an affordable and adequate supply of water must be available (World Health Organization (WHO) and United Nation Children Fund (UNICEF) 2012). Municipal or drinking water treatment plants (MWTP) are operated to remove bacteria, solids, micro-organisms, and other contaminants from untreated water through various approaches such as coagulation, filtration, etc. (Ogwueleka & Ogwueleka 2009). A satisfactory water treatment plant (WTP) is vital in order to avoid the discharge of pollutants and meet the required standards by law. The multi-parametric combination of physical, chemical, and biological characteristics are often the major factors affecting the operation and control of WTPs (Nourani *et al.* 2018).

Improper control for standard limits of different parameters in WTPs may lead to serious environmental and human health problems. As such, the proper control may be achieved by introducing a robust tool for modeling the WTP performance based on the importance of various key parameters. However, the WTP process is complex in nature due to the dilute mixture of numerous compositions, quality, and characteristic of the system which result in the difficulty of modeling WTP parameters (Belia *et al.* 2009). In addition, the biochemical and physical nature of WTPs exhibit non-linear phenomena which are too complicated to simulate by simple linear principles or mathematical models (Gaya *et al.* 2017). In the last decades, different linear models have been widely presented for managing

the overall performance of WTPs, but most of them have limitations and are incapable of meeting the standards of modeling non-linear systems (Falah Nezhad *et al.* 2016). On the other hand, non-linear artificial intelligence (AI) models such as artificial neural network (ANN), adaptive neurofuzzy inference system (ANFIS), and support vector machine (SVM) have shown merit in modeling non-linear systems of WTPs (Mashaly & Alazba 2017; Rahimi *et al.* 2019).

For instance, Al-Baidhani & Alameedee (2017) developed an ANN model to predict the effluent pH and turbidity using various measured input parameters such as pH, temperature, and dose turbidity. The results proved the suitability of ANN in modeling the WTP parameters. Wu & Lo (2008) used the ANN and ANFIS models to compute the real-time coagulant dosage in WTPs using the measured turbidity, pH, and color. The outcomes demonstrated that the ANFIS model was capable to accurately predict the coagulant dosage with regards to the ANN model. Zhang & Stanley (1999) employed the ANN model with process control for predicting the treated turbidity. Maleki *et al.* (2018) predicted the influent parameters in WTP using auto regressive integrated moving average (ARIMA) and neural network auto-regression (NNAR) models, and despite an acceptable performance of ARIMA model, the results observed better prediction performance for NNAR with regard to ARIMA. Gaya *et al.* (2017) demonstrated the application of ANN and Hammerstein-Wiener (HW) models for forecasting the influent turbidity in WTPs using different input parameters. The simulated results indicated that ANN could outperform the HW model and may serve as an acceptable tool for modeling the turbidity of WTPs.

According to the mentioned literature review, it is obvious that several studies using AI models have been conducted and show promising performance, each model for a specific case. Despite the fact there is no exceptional model that justifies superiority to others, applying the knowledge of ensemble techniques could lead to more promising outcomes. The ensemble techniques have been used in various fields of engineering (Sharghi *et al.* 2018). Nourani *et al.* (2018) demonstrated a combination of different AI-based single models in an ensemble modeling parameter to predict the performance of a wastewater treatment plant. The obtained results showed that ensemble techniques

outperform the AI-based models. Kim & Seo (2015) developed, respectively, ensemble methods using data mining techniques to simulate biological oxygen demand (BOD) and various water quality (WQ) parameters. However, since the pronouncement of the non-linear ensemble (NLE) approach, to the best of the authors' knowledge, there is no published research indicating the application of different AI-based models integrated with several non-linear ensemble techniques (NETs) in WTPs. Hence, the present work is focused on the application of four types of AI-based models (i.e., general regression neural network (GRNN), HW, non-linear autoregressive with exogenous (NARX) neural network, least square support vector machine (LSSVM) model for multi-parametric modeling in Tamburawa WTP, Kano, Nigeria. For this purpose, various combinations of the input parameters based on their significance are imposed into the models. In order to improve the overall performance of the prediction, NLE modeling is also employed using the outputs of the single models. In this way, three different types of black-box models, GRNN, LSSVM, HW, and NARX ensemble kernels (GRNN-E, LSSVM-E, HW-E, and NARX-E) are used and the results are compared together. Other feasible alternative single models may also be used, such as genetic programming, ARIMA models (Olyaie *et al.* 2015; Nabavi-Pelesaraei *et al.* 2017) and for an ensemble, wavelet-ANN (Shamshirband *et al.* 2019) and non-linear neural network ensemble (Elkiran *et al.* 2019) were also employed.

MATERIALS AND METHODS

Plant description and data used

Kano is the most populous state in Nigeria, in which drainage river water is conventionally treated in a plant exceeding the minimum standard WQ of WHO. The Tamburawa water treatment plant (TWTP) has a capacity of producing 150 ML potable water per day to cover communities in Kano city and the surroundings. The raw water from the source is pumped via pump station and enters a preliminary treatment unit where grits and some of the suspended solids are removed to avoid pump wear and pipe deterioration. The water is then clarified as shown in Figure 1(a), the clarifier retains and settles the water to

remove settleable organic and floatable solids by gravity. Figure 1(b) shows the schematic flow chart of the operational process. The operational process comprises the stages of rapid mix, coagulation/flocculation, sedimentation, filtration, disinfection, and final treated water, which can be distributed to different sources such as domestic, commercial, and institutional (Gaya et al. 2017).

The selected plant covers a significant profile interim of water demand and WQ across the state. The data and statistical analysis of the input–output parameters are essential steps in any AI-based modeling because such analysis demonstrates the type and strength of the relations between inputs and output. As in Equation (1), normalization of the data was first conducted before the model calibration steps:

$$y = 0.05 + \left(0.95 \times \left(\frac{x - x_{min}}{x_{max} - x_{min}} \right) \right) \quad (1)$$

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

The obtained weekly data set from the TWTP for the year 2015 are measured raw and treated pH (pH_R and pH_T), turbidity ($Turb_R$ and $Turb_T$) ($\mu\text{S}/\text{cm}$), conductivity ($Cond_R$ and $Cond_T$) (mS/cm), total dissolved solids (TDS_R and TDS_T) (mg/L), suspended solids (SS_R and SS_T) (mg/L), hardness ($Hardness_R$ and $Hardness_T$) (mg/L), chloride (Cl_R and Cl_T) (mg/L), and iron (Fe_R and Fe_T) (mg/L). Table 1 shows the descriptive statistical analysis used for this study of the data. The data set was divided into two parts, 75% for the calibration and the remaining 25% were used in the verification phase. As a very common and classic method of data splitting, we used simple random selection of train and test data samples. However, it is highly recommended that in order to have homogeneous data distribution in both training and test data sets, after splitting the data set, the statistical characteristics of both subsets should be checked to have almost the same statistics (mean, standard deviation, etc.) in both subsets to guarantee the generalization of the model. In order to see that the data splitting strategy does not affect the overall performance of the modeling, via the cross-validation method, such

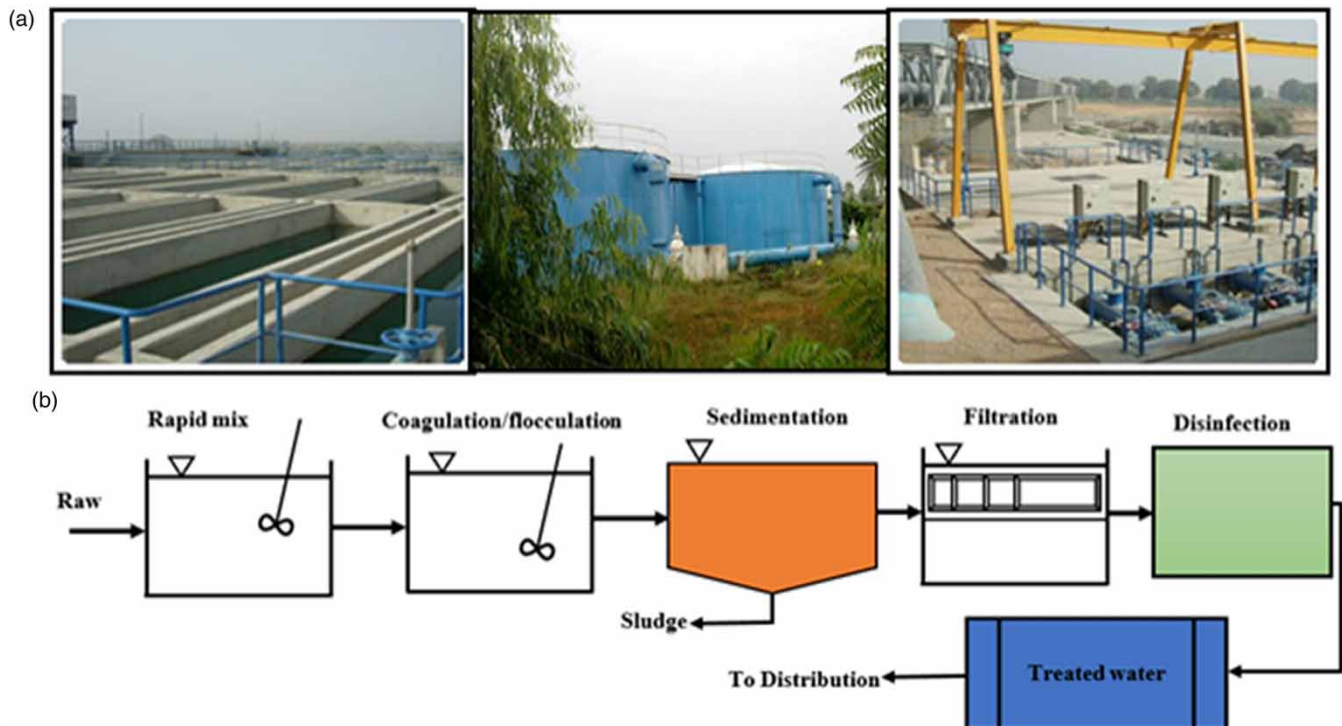


Figure 1 | (a) Clarifier, sedimentation tank and pump station, and (b) operational process of the plant.

Table 1 | Descriptive statistics of the data

Parameters	Mean	Max	Min	SD	Coef. skewness
pH _{Raw}	7.6813	11.5000	6.5000	0.6649	2.1368
Turb _{Raw} (mg/L)	200.7667	1,796.0000	51.0000	220.2917	4.7173
Cond _{Raw} (mS/cm)	116.5067	257.0000	53.0000	43.7466	0.5669
TDS _{Raw} (mg/L)	57.4705	106.400	17.9000	20.7953	0.2955
Hardness _{Raw} (mg/L)	36.3028	53.8700	24.6100	6.33222	0.8852
Cl _{Raw} (mg/L)	13.4560	33.5600	8.8800	4.1255	2.7859
SS _{Raw} (mg/L)	154.200	1,248.0000	34.0000	161.4377	4.1678
Fe _{Raw} (mg/L)	3.8977	32.0000	1.0400	7.2880	3.0713
pH _{Treated}	6.4370	8.8000	5.0000	0.6521	0.6055
Turb _{Treated} (mg/L)	0.8181	3.4100	0.2000	0.5390	2.7709
Cond _{Treated} (mS/cm)	160.8732	363.0000	8.1900	60.6630	0.8620
TDS _{Treated} (mg/L)	80.1022	181.1000	4.2000	29.8376	0.8334
Hardness _{Treated} (mg/L)	29.0826	44.8900	17.9600	5.1921	0.5919
Cl _{Treated} (mg/L)	37.8045	123.0000	11.8400	18.1758	1.7205
SS _{Treated} (mg/L)	0.8222	2.0000	0.000	0.6284	0.1481
Fe _{Treated} (mg/L)	0.1011	1.0100	0.0100	0.1155	5.7892

random data splitting is repeated several times to see whether there any large differences among the obtained performance metrics in different runs or not; if not, the model generalization is good and the quantity and quality of the data are appropriate.

Selection of the dominant inputs is one of the essential parts in AI-based modeling. As such, a suitable combination of inputs is determined (see Table 2). The holdout validation method was applied and considered as another version of k-fold cross-validation. In the validation process, different types of validation approach can be applied including cross-validation, which is called k-fold cross-validation, others are holdout, leave one out, and so on. The holdout method is also considered as the simple version of k-fold where the data are usually portioned randomly into two sets known as training and testing phase (Nourani et al. 2018).

The crux of selecting individual inputs, for example, hardness, turbidity, pH, and suspended solids parameters to simulate the performance of WTP is that these variables are usually utilized and considered when determining the performance of the treated water from the plant. The hardness is a common quality of water and contains dissolved compounds of calcium and magnesium and, sometimes, other divalent and trivalent metallic elements such as the

concentration of hydrogen ions that indicates the acidic and basic level of the solution. It is considered to be of great practical importance as it influences most of the chemical and biochemical reactions (Elkiran et al. 2018). Measuring suspended solids (SS) in water is used for control of various treatment processes and for the examination of wastewater quality. SS is considered to be one of the major pollutants that contributes to the deterioration of WQ, contributing to higher costs for water treatment, and the general esthetics of the water. Turbidity mostly provides cover and food for pathogens. If not effectively removed, turbidity can cause an outburst of waterborne diseases (Gaya et al. 2017).

The stability analysis is an important issue in both data-driven algorithms and numerical solution of initial-value problems. It has been studied extensively in the past for a variety of engineering problems, especially for input selection approach. The motivation behind exploring the stability is to provide evidence that the selected features are relatively robust to slight changes in the data (Křížek et al. 2007). However, different stability selection approaches are applied in an effort to reduce the noise and dimensionality of such data sets. Kalousis et al. (2007) measured the stability in terms of Pearson and Spearman correlation. Nourani & Fard (2012) introduced the stability

Table 2 | Pearson correlation matrix between raw and treated parameters

Parameters	pH _r	Turb _r	Cond _r	TDS _r	Hardn _r	Cl _r	SS _r	Fe _r	pH _T	Turb _T	Hardn _T	SS _T
pH _r	0.032525											
Turb _r	-0.33195	1										
Cond _r	-0.29748	-0.10229	1									
TDS _r	0.001972	0.438722	0.226554	1								
Hardn _r	0.048784	0.321302	-0.00621	0.031401	1							
Cl _r	0.005969	0.970237	-0.08517	-0.12697	0.365937	1						
SS _r	-0.34289	0.097055	0.166551	0.209041	-0.09534	0.076946	1					
Fe _r	0.461364	-0.15327	-0.01035	-0.02636	-0.04034	0.022556	-0.1392	1				
pH _T	-0.04921	0.448457	-0.05912	-0.10834	0.169559	0.200926	0.417396	0.092185	0.011003	1		
Turb _T	-0.04240	0.464929	0.201475	0.157669	0.581288	0.251558	0.407611	0.138094	-0.29883	0.26419	1	
Hardn _T	-0.07740	0.207289	-0.07127	0.015499	0.116308	0.051853	0.187615	0.100166	-0.03585	0.6448656	0.13846	1
SS _T												

features selection approach using sensitivity analysis. Similarly, Křížek et al. (2007), Nourani et al. (2019), Nagana et al. (2019), and Meinshausen & Bühlmann (2010) reported different input stability selection approaches. In this paper, both the Pearson and Spearman correlation and non-linear stability of input variables selection approach were assessed prior to model development.

Proposed model development

In this study, GRNN, HW, NARX, and LSSVM models were proposed for the prediction of TWTP parameters. As mentioned above, the normalized measured data were employed to simulate the treated hardness (Hardness_T) (mg/L), pH (pH_T), turbidity (Turb_T) (µs/cm) and suspended solids (SS_T) (mg/L). In addition, NLE models were proposed using four different algorithms (GRNN-E, LSSVM-E, HW-E, and NARX-E). Different combinations of input parameters for black-box AI-based models (GRNN, ANFIS, LSSVM) were examined in the modeling framework. Afterwards, a set of three different models (M1, M2, and M3) were derived on the basis of significant input variables as in Equation (2):

$$\left. \begin{aligned}
 M1 &= pH_R + Turb_R + Cond_R + TDS_R + Fe_R \\
 M2 &= pH_R + Turb_R + Cond_R + TDS_R + Fe_R \\
 &\quad + Hadness_R + Cl_R \\
 M3 &= pH_R + Turb_R + Cond_R + TDS_R + Fe_R \\
 &\quad + Hardness_R + Cl_R + SS_R
 \end{aligned} \right\} \tag{2}$$

It should be noted that the quality and quantity of the intake water can be drastically changed through different periods of the year due to climate/natural or anthropogenic issues. However, as shown by Equation (2), in this study, we tried to learn the pattern between the input and output parameters using available data (only one-year data), so that having the inputs, the model can estimate the output parameters (WQ parameters in outlet). Therefore, when a severe change occurred in the intake water parameters, via the trained (learned) models, proportionally the output will reflect such severe changes.

It should be mentioned that if we have a very long period of observed data set to use for training the models, the efficiency of such developed data-driven methods will be more robust, especially when facing such severe conditions.

Used data intelligence methods

For the purpose of this research, the employed data-driven intelligence, GRNN, HW, NARX, and LSSVM models and a brief description of the mathematical concept for each of these models as well as the related citations are provided in the Appendix, Figures A1, A2, and A3, in more detail.

Performance assessment of the models

For the implementation of the models, three indices were used to evaluate the performance of GRNN, HW, NARX, and LSSVM models. The outcomes of these models were compared with one another with respect to determination coefficient (DC), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). Other performance efficiency criteria can also be used, such as bias and mean absolute relative error (MARE). For a good analysis of any model the efficiency performance should include at least one goodness-of-fit (e.g., DC) and at least one absolute error measure (e.g., RMSE) (Legates & McCabe 1999):

$$DC = 1 - \frac{\sum_{j=1}^N [(Y)_{obs,j} - (Y)_{com,j}]^2}{\sum_{j=1}^N [(Y)_{obs,j} - \bar{(Y)}_{obs,j}]^2} \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N ((Y)_{obs,j} - (Y)_{com,j})^2} \quad (4)$$

$$MAE = \frac{1}{N} \sum_{j=1}^N |(Y)_{obs,j} - (Y)_{com,j}| \quad (5)$$

$$MAPE = \frac{1}{N} \left[\sum_{i=1}^N \left| \frac{(Y)_{obsj} - (Y)_{comj}}{(Y)_{obsj}} \right| \right] \times 100 \quad (6)$$

where N , Y_{obs} , Y_{com} , and \bar{Y} indicate the number of samples, observed data, computed data, and mean of observed values, respectively.

Ensemble learning technique

According to the ‘no-free lunch’ theorem, it is a plain and well-known fact that no single model proves the best for

all kinds of data set. The properties of data such as normality, size, linearity, etc. affect the performance accuracy of any model. In addition, several studies have already shown that even for the same type of data set, the performance results may vary for different models (Nourani et al. 2019). Hence, employing multiple models and combining them has been found to be effective for a variety of problems. The ensemble technique has been used with promising success in several fields including hydro-environmental sciences, data mining and statistics as an approach to improve the prediction skill. The main goal for this technique follows the concept of enhancing the performance of the single model, by combining the results of the various individual models. As such, the process of combining various single models will be an ensemble modeling parameter (Nourani et al. 2018). In this study, four different NET were used, GRNN-E, HW-E, NARX-E, and LSSVM-E, to compare and improve the performance of single models.

Non-linear ensemble techniques

As mentioned above, four black-box models, GRNN-E, HW-E, NARX-E, and LSSVM, are employed to perform the NLE modeling. In NLE techniques, the outputs of the individual models are built to form the input layer of the considered NLE, each of which is assigned to one neuron in the input layer. Sharghi et al. (2018) and Nourani et al. (2018) have used non-linear ANN ensemble and reported superiority of such technique over linear ensemble techniques. For the modeling procedure by GRNN-E, LSSVM-E, HW-E, and NARX-E, first the individual models are selected and the outputs of the individual models are used for the ensembling. Figure 2 shows a flowchart of the techniques employed for the ensemble modeling.

RESULTS AND DISCUSSION

The main context of data-driven and predictive intelligence models for system identification (i.e., GRNN, LSSVM, HW, and NARX) are evaluated numerically via various acceptable performance accuracy. Afterwards, NLE approaches (i.e., GRNN-E, LSSVM-E, HW-E, and NARX-E) are used to enhance the accuracy of the individual models.

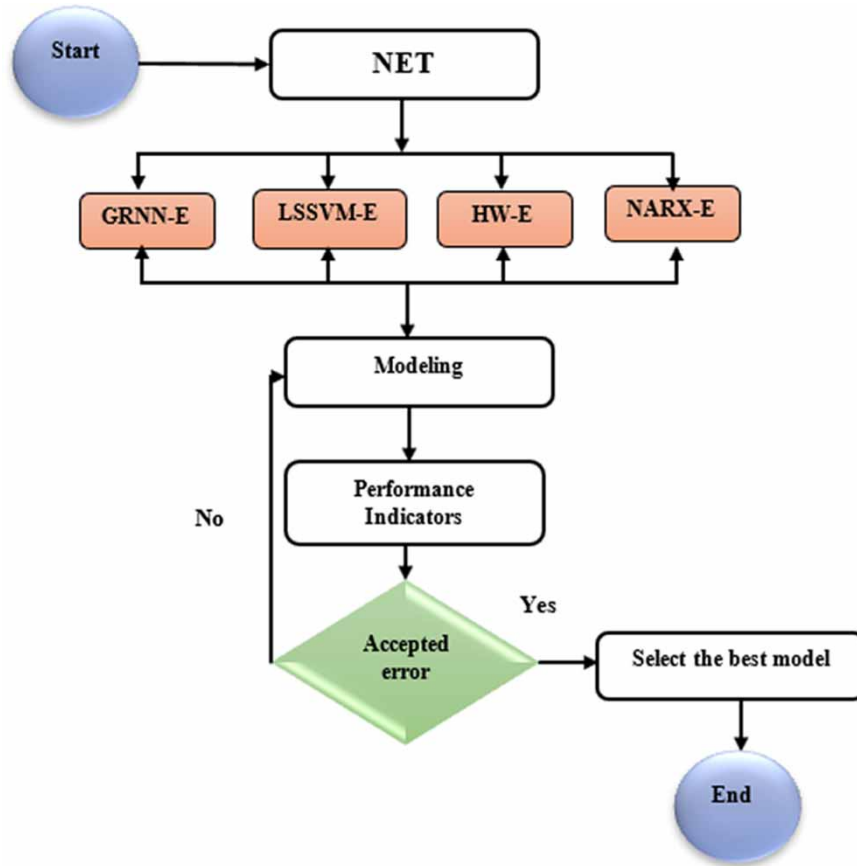


Figure 2 | General flow chart of the proposed NET employed for ensemble modeling.

The comparison of the individual models is presented in Table 3, according to Equations (3)–(6). As defined by Equation (2), three different models and input combinations were examined based on the number and types of input for the employed models in order to predict the treated hardness (Hardness_T) (mg/L), turbidity (Turb_T), pH (pH_T) ($\mu\text{s}/\text{cm}$), and suspended solids (SS_T) (mg/L).

Results of single models (GRNN, LSSVM, HW, and NARX)

It is essential to determine the appropriate range of the smoothing factor in any GRNN, due to its significant effect on the simulation ability of the model. The smoothing factor needs to be an average value because extremely large or small values will affect the regression result, so the smoothing factor was considered within the range of 0.01–1 (Ji *et al.* 2017). The LSSVM model with non-linear radial basis function (RBF) kernel function was created for

different conditions. For this purpose, different γ and σ values, through a grid search, were tried to get the best modeling result (Nourani *et al.* 2017). HW and NARX were built using the system identification toolbox of MATLAB based on the configuration of piecewise linear function (range 10–20) containing input and output non-linearity predictors in the case of the HW model. While for the NARX model, a specify delay and number of terms in standard regressors were considered as 1 and 4–8, respectively.

The direct comparison of results for all of the models shows that the HW model outperformed the other models in the prediction of hardness (see Table 3). The assessment of the results also indicates that for the HW model, the best model was acquired using the third input combination (M3). This can be proved using the performance indicators as the basis of identifying an optimal model for HW ((DC = 0.8272), (DC = 0.8316), and (DC = 0.9181) for M1, M2, and M3, respectively). According to the obtained DCs,

Table 3 | Performance indicators of the single models for modeling hardness, Turb, pH, and SS

	Models	Calibration				Verification				
		DC	RMSE	MAE	MAPE	DC	RMSE	MAE	MAPE	
Hardness	M1	GRNN	0.5789	0.1582	0.0241	0.0904	0.5528	0.1736	0.0305	0.2071
		HW	0.841	0.08	0.018	0.0674	0.8272	0.1208	0.0276	0.6264
		NARX	0.3949	0.1561	0.0098	0.0368	0.3552	0.126	0.0237	0.8231
		LSSVM	0.329	0.1691	0.0022	0.6001	0.3125	0.1526	0.0023	0.6907
		GRNN	0.6311	0.1219	0.0127	0.0478	0.6175	0.1317	0.0269	0.4883
	M2	HW	0.8445	0.0792	0.017	0.0639	0.8316	0.1269	0.0228	0.3639
		NARX	0.4644	0.1722	0.0661	0.2477	0.4257	0.1324	0.0926	0.6891
		LSSVM	0.5499	0.1482	0.003	0.3001	0.5014	0.1508	0.0021	0.3657
		GRNN	0.6387	0.1207	0.0034	0.0129	0.6181	0.1254	0.0211	0.2946
		HW	0.9851	0.0245	0.0048	0.018	0.9181	0.0254	0.0081	0.5639
	M3	NARX	0.5165	0.1396	0.0279	0.1046	0.4994	0.1432	0.2792	0.6885
		LSSVM	0.5167	0.1396	0.1423	0.6701	0.514	0.1384	0.1344	0.7387
		GRNN	0.6518	0.1517	0.0027	0.0104	0.6311	0.1079	0.0126	0.1456
		HW	0.9636	0.0362	0.003	0.0115	0.9263	0.0483	0.0129	0.1494
		NARX	0.9948	0.122	0.0103	0.0400	0.9937	0.0121	0.0057	0.0656
pH	M1	LSSVM	0.5392	0.1565	0.018	0.021	0.509	0.1245	0.0192	0.2214
		GRNN	0.7519	0.1504	0.0009	0.0034	0.7163	0.0947	0.0074	0.085
		HW	0.8395	0.0671	0.0081	0.0312	0.8066	0.1039	0.0265	0.3066
		NARX	0.9908	0.0027	0.0001	0.0016	0.9902	0.0027	0.0001	0.0022
		LSSVM	0.5421	0.1552	0.0134	0.0142	0.4911	0.1292	0.0127	0.1464
	M2	GRNN	0.793	0.1399	0.0036	0.0138	0.7786	0.0836	0.0331	0.3828
		HW	0.7703	0.1399	0.0036	0.0138	0.7555	0.0879	0.0068	0.0782
		NARX	0.9968	0.0025	0.0001	0.0011	0.9948	0.0027	0.0001	0.0016
		LSSVM	0.6033	0.1536	0.1055	1.973	0.6043	0.254	0.1088	1.2563
		GRNN	0.4358	0.1542	0.0005	0.0329	0.4144	0.0873	0.0108	0.2385
	M3	HW	0.7398	0.1136	0.0037	0.2407	0.7594	0.125	0.0206	0.4544
		NARX	0.5822	0.2681	0.0418	2.7325	0.5613	0.0757	0.0518	2.9404
		LSSVM	0.5384	0.1651	0.0027	0.1443	0.5478	0.0938	0.0086	0.1888
		GRNN	0.636	0.159	0.0175	1.1458	0.6552	0.0548	0.0149	1.3298
		HW	0.971	0.0012	0.0002	0.0101	0.9601	0.0002	0.0002	0.0043
Turb	M1	NARX	0.5959	0.1203	0.0198	0.1523	0.5693	0.0044	0.0045	0.0099
		LSSVM	0.7607	0.1641	0.0063	0.1474	0.7481	0.0999	0.0078	0.1715
		GRNN	0.7733	0.1644	0.0077	0.5047	0.751	0.0564	0.0131	0.2883
		HW	0.9166	0.0546	0.0021	0.138	0.9077	0.0283	0.0067	0.1472
		NARX	0.739	0.1235	0.0191	1.2451	0.7065	0.2911	0.1901	1.0086
	M2	LSSVM	0.761	0.1627	0.2053	3.4712	0.7239	0.4837	0.252	5.5602
		GRNN	0.5522	0.2141	0.0086	0.0307	0.5642	0.2131	0.0026	0.0271
		HW	0.754	0.2659	0.0362	0.0013	0.7447	0.0958	0.0354	0.3727
		NARX	0.5478	0.3221	0.03	0.0016	0.5011	0.1905	0.0401	0.0084
		LSSVM	0.4513	0.3226	0.004	1.928	0.408	0.4273	0.0529	1.5567
	M3	GRNN	0.5294	0.3374	0.0136	0.0485	0.508	0.2029	0.0527	0.555
		HW	0.997	0.0188	0.0004	0.0016	0.9897	0.0192	0.002	0.0209
		NARX	0.9687	0.0606	0.0074	0.0263	0.9546	0.0171	0.0095	0.0005
		LSSVM	0.672	0.3219	0.0832	0.44	0.649	0.2673	0.0526	0.5539
		GRNN	0.8176	0.3411	0.0868	0.0031	0.8167	0.173	0.0304	0.3203
SS	M1	HW	0.944	0.081	0.0103	0.0676	0.9355	0.0275	0.0004	0.0913
		NARX	0.9169	0.0607	0.0076	0.0127	0.9072	0.0317	0.0003	0.0166
		LSSVM	0.7126	0.3203	0.0047	2.0024	0.7793	0.5987	0.306	3.2214

addition of hardness, Cl, and SS increased the performance of the prediction as indicated in the third combination. This finding corresponds with that identified by Dan'azumi & Bichi (2010).

The M1, M2, and M3 for the HW model produced RMSE values of 0.1208, 0.1269, and 0.0245, respectively, and MAE of 0.0276, 0.3639, and 0.5639. Another important indicator is MAPE, which provides a better way to determine the predicting model (Gaya *et al.* 2017). The smaller the MAPE value, the more accurate the estimation model. Table 3 shows that the HW model has the smallest MAPE values with regards to LSSVM, and NARX models with the exception of GRNN. It is noteworthy that several evaluation metrics indicate the merits of the HW over the other models. To continue with additional examination of the estimating models, a series of scatter plots for the three input combinations for the HW model are shown in Figure 3(a). From the plots it can be seen that M3 was able to obtain the best agreement between the observed and computed hardness values. Assessment based on the time series estimated hardness is essential to provide a deeper and more comprehensive analysis, as Figure 4(a) indicates the best computed hardness values for the four models for the verification phase. A visual examination of the time series shows that the computed values of HW were closer to the observed hardness than GRNN, LSSVM, and NARX models.

Table 3 also depicts the obtained results for pH computed by GRNN, LSSVM, HW, and NARX using three different input combinations (M1, M2, and M3) and the best model emerged as NARX using the third input combination (M3). As shown in Table 3, M3 has the lowest RMSE and MAE and highest DC values in the verification phase with merit and suitable accuracy for two other input combinations. Considering the results of the four models in predicting pH, HW was the second-best model followed by GRNN and lastly LSSVM. Further examination of the models shows that when all the variables are used as inputs in pH modeling, the accuracy of the model is higher. Figures 3(b) and 4(b) show the scatter and time series plots for the computed pH, respectively. Turbidity (Turb) as the measure of WQ is also simulated and the obtained results are presented in Table 3. From Table 3, the HW model outperformed all models with reasonable accuracy in both calibration and verification steps.

According to the results, it can be justified that M2 with the combination of pH, Turb, Cond, TDS, hardness, and Cl demonstrated the best performance accuracy and therefore proved to be a reliable model for prediction of turbidity in TWTP. A survey of Turb results indicates that HW-M2 increases the prediction accuracy up to 20% and 5% with regards to HW-M1 and HW-M3, respectively. Comparison of results with respect to RMSE and MAE values indicates that the quantitative prediction of Turb can be arranged in the following order: HW > NARX > LSSVM > GRNN. Figures 3(c) and 4(c) show the scatter and time series plots for the Turb parameter in the verification phase, respectively. SS is considered to be one of the major pollutants that contributes to the deterioration of WQ, as such the obtained result for this variable is also presented in Table 3. It is noticed that the HW model is more satisfactory for predicting the SS in TWTP, supported by the values of DC, RMSE, MAE, and MAPE. Among the input combinations, M2 was found to be the best model and the simulated results showed a good level of satisfaction both in calibration and verification steps. Figures 3(d) and 4(d) show the scatter and time series plots for the SS parameter in the verification phase, respectively. Further investigation of SS results shows that HW-M2 increases the prediction accuracy up to 24% and 19% with regards to HW-M1 and HW-M3, respectively, according to computed DC. In addition, comparison of results with respect to RMSE and MAE values indicates that the quantitative prediction of SS can be arranged in the following order: HW > NARX > GRNN > LSSVM. The assessment of the different modeling paradigms (GRNN, HW, NARX, and LSSVM) with different input combinations (i.e., M1, M2, and M3) illustrates that the accurate predictions depend on the used model for a given input data set.

The comparison of multi-parametric simulation in this study proved that HW could serve as the best model for the simulation of hardness, Turb, and SS while the NARX model demonstrated high capability in predicting pH variable. It is apparent that system identification models (HW and NARX) verified the improvement of prediction capability better than GRNN and LSSVM models. Thus, the HW model clearly indicates the extent of effluent removal efficiency and plant performance in TWTP. As mentioned above and justified by Figure 4, there is no

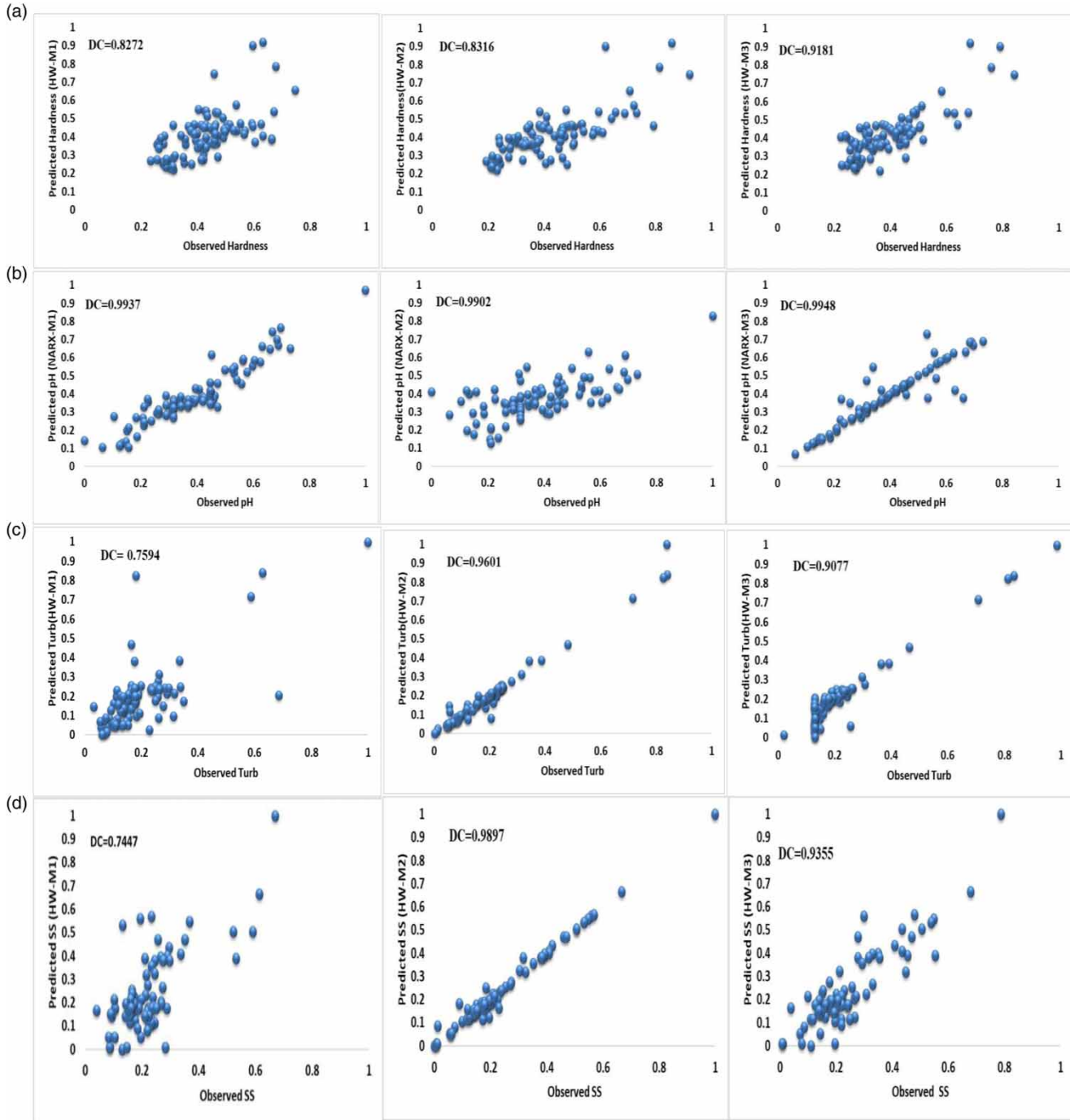


Figure 3 | Observed vs computed scatter plot for best single models for (a) hardness, (b) pH, (c) Turb, (d) SS treated water using the three sets of input combinations (M1, M2, and M3).

one exceptional model that delivers the best estimation performance. The best sets of model inputs were not the same for each of the explored predictive modeling techniques, demonstrating that the respective model types respond

differently to different input variable sets and the data patterns/characteristics in the historical input data. Therefore, combining the outputs of these models in the form of an ensemble context could increase the generalization

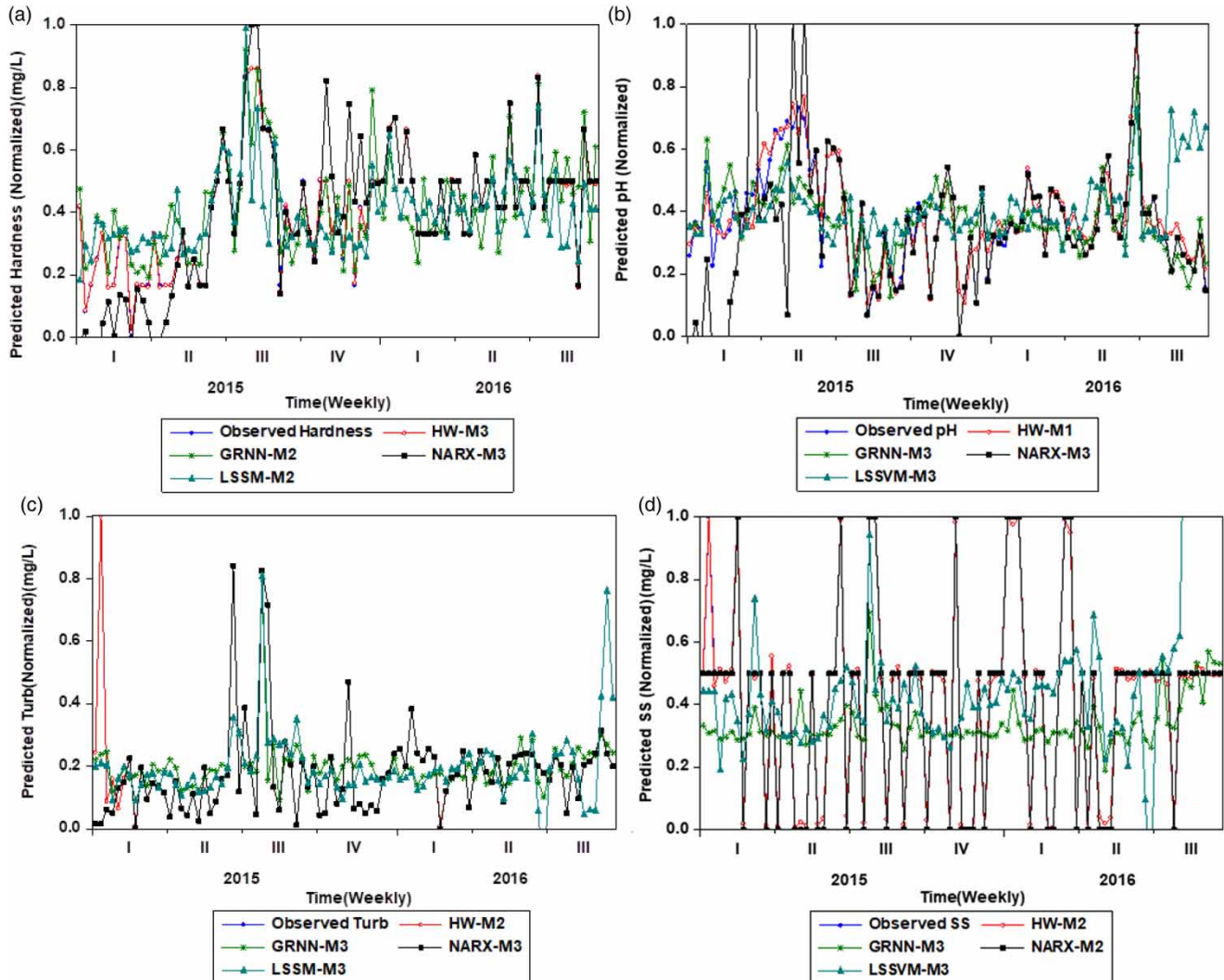


Figure 4 | Observed vs computed time series plots obtained by the best single models for (a) hardness, (b) pH, (c) Turb, and (d) SS treated water.

capability of the modeling by properly capturing the target and hence improving the prediction performance. As such, four ensemble techniques (GRNN-E, LSSVM-E, HW-E, and NARX-E) were employed to improve the predictive performance by combining the outputs of the single models.

Results of NLE models (GRNN-E, HW, NARX-E, and LSSVM-E)

As reported by Yaseen *et al.* (2016), no single model works best for all data sets, and different data characteristics such as data size, normality, linearity, and correlation all affect model

performance considerably. Nourani *et al.* (2019) also reported that even for the same data set, different models may excel in different aspects. Hence, gathering the multiple models in an ensemble approach has been found to be effective for a variety of problems. Different NLEs such as GRNN-E, LSSVM-E, HW-E, and NARX-E performed as shown in Figure 5 and the NLE results are presented in Table 4.

Table 4 demonstrates that among the NLE modeling, GRNN-E outperformed the other models with considerable accuracy, due to its robustness in dealing with complex interactions which perform prediction without requiring large samples. The accuracy of GRNN-E as a type of ANN can also be attributed to the promising predictive skills of

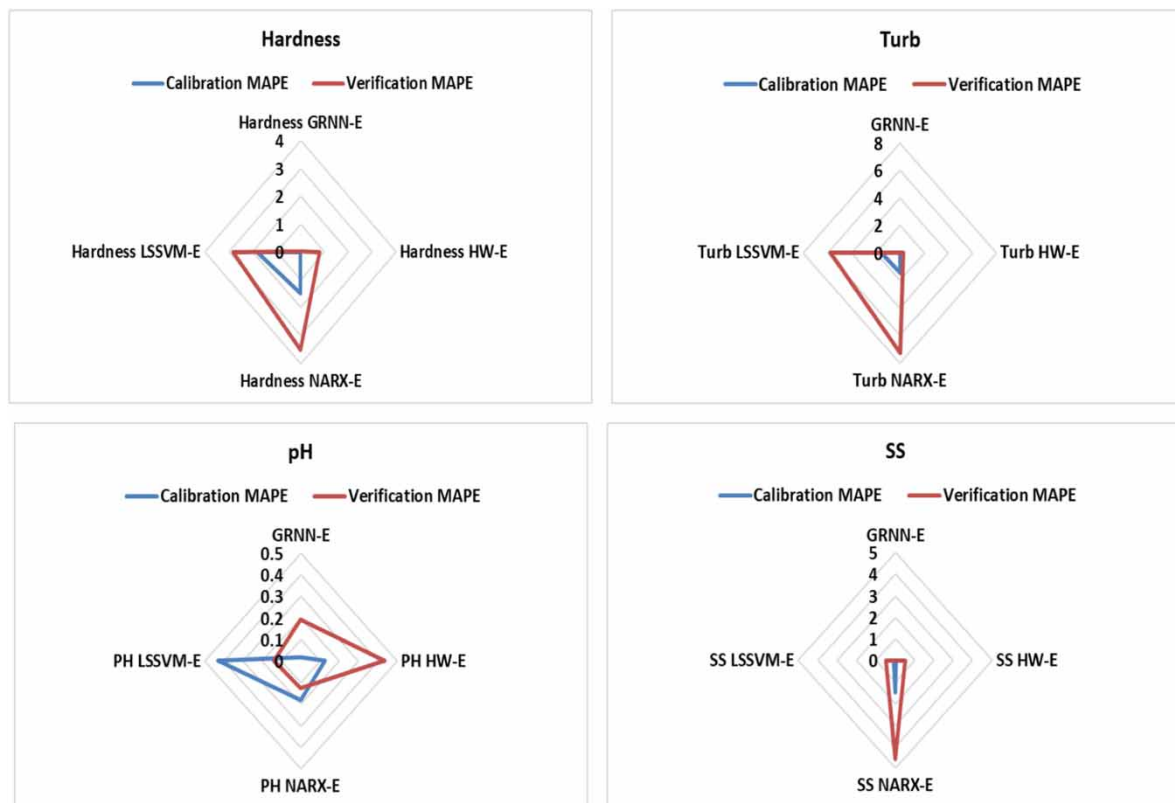


Figure 5 | Radar chart of MAPE for hardness, Turb, pH, and SS for both calibration and verification phases.

Table 4 | Results of non-linear ensemble models for estimation of hardness, Turb, pH, and SS

Parameter	Ensemble models	Calibration				Verification			
		DC	RMSE ^a	MAE	MAPE	DC	RMSE ^a	MAE	MAPE
Hardness _T	GRNN-E	0.9991	0.0091	0.0002	0.0007	0.9969	0.0085	0.0011	0.0093
	HW-E	0.6843	0.1576	0.0011	0.0041	0.6973	0.1365	0.0893	0.7884
	NARX-E	0.5179	0.4362	0.3873	1.4706	0.5195	0.5148	0.3975	3.5081
	LSSVM-E	0.8687	0.2821	0.5555	1.8235	0.8345	0.6771	0.0342	2.8374
pH _T	GRNN-E	0.9989	0.0439	0.0042	0.0164	0.9989	0.0663	0.0167	0.1927
	HW-E	0.6301	0.1537	0.0325	0.1255	0.6008	0.1721	0.0378	0.4369
	NARX-E	0.6779	0.3073	0.0464	0.1794	0.6258	0.2167	0.0109	0.1257
	LSSVM-E	0.8878	0.4091	0.1734	0.4301	0.7315	0.197	0.0238	0.1413
Turb _T	GRNN-E	0.9838	0.0241	0.0037	0.0281	0.9991	0.0008	0.0002	0.0052
	HW-E	0.9359	0.1923	0.0004	0.0029	0.9288	0.0596	0.0113	0.2483
	NARX-E	0.6022	0.2691	0.1913	1.4707	0.5349	1.6377	0.3308	7.2979
	LSSVM-E	0.8007	0.2529	0.2494	1.6101	0.6203	1.3691	0.5132	5.7450
SS _T	GRNN-E	0.9998	0.0008	0.0015	0.0053	0.9988	0.0002	0.0017	0.0043
	HW-E	0.5592	0.2995	0.0153	0.0546	0.5433	0.1797	0.0486	0.5112
	NARX-E	0.7292	0.5355	0.4118	1.4706	0.6837	0.5316	0.4336	4.5646
	LSSVM-E	0.6819	0.3609	0.0588	0.0924	0.6326	0.0341	0.1739	0.4645

^aSince all data are normalized, the RMSE has no dimension.

the overall ANN model. Peng & Kumar (2005) and Nourani *et al.* (2019) reported that in a few cases single models can be superior to the ensemble model in some instances, which was also observed by comparing Tables 3 and 4.

Figure 5 shows the radar chart showing the different percentage variation of MAPE. From Figure 5 it is clear

that the MAPE value is less than 10% for all the best models, which demonstrated the high accuracy of the model. Similarly, in terms of DC, GRNN-E increased the accuracy of the modeling up to 30% for hardness and Turb, 34% for pH, and 37% for SS with regards to the single models.

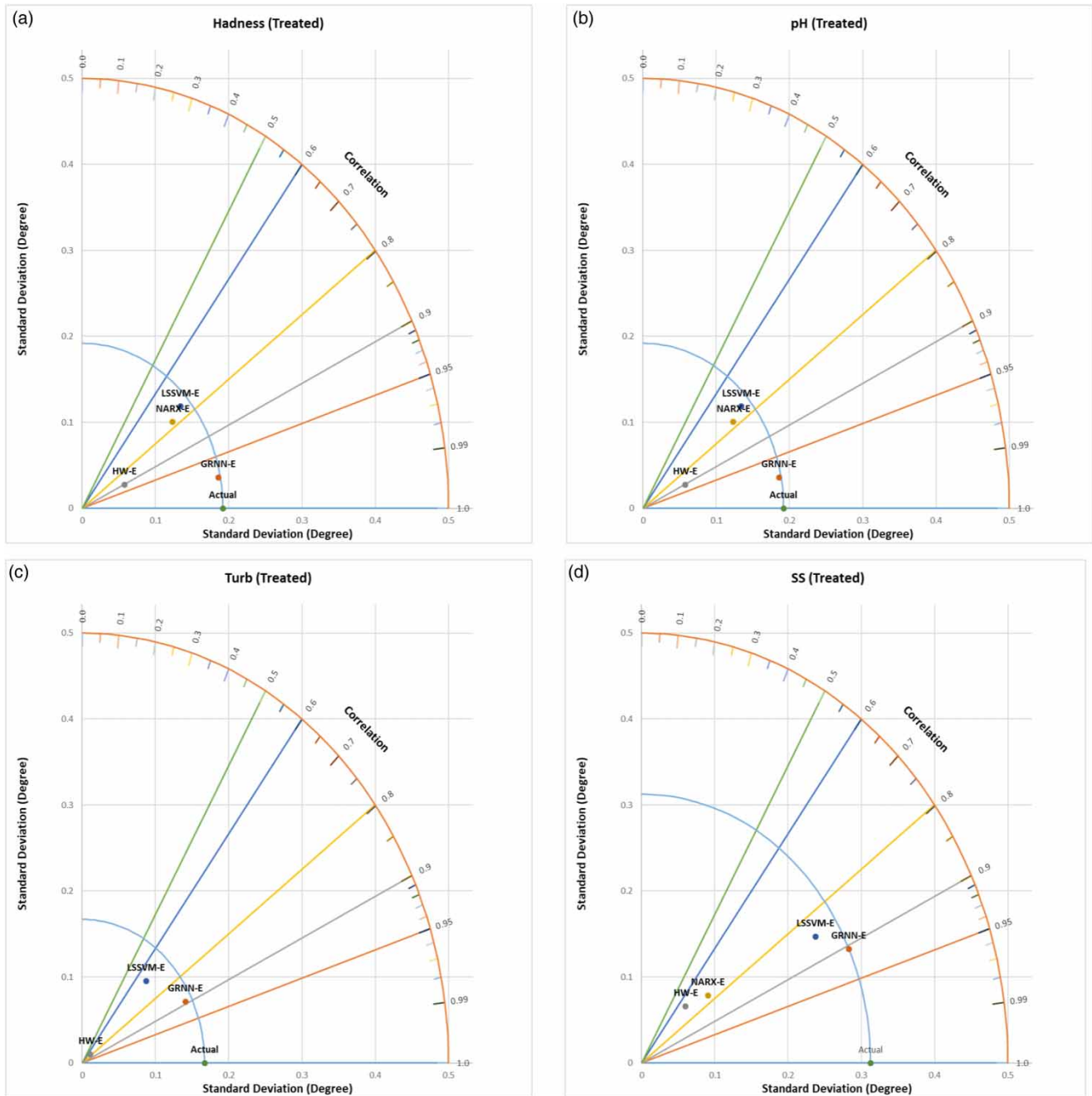


Figure 6 | Taylor diagram depicting the ensemble performance of GRNN-E, HW-E, NARX-E, and LSSVM-E models for (a) hardness, (b) pH, (c) Turb, and (d) SS during the verification phase.

In Figure 6, a Taylor diagram was used as the most widely recommended diagram for accuracy comparison (Kim et al. 2018; Zhu et al. 2019). The performances of the GRNN-E, HW-E, NARX-E, and LSSVM-E models were compared during verification phase, using the correlation coefficient (R) and the standard deviation (SD). From Figure 6 it can be seen that GRNN-E outperformed other models, because the observed points are closer to the computed points. This can also be proved by considering the high value of R which was attributed to the GRNN-E. Generally, if the SD of the computed values is higher than the SD of observed values, then it will result in overestimation and vice versa.

CONCLUSIONS

In this study, GRNN, HW, NARX, and LSSVM for modeling the treated hardness, turbidity, pH, and suspended solids were applied in Tamburawa WTP Kano, Nigeria. Subsequently, the NETs were also employed using the outputs of the single models to enhance the overall accuracy of the prediction. Weekly available data from the treatment plant were used and various statistical performance indicators were used to explore the predictive capability of the overall modeling techniques. The comparison results of multi-parametric simulation in this study showed that HW served as the best model for the simulation of Hardness_T, Turb_T, and SS_T while the NARX model demonstrated high capability in predicting pH_T variable. Yet, the HW and NARX as system identification techniques attained overall best predictive performance among the four modeling approaches. However, among the NLE models, GRNN-E proved of high merit and increased the accuracy of the best single models up to 30% for Hardness_T and Turb_T, 34% for pH_T, and 37% for SS_T with regards to the single models. Finally, such a model can serve as the background and be part of the portfolio of tools used by water engineers in decision-making at TWTP. The results also recommended that the application of other AI such as extreme learning machine (ELM), wavelet and the combination of hybrid models may also be combined in the proposed ensemble approach in order to integrate a set of models so as to create a new model which could yield higher precision than the used single models.

Just like any data-driven method and as the limitation of the study, the efficiency of the proposed model is markedly

related to the quality and quantity of the available data, which most of the time is a challenge in developing countries. Therefore, the used data should be checked and when gathering more data in the future, the models should be updated using the new data.

SUPPLEMENTARY DATA

The Supplementary Data for this paper is available online at <http://dx.doi.org/10.2166/aqua.2019.078>.

REFERENCES

- Al-Baidhani, J. H. & Alameedee, M. A. 2017 Prediction of water treatment plant outlet turbidity using artificial neural network. *International Journal of Current Engineering and Technology* 7 (4), 1559–1565.
- Belia, E., Amerlinck, Y., Benedetti, L., Johnson, B., Sin, G., Vanrolleghem, P. A., Gernaey, K. V., Gillot, S., Neumann, M. B., Rieger, L., Shaw, A. & Villez, K. 2009 *Wastewater treatment modelling: dealing with uncertainties*. *Water Science and Technology* 60 (8), 1929–1941.
- Dan'azumi, S. & Bichi, M. H. 2010 Industrial pollution and implication on source of water supply in Kano, Nigeria. *International Journal of Engineering and Technology* 10 (1), 101–109.
- Elkiran, G., Nourani, V., Abba, S. I. & Abdullahi, J. 2018 Artificial intelligence-based approaches for multi-station modelling of dissolve oxygen in river. *Global Journal of Environmental Science and Management* 4 (4), 439–450.
- Elkiran, G., Nourani, V. & Abba, S. I. 2019 *Multi-step ahead modelling of river water quality parameters using ensemble artificial intelligence-based approach*. *Journal of Hydrology* 577, 123962.
- Falah Nezhad, M., Mehrdadi, N., Torabian, A. & Behboudian, S. 2016 Artificial neural network modeling of the effluent quality index for municipal wastewater treatment plants using quality variables: south of Tehran wastewater treatment plant. *Journal of Water Supply: Research and Technology-Aqua* 65 (1), 18–27.
- Gaya, M. S., Yusuf, L. A., Mustapha, M., Muhammad, B., Sani, A., Tijjani, A. T. A., Wahab, N. A. & Khairi, M. T. M. 2017 *Estimation of turbidity in water treatment plant using Hammerstein-Wiener and neural network technique*. *Indonesian Journal of Electrical Engineering and Computer Science* 5 (3), 666–672.
- Ji, X., Shang, X., Dahlgren, R. A. & Zhang, M. 2017 *Prediction of dissolved oxygen concentration in hypoxic river systems using support vector machine: a case study of Wen-Rui Tang River*. *China. Environmental Science and Pollution Research* 24 (19), 16062–16076.
- Kalousis, A., Prados, J. & Hilario, M. 2007 *Stability of feature selection algorithms: a study on high-dimensional spaces*. *Knowledge and Information Systems* 12 (1), 95–116.

- Kim, S. E. & Seo, I. W. 2015 Artificial neural network ensemble modeling with conjunctive data clustering for water quality prediction in rivers. *Journal of Hydro-Environment Research* **9** (3), 325–339.
- Kim, J. S., Seo, I. W., Lyu, S. & Kwak, S. 2018 Modeling water temperature effect in diatom (*Stephanodiscus hantzschii*) prediction in eutrophic rivers using a 2D contaminant transport model. *Journal of Hydro-Environment Research* **19**, 41–55.
- Křifžek, P., Kittler, J. & Hlaváč, V. 2007 Improving stability of feature selection methods. In: International Conference on Computer Analysis of Images and Patterns. Springer, Berlin, Heidelberg, pp. 929–936.
- Legates, D. R. & McCabe, G. J. 1999 Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. *Water Resources Research* **35** (1), 233–241.
- Maleki, A., Nasser, S., Aminabad, M. S. & Hadi, M. 2018 Comparison of ARIMA and NNAR models for forecasting water treatment plant's influent characteristics. *KSCE Journal of Civil Engineering* **22** (9), 3323–3345.
- Mashaly, A. F. & Alazba, A. A. 2017 Artificial intelligence for predicting solar still production and comparison with stepwise regression under arid climate. *Journal of Water Supply: Research and Technology-Aqua* **66** (3), 166–177.
- Meinshausen, N. & Bühlmann, P. 2010 Stability selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **72** (4), 417–473.
- Nabavi-Pelesaraei, A., Bayat, R., Hosseinzadeh-Bandbafha, H., Afrasyabi, H. & Chau, K. W. 2017 Modeling of energy consumption and environmental life cycle assessment for incineration and landfill systems of municipal solid waste management – A case study in Tehran Metropolis of Iran. *Journal of Cleaner Production* **148**, 427–440.
- Nagana, S. R., Deka, P. C., Ghorbani, M. A., Biazar, S. M., Al-Ansari, N. & Yaseen, Z. M. 2019 Dew point temperature estimation: application of artificial intelligence model integrated with nature-inspired optimization algorithms. *Water* **11** (4), 742.
- Nourani, V. & Fard, M. S. 2012 Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes. *Advances in Engineering Software* **47** (1), 127–146.
- Nourani, V., Andalib, G. & Sadikoglu, F. 2017 Multi-station streamflow forecasting using wavelet denoising and artificial intelligence models. *Procedia Computer Science* **120**, 617–624.
- Nourani, V., Elkiran, G. & Abba, S. I. 2018 Wastewater treatment plant performance analysis using artificial intelligence – an ensemble approach. *Water Science and Technology* **78**, 2064–2076. DOI.org/10.2166/wst.2018.477.
- Nourani, V., Elkiran, G., Abdullahi, J. & Tahsin, A. 2019 Multi-region modeling of daily global solar radiation with artificial intelligence ensemble. *Natural Resources Research* **28** (4), 1217–1234.
- Ogwueleka, T. C. & Ogwueleka, F. N. 2009 Optimization of drinking water treatment processes using artificial neural network. *Nigerian Journal of Technology* **28** (1), 16–25.
- Olyaie, E., Banejad, H., Chau, K. W. & Melesse, A. M. 2015 A comparison of various artificial intelligence approaches performance for estimating suspended sediment load of river systems: a case study in United States. *Environmental Monitoring and Assessment* **187** (4), 189.
- Peng, P. & Kumar, A. 2005 A large ensemble analysis of the influence of tropical SSTs on seasonal atmospheric variability. *Journal of Climate* **18** (7), 1068–1085.
- Rahimi, M., Ebrahimi, K. & Araghinejad, S. 2019 A novel approach in determination of the effluent applications. *Journal of Water Supply: Research and Technology-Aqua* **68** (3), 157–170.
- Shamshirband, S., Nodoushan, E. J., Adolf, J. E., Manaf, A. A., Mosavi, A. & Chau, K. 2019 Ensemble models with uncertainty analysis for multi-day ahead forecasting of chlorophyll a concentration in coastal waters. *Engineering Applications of Computational Fluid Mechanics* **13**, 91–101.
- Sharghi, E., Nourani, V. & Behfar, N. 2018 Earthfill dam seepage analysis using ensemble artificial intelligence-based modeling. *Journal of Hydroinformatics* **20** (5), 1071–1084.
- World Health Organization (WHO); United Nation Children Fund (UNICEF) 2012 *Progress on Drinking Water and Sanitation*. WHO, Geneva and UNICEF, New York, pp. 122–125.
- Wu, G. D. & Lo, S. L. 2008 Predicting real-time coagulant dosage in water treatment by artificial neural networks and adaptive network-based fuzzy inference system. *Engineering Applications of Artificial Intelligence* **21** (8), 1189–1195.
- Yaseen, Z. M., Jaafar, O., Deo, R. C., Kisi, O., Adamowski, J., Quilty, J. & El-Shafie, A. 2016 Stream-flow forecasting using extreme learning machines: a case study in a semi-arid region in Iraq. *Journal of Hydrology* **542**, 603–614.
- Zhang, Q. & Stanley, S. J. 1999 Real-time water treatment process control with artificial neural networks. *Journal of Environmental Engineering* **125** (2), 153–160.
- Zhu, S., Heddad, S., Wu, S., Dai, J. & Jia, B. 2019 Extreme learning machine-based prediction of daily water temperature for rivers. *Environmental Earth Sciences* **78** (6), 202.

First received 29 May 2019; accepted in revised form 16 September 2019. Available online 22 October 2019