Wastewater treatment plant performance analysis using artificial intelligence – an ensemble approach

Vahid Nourani, Gozen Elkiran and S. I. Abba

ABSTRACT

In the present study, three different artificial intelligence based non-linear models, i.e. feed forward neural network (FFNN), adaptive neuro fuzzy inference system (ANFIS), support vector machine (SVM) approaches and a classical multi-linear regression (MLR) method were applied for predicting the performance of Nicosia wastewater treatment plant (NWWTP), in terms of effluent biological oxygen demand (BOD_{eff}), chemical oxygen demand (COD_{eff}) and total nitrogen (TN_{eff}). The daily data were used to develop single and ensemble models to improve the prediction ability of the methods. The obtained results of single models proved that, ANFIS model provides effective outcomes in comparison with single models. In the ensemble modeling, simple averaging ensemble, weighted averaging ensemble and neural network ensemble techniques were proposed subsequently to improve the performance of the single models. The results showed that in prediction of BOD_{eff}, the ensemble models of simple averaging ensemble (SAE), weighted averaging ensemble (WAE) and neural network ensemble (NNE), increased the performance efficiency of artificial intelligence (AI) modeling up to 14%, 20% and 24% at verification phase, respectively, and less than or equal to 5% for both COD_{eff} and TN_{eff} in calibration phase. This shows that NNE model is more robust and reliable ensemble method for predicting the NWWTP performance due to its non-linear averaging kernel.

Key words | artificial intelligence, black box model, ensemble learning, Nicosia wastewater treatment plant, wastewater

Vahid Nourani (corresponding author)
Department of Water Resources Engineering, Faculty of Civil Engineering,
University of Tabriz,
29 Bahman Ave., Tabriz 5166616471,
Iran
and
Faculty of Civil and Environmental Engineering,
Near East University, P.O. Box 99138,
Nicosia, North Cyprus, Mersin 10,
Turkey
E-mail: *nourani@tabrizu.ac.ir*

Gozen Elkiran
Faculty of Civil and Environmental Engineering,
Near East University,
Near East Boulevard 99138, Nicosia,
North Cyprus

S. I. Abba
Researcher Faculty of Civil and Environmental Engineering,
Near East University,
Near East Boulevard 99138, Nicosia,
North Cyprus

ABBREVIATIONS

%	Percentage	mg/L	Milligram per liter
°C	Degree centigrade	MLFF	Multi-layer feed forward
AI	Artificial intelligence	MBR	Membrane bioreactor
ANFIS	Adaptive neuro fuzzy inference system	MLR	Multi-linear regression
ANN	Artificial neural network	NWWTP	Nicosia wastewater treatment plant
ASCE	American society of civil engineer	NNE	Neural network ensemble (NNE)
BODinf	Influent biological oxygen demand	pHinf	Influent pH
BODeff	Effluents biological oxygen demand	pHeff	Effluent pH
BP	Back propagation	R	Correlation coefficient
Condinf	Influent conductivity	RMSE	Root mean square error
Condeff	Effluent conductivity	SVR	Support vector regression
CODinf	Influent chemical oxygen demand	SVM	Support vector machine
CODeff	Effluents chemical oxygen demand	SS	Suspended solids
DC	Determination coefficient	SAE	Simple averaging ensemble
FFNN	Feed forward neural network	TNinf	Influent total nitrogen

doi: 10.2166/wst.2018.477

TNeff	Effluent total nitrogen
TSS	Total suspended solids
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNDP	United Nations Development Programme
WAE	Weighted averaging ensemble
WWTPs	Wastewater treatment plants

INTRODUCTION

Wastewater treatment plants (WWTPs) consist of multiple equipment's and processes to treat wastewater which generally allows anthropogenic activities and industrial effluents to be disposed without danger to human health or undesirable damage to the natural environment (Gómez *et al.* 2017). According to the United Nations Educational, Scientific and Cultural Organization (UNESCO) (2015), wastewater treatment and environmental sanitation are of paramount importance for sustainable development and critical for human health ecosystems. A diluted mixture of different wastes from commercial, residential, industrial and other public environments is considered as wastewater or sewage, though its characteristics depend upon the discharge source but wastewater generally contains organic matter, inorganic matter and living organisms (Aydın Temel *et al.* 2018).

A WWTP is a highly complicated and dynamic system and so its proper operation and control is essential in order to safeguard the public and environmental health issues. The quality of raw and treated effluent has a great impact on the operation and performance of any WWTP (Ahmadi *et al.* 2017). In addition, the influent flow to a WWTP has a significant influence on the energy consumption and entire treatment process. It is difficult to measure some dominant variables; e.g. biological oxygen demand (BOD) requires 5-days' incubation, and this makes it hard to find and delay the timing process; it is also well known that high BOD concentration requires longer aeration time and supply of more oxygen. Thus, it is important to predict the effluent concentration at future time horizons in order to well manage the plant and control the effluent quality (Kim *et al.* 2006). Operational control of WWTPs is complex because of the complexity of mechanism in the treatment plant, influent quality and strength of the wastewater. Hence, modeling of such a complex system is quite difficult and most of the available traditional models are based on some rough expectations, linear approximations and assumptions (Gao *et al.* 2017; Hameed *et al.* 2017). The conventional processes used for determining the performance of WWTPs are based on the balance equations coupled with the mass flow rate equations for microbial growth and substratum consumption which involve non-linear interaction, time consuming and complex nature (American Society of Civil Engineer, ASCE Task Committee 2000). This complex behavior of WWTPs is difficult to predict or simply explained by classical linear models and, as such, development of reliable and convenient modeling tools plays an essential role in simulation and monitoring the performance of WWTPs and describing the overall phenomena occurring in the entire system.

Therefore, non-linear artificial intelligence (AI) based models have been applied recently to overcome the problems associated with traditional linear methods, due to their accuracy, flexibility and promising application in different engineering fields (Govindaraju 2000; Nourani 2017). AI based models such as adaptive neural fuzzy inference system (ANFIS), feed forward neural network (FFNN), support vector machine (SVM), are relatively new black box methods used at various aspects of water and environmental engineering as well as assessment of WWTP operations. To assess the WWTP performance there are certain important parameters that have been focused in the most of previous studies conducted using AI based models. These parameters include BOD, chemical oxygen demand (COD), total suspended solids (TSS) and total nitrogen (TN) (Hameed *et al.* 2017). Guo *et al.* (2015) used the influents of pH, temperature, COD and SS to predict the concentration of TN effluent from the WWTP in Ulsan, Korea, by employing ANN and SVM models and concluded that AI models could be reliable methods for prediction of the effluent conditions of the WWTPs. Granata *et al.* (2017) employed SVM and regression tree in order to forecast the fluent concentration of BOD, COD, TSS and TDS in WWTP, the SVM showed a bit better performance than regression tree which is a kind of decision tree. Civelekoglu *et al.* (2009) applied ANN and ANFIS methods to model the COD removal in biological WWTP, the overall results indicated that ANFIS is a suitable model for prediction of the WWTP performance. Hamed *et al.* (2004) used the BOD and TSS values recorded at various positions as input parameters and outlet BOD and TSS as target variables to predict the performance of WWTP using ANN model. The results proved the ability of ANN model for predicting WWTP performance.

As the literature review shows, there is no unique model to be superior to others in all cases and the performances of different models may be different according to condition of each WWTP. Therefore, it is verified that the combination of outputs (from different models) through an ensemble method may lead to more accurate results. The idea of such an ensemble model has been already used at different fields of engineering, environmental and water quality modelling (Cloke & Pappenberger 2009; Sharghi et al. 2018). However, since the pronouncement of ensemble methods in engineering, to the best of the authors' knowledge, there is no published study in the technical literature indicating the application of AI based ensemble approach in WWTP modeling. The current study aimed at the following. (i) To develop and compare the potential of some AI based models (FFNN, SVM and ANFIS) and conventional multilinear model (MLR) for prediction of the Nicosia WWTP performance considering four different combinations of input parameters. Other feasible alternative models may also be used (e.g. genetic programing, ARIMA models, Olyaie et al. 2017), but these models were adopted here due to their outstanding performances in various hydro-environmental studies. (ii) To establish and apply three ensemble techniques using the outputs of aforementioned single models in order to improve the overall efficiency of the prediction performance. In this way, simple linear averaging, weighted linear averaging and non-linear neural ensemble techniques are proposed to combine the outputs of the methods.

MATERIALS AND METHODS

Plant description and used data

The new WWTP of Nicosia (NWWTP) was planned to take care with 270,000 populations with the project horizon year 2025. The implementation of stage has been put in place and considered so as to sidestep general and extensive capacity surcharge of the consumer for unexploited capacity. The volume of stage 1 was conventionally established as 30,000 m^3/day and stage 2, on the other hand, will be executed to accomplish the final design volume of 45,000 m^3/day. The new plant has been planned in light of membrane bioreactor (MBR) technology. It has high condition of workmanship and state of art technology for a WWT and the new plant has been designed as an MBR plant with advanced biological nutrient removal. It is currently the second biggest WWTP in Europe that uses MBR technology which serves the needs of both Turkish Cypriot and Greek Cypriot (bio-communal). About 10 million m^3 of treated water every year can be reused for agricultural purpose. Depending on the crop type and the rotation approach, approximately 500 hectares can be irrigated, so the treated water can significantly reduce the over-extraction of groundwater in the area, and enhancing water resources and water conservation. More than 3,000 tons of dry solids suitable for use as natural fertilizers are also produced every year. The NWWTP has anaerobic sludge digesters which are equipped and capable of producing electricity from biogas. The operation of the plant will therefore be partly powered by renewable energy (10–20% on average), reducing its carbon dioxide (CO_2) emissions. Figure 1 shows the map of the study area and schematic of the NWWTP process; this process presents only the treatment process from raw sewage to treated effluents.

For this study, the available daily data were obtained from the new NWWTP (United Nations Development Programme (UNDP) 2014). The measurement of the selected parameters covers all the seasonal variations and consists of various sets of inputs and outputs parameters. The daily measured data set comprising of influent of pH_{inf}, conductivity ($Cond_{inf}$), BOD_{inf}, COD_{inf}, total nitrogen (TN_{inf}) as the inputs of models, and three effluent values like: $BOD_{eff,}$ COD_{eff} and TN_{eff} as targets over the period of 2015–2016. The available data set was divided into two parts: 75% of the data were employed for the calibration and the remaining data were used for the verification purpose. For the validation purpose there are several methods which can be implemented such as holdout, leave one out, cross validation and so on. The method we employed here is holdout or sometimes called leave-group-out. In the holdout method, the data are randomly assigned to two sets, usually called training and testing, and can be considered as the simplest version of k- fold cross validation (Sargent 2009; Tsioptsias et al. 2016).

At the initial stage before the training of the model, both input and output data should be normalized (e.g. within the range of 0 and 1) as (Nourani et al. 2012a, 2012b):

$$X_i = \frac{x_u - x_{min}}{x_{max} - x_{min}} \tag{1}$$

where 'X_i' is the normalized data value, 'x_u' is observed (measured) data, 'x_{min}' is the minimum and 'x_{max}' is the maximum values of the data set. The statistical analysis of the input–output parameters is an essential step in any AI based modeling because such analysis identifies the type and strength of the relations between inputs and outputs. Descriptive statistics of the selected parameters are presented in Table 1. Figure 2 shows the observed influent (inf) and treated effluent (eff) concentrations of BOD, COD and TN at the entrance and outlet of the plants. The box and whiskers plots of each parameter are also indicated

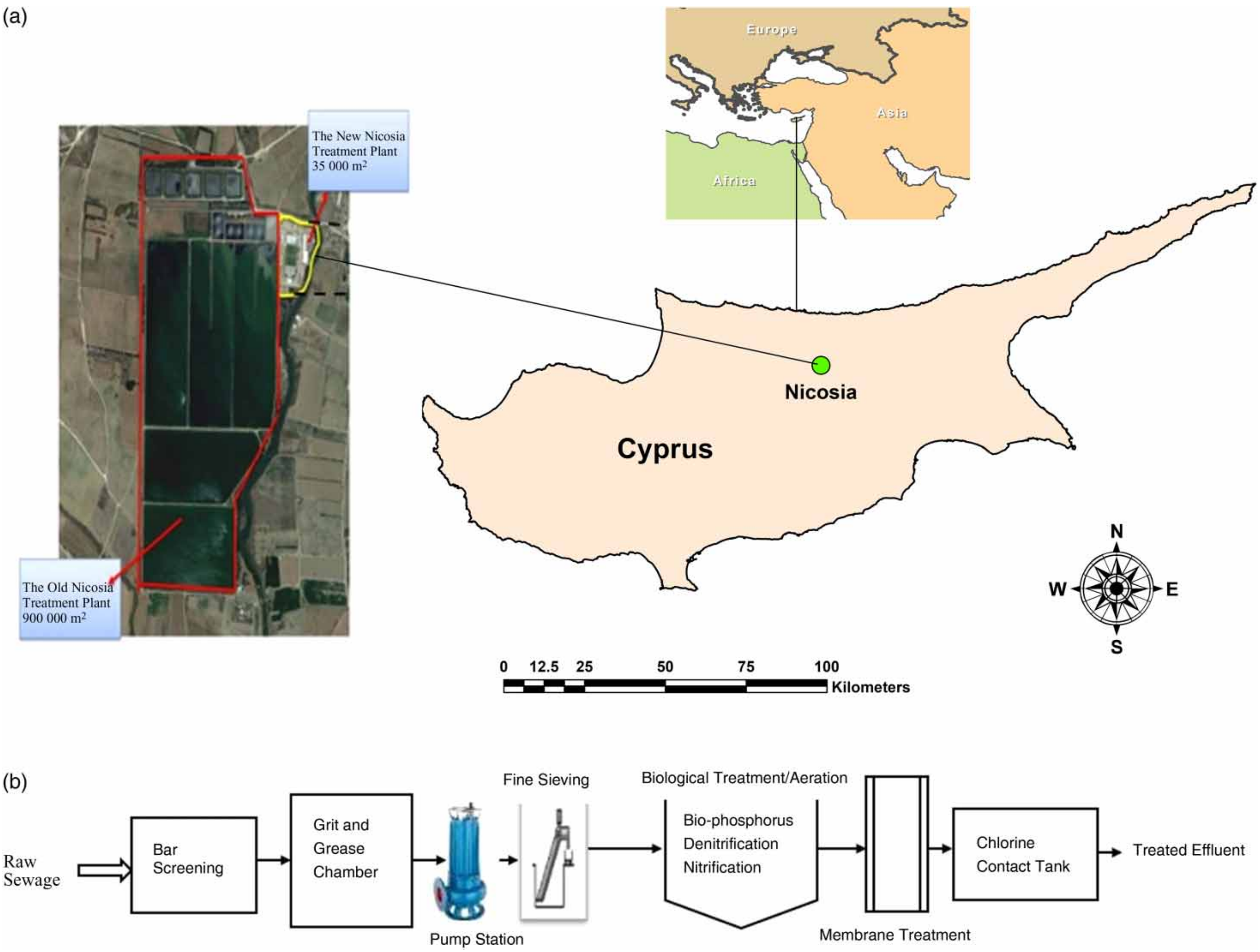

Figure 1 | (a) Map of the study location and (b) schematic of the Nicosia WWTP process.

in Figure 2. The plot signifies the extent of outliers in each parameter and subsequently the efficiency of the treatment plant and range of each variable.

Table 1 | Descriptive statistics of the used data parameters

Parameters	Minimum	Maximum	Mean	Median	Std. deviation
pH_{inf}	5.60	8.20	7.58	7.6	0.22
$Cond_{inf}$ (mS/cm)	1.40	4.00	3.25	3.4	0.44
BOD_{inf} (mg/L)	156	11,685	401.25	410	1,591.83
COD_{inf} (mg/L)	100	1,463	876.24	948	299.86
TN_{inf} (mg/L)	50.0	121	85.56	85.0	10.24
BOD_{eff} (mg/L)	0.20	27.00	1.26	4.20	2.74
COD_{eff} (mg/L)	2.00	55.00	21.94	21.0	3.93
TN_{eff} (mg/L)	0.30	75.00	8.21	8.00	3.89

It could be observed from Table 1 that the inflow characteristics of some parameters in WWTP are slightly higher than the design values, for example: BOD_{inf}, COD_{inf} and TN_{inf} have the design values of 538, 1,202 and 121 mg/L, respectively (UNDP 2014). It is clear that for all kind of data driven methods (e.g. AI methods) if the amount of dispersion (standard deviation) of data is low (this indicates the closeness of the data to the mean) it is expected to get less biased outputs from the models.

In the descriptive statistical analysis, correlation coefficient (R) as the most commonly used techniques was calculated to measure the strength and degree of linear relationship between two variables, which can be served as the preliminary indication of probable linear correlation between a set of variables. However, the weakness of the computed R values depicts that the application of conventional linear techniques to handle such non-linear

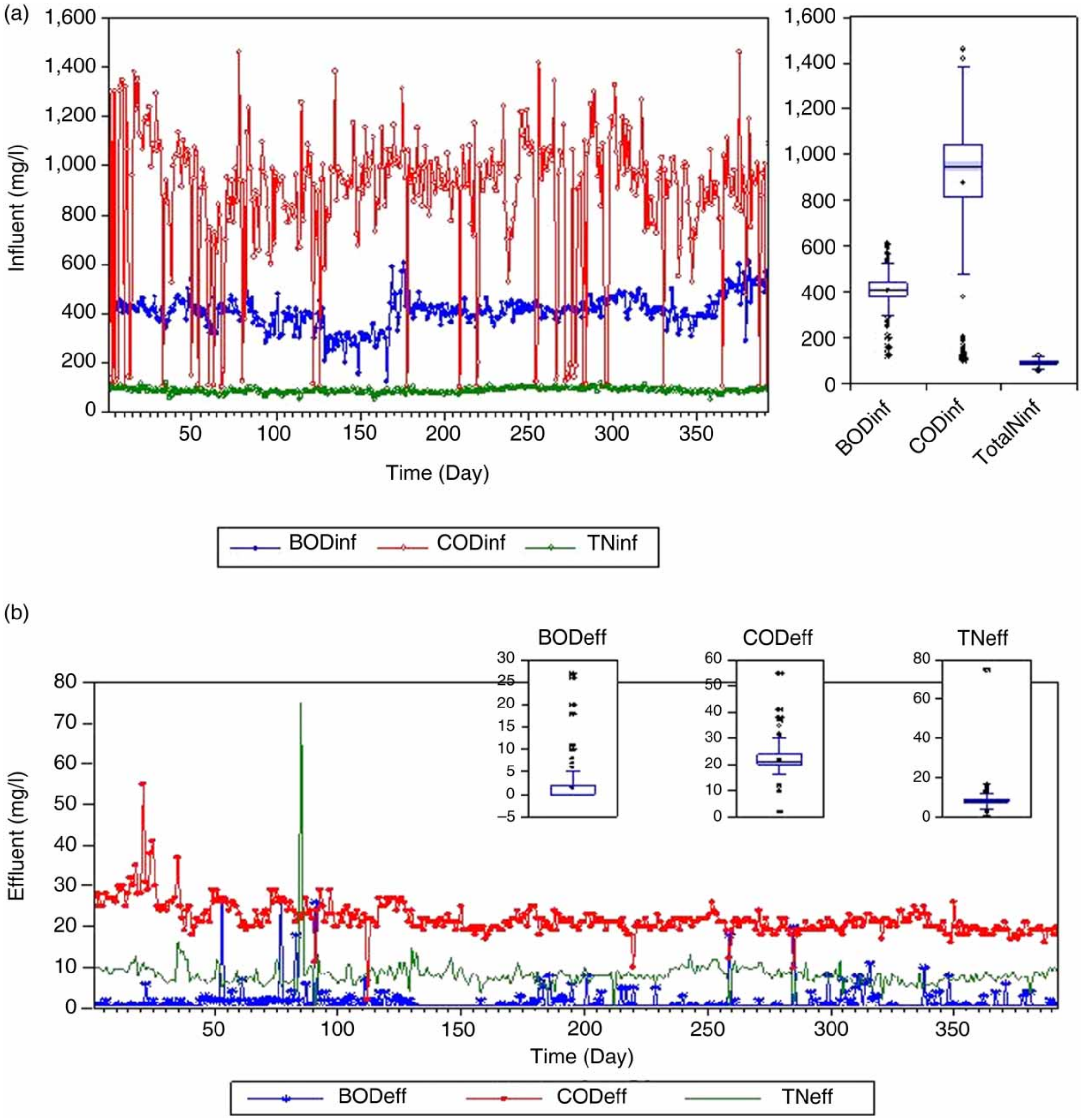

Figure 2 | BOD, COD and TN concentrations of the plant at the (a) inlet (influent) and (b) outlet (effluent).

complex interactions cannot be recommended and there is a great need to introduce more robust non-linear tools. Therefore, instead of some other studies which used linear correlation coefficient between input and output parameters to select the dominant inputs of non-linear, different combinations of input parameters are examined through the used methods (FFNN, ANFIS, SVM and MLR) in this study.

Proposed methodology

In this study, FFNN, ANFIS, SVM and MLR models were proposed for simulation of the performance of NWWTP. The used data were normalized and partitioned into calibration and verification sets to simulate BOD_{eff}, COD_{eff} and TN_{eff}. In addition, ensemble models were proposed using simple average, weighted average and neural average techniques. Different combinations of input parameters for black box AI based models (FFNN, ANFIS, SVM) and linear model (MLR) were examined in the modeling framework. The overall schematic of the proposed methodology is presented in Figure A1 at Appendix A (available with the online version of this paper).

Used data driven methods and efficiency criteria

As mentioned, three AI based methods of FFNN, ANFIS and SVM and one classic MLR approach were employed for single and ensemble modeling in this study. A brief description of mathematical concept for each of these models as well as the related citations are provided in Appendix B for more details (available with the online version of this paper). Furthermore, determination coefficient (DC) and root mean square error (RMSE) criteria. Other performance efficiency of the model can also be used such as bias or mean absolute error (MAE). 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).

Ensemble learning techniques

Ensemble learning is a machine learning to combine the process of multiple predictors in order to enhance the final performance (Sharghi *et al.* 2018). Ensemble techniques were proved to produce more accurate results than a single model. Ensemble techniques have already applied in several fields such as web ranking, classification and clustering, time series and regression modeling (Kazienko *et al.* 2013). In this study, two linear and one non-linear ensemble techniques were used as follows to improve the performance of the single models.

Technique 1: simple averaging ensemble (SAE)

In the proposed SAE technique, first the FFNN, ANFIS, SVM and MLR models are trained and tested separately, then the average of FFNN, ANFIS, SVM and MLR outputs is compared and tested against the test observed values (see Figure 3). The general formula for SAE is given as:

$$P_{(t)} = \frac{1}{N}\sum_{i=1}^{N} p_i(t) \qquad (2)$$

where N is the number of learners (here $N = 4$) and p_i denotes to the output of single model (i.e. FFNN, ANFIS, SVM and MLR) at time t.

Technique 2: weighted average ensemble (WAE)

Weighted averaging is predicted by assigning different weights to the individual outputs based on the relative significance of the outputs (see Figure 3). The weight is assigned to each output based on relative importance which is not in the case of simple outputs. The weighted averaging model is expressed as:

$$P_{(t)} = \sum_{i=1}^{N} w_i p_i(t) \qquad (3)$$

where w_i is the applied weight on output of i^{th} model which can be determined based on the model performance as:

$$w_i = \frac{DC_i}{\sum_{i=1}^{N} DC_i} \qquad (4)$$

DC_i is the performance efficiency of i^{th} single model.

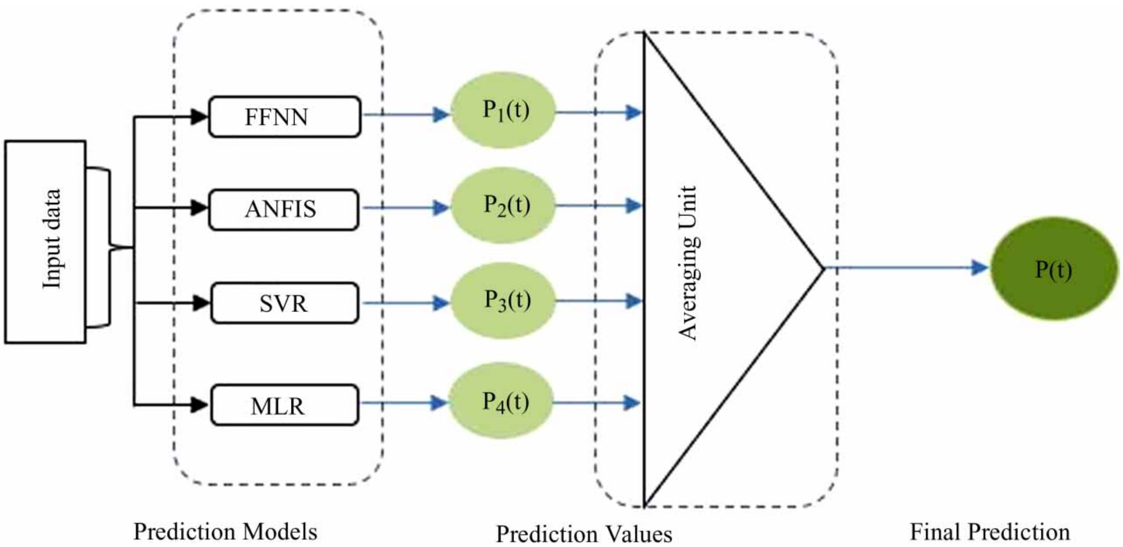

Figure 3 | Schematic of simple and weighted averaging ensemble techniques.

Technique 3: non-linear neural ensemble (NNE)

In the neural ensemble techniques, non-linear averaging is performed by training another neural network. The input layer of the neural ensemble model is fed by the outputs of the considered models, each of which is assigned to one neuron in the input layer. In a neural ensemble model like single FFNN, considering tangent sigmoid as activation functions of hidden and output layers, the network can be trained using BP algorithm and the best structure and epoch number of the ensemble network can be determined through the trial–error procedure. It is clear that, other non-linear kernels (e.g. ANFIS, SVM, etc.) can be alternatively used for such a non-linear ensembling, but FFNN was used in this study since it is the most common AI based method.

RESULTS AND DISCUSSION

Results of single model predictions of BOD_{eff}, COD_{eff} and TN_{eff}

It is worth mentioning that, determining the number of hidden neurons, training epoch number and transfer functions are essential aspects in designing the FFNN model. Lavenberg Marquardt was chosen and used in this study as the BP training algorithm due to its fast learning and high-performance accuracy. Four different FFNNs for each output were trained considering different input combinations (as presented in Table 2) and the best architecture was determined through a trial–error process. FFNN (5-1) which include five inputs and one output neurons found to be the best for all three simulated variables as shown in Table 3. It is noticed that obtained results of FFNN are satisfactory for predicting the NWWTP performance, which are supported by the values of DC and RMSE displayed in Table 3, except for the simulation of BOD_{eff}. For the determination of appropriate ANFIS model, different types of membership functions were examined by trial–error process. ANFIS models were trained for each model using hybrid algorithms (Table 3). According to Table 4, ANFIS (5-1) was found to be the best model and the simulated results showed a good level of satisfaction. In addition, Table 3 confirms that increase in performance is considerable due to the increase in input parameters. ANFIS performance was increased in verification phase up to 10% for $BOD_{eff,}$ with no or negligible increase for COD_{eff} and TN_{eff} (Table 3). This shows that for predicting the models of NWWTP, ANFIS is more recommended for BOD_{eff} modeling while both FFNN and ANFIS could be applied for the simulation of COD_{eff} and TN_{eff}. For SVM modeling, the optimum model was obtained by adjusting two parameters, i.e. squared kernel and regularization constant parameter until the desired output set was achieved. Table 3 shows the simulated results of BOD_{eff}, COD_{eff} and TN_{eff} by SVM model. Table 3 proves that, for predicting the NWWTP performance, the use of SVM model is recommended for simulation of COD_{eff} and TN_{eff}, with regards to BOD_{eff} which the result was not so reliable. In case of BOD_{eff}, SVM (4-1) was found to be the best while for COD_{eff} and TN_{eff}, SVM (5-1) could lead to better outcomes. MLR model was also applied as a conventional method to predict the performance of NWWTP. Four different models were used to simulate the BOD_{eff}, COD_{eff} and TN_{eff} (Table 3). The best model was determined according to the DC and RMSE criteria. The efficiency criteria of the BOD_{eff} modeling ranged between 0.5941–0.6212, 0.0077–0.0193 for DC and RMSE, respectively. However, for COD_{eff} those ranged between 0.7013–0.8669, 0.0012–0.0065 and for TN_{eff} ranged between 0.7055–0.8392, 0.0006–0.0034 for DC and RMSE, respectively. The best model was found to be MLR (5-1) for BOD_{eff} and COD_{eff} while MLR (4-1) for TN_{eff}. Presented results indicate the improved performance of SVM in verification phase up to 10%, 4% and 11% for BOD_{eff}, COD_{eff} and TN_{eff} modeling, respectively, over MLR model. It is apparent that SVM model slightly demonstrated the enhancement of prediction capability than MLR model, and the MLR also clearly indicates the extent of effluent removal efficiency and plant performance in NWWTP.

Table 3 also justifies that, when all the variables are fed into the models to simulate the outputs, the prediction turns to improve in terms of performance criteria. It can be observed from Table 3 that ANFIS is the best model among all applied models due to capability of the fuzzy concept to handle uncertainty in the process. Meanwhile, COD_{eff} and TN_{eff} provide reliable accuracy while BOD_{eff}

Table 2 | Input variables in FFNN, ANFIS, SVM and MLR models

Model input variables	Model output variable(s)
$pH_{inf,}$ $Cond_{inf}$	
$pH_{inf,}$ $Cond_{inf}$, BOD_{inf}	CODeff or TNeff or *BODeff*
$pH_{inf,}$ $Cond_{inf}$, $BOD_{inf,}$ COD_{inf}	
$pH_{inf,}$ $Cond_{inf}$, $BOD_{inf,}$ $COD_{inf,}$ TN_{inf}	

Table 3 | Performance efficiency of single FFNN, ANFIS, SVM and MLR models

	FFNN			
	Calibration		**Verification**	
	DC	**RMSE[a]**	**DC**	**RMSE[a]**
BOD_{eff}				
FFNN (2 – 1)	0.5889	0.0081	0.6721	0.0106
FFNN (3 – 1)	0.6304	0.0076	0.6088	0.0105
FFNN (4 – 1)	0.5788	0.0071	0.5046	0.0105
FFNN (5 – 1)	**0.6779**	**0.0065**	**0.6600**	**0.0102**
COD_{eff}				
FFNN (2 – 1)	0.9081	0.0014	0.9005	0.0062
FFNN (3 – 1)	0.9297	0.0012	0.9104	0.0062
FFNN (4 – 1)	0.9102	0.0011	0.9100	0.0059
FFNN (5 – 1)	**0.9328**	**0.0014**	**0.9363**	**0.0053**
TN_{eff}				
FFNN (2 – 1)	0.9367	0.0006	0.8943	0.0034
FFNN (3 – 1)	0.9325	0.0006	0.8949	0.0034
FFNN (4 – 1)	0.9258	0.0007	0.896	0.0033
FFNN (5 – 1)	**0.9343**	**0.0004**	**0.9022**	**0.0034**

	ANFIS			
	Calibration		**Verification**	
	DC	**RMSE[a]**	**DC**	**RMSE[a]**
BOD_{eff}				
ANFIS (2 – 1)	0.6068	0.0079	0.6766	0.0106
ANFIS (3 – 1)	0.6806	0.0066	0.7454	0.0103
ANFIS (4 – 1)	0.7512	0.0074	0.7183	0.0093
ANFIS (5 – 1)	**0.7828**	**0.0053**	**0.7640**	**0.0083**
COD_{eff}				
ANFIS (2 – 1)	0.9087	0.0013	0.9060	0.0062
ANFIS (3 – 1)	0.9279	0.0014	0.9020	0.0059
ANFIS (4 – 1)	0.9492	0.0013	0.9256	0.0054
ANFIS (5 – 1)	**0.9388**	**0.0012**	**0.9260**	**0.0037**
TN_{eff}				
ANFIS (2 – 1)	0.8365	0.0006	0.7946	0.0034
ANFIS (3 – 1)	0.9389	0.0006	0.8985	0.0033
ANFIS (4 – 1)	0.9383	0.0006	0.8113	0.0031
ANFIS (5 – 1)	**0.9571**	**0.0005**	**0.9410**	**0.0010**

	SVM			
	Calibration		**Verification**	
	DC	**RMSE[a]**	**DC**	**RMSE[a]**
BOD_{eff}				
SVM (2 – 1)	0.59607	0.0084	0.5635	0.0107
SVM (3 – 1)	0.62809	0.0081	0.6393	0.0106
SVM (4 – 1)	**0.6554**	**0.0080**	**0.6119**	**0.0106**
SVM (5 – 1)	0.6013	0.6081	0.1739	0.0106
COD_{eff}				
SVM (2 – 1)	0.9009	0.0039	0.9000	0.0063
SVM (3 – 1)	0.9051	0.0053	0.9007	0.0061
SVM (4 – 1)	0.9091	0.0048	0.9109	0.0061
SVM (5 – 1)	**0.9096**	**0.0047**	**0.9018**	**0.0060**
TN_{eff}				
SVM (2 – 1)	0.8966	0.0010	0.7055	0.0034
SVM (3 – 1)	0.8929	0.0010	0.7957	0.0033
SVM (4 – 1)	0.8892	0.0010	0.7958	0.0033
SVM (5 – 1)	**0.8642**	**0.0013**	**0.8050**	**0.0032**

	MLR			
	Calibration		**Verification**	
	DC	**RMSE[a]**	**DC**	**RMSE[a]**
BOD_{eff}				
MLR (2 – 1)	0.5999	0.0083	0.61003	0.0109
MLR (3 – 1)	0.6146	0.0078	0.5941	0.0193
MLR (4 – 1)	0.6007	0.0077	0.6000	0.0110
MLR (5 – 1)	**0.6212**	**0.0077**	**0.6037**	**0.0109**
COD_{eff}				
MLR (2 – 1)	0.8062	0.0012	0.7013	0.0065
MLR (3 – 1)	0.8455	0.0013	0.8187	0.0064
MLR (4 – 1)	0.8585	0.0013	0.7090	0.0064
MLR (5 – 1)	**0.8669**	**0.0014**	**0.8591**	**0.0064**
TN_{eff}				
MLR (2 – 1)	0.7375	0.0006	0.7029	0.0034
MLR (3 – 1)	0.7377	0.0006	0.7117	0.0034
MLR (4 – 1)	**0.8392**	**0.0006**	**0.7930**	**0.0034**
MLR (5 – 1)	0.7956	0.0007	0.7227	0.0034

[a]Since all data are normalized, the RMSE has no dimension.

Table 4 | Results of the proposed ensemble techniques

Ensemble technique[a]	Predicted variables	Calibration DC	Calibration RMSE[b]	Verification DC	Verification RMSE[b]
SAE	BOD_{eff}	0.884	0.006	0.860	0.008
	COD_{eff}	0.909	0.004	0.903	0.004
	TN_{eff}	0.897	0.009	0.873	0.002
WAE	BOD_{eff}	0.891	0.005	0.806	0.003
	COD_{eff}	0.919	0.004	0.900	0.009
	TN_{eff}	0.947	0.006	0.934	0.002
NNE	BOD_{eff}	0.902	0.085	0.899	0.053
	COD_{eff}	0.958	0.052	0.947	0.024
	TN_{eff}	0.979	0.020	0.968	0.015

[a]The result has been presented for the best structure.
[b]Since all data are normalized, the RMSE has no dimension.

is found to be the worst in all models. The fair BOD indicates the presence of organic matter and bacteria, the main focus of NWWTP is to reduce the BOD, COD and TN in the effluent before discharging to natural waters. However, for the simulation of BOD_{eff} in verification phase, ANFIS performance showed increase up to 10%, 15% 16%, for FFNN, SVM and MLR, respectively. For all models, the verification and calibration results were used to assess the accuracy and efficiency of the algorithm in terms of DC and RMSE. Nevertheless, in hierarchical comparison to other models, FFNN ranked second best, followed by SVM and finally MLR model. It is also clear that the ratio between the various components in wastewater has a significant influence on the selection and functioning of NWWTP processes. In NWWTP, COD/BOD ratio is high which indicates that substantial part of organic matter will be difficult to degrade biologically, leading to fair result of BOD_{eff}. The result of BOD_{eff} modeling also proved that the pollutions load is mostly contributed from the households and institutions with low significant contribution from the industrial catchment. Also, the variations and compositions of NWWTP are contributed by the amount of organic waste produced by domestic, institutional and commercial areas. Figure 4 shows the time series plot of the fitted models for three simulated outputs via ANN, SVM, ANFIS and MLR.

As seen in Figure 2, there are some missing data, mostly for COD. Considering the presented input layers in Table 2, different alternatives may be used as input parameters of models. For example, in the case of several missing data of COD, the input combination which does not include COD may be used for modeling. Of course, the performance will

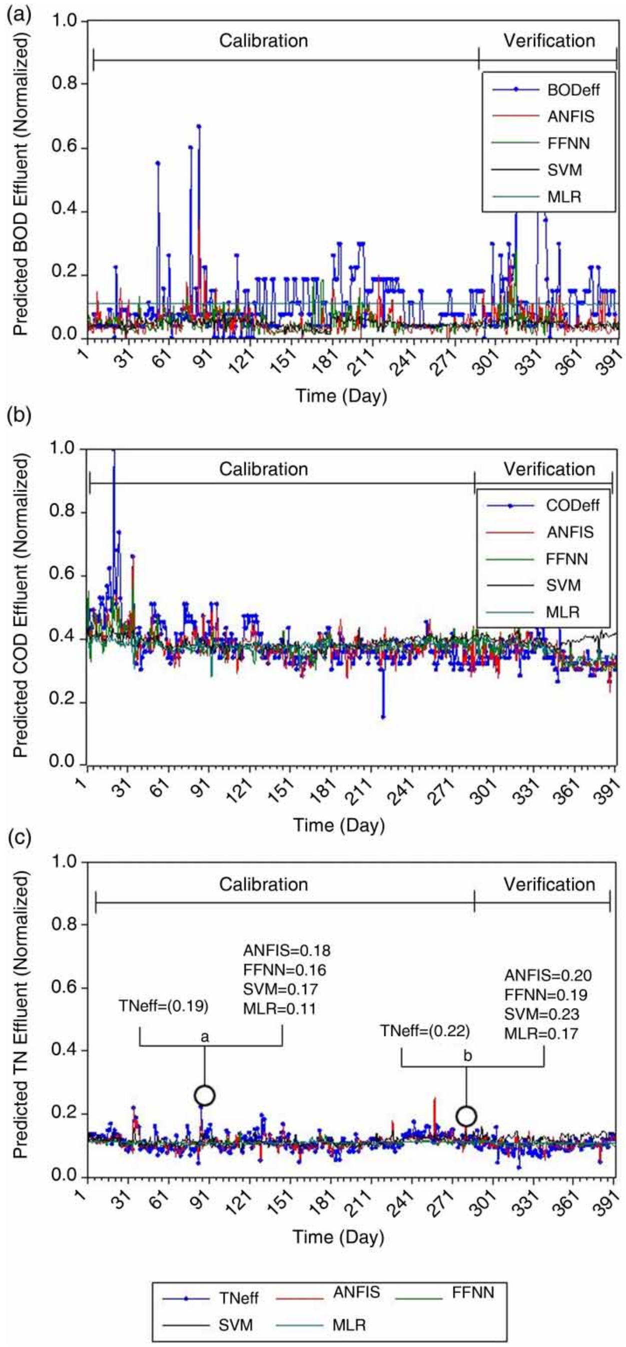

Figure 4 | Observed vs predicted time series obtained by best single models for (a) BOD_{eff}, (b) COD_{eff} and (c) TN_{eff}.

be lower than when we have COD data in input layer, but the data availability dictate this to the modeler.

Also, some appropriate methods may be used to fill in the missing data first and then to use the synthetic data in modeling, but since in this case the synthetic data may include larger amount of error, then the developed AI using such data may not lead to appropriate results (Nourani *et al.* 2012a, 2012b).

Figure 4 depicts plot of predicted TN_{eff} by different methods highlighting two sample points (a) and (b). From this figure, it is clear that for sample point (a) ANFIS model could lead to a bit better performance than FFNN and SVM models. On the other hand, for sample point (b), FFNN and SVM models are better than ANFIS model. Therefore, although overall performance of one of the models may be better for whole time series, at different spans of time series, the performance of the models may be different. As such, at different conditions, different methods may lead to different outcomes and so it is a logical idea to ensemble the outcomes of different methods to get more accurate results for the future predictions.

Results of ensemble predictions for BOD_{eff}, COD_{eff} and TN_{eff}

The ensemble of outputs from FFNN, ANFIS, SVM, and MLR were carried out based on proposed SAE, WAE and NNE to improve the overall prediction accuracy of the single models. Table 4 shows the obtained results by SAE, WAE, and NNE techniques. The obtained DC and RMSE values for both calibration and verification phases show improvement in the modeling efficiency with regards to the single models. The performance of ensemble techniques depends on the accuracy of each individual model as each model has its own drawback and merit in the modeling process. The results also proved that, for the prediction of BOD_{eff} in NWWTP, ensemble methods could lead superior results with regard to single models (Table 4). This is because the integration of single model's outputs reduces the variance, bias and improves performance of the overall modeling. Despite a reliable result for all the ensemble techniques, NNE found to be more accurate, followed by WAE and lastly SAE. In the verification phase, SAE, WAE and NNE increased the efficiency performance of AI modeling up to 14%, 20% and 24%, respectively, for predictions of BOD_{eff} and up to about 5% for modeling both COD_{eff} and TN_{eff} parameters. This proved remarkable improvement in the prediction of BOD_{eff} which was found poor using single models. Ensemble methods aimed primarily not only to integrate a set of models but also to decrease the weaknesses of each single model and come up with the enhanced and composite model, which is feasible, and more reliable with high accuracy than single models. According to Table 4, the results of WAE slightly outperformed SAE due to the fact that weight is assigned to each parameter based on relative importance which is not in the case of simple averaging. The performance of NNE is better than two ensemble techniques in both calibration and verification steps, because of the robustness of NNE in handling non-linear interactions, and able to back propagate the produced error during calibration phase until the desire result is achieved. Figure 5 shows the results obtained by NNE as scatter plot and time series plots for BOD_{eff}, COD_{eff} and TN_{eff} versus observed values.

In non-linear method (e.g. proposed neural ensemble technique), the errors (or noise of data) of the single predictions may be non-linearly magnified. Therefore, for only a few (two or three) extreme values which include higher values and consequently may have a higher degree of errors, the non-linear method may not provide good agreement with the measured values. But this can be true only for two or three samples (as seen in Figure 5) and for most of other samples which include lower error, the errors are not significantly magnified and so, the overall performance (measured by RMSE and DC) will be okay. Actually to handle such problems, linear method is also embedded in the modeling (as already proposed in this study) but sometimes it is not enough and, in this case, a data preprocessing approach such as de-noising method (e.g. see Nourani *et al.* 2014) can be more helpful.

CONCLUSIONS

In this paper, the performance of NWWTP was modeled by different AI models of FFNN, ANFIS, SVM and a conventional MLR. Simple averaging, weighted and neural network ensemble techniques were subsequently employed to enhance the prediction performance of single models. For this purpose, daily data from NWWTP were obtained and DC and RMSE were used in order to determine the prediction performance. The comparison of single models showed that ANFIS was better than other single models in both calibration and verification phases. According to the results, SVM was found to be more reliable than the MLR model. Also, in verification step of BOD_{eff} modeling, the models showed more accurate performance up to 10%, 15% and 16% with regards to FFNN, SVM and MLR models, respectively. In the verification phase of ensemble predictions, SAE, WAE and NNE increased the efficiency of AI modeling up to 14%, 20% and 24%, respectively, in the BOD_{eff} and about 5% for both COD_{eff} and TN_{eff} predictions. Among ensemble techniques, NNE was found to be

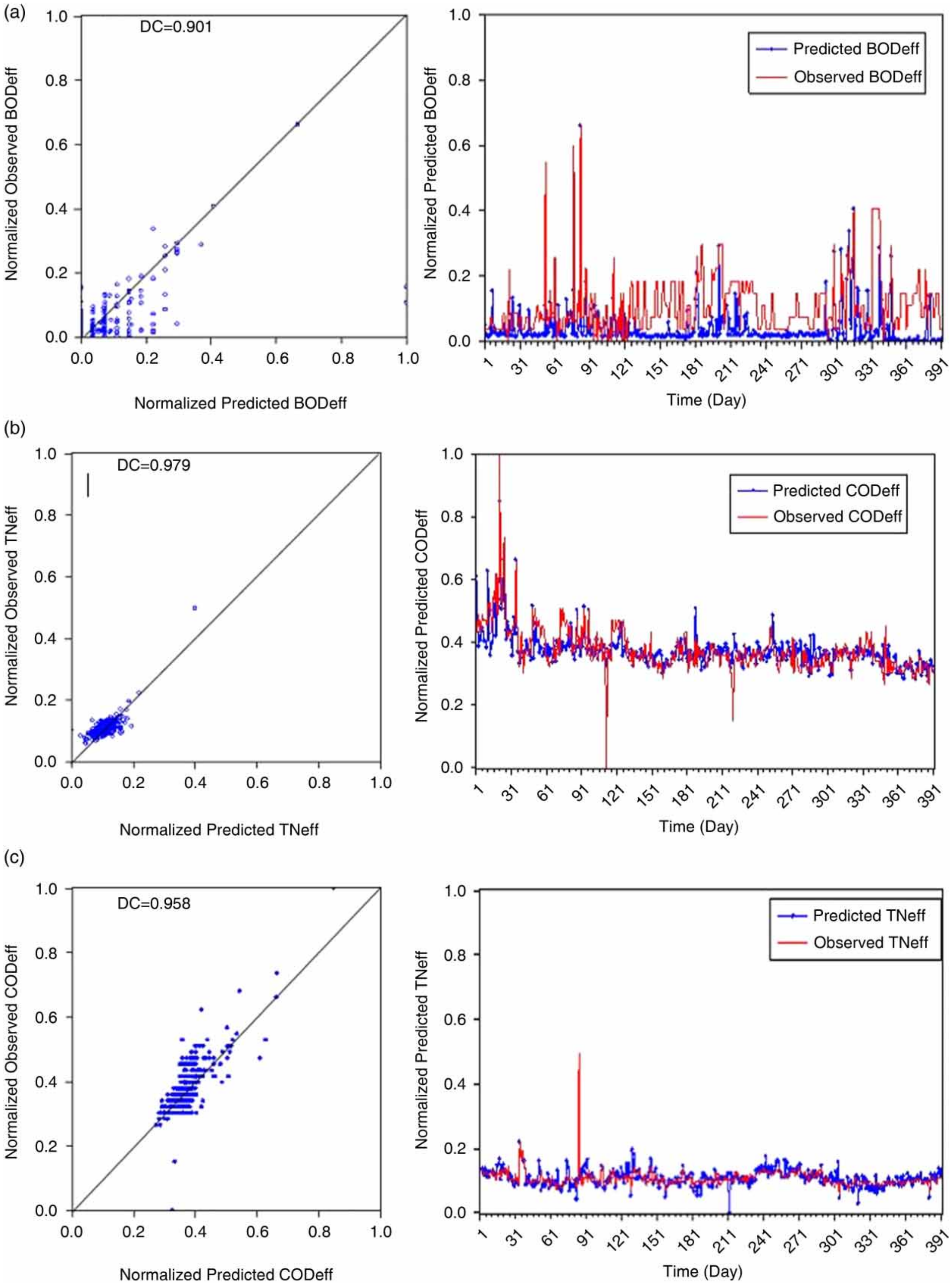

Figure 5 | Scatter plot and time series plots of results obtained by NNE techniques for (a) BOD_{eff}, (b) COD_{eff} and (c) TN_{eff}.

more robust and efficient method of combination and could improve the performance of AI modeling up to 24%. The benefit of the NNE was due to the fact that FFNN model has the ability of handling non-linear behavior in the system. According to the results obtained so far, firstly, single models should not be considered as reliable models for the simulation of BOD_{eff} in NWWTP, as it proved fair results for all AI models. Secondly, all AI and classical models employed in this study were found to be satisfactory and, therefore, recommended for the simulation of COD_{eff} and TN_{eff}. Thirdly, the NWWTP performance indicated high quality of treated effluent which can be used for irrigation and other re-use purposes and the ensemble results provide more reliable and promising results than the single models. Finally, the study may serve as the background for researchers carrying out further studies in NWWTP. The outcomes also suggested that for the application of these models in the real world, the uncertainty involved in the process could be addressed. As such, the application of other AI tools may also be combined in the proposed ensemble approach in order to integrate a set of models so as to come up with a new model which could produce higher accuracy and more reliable estimates than the single models.

In this study, the performance of WWTP was investigated according COD, BOD and TN parameters at outlet due to the availability of the data, but the method may be applied similarly (of course with the new trained AI structures) for other important parameters of wastewater such as ammonia.

REFERENCES

Ahmadi, M. M., Mahdavirad, H. & Bakhtiari, B. 2017 Multi-criteria analysis of site selection for groundwater recharge with treated municipal wastewater. *Water Science and Technology* **76** (4), 909–919.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology 2000 Artificial neural networks in hydrology. II: hydrologic applications. *Journal of Hydrologic Engineering* **5** (2), 124–137.

Aydın Temel, F., Özyazıcı, G., Uslu, V. R. & Ardalı, Y. 2018 Full scale subsurface flow constructed wetlands for domestic wastewater treatment: 3 years' experience. *Environmental Progress and Sustainable Energy* **37** (4), 1348–1360.

Civelekoglu, G., Yigit, N. O., Demopoulos, E. & Kitis, M. 2009 Modelling of COD removal in a biological wastewater treatment plant using adaptive neuro-fuzzy inference system and artificial neural network. *Water Science and Technology* **60** (6), 1475–1487.

Cloke, H. L. & Pappenberger, F. 2009 Ensemble flood forecasting: a review. *Journal of Hydrology* **375** (3–4), 613–626.

Gao, F., Nan, J. & Zhang, X. 2017 Simulating a cyclic activated sludge system by employing a modified ASM3 model for wastewater treatment. *Bioprocess and Biosystems Engineering* **40** (6), 877–890.

Gómez, T., Gémar, G., Molinos-Senante, M., Sala-Garrido, R. & Caballero, R. 2017 Assessing the efficiency of wastewater treatment plants: a double-bootstrap approach. *Journal of Cleaner Production* **164**, 315–324.

Govindaraju, R. S. 2000 Artificial neural networks in hydrology. I: preliminary concepts. *Journal of Hydrologic Engineering* **5** (2), 115–123.

Granata, F., Papirio, S., Esposito, G., Gargano, R. & de Marinis, G. 2017 Machine learning algorithms for the forecasting of wastewater quality indicators. *Water* **9** (2), 105.

Guo, H., Jeong, K., Lim, J., Jo, J., Kim, Y. M., Park, J. P. & Cho, K. H. 2015 Prediction of effluent concentration in a wastewater treatment plant using machine learning models. *Journal of Environmental Sciences* **32**, 90–101.

Hamed, M. M., Khalafallah, M. G. & Hassanien, E. A. 2004 Prediction of wastewater treatment plant performance using artificial neural networks. *Environmental Modelling and Software* **19** (10), 919–928.

Hameed, M., Sharqi, S. S., Yaseen, Z. M., Afan, H. A., Hussain, A. & Elshafie, A. 2017 Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia. *Neural Computing and Applications* **28** (1), 893–905.

Kazienko, P., Lughofer, E. & Trawiński, B. 2013 Hybrid and ensemble methods in machine learning. *Journal of Computer Science* **19** (4), 457–461.

Kim, J. R., Ko, J. H., Im, J. H., Lee, S. H., Kim, S. H., Kim, C. W. & Park, T. J. 2006 Forecasting influent flow rate and composition with occasional data for supervisory management system by time series model. *Water Science and Technology* **53** (4–5), 185–192.

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.

Nourani, V. 2017 An emotional ANN (EANN) approach to modeling rainfall-runoff process. *Journal of Hydrology* **544**, 267–277.

Nourani, V., Hakimzadeh, H. & Amini, A. B. 2012a Implementation of artificial neural network technique in the simulation of dam breach hydrograph. *Journal of Hydroinformatics* **14** (2), 478–496.

Nourani, V., Hosseini Baghanm, A. & Gebremichael, M. 2012b Investigating the ability of Artificial Neural Network (ANN) models to estimate missing rain-gauge data. *Journal of Environmental Informatics* **19** (1), 38–50.

Nourani, V., Hoseini Baghanam, A., Rahimi, A. Y. & Nejad, F. H. 2014 Evaluation of wavelet-based de-noising approach in hydrological models linked to artificial neural networks. In: *Computational Intelligence Techniques in Earth and Environmental Sciences*, Vol. 12 (T. Islam, P. Srivastava, M. Gupta, X. Zhu & S. Mukherjee, eds). Springer, Dordrecht, The Netherlands, pp. 209–241.

Olyaie, E., Abyaneh, H. Z. & Mehr, A. D. 2017 A comparative analysis among computational intelligence techniques for dissolved oxygen prediction in Delaware River. *Geoscience Frontiers* **8** (3), 517–527.

Sargent, R. G. 2009 Verification and validation of simulation models. In: *Simulation Conference (WSC), Proceedings of the 2009 Winter* (pp. 162–176). IEEE.

Sharghi, E., Nourani, V. & Behfar, N. 2018 Earthfill dam seepage analysis using ensemble artificial intelligence-based modeling. *Journal of Hydroinformatics* **20** (5), 1071–1084.

Tsioptsias, N., Tako, A. A. & Robinson, S. 2016 *Model Validation and Testing in Simulation: A Literature Review*.

UNDP 2014 *New Nicosia Waste Water Treatment Plant*. Nicosia, Northern Part Of Cyprus

UNESCO 2015 *Water Supply, Sanitation And Health*. Available at: Http://Www.Unesco.Org/New/En/Natural-Sciences/Environment/Water/Wwap/Facts-Andfigures/Water-Supply-Sanitation-And-Health/.

First received 24 July 2018; accepted in revised form 8 November 2018. Available online 20 November 2018