Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2021.27049

221 (2021) 31–40
May

Metro-environmental data approach for the prediction of chemical oxygen demand in new Nicosia wastewater treatment plant

A.S. Mubarak[a], Parvaneh Esmaili[a], Z.S. Ameen[b], R.A. Abdulkadir[c], M.S. Gaya[c], Mehmet Ozsoz[b], Gaurav Saini[d], S.I. Abba[e,]*

[a]*Department of Electrical and Electronic Engineering, Near East University, Lefkoşa–North Cyprus, via Mersin 10 Turkey, emails: mubarakauwal@gmail.com (A.S. Mubarak), parvaneh.esmaili@neu.edu.tr (P. Esmaili)*
[b]*Department of Biomedical Engineering, Near East University, Lefkoşa–North Cyprus, via Mersin 10 Turkey, emails: zubaidas11@gmail.com (Z.S. Ameen), mehmet.ozsoz@neu.edu.tr (M. Ozsoz)*
[c]*Department of Electrical Engineering, Kano University of Science and Technology, Wudil, Kano, Nigeria, emails: rabiukk@gmail.com (R.A. Abdulkadir), muhdgayasani@gmail.com (M.S. Gaya)*
[d]*Department of Civil Engineering, Sharda University, Greater Noida, Uttar Pradesh, India, email: gaurav.saini@sharda.ac.in*
[e]*Department of Civil Engineering, Baze University, Abuja, Nigeria, email: saniisaabba86@gmail.com*

Received 14 May 2020; Accepted 16 January 2021

ABSTRACT

This study aimed at employing three data-driven models, namely the Hammerstein–Weiner (HW) model, support vector machine (SVM), and feedforward back propagation neural network (FFBPNN) and traditional multi-linear regression, as well as two non-linear ensemble techniques viz: HW-ensemble and FFBPNN-ensemble, were employed to predict chemical oxygen demand (COD_{eff}). For the prediction of the COD_{eff}, two types of data were used, the first one being environmental data from the new Nicosia waste water treatment plant conductivity ($Cond_{inf}$), including total nitrogen (TN_{inf}), total phosphorus (TP_{inf}) and one-effluent parameter COD_{eff} as M1, where the second was meteorology data from the National Aeronautics and Space Administration (NASA) (at 2 m above the Earth's surface), such as relative humidity (R2H), maximum temperature (T2M_M) and mean temperature (T2M) as M2, in a hybrid model M3, which was a combination of both the meteorology and environmental data M1 and M2. According to the performance criteria RMSE and DC of the single models, values of HW-M1 (0.0308 and 0.9686), HW-M2 (0.0322 and 0.9093) and SVM-M3 (0.025 and 0.9486) were recorded. The ensemble technique improved the performance of the single models in the verification phase by 12% and 19% for HW-E and FFBPNN-E, respectively.

Keywords: Artificial intelligence; Nicosia; Wastewater; Chemical oxygen demand; Ensemble learning

1. Introduction

With expanding urbanization, climate change and industrialization, water quality is being reduced due to natural processes and anthropogenic activities. Natural processes such as climate change, volcanic eruptions and forest fires are making significant contributions to the reduction in water quality, along with anthropogenic activities such as industrial effluent, urban development, mining domestic discharge and agricultural drainage and so on. Seasonal variations play essential roles in determining the quality of wastewater before it is being discharged for human consumption and agriculture [1,2]. To deal with industrial effluent, as well as domestic and environmental pollutants, a wastewater treatment plant (WWTP) as a complex dynamic system plays the role of processing the influent diluted

* Corresponding author.

mixture of waste and processes it to reduce the effluent concentration so that it will not cause changes in the natural environment or affect human health, which is paramount for the sustainable development and protection of ecosystems according to the United Nations Educational, Scientific and Cultural Organization (UNESCO) [3–5].

Climate change and meteorology patterns can affect the quality of the influent of the water treatment based on parameters such as precipitation, relative humidity and temperature. At the same time, the effluent can also be affected by the same factor [1,2]. To safeguard the environment and public health, proper control and operation of WWTPs as complicated and dynamical systems is essential. The quality of treated and raw effluent has a considerable impact on the performance and operation of any WWTP [6]. The influent flow has a significant influence on the entire WWTP and energy consumption; however, it is difficult to measure some dominant variables such as chemical oxygen demand (COD) and biological oxygen demand (BOD) as they require a 5-d incubation period. To control and manage the effluent quality, it is crucial to predict the effluent at future times [7].

The complexity of the mechanism of WWTP makes its operational control difficult. Traditionally linear approximation methods are adopted for modelling of the complex treatment system. Reliable and sustainable methods are required to determine the performance of WWTP based on mass flow rate coupled with balanced equations for microbial growth substratum consumption, which is time-consuming, complex and involves non-linear interactions [8]. The prediction of complex WWTP by classical methods is difficult; as such, convenient and reliable modelling tools can play an important role in monitoring and simulating the overall performance of the WWTP. To overcome the drawbacks of the traditional linear methods, non-linear artificial intelligence methods have been recently applied due to their flexibility, accuracy and promising applications in different fields such as engineering, medicine, and sciences [9–11].

As reported by Zhao et al. [2], seasonal variations have an impact on the parameters of wastewater. Based on the multi-pollution source water quality model (MPSWQM) integrated with Bayesian statistics to improve water quality management using load reduction, the curve of decay rate (k) was estimated for the whole year from multi-source; Heilongjiang Province and Harbin City, which are being considered as the key factor in MPSWQM. To improve robustness, key decision-makers could pick up values of k to attain the quality goal at any specific time. Based on surface water samples collected from three different stretches of the river, it is reported that seasonal variations and spatial-temporal changes affect the water quality. Different parameters such as the temperature of the water, pH, BOD, COD, electrical conductivity, total dissolved solids (TDS) and total alkalinity mean were considered and compared with WHO and ISI standards; according to the research, the water is not safe for human consumption. The two-way ANOVA was used to analyze the parameters based on seasonality and location. It was observed that high-temperature decrease solubility of BOD, elevated level of COD and lowered the concentration of DO. Therefore, with long-term effect of temperature, the quality of water in the river will be affected [1]. Gaya et al. [12] compared ANN with HW in the prediction of influent turbidity of WTP using different input parameters. The results showed that ANN outperformed HW, thus indicating that ANN can serve as an acceptable tool for the prediction of turbidity in WTP.

Similarly, De Wu and Lo [13] employed ANN and ANFIS models with pH and colour as inputs to predict the real-time coagulant dosage in a WTP. The prediction performance of ANN was lower in terms of accuracy compared with ANFIS. Al-Asheh et al. [14] predicted the influent parameters in WTP using auto-regressive integrated moving average (ARIMA) and neural network auto-regression (NNAR) models for the prediction of influent WTP parameter. While ARIMA performed well in terms of prediction in their research, the performance of NNAR was better. Abdulkadir et al. [10] forecasted the daily rainfall at Ercan airport in Northern Cyprus using ANFIS, ANN and multi-linear regression (MLR), where ANFIS was found to be the most reliable.

More recent state-of-art studies were carried in the field of WWTP for instance, Kang et al. [15] proposed ANN model for the simulation of odour concentration in WWTP using different input variables (BOD, pH, DO). The results indicated that odour concentration can be successfully predicted using the most utilized AI model (ANN). Yaqub et al. [16] adopted long short-term memory (LSTM) for modelling and removal of ammonium (NH_4–N), total nitrogen (TN), and total phosphorus (TP) in an anaerobic membrane bioreactor (MBR) using various influent parameters. Based on the evaluation criteria, the proposed deep learning state of art model (LSTM) displayed the promising ability with regards to the removal and the prediction of NH_4–N, TN, and TP. Ansari et al. [17] employed neuro-fuzzy logic model coupled with GA and optimization of particle swam algorithms (PSO), the study served as the multi-parametric modelling including BOD, COD, NH_3–N, pH, oil and grease (OG), and SS. The outcomes indicated that both GA-FIS, and POS-FIS outperformed ANFIS model in terms of error estimate. Anter et al. [18] proposed a new algorithm based on the updated version of whale optimization algorithms (WOA) integrated with feature input selection called chaos theory and fuzzy logic (CF-BWOA). The model proved the capability of detecting sensor faults in WWTP with promising accuracy. However, the literature contains studies [15,17–20] that proposed a predictive approach using a novel ensemble learning model namely; ada boost regression (ABR), gradient boost regression (GBR) and random forest regression for the estimation of TDS, BOD, and COD in Qom industrial wastewater treatment plant, Iran. Based on the performance matrix, the obtained results indicated the prediction skill of ABR model for TDS, GBR model for both COD and BOD. The main contributions of this research are as follows: (1) To propose and compare three AI-based models, namely HW, SVM, FFBPNN, and one traditional MLR model for the simulation of COD_{eff} from new Nicosia WWTP using meteorology, environmental and hybrid meteorology and environmental data; (2) To improve the performance of the single model using two different ensemble learning approaches. (3) To show the correlation and effect of meteorology data in the WWTP process.

2. Materials and methods

2.1. Plant description and used data

The New Nicosia WWTP is the second largest WWTP equipped with MBR technology that will serve both the needs of both Turkish Cypriots and Greek Cypriots (bio-communal). The planned project is aimed to serve a population of 270,000 populations by 2025. The plant will have a 30,000 m^3/d volume in the first stage and 45,000 m^3/d volume for the second stage of capacity, based on the rotation approach. About 10 million m^3 of treated water can be reused for agricultural purpose every year. From the plant, about 3,000 tons of fertilizer is produced and 10%–20% of the electricity generated in the plant will be powered by renewable energy generated by the anaerobic sludge digesters, reducing its carbon dioxide (CO_2) emissions (Fig. 1). The daily data obtained from the new Nicosia WWTP were conductivity ($Cond_{inf}$), total nitrogen (TN_{inf}), total phosphorus (TP_{inf}) and one effluent parameter COD_{eff}. Daily meteorology data were collected from the National Aeronautics and Space Administration (NASA) (at 2 m above the Earth's surface) and included relative humidity (R2H), maximum temperature (T2M_M) and mean temperature (T2M). The available data set was divided into two parts, where 75% of the data were employed for the calibration and the remaining data were used for verification purposes. The output and the input data were normalized to a range of 0–1 before dividing the data into 75% for calibration and 25% for verification [21]. The description statistic of the input–output variables is presented in Table 1.

2.2. Proposed model development

The models proposed for the prediction of $COD_{eff.}$ in this study for the new Nicosia WWTP were HW, SVM, FFBPNN and MLR. The normalized measured data mentioned previously were employed to simulate the Nicosia meteorology data including relative humidity (R2H), maximum temperature (T2M_M) and mean temperature (T2M), while environmental data parameters such as $Cond_{inf}$, TN_{inf}, and TP_{inf} were used to predict COD_{eff}. In addition, non-linear ensemble models were proposed using two different algorithms (HW-E and FFBPNN-E). Different combinations of input parameters for black box data-driven models (HW, SVM and FFBPNN) were examined in the modelling framework, while for the traditional linear regression, MLR was also used. Subsequently, a set of three different models (M1, M2 and M3) was derived based on the correlation between the input and output variables as in Eq. (1).

$$\left.\begin{aligned} M1 &= TP_{inf} + TN_{inf} + Cond_{inf} \\ M2 &= R2H + T2M_M + T2M \\ M3 &= TP_{inf} + TN_{inf} + Cond_{inf} + R2H + T2M_M + T2M \end{aligned}\right\} \quad (1)$$

2.3. Performance criteria

The performances of the predictive models and ensemble techniques are determined after the desire training and

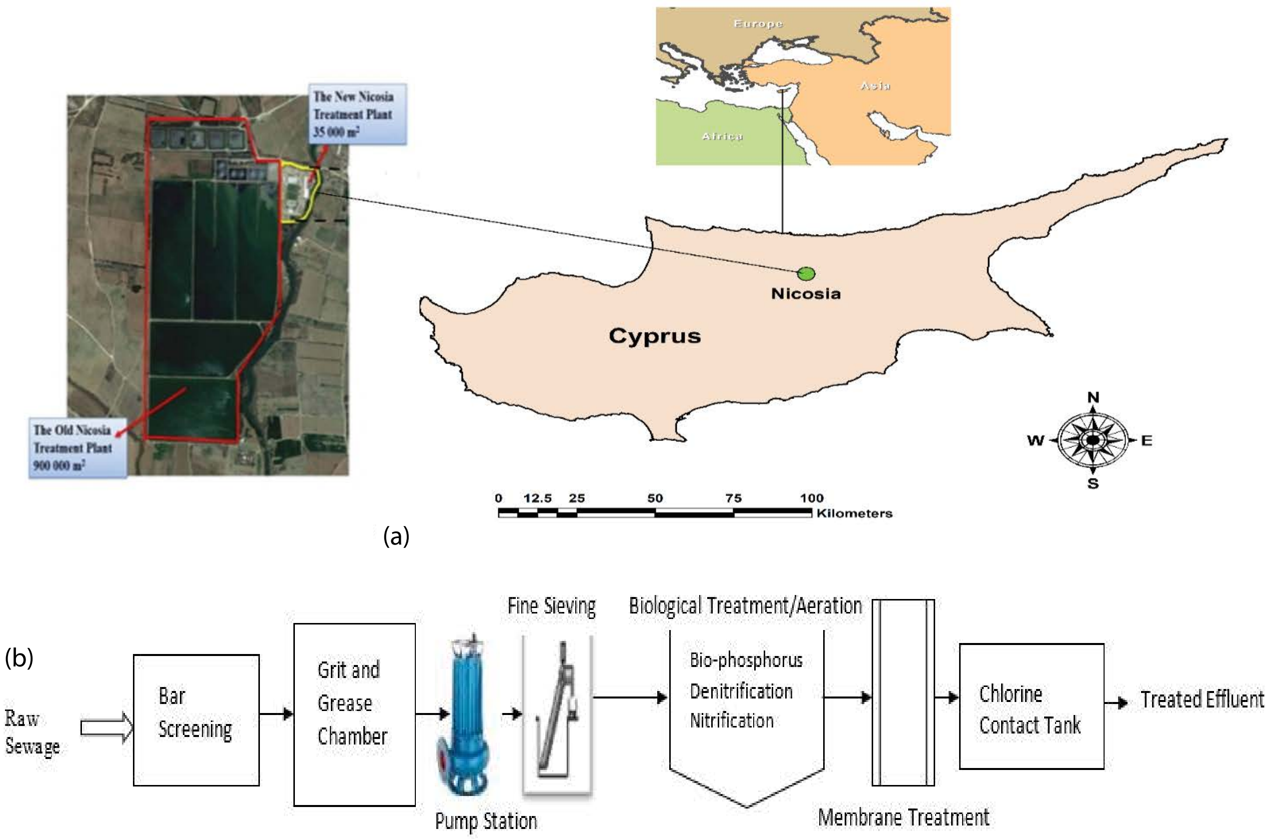

Fig. 1. (a) Map of the study location and (b) diagram of the new Nicosia WWTP process.

testing phase is achieved. The computed values are evaluated using the following measures of goodness-of-fit, determination coefficient (DC), root mean square error (RMSE), mean square error (MSE) and mean absolute error (MAE).

$$\mathrm{DC} = 1 - \frac{\sum_{i=1}^{N}\left(S_i - \hat{S}_i\right)^2}{\sum_{i=1}^{N}\left(S_i - \bar{S}\right)^2} \tag{2}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(S_i - \hat{S}_i\right)^2}{N}} \tag{3}$$

$$\mathrm{MSE} = \frac{\sum_{i=1}^{N}\left(S_i - \hat{S}_i\right)^2}{N} \tag{4}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N}\left[S_i - \hat{S}_i\right]}{N} \tag{5}$$

where N = number of data used, $\bar{S}$ = average observed data, $\hat{S}_i$ = model computed value, S_i = observed value.

2.2.1. Hammerstein–Weiner model (HW)

Hammerstein–Weiner (HW) model is a black box model employed as a non-linear system identification tool. In the HW model, the non-linearity follows and preceded by a linear dynamic system. The estimator configuration of non-linearity estimator for the input and output is a piecewise linear function set at 10 units. The number of units is proportional to the complexity of the system. The intersection of the HW model was characterized as an appropriate illustration with a more precise and understandable relationship to the linear and non-linear systems than the other traditional ANN [22]. Also, the HW model involves a flexible and straightforward process of finding parametric specifications for non-linear models and functionally captures physical knowledge about the system characteristics [23].

2.2.2. Feed forward back propagation neural network

Artificial neural networks as mathematical models have proven to be effective in handling non-linear relationships between input data and target data. ANNs show good performance in various fields when applied in problems such as classification, pattern recognition, forecasting and control systems. A widely applied ANN is FFBPNN, in which generated errors in this network are propagated back until the desired output is achieved. The FFBPNN consists of three layers, as shown in Fig. 2 [8,21].

In this study, the input layer consisted of a combination of influent parameters of the WWTP and meteorology data, while the target was the effluent parameters (as shown later in Table 3). The learning rate determines the intelligence of the network, which overcomes the problem of local minimum and convergence of the network by using trial and error for both the learning rate and architecture (number of neurons, number of layers, transfer function) [24].

Table 1
Descriptive statistics of the data

Parameters	RH2M	T2M_MAX	T2M	$Cond_{inf}$	Total N_{inf}	Total P_{inf}	COD_{eff}
Mean	66.33	22.86	20.04	3.25	85.57	11.12	21.94
Standard deviation	8.22	7.24	6.60	0.44	10.24	1.40	3.93
Kurtosis	–0.53	–1.14	–1.14	1.74	0.78	6.15	15.98
Skewness	0.07	0.05	0.06	–1.48	–0.04	1.54	2.13
Minimum	46.41	7	5.04	1.4	50	7	2
Maximum	87.26	37.43	32.78	4	121	19	55

$\bar{X}$, X_{max}, X_{min}, σ and Csx indicate the mean, maximum, minimum, standard deviation and skewness coefficient, respectively.

Table 2
Correlation analysis of data

Parameters	RH2M	T2M_MAX	T2M	$Cond_{inf}$	Total N_{inf}	Total P_{inf}	COD_{eff}
RH2M	1						
T2M_MAX	–0.51467	1					
T2M	–0.49942	0.995992	1				
$Cond_{inf}$	–0.11865	0.329437	0.326123	1			
Total N_{inf}	–0.02247	0.118767	0.137969	0.143819	1		
Total P_{inf}	0.045282	–0.08983	–0.08013	0.235745	0.460211	1	
COD_{eff}	0.242402	–0.26973	–0.28151	0.239672	0.165253	0.289567	1

2.2.3. Support vector machine

SVM is designed based on machine learning data-driven model, which has achieved good performance in regression, classification and pattern recognition [25]. The difference between ANN and SVM is that in SVM, structural risk minimization and a statistical learning approach are employed to reduce the error, complexity and increase the performance of the network. The two basic structural layers of the SVM are the first layer, which weights the kernel function of the input, and the second layer, which weights the sum of the kernel outputs [25]. In SVM, a linear regression is first fitted to the data and then the outputs go through a non-linear kernel to catch the non-linear pattern of the data. Given a set of calibration data $\{(x_i, d_i)\}_i^N$ (x_i is the input vector, d_i is the actual value and N is the total number of data patterns), the general SVM function is given as:

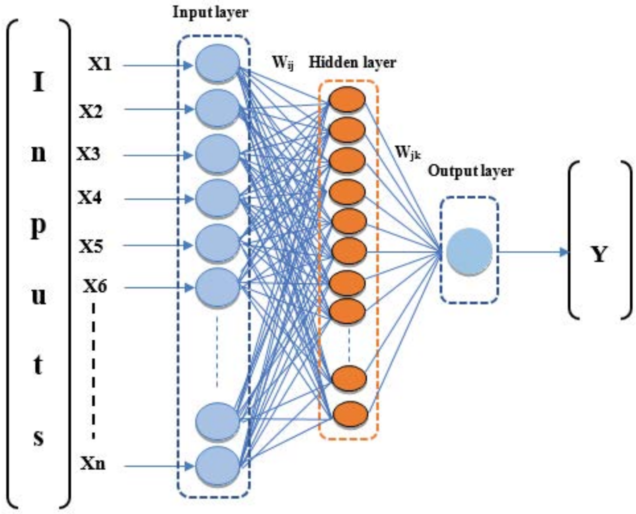

Fig. 2. Schematic of FFBPNN showing input, hidden and output layers.

$$y = f(x) = w\phi(x_i) + b \tag{6}$$

where $\phi(x_i)$ indicates feature spaces, non-linearly mapped from input vector x [25]. The general conceptual model structure of SVM is illustrated in Fig. 3.

2.2.4. Multilinear regression

MLR is a traditional method based on the least square concept used to evaluate the linear relationship between dependent and independent variables, which is the value of the predicted parameter expressed as a linear function. The SVM can be categorized into two simple linear regressions: SLR and ML, where the SLR predicts the linear relationship or correlation between one predictor and one criterion variable, while the MLR forecasts the linear correlation between two or more predictors. The general form of MLR can be represented as follows:

$$\hat{Y} = a_0 + \sum_{j=1}^{m} a_j X_j \tag{7}$$

where $\hat{Y}$ is the model's output, X_j are the independent input variables to the model, and $a_0, a_1, \ldots, a_m$ are partial regression coefficients.

3. Results and discussion

It is essential to determine the appropriate range of the smoothing factor in any FFBPNN, as it can have a significant effect on the simulation ability of the model. The smoothing factor needs to have an average value because extremely large or small values will affect the regression results. Therefore, the smoothing factor was considered within the range of 0.01–1 [26]. The SSVM model with non-linear radial basis function kernel function was created for different conditions. For this purpose, different γ and σ values were tried through a grid search to obtain the best modelling result [27]. Moreover, HW was built

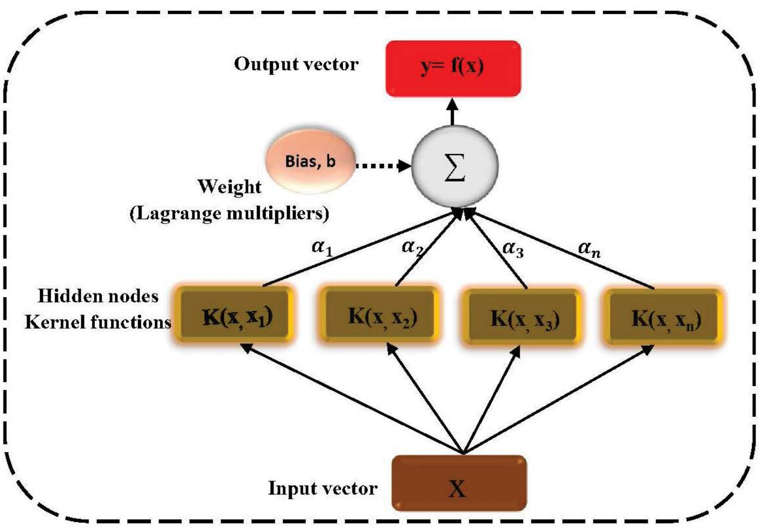

Fig. 3. Conceptual architecture of SVM algorithm.

Table 3
Performance evaluation of HW, SVM, FFBP and MLR

	Calibration				Verification			
Models	MSE	RMSE	MAE	DC	MSE	RMSE	MAE	DC
HW-M1	0.0009	0.0308	0.0238	0.9686	0.0025	0.0501	0.0311	0.9412
HW-M2	0.0010	0.0322	0.0252	0.9149	0.0034	0.0581	0.0399	0.8951
HW-M3	0.0032	0.0569	0.0464	0.9099	0.0042	0.0652	0.0461	0.8738
SVM-M1	0.0063	0.0797	0.0500	0.8149	0.0063	0.0797	0.0500	0.7111
SVM-M2	0.0028	0.0534	0.0406	0.8506	0.0048	0.0696	0.0430	0.7113
SVM-M3	0.0007	0.0259	0.0181	0.9486	0.0038	0.0620	0.0359	0.8111
FFBPNN-M1	0.0014	0.0378	0.0292	0.7303	0.0056	0.0751	0.0504	0.7676
FFBPNN-M2	0.0034	0.0581	0.0438	0.7144	0.0045	0.0671	0.0428	0.8766
FFBPNN-M3	0.0010	0.0318	0.0232	0.8152	0.0044	0.0663	0.0431	0.8742
MLR-M1	0.0015	0.0382	0.0313	0.6105	0.0060	0.0773	0.0521	0.6032
MLR-M2	0.0038	0.0614	0.0513	0.6372	0.0052	0.0722	0.0463	0.6175
MLR-M3	0.0015	0.0389	0.0320	0.7096	0.0044	0.0663	0.0431	0.7504

using the system identification toolbox of MATLAB based on the configuration of a piecewise linear function (range between 10 and 100) for both input and output non-linearity predictors in the case of HW model.

In Table 3, the performance criteria for each model are presented. The performance criteria for the calibration stage are MSE, RMSE, MAE and DC. HW-M1, HW-M2 and SVM-M3 show great performance for the single models in terms of predicting COD_{eff}, with values of 0.0009, 0.0308, 0.0238 and 0.9686 for MSE, RMSE, MAE and DC, respectively, for HW-M1, 0.0010, 0.0322, 0.0252 and 0.9149 for MSE, RMSE, MAE and DC, respectively, for HW-M2 and 0.0007, 0.0259, 0.0181 and 0.9486 for MSE, RMSE, MAE and DC, respectively. It was observed that HW-M1, HW-M2 and SVM-M3 performed greatly for the single models. Analysis of the M3 results from each AI and the traditional technique shows that the RMSE and the DC, M3 combination can be adopted as a satisfying combination for the prediction of COD_{eff}.

To determine the performance on the models, the performance of the verification state RMSE was examined to determine the hierarchy of the performance of the models based on the techniques. For HW, it was observed that the performance of HW-M1 based on the RMSE is up to 15% and 30% greater than HW-M2 and HW-M3, respectively. Also, the performance of SVM-M3 based on RMSE is up to 14% and 20% greater than SVM-M2 and SVM-M3, respectively. And finally, FFBPNN, FFBPNN-M3 outperformed FFBPNN-M2 with up to 1% and up to 13% for FFBPNN-M1. The traditional MLR technique MLR-M3 showed the best performance of RMSE compared with MLR-M2 and MLR-M1 based on the errors, as MLR-2 and MLR-M1 attained 8% and 16% more error, respectively, compared with MLR-M3 (Figs. 4 and 5). The scatter plot for the best single models is presented in Fig. 6.

3.1. Ensemble techniques

In the ensemble techniques, the outputs of the best models were selected and then fed into the ensemble model as inputs. The models employed in this research were non-linear HW-E and FFBPNN. Based on the performance criteria DC and RMSE, it was observed that HW-E outperformed FFBPNN-E, as shown in Table 4, the values of RMSE and DC are obtained as 0.0246 and 0.984, and 0.0293 and 0.961 for HW-E, and FFBPNN-E respectively, the performance compared with the single model's verification state have improved by 12% RMSE for HW-E and 19% for the FFBPNN-E (Fig. 8). This is due to the robustness of FFBPNN in handling non-linear interactions, and its ability to back propagate the error produced during the calibration phase until the desired result is achieved. Consequently, the prediction can be improved using ensemble techniques for the better prediction of COD_{eff}; hence the quality of wastewater can be determined before being released into water bodies for human consumption and agricultural purposes. Fig. 9 shows the scatter plots of the two different ensemble techniques.

However, several factors affect the model's performance such as overfitting in the case of AI (ANN), class in balance, systematic noise associated with the data, pre-processing, model types and randomness of the data. According to the literature [28–31] for a good analysis of any data intelligence model, the efficiency performance should include at least one goodness-of-fit (e.g., R^2) and at least one absolute error measure (e.g., RMSE). Also, several studies have already shown that even for the same type of data set, the performance results may deviate from one model performance to another. For example, R^2 does not take into consideration any biases that might be present in the data. Therefore, a good model might have low R^2 value or a model that does not fit the data might have a high R^2 value. Hence, combining the goodness-of-fit, the error measure and biases measure could lead to promising and reliable simulation.

4. Conclusion

In this work, the prediction of COD_{eff} was performed using conventional MLR and three data-driven models, namely HW, SVM and FFBPNN. Three single models were considered, namely M1, M2 and M3, where M1 consists

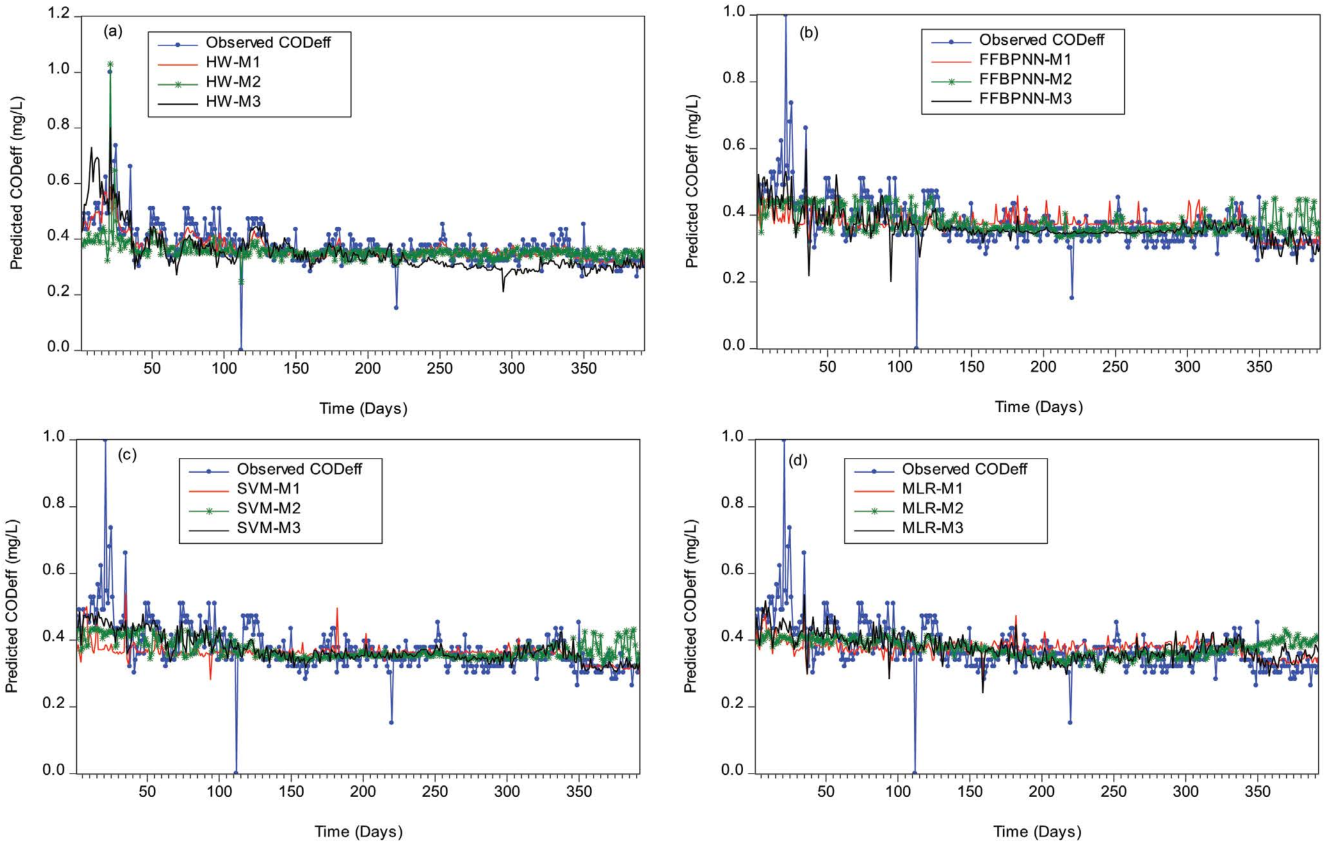

Fig. 4. Time series of the observed and predicted values for (a) HW, (b) SVM, (c) FFBP, and (d) MLR.

of environmental data from the Nicosia waste water treatment plant, M2 consists of meteorology data, while M3 is a combination of M1 and M2. The idea behind the data selection is to prove the effectiveness of the meteorology data in the WWTP process, while two ensemble models were also compared to determine the best model that will improve the prediction of COD_{eff} before being discharged into water bodies for other uses. Based on the single models, HW-M1, HW-M2 and SVM-M3 proved to be the best models based on the RMSE and DC values of 0.0308 and 0.9686 for HW-M1, 0.0322 and 0.9149 for HW-M2, and 0.0259 and 0.9486 for SVM-M3, respectively. In terms of the ensemble techniques, FFBPNN-E proved to be the best with an improvement of up to 19% in RMSE, while the HW-E improved the performance by 12%. In conclusion, the concept of the hybrid data technique for the prediction of COD_{eff} can be adopted into the WWTP (Fig. 7).

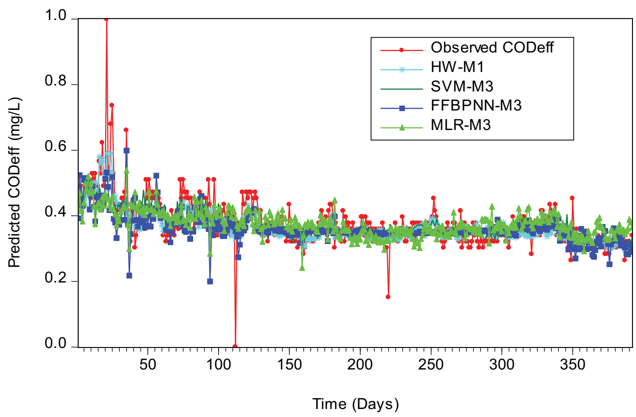

Fig. 5. Time series of the observed and best predicted values for best models.

Table 4
Performance evaluation of HW-E and FFBPNN-E

	Calibration				Verification			
	MSE	RMSE	MAE	DC	MSE	RMSE	MAE	DC
HW-E	0.0006	0.0246	0.0202	0.9844	0.0020	0.0446	0.0312	0.9768
FFBPNN-E	0.0009	0.0293	0.0228	0.9619	0.0017	0.0418	0.0279	0.9527

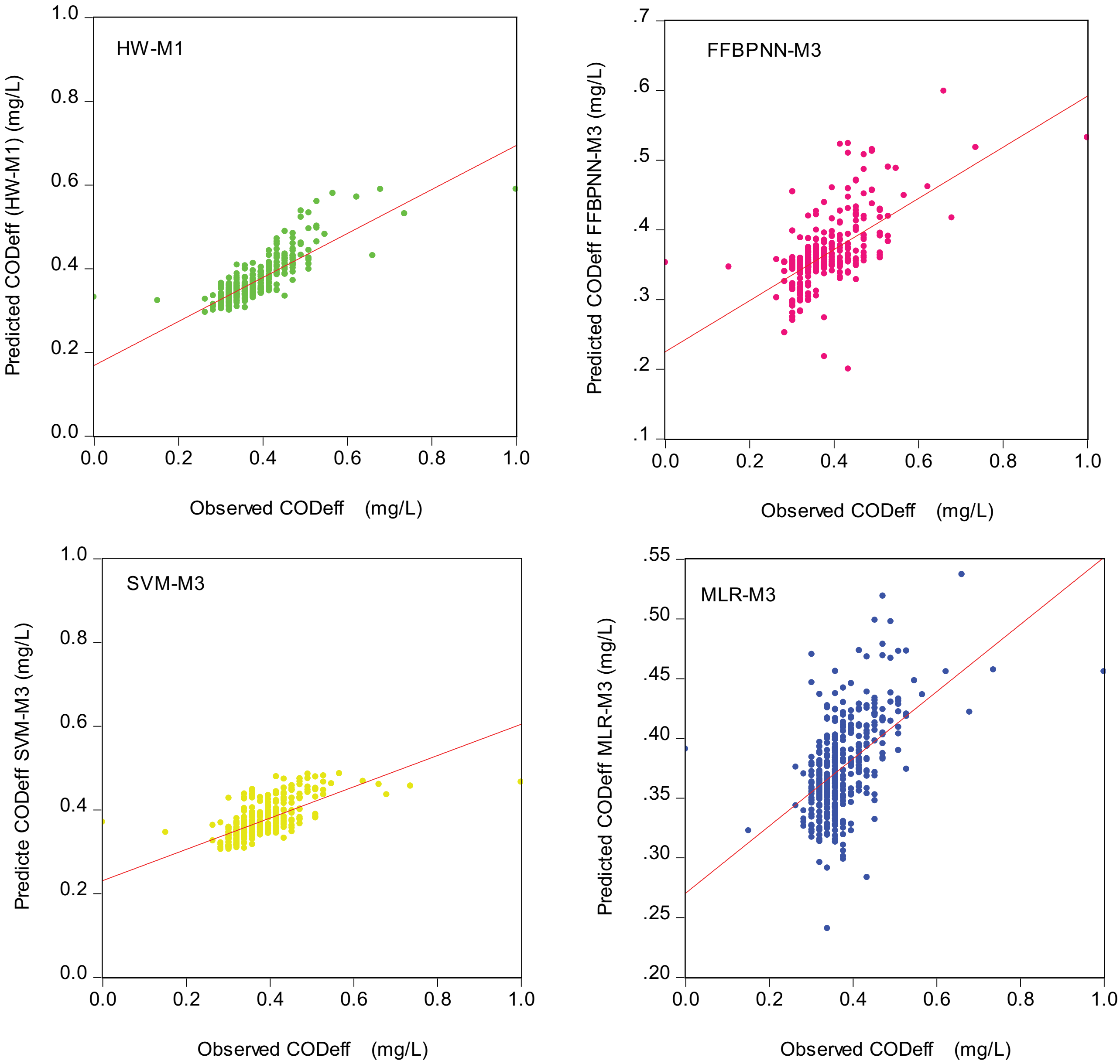

Fig. 6. Scatter plot of results best single models.

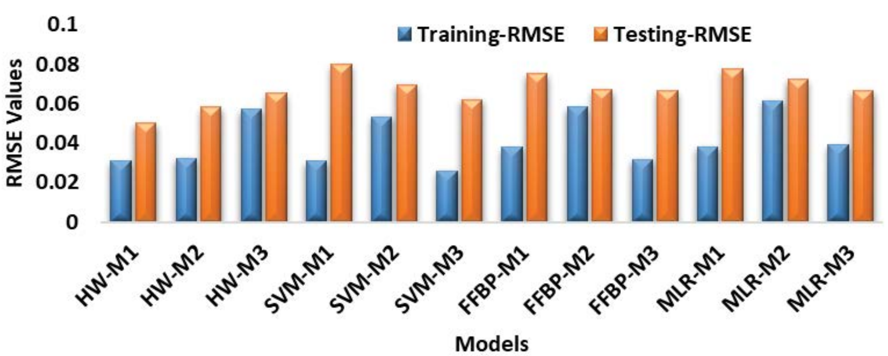

Fig. 7. Performance accuracy in terms of RMSE.

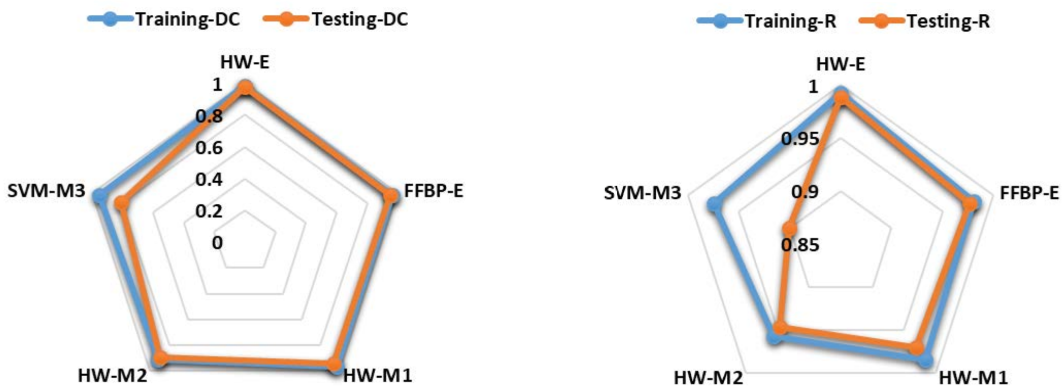

Fig. 8. Performance accuracy for single and ensemble models in terms of DC and R.

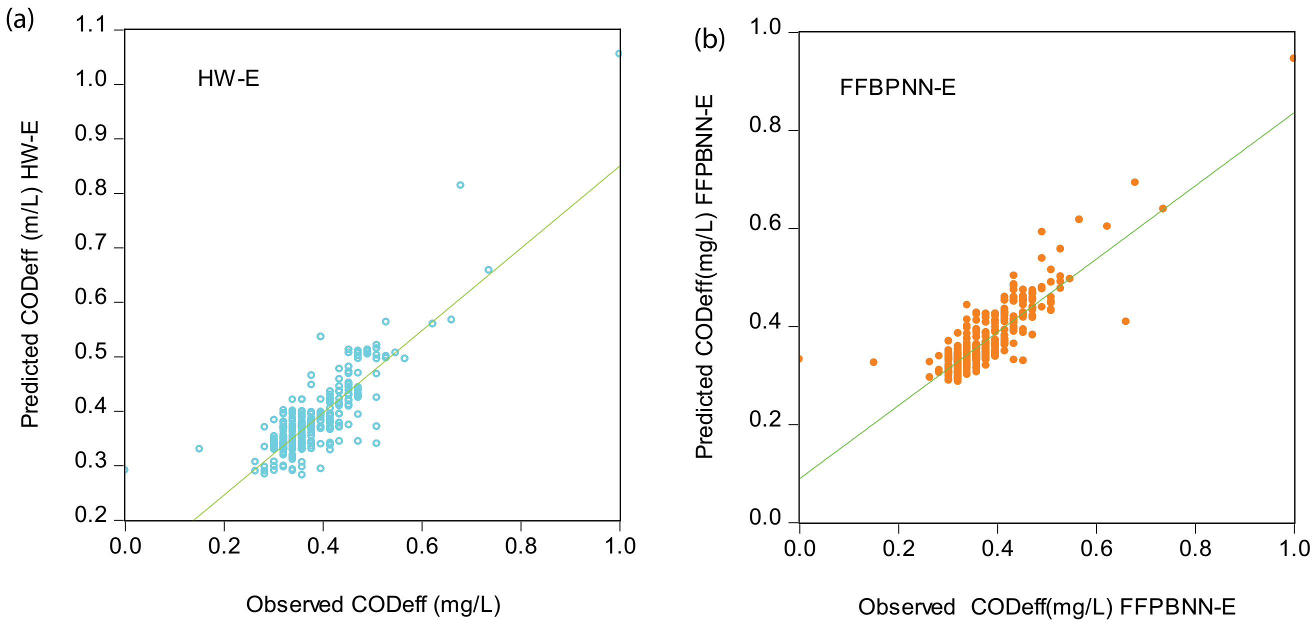

Fig. 9. Scatter plot of ensemble techniques for (a) HW-E and (b) FFBPNN-E.

References

[1] T. Hassan, S. Parveen, B.N. Bhat, U. Ahmad, Seasonal variations in water quality parameters of river Yamuna, India, Int. J. Curr. Microbiol. Appl. Sci., 6 (2017) 694–712, 2017.

[2] Y. Zhao, A. Sharma, B. Sivakumar, L. Marshall, P. Wang, J. Jiang, A Bayesian method for multi-pollution source water quality model and seasonal water quality management in river segments, Environ. Model. Softw., 57 (2014) 216–226.

[3] B.P.K.J.K. Jain, Wastewater Engineering (Including Air Pollution), Second, Laxmi Publications (P) LTD., India, 2014 , 1–655.

[4] S.I. Abba, G. Elkiran, Effluent prediction of chemical oxygen demand from the astewater treatment plant using artificial neural network application, Procedia Comput. Sci., 120 (2017) 156–163.

[5] T. Gómez, G. Gémar, M. Molinos-Senante, R. Sala-Garrido, R. Caballero, Assessing the efficiency of wastewater treatment plants: a double-bootstrap approach, J. Cleaner Prod., 164 (2017) 315–324.

[6] M.M. Ahmadi, H. Mahdavirad, B. Bakhtiari, Multi-criteria analysis of site selection for groundwater recharge with treated municipal wastewater, Water Sci. Technol., 76 (2017) 909–919.

[7] H. Bae, S. Kim, Y.J. Kim, Decision algorithm based on data mining for coagulant type and dosage in water treatment systems, 53 (2006) 321–329.

[8] ASCE, Application of artificial neural networks in hydrology 2000. Artificial neural networks in hydrology. 2: Hydrology applications, J. Hydrol. Eng., 5 (2000) 124–137.

[9] R.S. Govindaraju, Artificial neural networks in hydrology. II: hydrologic applications, J. Hydrol. Eng., 5 (2000) 124–137.

[10] R.A. Abdulkadir, S.I.A. Ali, S.I. Abba, P. Esmaili, Forecasting of daily rainfall at Ercan Airport Northern Cyprus : a comparison of linear and non-linear models Forecasting of daily rainfall at Ercan Airport Northern Cyprus : a comparison of linear and non-linear models, February, pp. 9–10, 2020, doi: 10.5004/dwt.2020.25321.

[11] A.S. Mubarak, A. Süleyman, O. Mehmet, Development of CNN Model for Prediction of CRISPR/Cas12 Guide RNA Activity, International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions, Springer, Cham, 2019 August, pp. 697–703.

[12] M.S. Gaya, M.U. Zango, L.A. Yusuf, M. Mustapha, B. Muhammad, A. Sani, A. Tijjani, N.A Wahab, M.T.M. Khairi, Estimation of turbidity in water treatment plant using

Hammerstein-Wiener and neural network technique, Indones. J. Electr. Eng. Comput. Sci., 5 (2017) 666–672.

[13] G. De Wu, S.L. Lo, Predicting real-time coagulant dosage in water treatment by artificial neural networks and adaptive network-based fuzzy inference system, Eng. Appl. Artif. Intell., 21 (2008)1189–1195.

[14] S. Al-Asheh, F.S. Mjalli, H.E. Alfadala, Forecasting Influent–Effluent Wastewater Treatment Plant Using Time Series Analysis and Artificial Neural Network Techniques, Chemical Product and Process Modeling, 2 (2007).

[15] J.-H. Kang, J. Song, S.S. Yoo, B.-J. Lee, H.W. Ji, Prediction of odor concentration emitted from wastewater treatment plant using an artificial neural network (ANN), Atmosphere (Basel), 11 (2020) 784.

[16] M. Yaqub, H. Asif, S. Kim, W. Lee, Modeling of a full-scale sewage treatment plant to predict the nutrient removal efficiency using a long short-term memory (LSTM) neural network, J. Water Process Eng., 37 (2020) 101388.

[17] M. Ansari, F. Othman, A. El-Shafie, Optimized fuzzy inference system to enhance prediction accuracy for influent characteristics of a sewage treatment plant, Sci. Total Environ., 722 (2020), doi: 10.1016/j.scitotenv.2020.137878.

[18] A.M. Anter, D. Gupta, O. Castillo, A novel parameter estimation in dynamic model via fuzzy swarm intelligence and chaos theory for faults in wastewater treatment plant, Soft Comput., 24 (2020) 111–129.

[19] N. Patel, J. Ruparelia, J. Barve, Prediction of total suspended solids present in effluent of primary clarifier of industrial common effluent treatment plant: mechanistic and fuzzy approach, J. Water Process Eng., 34 (2020) 101146, doi: 10.1016/j.jwpe.2020.101146.

[20] A. Sharafati, S.B.H.S. Asadollah, M. Hosseinzadeh, The potential of new ensemble machine learning models for effluent quality parameters prediction and related uncertainty, Process Saf. Environ. Prot., 140 (2020) 68–78.

[21] V. Nourani, H. Hakimzadeh, A.B. Amini, Implementation of artificial neural network technique in the simulation of dam breach hydrograph, J. Hydroinf., 14 (2012) 478.

[22] X. Zhang, F.W. Zwiers, G. Li, H. Wan, A.J. Cannon, Short-Duration Rainfall, 2017, doi: 10.1038/NGEO2911.

[23] F. Guo, G. Bretthauer, L.E. Phdqv, R.I. Vlpxodwlrq, Identification of Miso wiener and Hammerstein Systems, European Control Conference, 2003, pp. 2144–2149, doi: 10.23919/ECC.2003.7085284.

[24] K. Yetilmezsoy, B. Ozkaya, M. Cakmakci, Artificial Intelligence-Based Prediction Models for Environmental Engineering, Neural Network World, 21 (2011) 193.

[25] V. Vapnik, The Nature of Statistical Learning Theory, p. 188, 1995, doi: 10.1007/978–1-4757–2440–0.

[26] X. Ji, X. Shang, R.A. Dahlgren, 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., 24 (2017) 16062–16076.

[27] V. Nourani, S. Mousavi, F. Sadikoglu, V.P. Singh, Experimental and AI-based numerical modeling of contaminant transport in porous media, J. Contam. Hydrol., 205 (2017) 78–95.

[28] Q.B. Pham, S. I. Abba, A.G. Usman, N.T.T. Linh, V. Gupta, A. Malik, R. Costache, N.D. Vo, D.Q. Tri, Potential of hybrid data-intelligence algorithms for multi-station modelling of rainfall, Water Resour. Manage., 33 (2019) 5067–5087.

[29] D.R. Legates, G.J. McCabe Jr., Evaluating the use of 'Goodness of Fit' measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 5 (1999) 233–241.

[30] M. Alas, S.I. Albrka Ali, Y. Abdulhadi, S.I. Abba, Experimental evaluation and modeling of polymer nanocomposite modified asphalt binder using ANN and ANFIS, 32 (2020) 1–11.

[31] S.I. Abba, Q. Bao Pham, G. Saini, N. Thi Thuy Linh, N. Ahmed, M. Mohajane, M. Khaledian, R. Aliyu Abdulkadir, Q.-V. Bach, Implementation of data intelligence models coupled with ensemble machine learning for prediction of water quality index, Environ. Sci. Pollut. Res., 2020, doi: 10.1007/s11356-020-09689-x.