Journal of Water Process Engineering 44 (2021) 102380

Contents lists available at ScienceDirect

Journal of Water Process Engineering

journal homepage: www.elsevier.com/locate/jwpe

Forecasting effluent and performance of wastewater treatment plant using different machine learning techniques

Mustafa El-Rawy [a,b,*], Mahmoud Khaled Abd-Ellah [c], Heba Fathi [d], Ahmed Khaled Abdella Ahmed [e]

[a] Civil Engineering Department, Faculty of Engineering, Minia University, Minia 61111, Egypt
[b] Civil Engineering Department, College of Engineering, Shaqra University, Dawadmi 11911, Ar Riyadh, Saudi Arabia
[c] Faculty of Engineering, Egyptian Russian university, Cairo, Egypt
[d] ECH Consultant Office, Cairo 11568, Egypt
[e] Department of Civil Engineering, Sohag University, Sohag 82524, Egypt

A R T I C L E I N F O

Keywords:
Wastewater treatment plant
Machine learning techniques
Artificial neural network
Deep learning techniques

A B S T R A C T

Expectation of wastewater quality in wastewater treatment plants (WWTPs) is significant and can decrease the sampling number, cost, decision time, and energy. This paper presents two methods for predicting and forecasting the removal efficiency of total suspended solids (TSS), chemical oxygen demand (COD), biological oxygen demand (BOD_5), ammonia, and sulphide at El-Berka wastewater treatment plant, Egypt. The first method uses different prediction models, includes the use of traditional feed-forward (TF), deep feed-forward backpropagation (DFB), and deep cascade-forward backpropagation (DCB) networks. The TF was generated in three layers: input, hidden, and an output layer. The DFB network comprised of six layers: an input, four hidden, and an output layer. The DCB network was created with six layers (with skip connections): an input, four hidden, and an output layer. The other method is a forecasting model by using deep learning time series forecasting (DLTSF) with a long short-term memory (LSTM) network. The developed models were trained, validated, and tested on a real-life dataset over eight years. The results indicated that the presented models could effectively predict and forecast the future series values of the removal efficiency of the El-Berka WWTP. The DCB network achieved the highest accuracy as compared to those exhibited by the TF and DFB networks. The RMSE and R-squared for training with the DCB model are 1.95 and 0.90, respectively. The RMSE of DLTSF was 0.85 for forecasting of BOD_5. Thus, the DCB and DLTSF models are recommended for evaluating and predicting the performance of WWTP.

1. Introduction

Wastewater treatment is one of the non-conventional water resources used in the irrigation sector in many countries, particularly in arid countries. Owing to this and the development of ecological insurance, the need to evaluate, predict, and redesign wastewater treatment

Abbreviations: ABC, artificial bee colony; ANFIS, adaptive neuro-fuzzy inference system; ANN, artificial neural networks; ARIMA, autoregressive integrated moving average; BOD_5, biochemical oxygen demand; BWWTP, El-Berka WWTP; COD, chemical oxygen demand; CRA, classical regression analysis; DCB, deep cascade-forward backpropagation; DFB, deep feed-forward backpropagation; DLTSF, deep learning time series forecasting; FFNN, feed forward artificial neural network; FL, fuzzy logic; FS, feature selection; GBM, gradient boosting machine; GRU, gated recurrent units network; LCA, life cycle assessment; LS-SVM, least square support vector machine; LSTM, long short-term memory network; MAE, mean absolute error; MAPE, mean absolute percentage error; MARS, multivariate adaptive regression spline; MLP, multilayer perceptron; MLR, multiple linear regression; MSA, multivariate statistical analysis; MSE, mean-square error; PCA, principal component analysis; PLS, partial least analysis; R, connection coefficient; RBF, radial basis function; RF, random forest; RMSE, root-mean-square error; RNNs, recurrent neural networks; R-squared, coefficient of determination; SDAE, stacked denoising auto-encoders; SS, suspended solids; SVM, support vector machine; TDS, total dissolved solids; TF, traditional feed-forward; TLBO, teaching-learning based optimization; TSS, total suspended solids; UASB, up-flow anaerobic sludge blanket reactor; WWTPs, wastewater treatment plants.

* Corresponding author at: Civil Engineering Department, Faculty of Engineering, Minia University, Minia 61111, Egypt.
E-mail addresses: mustafa.elrawy@mu.edu.eg, mustafa.elrawy@su.edu.sa (M. El-Rawy), Mahmoud-khaled@eru.edu.eg (M.K. Abd-Ellah), dr.ahmed_khaled@eng.sohag.edu.eg (A.K.A. Ahmed).

https://doi.org/10.1016/j.jwpe.2021.102380
Received 12 July 2021; Received in revised form 9 October 2021; Accepted 12 October 2021

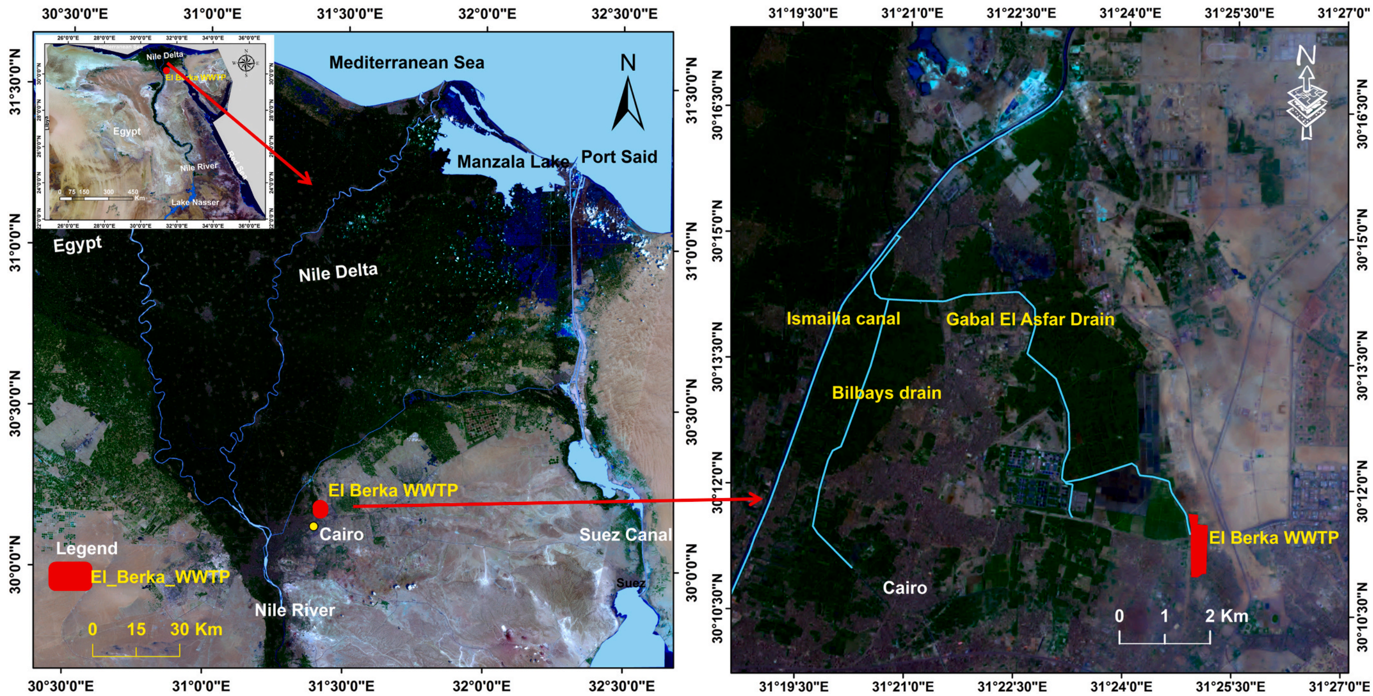

Fig. 1. Location of El Berka WWTW, Ismailia canal, Jabal EL-Asfar and Bilbays drains.

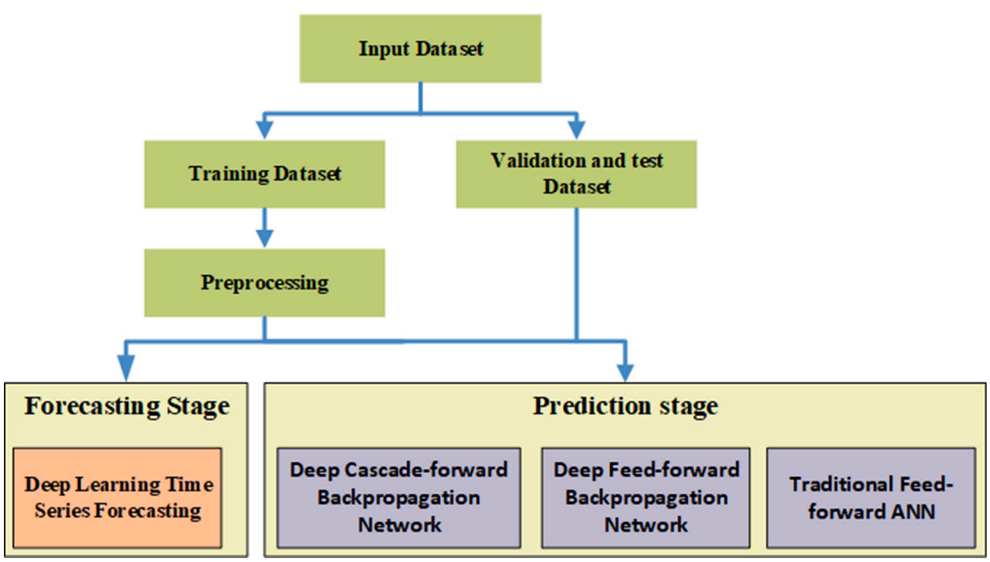

Fig. 2. Structure of the proposed model.

plants (WWTPs) is expanding [1].

Although all WWTP data are easy to assemble, it is still difficult to make an effective decision because of the sophisticated relationships of the data. Consequently, it is necessary to find an appropriate methodology for WWTP modelling and evaluating plant efficiency [2]. Moreover, owing to the territorial distinction of the wastewater quality and different human habits, it is difficult to establish a coordinated relationship that can portray wastewater activity through all anticipated circumstances [3]. Furthermore, the sophisticated relations between the chemical, physical, and biological treatment techniques applied in WWTPs present nonlinear practices that are difficult to depict using straight numerical models [4]. Therefore, using various statistical models can provide better spatial and transient connections to improve the comprehension of wastewater treatment [5]. Recently, a few investigations on wastewater treatments were performed, while focussing on keen strategies concerning the expectations of the WWTP boundaries, implementation of control measures for WWTPs, and assessment of the attributes of the emanate as per the WWTP influent [6].

The most common methodologies and frameworks that have been published in the field of WWTP evaluation are as follows:

1) Life cycle assessment (LCA) is a useful technique to survey the environmental effects of wastewater treatment [7–10].
2) Multivariate statistical analysis (MSA) was used to establish reliable and straightforward predictive models to predict the quality of treated and untreated wastewater [5,11].
3) Partial least analysis (PLS) and principal component analysis (PCA) were developed to obtain data on occasional impact and compositional contrasts in sewage created by industrial and domestic waste [12,13].

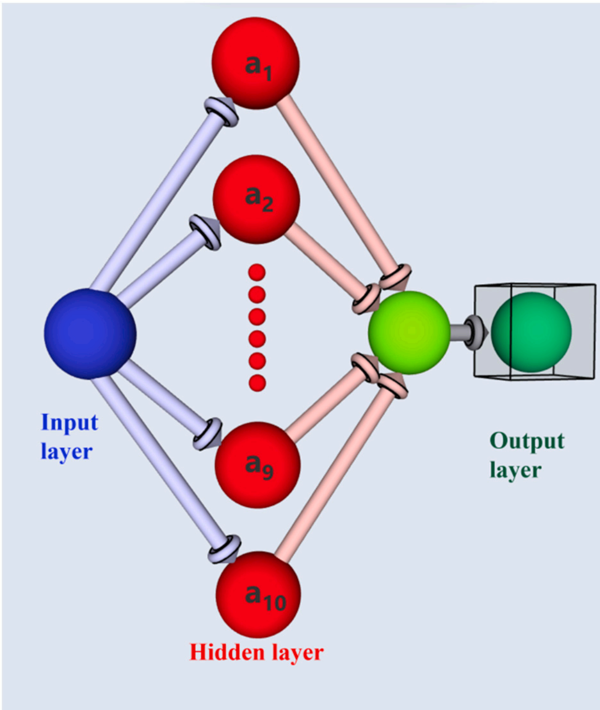

Fig. 3. Structure of traditional feed-forward artificial neural network.

Table 1
Network properties of prediction stage.

Network properties			
Network inputs:	Influent: TSS, BOD_5, COD, ammonia, and sulphide		
Network outputs:	Effluent: TSS, BOD_5, COD, ammonia, and sulphide		
Type of the network:	Feed-forward back-propagation	Deep feed-forward backpropagation	Deep cascade-forward backpropagation
Training function:	TRAINLM		
Adaption learning function:	LEARNGDM		
Performance function:	MSE		

4) Artificial neural networks (ANN) are proven as useful analysis tools to comprehend and mimic the nonlinear conduct of the WWTP and have been implemented as execution appraisal practices by plant administrators and leaders [14–16].

Vyas et al. [17] used ANNs to forecast effluent BOD_5 using 3-year data sets. Jami et al. [18] developed an ANN model for a sewage treatment plant in Malaysia by considering the variations in three different WWTP parameters (BOD_5, SS, and COD), thereby solving the difficulties faced in monitoring and processing activities. Li et al. [19] proposed a model for a modified sequencing batch reactor in a WWTP using the Bayesian approach. Guo et al. [20] used machine learning to develop two different models for predicting the effluent total nitrogen of a wastewater treatment plant in Ulsan, Korea. Tumer & Edebali [3] developed models for the Konya WWTP using ANNs with various MATLAB programming designs to anticipate the plant performance, with a high coincidence between the anticipated and observed variables. Vijayan and Mohan [21] used an ANN to predict the performance of the dairy industry; they focused on estimating the root mean square error (RMSE) from the inputs and outputs given to the ANN, and they revealed a high agreement between the expected and measured variables. Granata et al. [22] applied a regression tree to forecast the total suspended solids (TSS), BOD_5, COD, and total dissolved solids (TDS) in WWTPs. Nourani et al. [23] applied different artificial intelligence-based nonlinear models (FFNN, ANFIS, SVM, and MLR) to predict the effluent BOD_5, COD, and total nitrogen in the Nicosia WWTP. Hamada et al. [16] used several simulations to train radial basis function (RBF) and multilayer perceptron (MLP) networks to choose the best model for forecasting the water quality in the Gaza WWTP. The results showed that the ANN model with minor input data could successfully predict BOD_5, COD, and TSS effluent concentrations. Nadiri et al. [24] supported the ANN model, providing a higher quality model for the Tabriz WWTP by using a nonlinear assortment of fuzzy logic (FL) models under supervision. Khatri et al. [25] used a feed-forward artificial neural network to predict the treatment quality of WWTP in Jamnagar, India. To decrease the public health risk in a WWTP in Jamnagar, India, Khatri et al. [26] used ANN-based models to anticipate fecal and total coliform removal. Arismendy et al. [27] developed an intelligent system to facilitate decision-making, with respect to the activity of a wastewater treatment plant (WWTP).

Wang et al. [28] created a machine learning framework to enhance effluent quality control of wastewater treatment plants. Bagherzadeh et al. [29] investigated the impact of seven different feature selection (FS) approaches on the accuracy of total nitrogen forecasting in WWTP influent flow. Three different Machine Learning (ML) approaches were compared to the four FS-suggested scenarios. Because of the findings, it was discovered that Random Forest (RF) and Gradient Boosting Machine

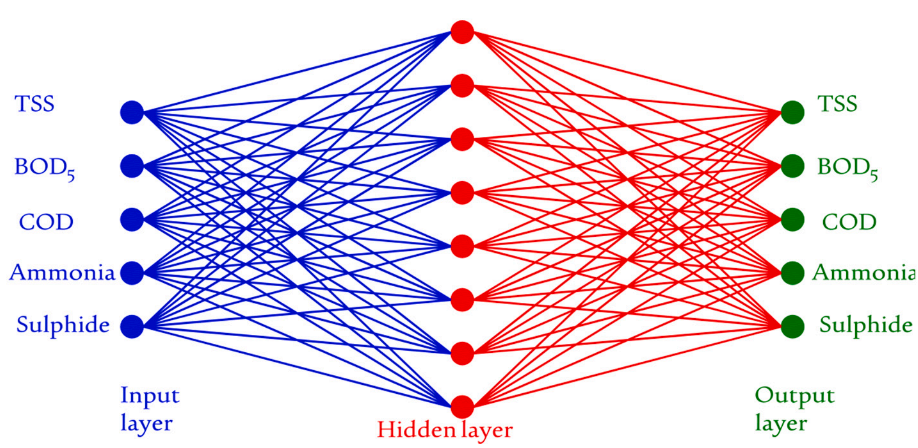

Fig. 4. Structure of the detailed traditional feed-forward (TF) model.

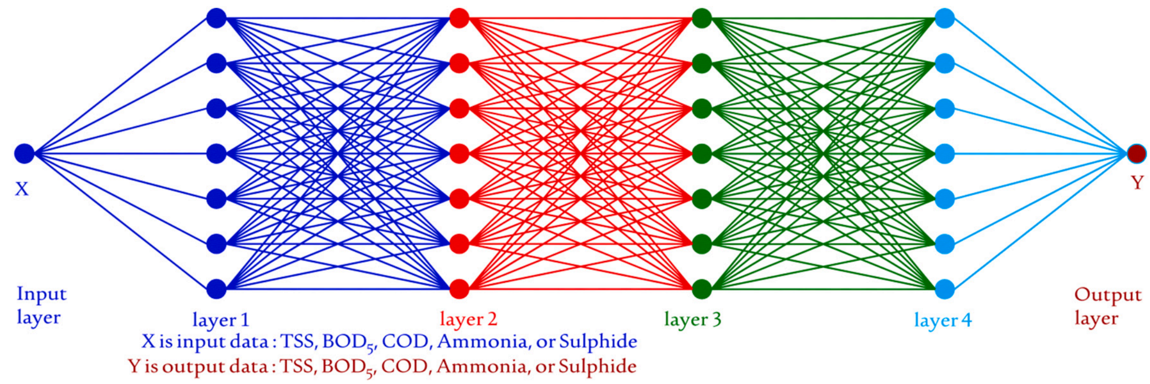

Fig. 5. Structure of the deep feed-forward backpropagation (DFB network.

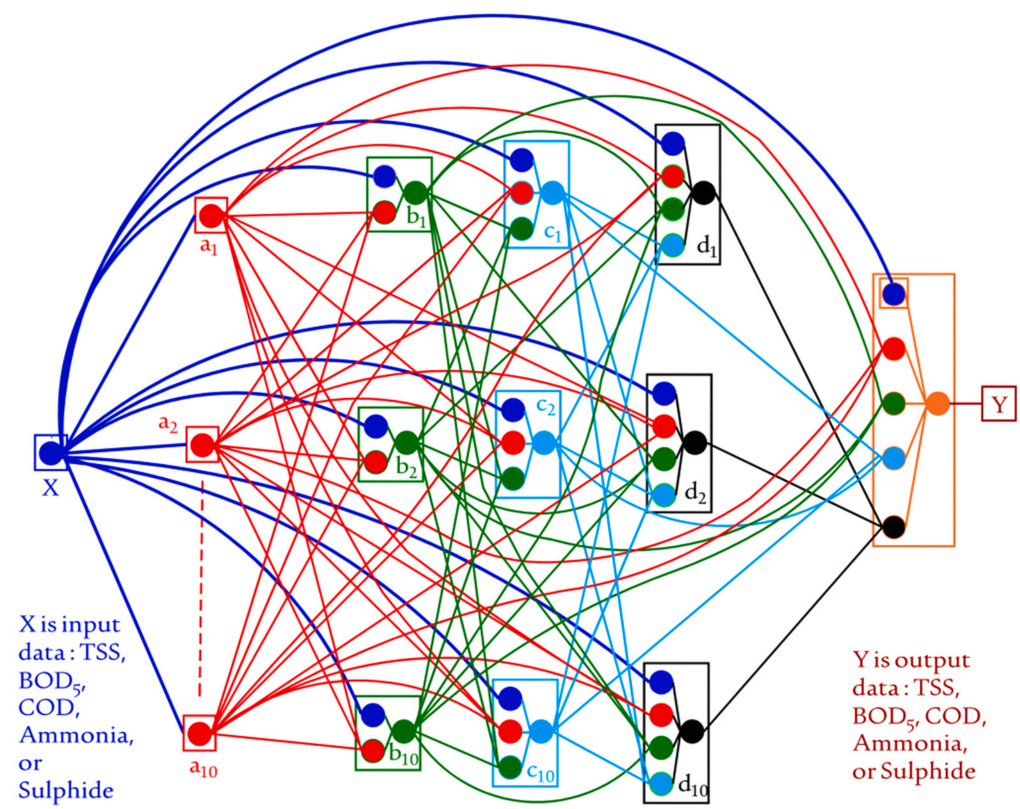

Fig. 6. Structure of deep cascade-forward backpropagation (DCB) network.

(GBM) outperformed ANN. Alsulaili and Refaie [30] investigated the application of ANNs in predicting the influent BOD_5, COD, and TSS concentrations and the performance of WWTPs in Kuwait. Picos-Benítez et al. [31] combined genetic algorithms (GAs) with artificial neural networks (ANNs) to develop an artificial intelligence (AI) model for predicting the performance of up-flow anaerobic sludge blanket reactor (UASB). The developed AI model was proved as a successful tool for performance assessment of the studied reactor [31]. An ANN was developed to predict the effluent COD at WWTP. A comparison between the ANN model and the Multilinear regression analysis (MLR) had been analyzed, and therefore the potency disclosed that ANN model showed the distinguished accuracy in forecasting the effluent COD [32]. Two different machine learning system – multivariate adaptive regression spline (MARS) and least square support vector machine (LS-SVM) – have been proposed to anticipate BOD_5 and COD in rivers. These models were more efficient in estimating BOD_5 and COD in natural streams than adaptive neuro-fuzzy inference system (ANFIS), ANN, and multiple regression equations [33].

A comparison between the efficiency of nonlinear and linear models in predicting effluent parameters of WWTP was performed. The compared models were autoregressive integrated moving average (ARIMA), ANFIS, support vector regression (SVR), and feed forward neural network (FFNN). The results showed the priority of ANFIS, SVR, and FFNN than the ARIMA model in predicting effluent biological oxygen demand and chemical oxygen demand of the WWTP. This was returned to the high rendering quality of artificial intelligence nonlinear models (ANFIS, SVR, and FFNN) compared to that for linear model (ARIMA) [34]. A hybrid model was suggested by merging data-driven model and mechanistic model to provide a more precise model in predicting effluent quality in WWTP [35]. A new deep learning network model was introduced to forecast the biofilm process through oxic/anaerobic of wastewater treatment system. The new model was called

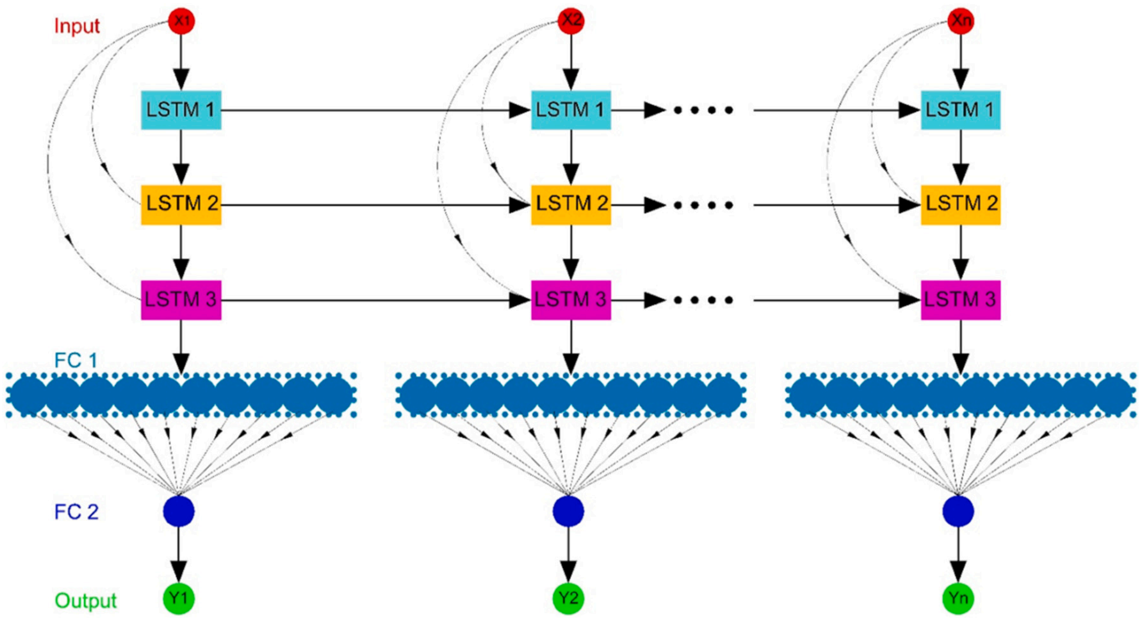

Fig. 7. Structure of deep learning time series forecasting (DLTSF).

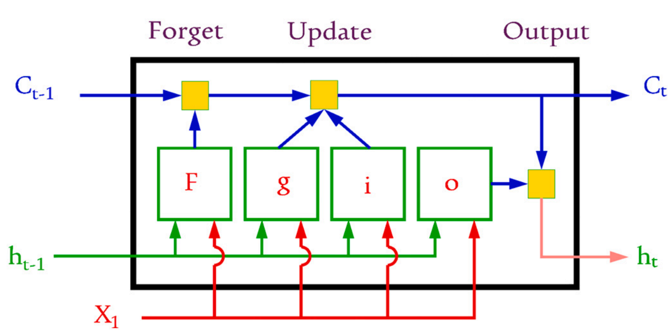

Fig. 8. Flow of data at time step t.

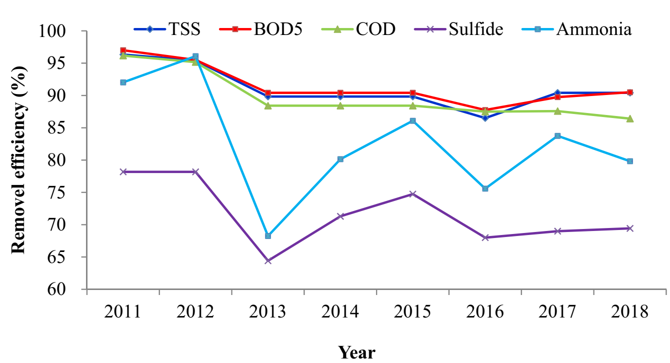

Fig. 9. Removal efficiency of TSS, BOD_5, COD, ammonia, and sulphide at El-Berka wastewater treatment plant (WWTP).

stacked denoising auto-encoders (SDAE). The SDAE model showed a prominent accuracy to predict the biofilm process [36]. Four different models have been compared in forecasting biological oxygen demand at WWTP. The adaptive regression splines (MARS) showed a higher reliability and better performance compared to multivariate artificial bee colony (ABC), classical regression analysis (CRA), and teaching-learning based optimization (TLBO) [37]. A long short-term memory (LSTM) was proved as an appropriate fulfillment tool to anticipate the nutrient

Table 2
Statistical summary of laboratory results of the El-Berka wastewater treatment plant (WWTP) effluent.

Parameters	Unit	Minimum	Maximum	Average	Egyptian law
Average flow	m^3/day	374,613	542,891	475,018	–
TSS	mg/l	12	58	30	20
BOD_5	mg/l	13	77	30	20
COD	mg/l	20	120	60	30
Sulphide	mg/l	0	7	1	–
Ammonia	mg/l	1	20	5	0.5

removal efficiency of an oxic-anoxic-anaerobic membrane bioreactor [38]. Cheng et al. [39] adopted six deep learning-based models (LSTM, GRU, adaptive version of LSTM, exponentially smoothed LSTM, and smoothed LSTM) to forecast the key features of a WWTP (influent temperature, influent flow, influent and effluent BOD_5, effluent chloride, and power utilization).

According to mentioned references, all previous models proposed to provide better anticipation of the treated effluents of WWTPs. They focused on predicting the number of pollutants in treated wastewater based on previously monitored values over a limited period (not exceeding three years). Also, they have used a simple design of TF, DFB, DCB, and DLTSF in the WWTP field.

Furthermore, only a few parameters were taken into account in each investigation. Based on the previous discussion, various issues should be addressed, including the consideration of field data over a long period of time, the development of improved machine learning technique systems to give accurate models, and the assessment of more parameters in the WWTP in the same study.

Due to the lack of the WWTP time series forecasting models in the literature, there is a need for more solid and effective forecasting models for all sorts of comparable difficulties. This study was carried out in order to address some of the above discussed issues. Applied neural network models were developed using five wastewater parameters (TSS, COD, BOD_5, ammonia, and sulphide) associated with WWTP effluents in Egypt (Al-Berka WWTP). The DCB, DFB, and DLTSF network structures have all been enhanced, resulting in new network designs that improve the outcomes. The models were developed based on the data obtained over eight years to achieve more accurate results. Two methods were proposed: The first method provides a comparison between three

Table 3
: Correlation matrix for removal efficiency and average flow.

Parameters	Average flow	TSS	BOD_5	COD	Sulphide	Ammonia
Average flow	**1**					
TSS	0.554	**1**				
BOD_5	**0.613**	**0.985**	**1**			
COD	0.550	**0.907**	**0.949**	**1**		
Sulphide	−0.121	−0.282	−0.248	−0.322	**1**	
Ammonia	0.287	**0.786**	**0.760**	**0.752**	−0.282	**1**

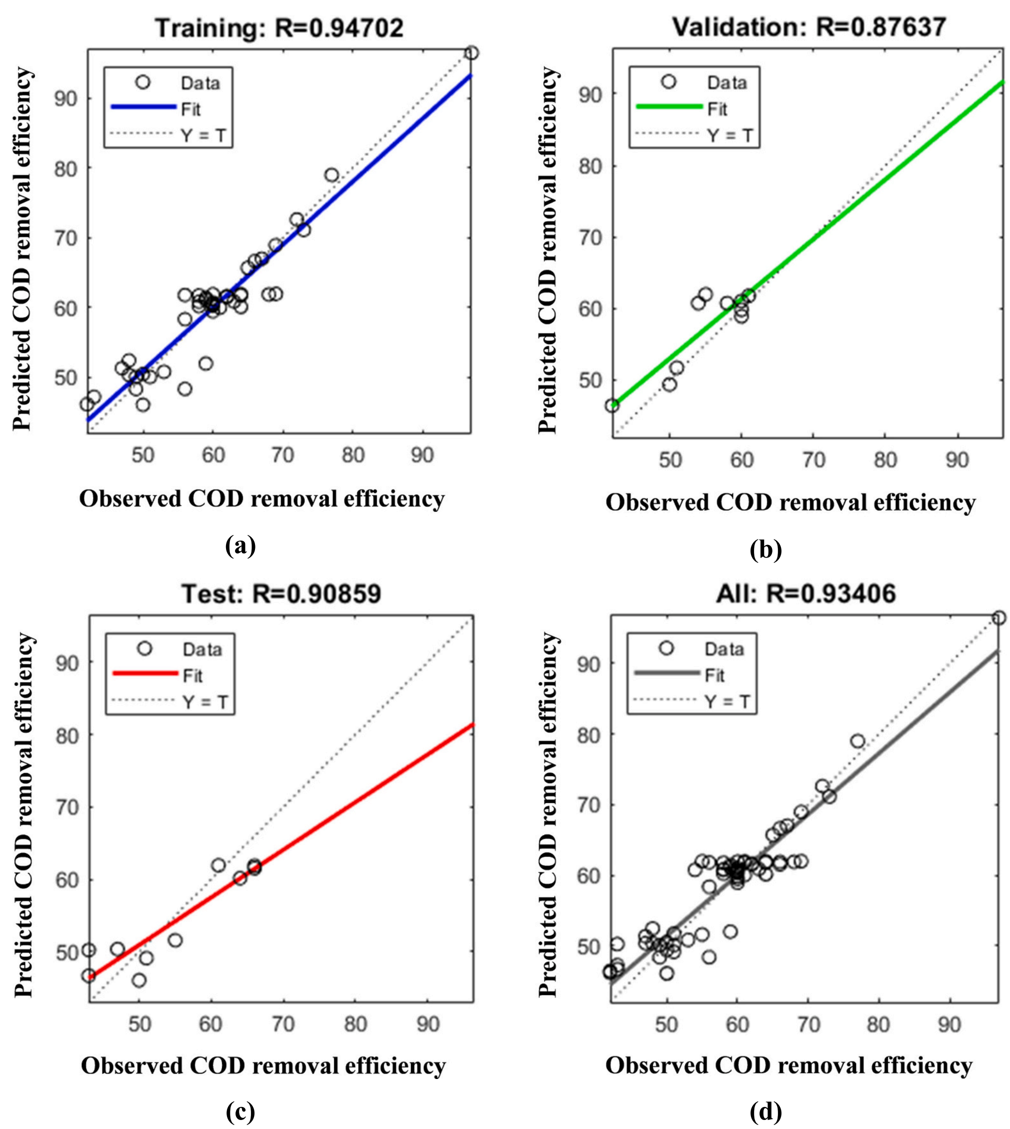

Fig. 10. Traditional feed-forward model results showing observed chemical oxygen demand (COD) versus the corresponding predicted COD and the regression factor (R) value for: a) training; b) validation; c) testing; and d) all the processes.

prediction models: the traditional feed-forward (TF), deep feed-forward backpropagation (DFB), and deep cascade-forward backpropagation (DCB) networks. The second method implements deep learning time series forecasting (DLTSF) using a long short-term memory (LSTM) network for wastewater treatment.

2. Materials and evaluation metrics

2.1. Plant description

The selected plant for this study was the El-Berka WWTP (BWWTP) located in northeast Cairo in the eastern desert of Egypt (Fig. 1). The plant uses activated sludge as a biological treatment technology for wastewater treatment; it receives 600,000 m^3/day of wastewater, which is equivalent to the waste from approximately 3 million people. The final effluent discharges into the Gabal El Asfar drain that flows to the Bilbays drain, and then into the Bahr EI Baqar Drain and Lake Manzala; finally, it flows into the Mediterranean (Fig. 1) [40].

2.2. Data collection

Daily records over eight years (from 2011 to 2018) were obtained

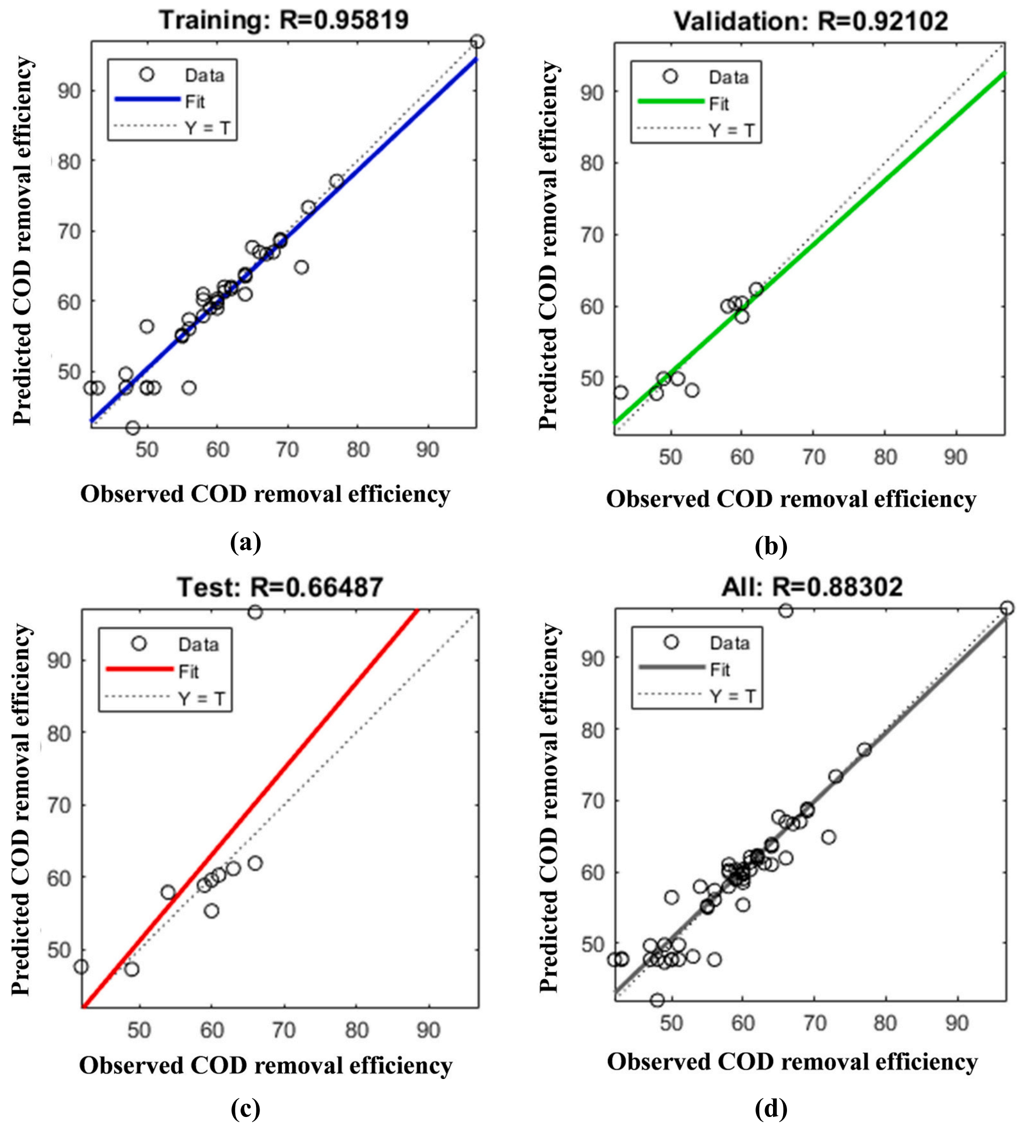

Fig. 11. Deep feed forward backpropagation model results showing observed chemical oxygen demand (COD) versus corresponding predicted COD and the regression factor (R) value for: a) training; b) validation; c) testing; and d) all the processes.

from the treatment plant, and a comprehensive analysis of the influent wastewater and effluent treated water was carried out for five parameters: BOD_5, COD, TSS, ammonia, and sulphide. Table S1 in the Supplementary information (SI) presents a detailed list of the field data used in this assessment. All parameters were analyzed according to standard methods [41,42].

2.3. Evaluation metrics

Generally, networks store the information regarding their corresponding loads. The activity of deciding network loads is called training, learning, or preparing, wherein a set of information is converted to known yield information [43]. A generous hindrance in the preparation cycle is minimizing errors between the observed and anticipated yield in the layer through numerous factual standards [44]. The statistical assessment measurements utilized for network preparation in the current work were mean-square error (MSE), mean absolute error (MAE), root-mean-square error (RMSE), mean absolute percentage error (MAPE), the coefficient of determination (R-squared) and coefficient of connection (R) [45]. These parameters were determined using the following equation:

$$\mathrm{MAE} = \frac{1}{N}\sum_{a}^{N}|X_a - Y_a| \tag{1}$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{a}^{N}(X_a - Y_a)^2 \tag{2}$$

$$\mathrm{RMSR} = \sqrt{\frac{1}{N}\sum_{a}^{N}(X_a - Y_a)^2} \tag{3}$$

$$\mathrm{MAPE} = \frac{1}{N}\sum_{a}^{N}\frac{|X_a - Y_a|}{X_a} \times 100 \tag{4}$$

$$\mathrm{R-squared} = \left(\frac{n\left(\sum_{a}^{N}X_aY_a\right) - \left(\sum_{a}^{N}X_a\right)\left(\sum_{a}^{N}Y_a\right)}{\sqrt{\left[n\sum_{a}^{N}X_a^2 - \left(\sum_{a}^{N}X_a\right)^2\right]\left[n\sum_{a}^{N}Y_a^2 - \left(\sum_{a}^{N}Y_a\right)^2\right]}}\right)^2 \tag{5}$$

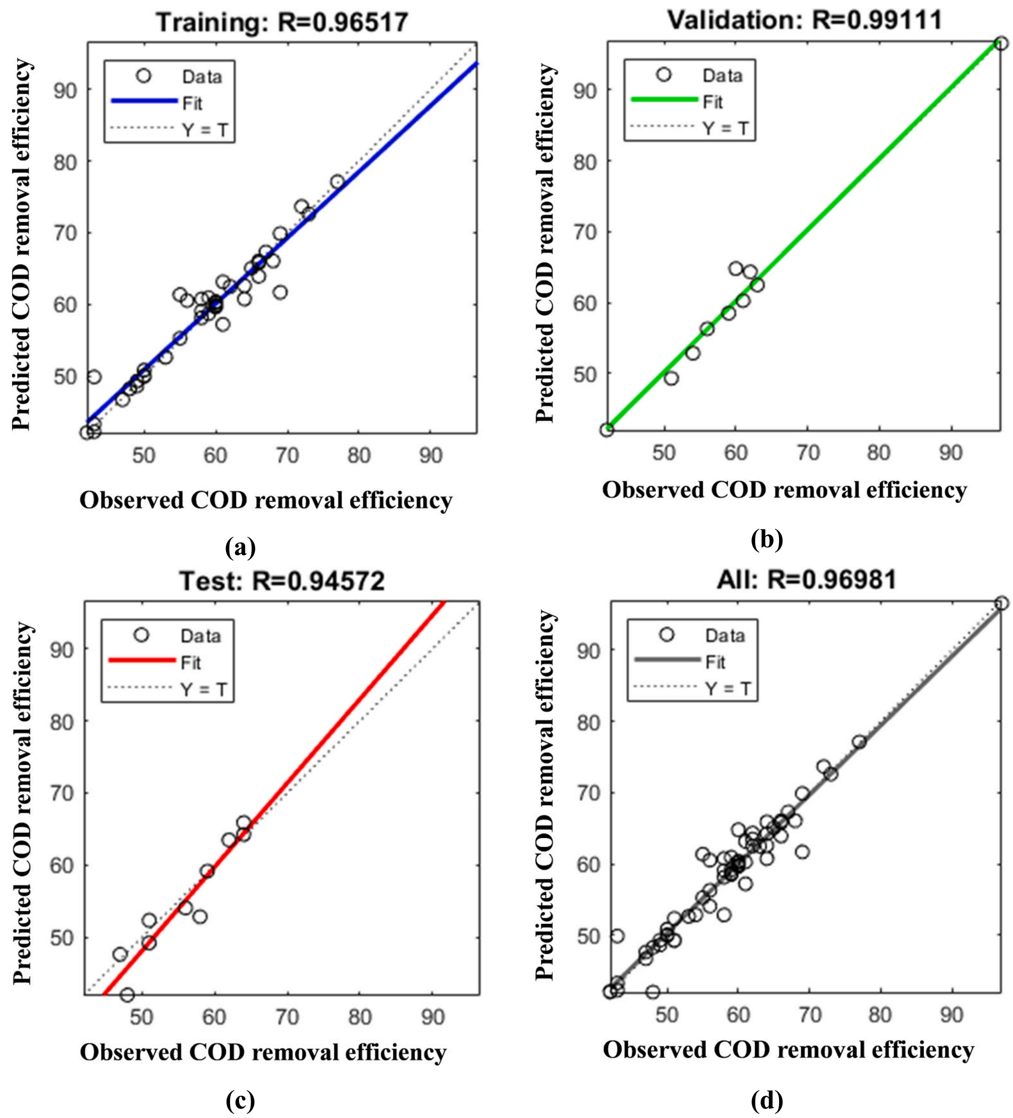

Fig. 12. Deep cascade-forward backpropagation model results showing observed chemical oxygen demand (COD) versus corresponding predicted COD and the regression factor (R) value for: a) training; b) validation; c) testing; and d) all the processes.

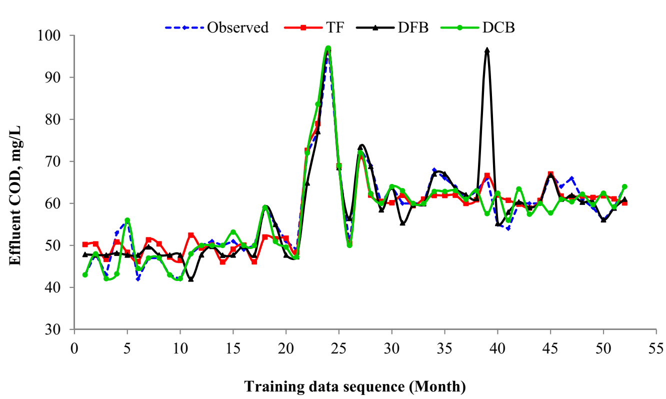

Fig. 13. Observed and output values associated with traditional feed-forward (TF), deep feedforward backpropagation (DFB, and deep cascade-forward backpropagation (DCB) models for the effluent chemical oxygen demand (COD).

$$R = \frac{\sum_{a}^{N}\left(X_a - \overline{X}\right)\left(Y_a - \overline{Y}\right)}{\sqrt{\sum_{a}^{N}\left(X_a - \overline{X}\right)^2 \sum_{a}^{N}\left(Y_a - \overline{Y}\right)^2}} \tag{6}$$

where N is the total number of measurements, X_a are the actual values, Y_a are the corresponding estimated values, $\overline{X}$ is the mean of the actual values of the X variables, and $\overline{Y}$ is the mean of the estimated values in Y variables.

3. Methodology

The design of the proposed model is established using learning-based, data-driven multi-classification, and time-series forecasting. The proposed model structure has two stages: prediction and forecasting (Fig. 2). This was decided based on the findings in the literature as most researchers work in one stage and one or two parameters. We didn't find similar techniques which give us the advantages over the literature. The prediction stage comprises three different networks: TF, DFB, and DCB networks. Here, a comparison of the three networks was performed. Subsequently, DLTSF was used to forecast future time step values.

Table 4
Comparison between the performance indices of traditional feed-forward (TF), deep feed forward backpropagation (DFB), and deep cascade-forward backpropagation (DCB) models for effluent chemical oxygen demand (COD).

Performance indices	TF		DFB		DCB	
	Training	Verification	Training	Verification	Training	Verification
MSE	0.01	−1.38	−0.19	−0.96	0.53	0.04
MAE	2.81	1.60	2.72	1.04	1.95	0.46
RMSE	3.56	1.97	5.23	1.51	3.35	0.77
MAPE	5.16	2.74	4.98	1.76	3.26	0.77
R-squared	0.87	0.70	0.78	0.77	0.90	0.89

Several stages were completed during the model advancement process, as follows:

3.1. Data collection and pre-processing

Before running the proposed networks, the collection data were divided into 'training', 'validation', and 'testing'. The dataset was split into two parts; that is, the data for first seven years were used as the training (52 of the dataset) and validation (32 of the dataset) set, and the data for the last year were used as the verification test (12 of the dataset) set.

3.2. Prediction stage

In this study, the network structure/training rate in wastewater treatment of El Berka plant was investigated for five key descriptive variables: TSS, COD, BOD_5, ammonia, and sulphide. The WWTP data samples from 2011 to 2017 were used as input data in the models, while the data from 2018 was kept for model verification.

3.2.1. Traditional feed-forward (TF) model design

The output node is termed the output neuron, and all output combined neurons are called the output layer (If only a single output is considered, the output layer consists of one neuron). Each of the M nodes in the center is called a hidden neuron; all collectively hidden neurons are called a hidden layer. Inputs are often referred to input neurons, and all input neurons are referred to the input layer (Fig. 3) [46]. There are numerous types of ANNs; perhaps, the most widely recognized neural network used in this work is the traditional feed-forward (TF) network, wherein the data are communicated in a forward manner only. The developed neural network model in MATLAB presents a platform to simulate any conditions, and it successfully predicts peak concentrations. The network properties of the prediction stages are summarized in Table 1.

Karunanithi et al. [47] confirmed that the number of hidden layers depends on the problem conditions. On the off chance that the number of nodes in the hidden layer is less, the network might not have enough levels of autonomy to finish the interaction effectively; however, if the number is too high, the preparation will take a lot of time, and the network may (now and then) over-fit the data. Therefore, in this study, different values of hidden layers and neurons in each hidden layer were applied to discover the ideal model, as shown in Fig. 4.

There are no norms or rules to decide the most reasonable network; a few preliminaries for each gathering have been directed to an appropriate learning rate. Some hidden layers and the number of neurons per hidden layer were also obtained.

3.2.2. Deep feed forward back propagation (DFB) network

The DFB network is a multilayer neural network; a graphical representation of this network is presented in Fig. S2. The DFB network includes six independent layers: one input layer, four hidden layers, and one output layer. The input layer obtains the input values as a matrix representation of the neuron that enters the network (Fig. 5), while the output layer provides the predicted output [48]. The output equation of layer N is given below:

$$a^N = f^N\left[W^N a^{N-1} - b^N\right], N = 1, \ldots, M \tag{6}$$

where W, a^{N-1}, b, and a^N represent the neuron vector weights, input, bias, and output, respectively, and the number of nodes in the four hidden layers is ten nodes each. The four hidden layer activation levels n_i^N of the N layer depend on the biases and weights of the layer, as presented in the following equation:

$$n_i^N = \sum_{j=1}^{S^{N-1}} w_{ij}^N a_j^{N-1} - b_i^N \tag{7}$$

3.2.3. Deep cascade-forward backpropagation network (DCB)

The DCB network is a multi-layer neural network with a direct connection between the input, hidden, and output layers. Each layer is connected directly with the primary layer, providing a skip connection. The graphical representation is shown in Fig. 6. The DCB network includes six layers: one input layer, four hidden layers, and one output layer. The input layer receives the input values as a matrix representation of the neuron that enters the network, whereas the output layer provides the predicted output [48]. Each hidden layer has ten nodes and one bias, and the number of weights in each layer is equal to the number of previous layers.

3.3. Forecasting stage

A DLTSF was proposed for simulating the El-Berka WWTP, based on a LSTM. In recent years, recurrent neural networks (RNNs) have been considered the most effective tools in machine learning for solving different forecasting tasks. LSTM is used in different applications designed by Hochreiter and Schmidhuber [49]. The network includes seven layers: the input layer, three LSTM layers, two fully connected layers, and a regression output layer (Fig. 7).

The LSTM architecture comprises an RNN, in which the evolution state depends on the previous time step entry and current entry. LSTMs learn from previous experiences and use method corresponding to data in the computer's memory. Data can be saved, read, or written from a network cell. Moreover, this type of design can limit error propagation in different layers over time. This technology allows the network to extend its learning process over various periods [50]. Fig. 8 displays how the gate neglects, updates, and yields the cell and hidden states. The LSTM layer learnable weights are the input weights W and bias b. The matrices W and b are concatenations of the input weights and bias of each element, respectively. The cell state equation is given as:

$$c_t = f_t \odot c_{t+1} + i_t \odot g_t \tag{8}$$

where $\odot$ denotes the element-wise multiplication of vectors. i, f, g, and o denote the input gate, forget gate, cell candidate, and output gate, respectively. The hidden state at time step t is given as follows:

$$\mathbf{h}t = ot \odot \sigma c(\mathbf{c}t), \tag{9}$$

where σ_c denotes the state activation function.

Table 5
Comparison between the proposed model and recent similar studies in the literature.

Ref.	Parameters	Models	Performance Indices					Remarks
			MAE	MSE	RMSE	MAPE	R^2	
[16]	BOD_5, COD and TSS	MLP	–	–	59.48	26.29	–	MLP showed better results rather than the RBF and MLR. The best performance for predicting TSS.
		RBF			63.39	28.77		
		MLR			68.92	68.92		
[20]	TSS, COD, T-P, and T-N	SVM, ANN	–	–	–	–	0.47	ANN showed better results rather than the SVM algorithm. The best performance for predicting TN
[27]	BOD_5, COD, or TSS	ANN	–	–	–	10.8	–	The best performance for predicting COD
[29]	NH_4-N, COD, BOD_5, DO	ANN, GBM, RF	0.017	0.084	0.92	–	0.58	GBM model showed the best performance on training and test data-set and less vulnerable to add or remove extra features. Also, the Mutual Information feature selection method suggested the best features.
[30]	TSS, BOD_5, and COD	ANN	–	–	–	–	0.754	the developed prediction model for the influent BOD_5 concentration
[31]	COD	ANN	–	–	6.376	9.226	–	experimental phase required a long time session
[32]	COD	ANN (4-input)	–	–	19.97	–	0.4142	The result demonstrated that COD_{eff} is more effective for all the eight input parameters than considering the four input parameters
		ANN (8-input)			1.08		0.7034	
[33]	BOD and COD	LS-SVM-RBF	3.165	–	4.461	–	–	It takes a huge amount of time. Some water quality parameters such as TDS, TS, and TSS were not considered as input or output variables. Combine LS-SVM models with RBF model for increasing the model performance
		LS-SVM-Poly	3.399		4.491			
		MARS	4.045		5.306			
		ANN	4.821		5.688			
		ANFIS	7.554		10.703			
		MLR	18.93		19.76			
		MNLR	6.067		8.221			
[34]	BOD_5 and COD	FFNN	–	–	1.53	–	–	Ensemble models provided a better approximation than single models and model combination methods. The best performance for predicting COD
		ANFIS			1.56			
		SVR			1.57			
		ARIMA			1.85			
[35]	TN, TP, COD, BOD_5, TSS	parallel-serial hybrid	–	–	–	–	0.81	Combine ML models with mechanistic models (biological simulation) for increasing the model performance
[36]	COD, NH_4^+-N, TN	SDAE, SVR, BNN, GBM, SAE	4.80	1.58	5.94	–	0.05	Stacked denoising auto-encoders (SDAE) showed the best performance for predicting TN
[37]	BOD_5	CRA	54.3939	–	58.4138	–	0.7313	The MARS method gave a better prediction of the peak BOD_5 than the CRA, ABC, and TLBO methods.
		ABC	38.2109		46.0179		0.7820	
		TLBO	56.8979		71.3260		0.7690	
		MARS	31.4827		44.1090		0.7966	
This study	TSS, COD, BOD_5, ammonia, sulphide	TF	1.60	−1.38	1.97	2.74	0.70	DCB model showed the best performance on training and test data-set. The best performance for predicting COD, TSS, and BOD5
		DFB	1.04	−0.96	1.51	1.76	0.77	
		DCB	0.46	0.04	0.77	0.77	0.89	

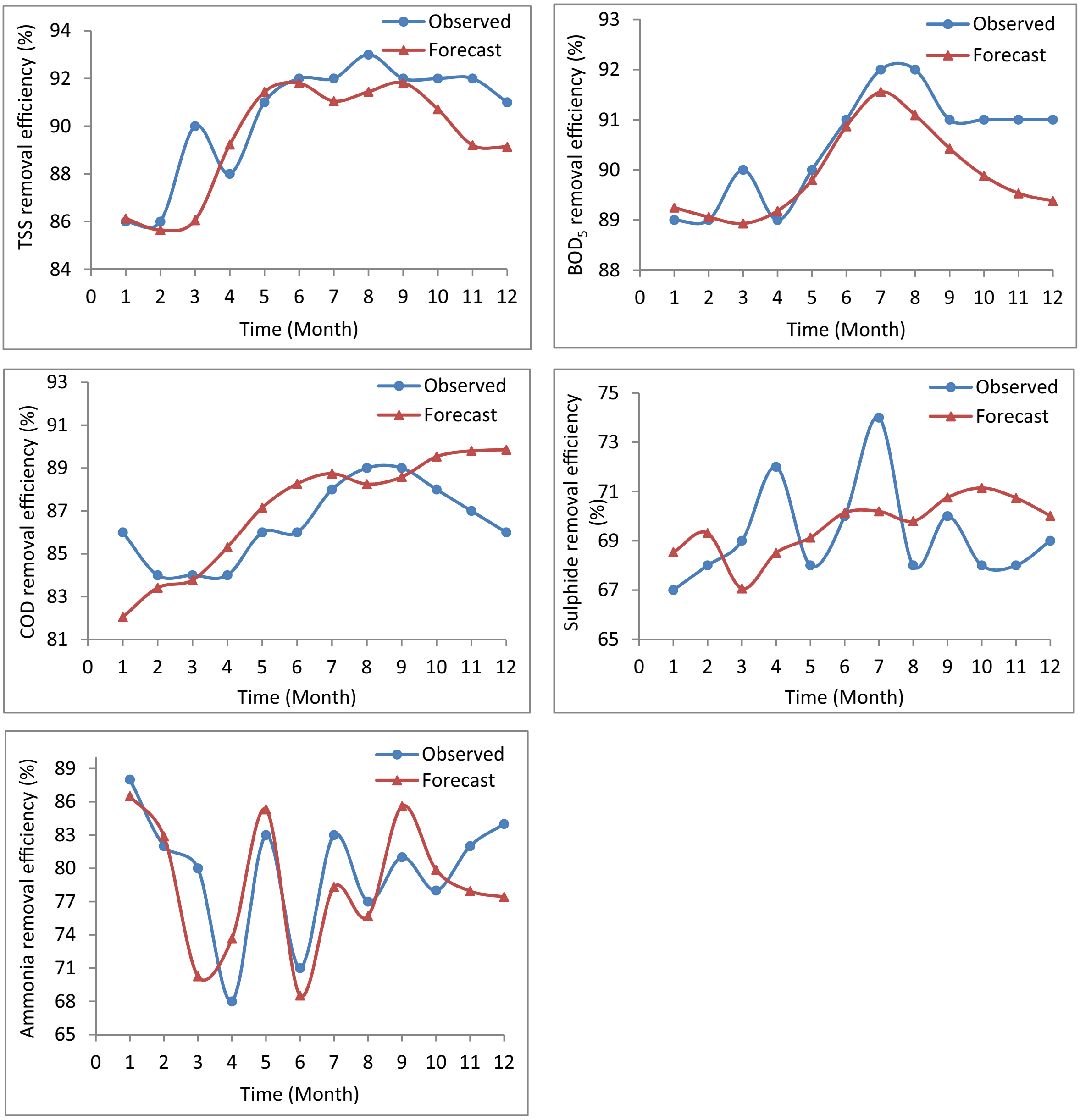

Fig. 14. Observed and deep learning time series forecasting output of total suspended solids (TSS), chemical oxygen demand (COD), biological oxygen demand (BOD_5), sulphide, and ammonia.

4. Results and discussions

4.1. Field data analysis

Fig. 9 shows the fluctuation in the removal efficiency for TSS, BOD_5, COD, ammonia, and sulphide. Over the years, these fluctuations may return to the various influent concentrations of the parameters during summer and winter over the time. Table 2 illustrates a statistical summary of the effluent monitored parameters, which indicates the necessity to raise the WWTP efficiency for achieving the permissible limits of the treated wastewater parameters. This would save the surrounding surface water or groundwater resources from the continuous accumulation of disposal pollutants.

From the previous eight years of daily reports on the El-Baka WWTP [51], it was found that the efficiency of the plant decreased in the summer months, during which the amount of influent raw sewage increased; therefore, a statistical analysis is required to determine the correlation between the removal efficiency and influent flow. Statistical tests were performed by computing Pearson's correlation coefficient (r) values. Table 3 shows the correlation matrix for the removal efficiencies of the five parameters and the average influent flow. The matrix showed a positive correlation of BOD_5, COD, and TSS with the average flow ($r > 0.50$). This also confirms that the plant's design capacity must be increased by building an extended area or converting the extra flow to a new plant, for avoiding a negative impact on the environment around the plant.

4.2. Prediction models results

In this study, the proposed methods for modelling the removal ratio of the BOD_5, COD, TSS, ammonia, and sulphide were trained using three methods: the TF, DFB, and DCB. MATLAB software was used to implement and execute the proposed model. The regression factor R for the proposed models with the three training methods, TF, DFB, and DCB, were plotted for modelling the BOD_5, COD, TSS, ammonia, and sulphide contents.

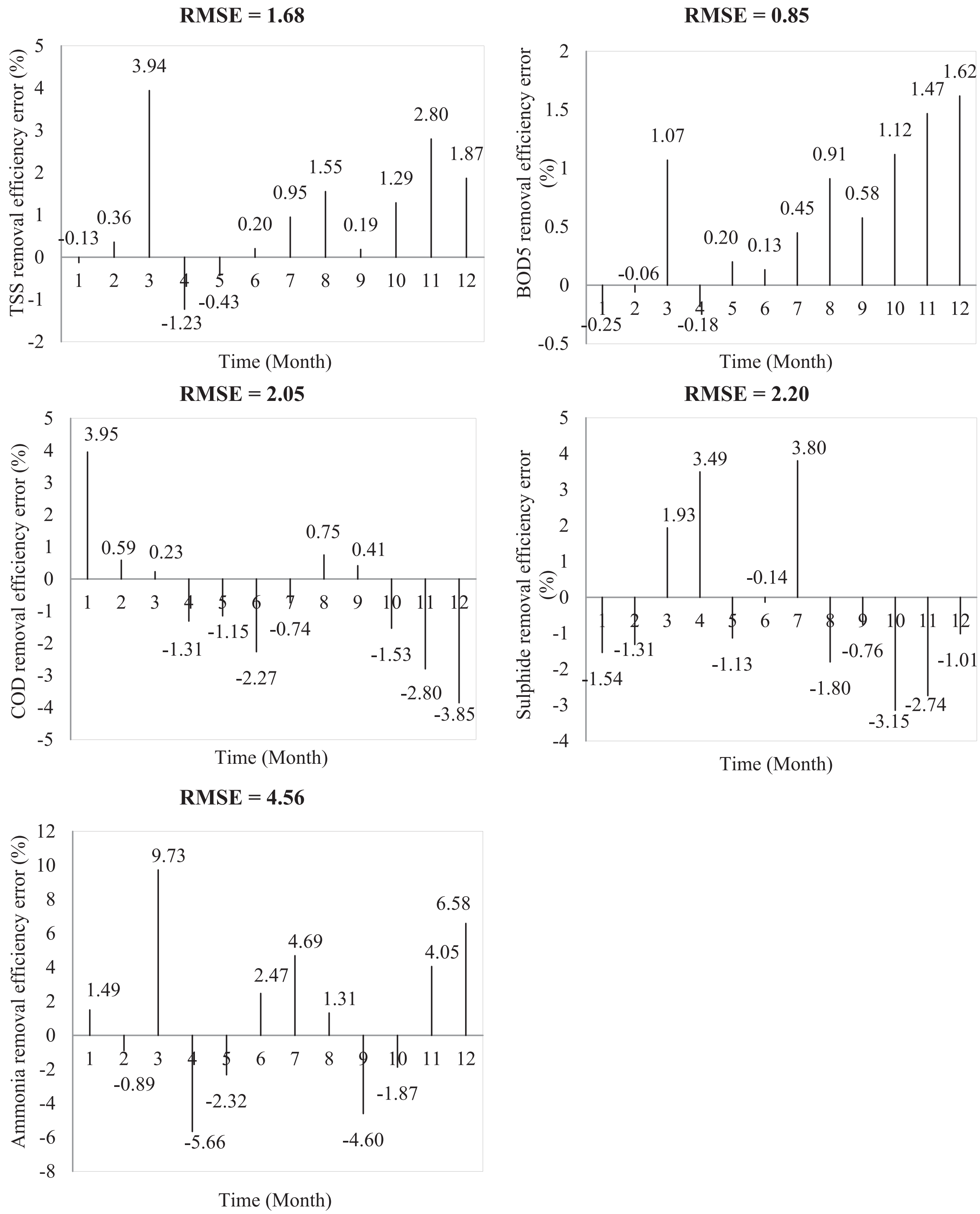

Fig. 15. Deep learning time series forecasting (DLTSF) removal efficiency error (%) and root-mean-square error (RMSE) of total suspended solids (TSS), chemical oxygen demand (COD), biological oxygen demand (BOD_5), sulphide, and ammonia.

Table 6
Root-mean-square error (RMSE) of the deep learning time series forecasting (DLTSF) model.

Method	TSS	BOD_5	COD	Sulphide	Ammonia
DLTSF	2.63	1.92	3.36	2.35	6.45

The COD regression factor results for the training, validation, and test datasets for the TF, DFB, and DCB methods are shown in Figs. 10–12, respectively. The yield effectively tracked the objectives associated with the training of COD, with R-values being 0.94702, 0.95819, and 0.96517 for TF, DFB, and DCB, respectively. The R-values for the three models were acceptable, and it highlighted the ability of the DCB model to give the closest R-value to 1, which indicated the superiority of the DCB model over the TF and DFB models. The regression factors for other parameters (BOD_5, TSS, ammonia, and sulphide) for the three proposed methods are presented in the supplementary materials (SI; see Figs. S2 to

S13). The illustrated regression factors showed that the fitting results and conversion were acceptable for all the models. However, optimizing the parameters with the DCB model generated regression factors much closer to 1, as compared to those generated using the TF and DFB methods.

Fig. 13 shows a comparison between the measured COD removal ratio and the output COD removal ratio from the TF, DFB, and DCB models. This figure demonstrates that the COD removal ratios obtained from all models were close to those of the measured removal ratios, thereby indicating that all the models were accurate and valid. Furthermore, the proposed DCB-model helps to minimise the distinction between the model and measured outputs, resulting in a minimal relative error as compared to those obtained using TF and DFB models. Figs. S14–S17 (in SI File) show a comparison between the measured BOD_5, TSS, sulphide, and ammonia removal ratios, along with their output removal ratios obtained from the TF, DFB, and DC models, respectively. The deviations between the BOD_5 measured values and the output values of the three models are presented in Fig. S14. The results of the three models show acceptable closeness to the measured BOD_5. However, a better consistency may be needed. This could be achieved by further training or through the application of another model. For TSS, the DFB model coincided with the measured values better than the TF and DCB models (see Fig. S15). Notably, the DCB model presented the best coincidence for sulphide and ammonia (see Figs. S16 and S17).

The statistical analyses used to check the exactness of the proposed TF, DFB, and DCB models for modelling the COD, BOD_5, TSS, sulphide, and ammonia removal ratios were MSE, MAE, RMSE, MAPE, and R-squared (Table 4). The rendering of the proposed TF, DFB, and DCB models for COD removal ratio was analyzed using the four indices [MAE, RMSE, MAPE, and R-squared]. From the results presented in Table 4, when the TF was used, the performance indices (the MAE, RMSE, MAPE, and R-squared) were 0.01, 2.81, 3.56, 5.15, and 0.87, respectively. In the instance of the DFB model, the MAE, RMSE, MAPE, and R-squared were −0.19, 2.72, 5.23, 4.98, and 0.78, respectively. While the MAE, RMSE, MAPE, and R-squared for training with the DCB model were 0.53, 1.95, 3.35, 3.26, and 0.90, respectively. The DCB model's verification findings were also well, with R-squared and RMSE of 0.89 and 0.77, respectively. From these results, it can conclude that the performance of the DCB model was better than that of the TF and DFB models. Table S2 in the SI also confirms the priority of the DCB model over the TF and DFB models in modelling BOD_5, TSS, ammonia, and sulphide. This result may be associated with the ability of each layer in this model to connect directly with the all-followed layer with a skip connection, which improves the model to simulate the measured data in different layers.

The correlations between the observed and predicted values of COD, BOD_5, TSS, sulphide, and ammonia for 2018 are shown in Fig. S18. Several additional discussions can be made. For example, the dataset remains convergent most of the year, even when recreation is employed for entering new data sources; therefore, the model created in this work exhibits satisfactory speculation ability and gives exact results. An additional benefit of the model is that there is no need for any supposition about the scope of discharge or temperature. Notwithstanding, the information data should be reliable, and the controlling variables for preparing and testing data should be very similar.

There is no single model applicable for all types of similar issues due to the complicated nature of various WWTP processes. As a result, developing more robust and effective models was required based on the data provided. Table 5 summarizes and compares our models to previous studies on WWTP prediction. The performance of several models was evaluated using MAE, MSE, MAPE, RMSE, and R-squared calculations, as shown in Table 5. The DCB model showed the highest R-squared of 0.89 for COD, which is substantially higher than that for other models [20,29,30,32,35,36,37]. In the DFB model, R-squared for the COD was 0.77, which is close to other research [35,37]. Hvala and Kocijan [35] created a parallel-serial hybrid model to predict COD with an R-squared of 0.81, much lower than this study. Also, the current DCB model achieved an RMSE of 0.77, which is lower than previous literature models. The predicted value is closer to the actual value with lower RMSE values. As a result, the proposed method is a high-precision prediction model based on machine learning models, as seen in recent similar studies.

4.3. Forecasting model results

The primary reason for this section is to comprehend the capacities of the DLTSF model for forecasting a future condition of the WWTP. The data for eight years from the El-Berka WWTP were used. The dataset was divided into two parts: the first seven years were used as the training and validation sets, and the last year was used as the verification test set. The training curves of the TSS, BOD_5, COD, sulphide, and ammonia contents are shown in Fig. S19. This figure shows the evolution of the DLTSF as a function of the iteration number. It can be observed that the training data were fit and worked efficiently using the DLTSF model. Fig. 14 presents the observed and DLTSF forecasting outputs of the TSS, BOD_5, COD, sulphide, and ammonia contents. The results show satisfactory competence and accuracy in forecasting the TSS, BOD_5, COD, and ammonia concentrations in the WWTP wastewater. However, sulphide forecasting data is not fully coincided with the observed sulphide dataset. This may need a special model modification to be only applicable for sulphide forecasting process. Notably, the DLTSF forecasting output data (especially; TSS, BOD_5, COD, and ammonia contents) were consistent with the measured dataset from the El-Berka WWTP. This confirms the validation of the model to forecast future data for wastewater treatment plants.

Fig. 15 displays the forecasting errors for the DLTSF model for each variable (i.e., TSS, BOD_5, COD, sulphide, and ammonia). The anticipating mistake is defined as the distinction between the measured and forecasted worth. The removal efficiency error (%) of total suspended solids (TSS), chemical oxygen demand (COD), biological oxygen demand (BOD_5), sulphide, and ammonia is ranging between −0.06 and 3.95. The analysis of the error reveals that the forecasting precision obtained by the DLTSF model can fulfil pragmatic necessities and be applicable for full-scale WWTPs. The DLTSF model for each variable (i.e., TSS, BOD_5, COD, sulphide, and ammonia) was assessed using the RMSE, as shown in Table 6. The RMSE value for TSS, BOD_5, COD, sulphide, and ammonia are 1.68, 0.85, 2.05, and 4.56, respectively. The RMSE indicates the reliability of the DLTSF as a valuable forecasting method for all the five parameters. The minimum RMSE was 0.85 for BOD_5, and the maximum was 4.56 for ammonia. This may be due to the high fluctuation of ammonia in the measured data, which may require more datasets for training. However, the BOD_5 dataset has the best consistency as compared to the other four parameters (i.e., TSS, COD, sulphide, and ammonia). This helps to finalize the best model for the available dataset.

RMSE is used to analyze the performance of models. Only two pieces of research for WWTP time series forecasting have been founded in the literature related to our research. Yaqub et al. [38] applied a simple LSTM model for COD forecasting, which gained an RMSE of 10.89. Also, Cheng et al. [39] introduced the LSTM model for BOD_5 forecasting that gained an RMSE of 2.02. The proposed model performance indicated a higher RMSE over the literature model with RMSE of 3.36 for COD and 1.92 for BOD_5.

5. Conclusions

The relevance of utilizing the appropriate machine learning techniques as a prediction booster and time series forecasting for El-Berka WWTP was assessed in this study. The following findings have been derived:

- Considering the outcome of recent literature for WWTP data prediction, the current study introduced more powerful prediction models (TF, DFB, and DCB) for the WWTP. Also, it presented a convenient and efficient time series forecasting model (DLTSF) for the WWTP.

- The models result indicated a positive correlation between the measured data (BOD_5, COD, and TSS) and the average flow of the WWTP.
- A high relationship coefficient (R) existed between the measured and anticipated factors (0.94702, 0.95819, and 0.96517), with three types of prediction models: TF, DFB, and DCB, respectively.
- The R-squared values for the three models were acceptable, and it highlighted the ability of the DCB model to give the highest R-squared value to 0.89 for COD.
- All proposed designed models had an acceptable outcome. Also, the DCB model introduces the most effective model for predicting the removal efficiency of the WWTP, and it is highly recommended as an essential performance appraisal instrument for plant decision-makers.
- A DLTSF was also used to simulate the removal efficiency of BOD_5, COD, TSS, ammonia, and sulphide for the WWTP; it showed high proficiency in time series forecasting of the five tested parameters over the literature with the smallest RMSE of 1.92, 3.36, 2.63, 6.45, and 2.35, respectively.

Therefore, the key results of this study can be used to fathom and anticipate plant behavior for facilitating plant management, enhancing system reliability, diminishing operational expenses, and improving plant performance.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

The authors thank the editor and co-editor of the journal and the anonymous reviewers for their valuable comments and suggestions for enhancing the manuscript's quality.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jwpe.2021.102380.

References

[1] N.E. El Arabi, M.A. Dawoud, Groundwater aquifer recharge with treated wastewater in Egypt: technical, environmental, economical and regulatory considerations, Desalin. Water Treat. 47 (2012) 266–278.
[2] O. Mustafa, Impact of Sewage Waste Water on the Environment of Tanjero River and Its Basin Within Sulaimani City/NE-Iraq, Geology MSc Thesis, College of Science, University of Baghdad, Baghdad, 2006.
[3] A.E. Tumer, S. Edebali, An artificial neural network model for wastewater treatment plant of Konya, Int. J. Intell. Syst. Appl. Eng. 3 (2015) 131–135.
[4] M.F. Hamoda, I.A. Al-Ghusain, A.H. Hassan, Integrated wastewater treatment plant performance evaluation using artificial neural networks, Water Sci. Technol. 40 (1999) 55–65.
[5] M. Ebrahimi, E.L. Gerber, T.D. Rockaway, Temporal performance assessment of wastewater treatment plants by using multivariate statistical analysis, J. Environ. Manag. 193 (2017) 234–246.
[6] L. Belanche, J.J. Valdés, J. Comas, I. Roda, M. Poch, Prediction of the bulking phenomenon in wastewater treatment plants, Artif. Intell. Eng. 14 (2000) 307–317.
[7] A. Hospido, M.T. Moreira, M. Fernández-Couto, G. Feijoo, Environmental performance of a municipal wastewater treatment plant, Int. J. Life Cycle Assess. 9 (2004) 261–271.
[8] A.P. Machado, L. Urbano, A. Brito, P. Janknecht, J. Salas, R. Nogueira, Life cycle assessment of wastewater treatment options for small and decentralized communities, Water Sci. Tech. 56 (2007) 15–22.
[9] H. Wenzel, H.F. Larsen, J. Clauson-Kaas, L. Høibye, B.N. Jacobsen, Weighing environmental advantages and disadvantages of advanced wastewater treatment of micro-pollutants using environmental life cycle assessment, Water Sci. Tech. 57 (2008) 27–32.
[10] X. Flores-Alsina, A. Gallego, G. Feijoo, I. Rodriguez-Roda, Multiple-objective evaluation of wastewater treatment plant control alternatives, J. Environ. Manag. 91 (2010) 1193–1201.
[11] O. Gundaz, C. Simsek, Assessment of three wastewater treatment plants in Turkey, in: Wastewater Reuse–Risk Assessment, Decision-Making and Environmental Security, Springer, 2007, pp. 159–167.
[12] A. Ouali, C. Azri, K. Medhioub, A. Ghrabi, Descriptive and multivariable analysis of the physico-chemical and biological parameters of Sfax wastewater treatment plant, Desalination 246 (2009) 496–505.
[13] Z. Zhang, F. Tao, J. Du, P. Shi, D. Yu, Y. Meng, Y. Sun, Surface water quality and its control in a river with intensive human impacts—a case study of the Xiangjiang River, China, J. Environ. Manage. 91(2010) 2483–2490.
[14] M. M. Hamed, M. G. Khalafallah, E. A, Hassanien, Prediction of wastewater treatment plant performance using artificial neural networks, Environ. Model. Softw. 19 (2004) 919–28.
[15] M.S. Nasr, M.A. Moustafa, H.A. Seif, G. El Kobrosy, Application of artificial neural network (ANN) for the prediction of El-Agamy wastewater treatment plant performance-Egypt, Alex. Eng. J. 51 (2012) 37–43.
[16] M. Hamada, H. Adel Zaqoot, A. Abu Jreiban, Application of artificial neural networks for the prediction of Gaza wastewater treatment plant performance-Gaza strip, J. App. Res. Water Wastew. 5 (2018) 399–406.
[17] M. Vyas, B. Modhera, V. Vyas, A. Sharma, Performance forecasting of common effluent treatment plant parameters by artificial neural network, ARPN J. Eng. Appl. Sci. 6 (2011) 38–42.
[18] M.S. Jami, I.A. Husain, N.A. Kabashi, N. Abdullah, Multiple inputs artificial neural network model for the prediction of wastewater treatment plant performance, Aust. J. Basic Appl. Sci. 6 (2012) 62–69.
[19] D. Li, H.Z. Yang, X.F. Liang, Prediction analysis of a wastewater treatment system using a Bayesian network, Environ. Model. Softw. 40 (2013) 140–150.
[20] H. Guo, K. Jeong, J. Lim, J. Jo, Y.M. Kim, J.P. Park, K.H. Cho, Prediction of effluent concentration in a wastewater treatment plant using machine learning models, J. Environ. Sci. 32 (2015) 90–101.
[21] A. Vijayan, G.S. Mohan, Prediction of effluent treatment plant performance in a diary industry using artificial neural network technique, J. Civil Environ. Eng. 6 (2016) 1–4.
[22] F. Granata, S. Papirio, G. Esposito, R. Gargano, G. De Marinis, Machine learning algorithms for the forecasting of wastewater quality indicators, Water 9 (2017) 105.
[23] V. Nourani, G. Elkiran, S.I. Abba, Wastewater treatment plant performance analysis using artificial intelligence—an ensemble approach, Water Sci. Techn. 78 (2018) 2064–2076.
[24] A.A. Nadiri, S. Shokri, F.T.-C. Tsai, A.A. Moghaddam, Prediction of effluent quality parameters of a wastewater treatment plant using a supervised committee fuzzy logic model, J. Clean. Prod. 180 (2018) 539–549.
[25] N. Khatri, K.K. Khatri, A. Sharma, Prediction of effluent quality in ICEAS-sequential batch reactor using feedforward artificial neural network, Water sci. Technol. 80 (2019) 213–222, https://doi.org/10.2166/wst.2019.257.
[26] N. Khatri, K.K. Khatri, A. Sharma, Artificial neural network modelling of faecal coliform removal in an intermittent cycle extended aeration system-sequential batch reactor based wastewater treatment plant, J. Water Process Eng. 37 (2020), 101477, https://doi.org/10.1016/j.jwpe.2020.101477.
[27] L. Arismendy, C. Cárdenas, D. Gómez, A. Maturana, R. Mejía, M.C.G. Quintero, Intelligent system for the predictive analysis of an industrial wastewater treatment process, Sustainability 12 (2020) 6348.
[28] D. Wang, S. Thunéll, U. Lindberg, L. Jiang, J. Trygg, M. Tysklind, N. Souihi, A machine learning framework to improve effluent quality control in wastewater treatment plants, Sci. Total Environ. 784 (2021), 147138, https://doi.org/10.1016/j.scitotenv.2021.147138.
[29] F. Bagherzadeh, M.J. Mehrani, M. Basirifard, J. Roostaei, Comparative study on total nitrogen prediction in wastewater treatment plant and effect of various feature selection methods on machine learning algorithms performance, J. Water Process Eng. 41 (2021), 102033, https://doi.org/10.1016/j.jwpe.2021.102033.
[30] A. Alsulaili, A. Refaie, Artificial neural network modeling approach for the prediction of five-day biological oxygen demand and wastewater treatment plant performance, Water Supply 21 (2021) 1861–1877.
[31] A.R. Picos-Benítez, J.D. López-Hincapié, A.U. Chávez-Ramírez, A. Rodríguez-García, Artificial intelligence based model for optimization of COD removal efficiency of an up-flow anaerobic sludge blanket reactor in the saline wastewater treatment (in Eng), Water Sci. Technol. 75 (2017) 1351–1361.
[32] S.I. Abba, G. Elkiran, Effluent prediction of chemical oxygen demand from the wastewater treatment plant using artificial neural network application, Procedia Comput. Sci. 120 (2017) 156–163.
[33] M. Najafzadeh, A. Ghaemi, Prediction of the five-day biochemical oxygen demand and chemical oxygen demand in natural streams using machine learning methods, Environ. Monit. Assess. 191 (2019) 380.
[34] V. Nourani, P. Asghari, E. Sharghi, Artificial intelligence based ensemble modeling of wastewater treatment plant using jittered data, J. Clean. Prod. 291 (2021), 125772.

[35] N. Hvala, J. Kocijan, Design of a hybrid mechanistic/Gaussian process model to predict full-scale wastewater treatment plant effluent, Comput. Chem. Eng. 140 (2020), 106934.
[36] S. Shi, G. Xu, Novel performance prediction model of a biofilm system treating domestic wastewater based on stacked denoising auto-encoders deep learning network, Chem. Eng. J. 347 (2018) 280–290.
[37] O.T. Baki, B. Akdeml, R.E. Aras, U.O. Akdemir, B.J.D. Yilmaz, W. Treatment, Biochemical oxygen demand prediction in wastewater treatment plant by using different regression analysis models, Desalin. Water Treat. 157 (2019) 79–89.
[38] 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, https://doi.org/10.1016/j.jwpe.2020.101388.
[39] T. Cheng, F. Harrou, F. Kadri, Y. Sun, T. Leiknes, Forecasting of wastewater treatment plant key features using deep learning-based models: a case study, IEEE Access 8 (2020) 184475–184485.
[40] M. Lasheen, N. Ammar, Assessment of metals speciation in sewage sludge and stabilized sludge from different wastewater treatment plants, Greater Cairo, Egypt, J. Hazard. Mater. 164 (2009) 740–749.
[41] F.H. Rainwater, L.L. Thatcher, Methods for Collection and Analysis of Water Samples, US Government Printing Office, 1960.
[42] W. E. Federation, APH Association, Standard Methods for the Examination of Water and Wastewater, American Public Health Association (APHA), Washington, DC, USA, 2005.
[43] T. Pai, Y. Tsai, H. Lo, C. Tsai, C. Lin, Grey and neural network prediction of suspended solids and chemical oxygen demand in hospital wastewater treatment plant effluent, Comput. Chem. Eng. 31 (2007) 1272–1281.
[44] C. Kriger, R. Tzoneva, Neural networks for prediction of wastewater treatment plant influent disturbances, in: Proceedings of the AFRICON 200, IEEE, 2007, pp. 1–7.
[45] T. Pai, T. Wan, S. Hsu, T.C. Chang, Y.P. Tsai, C.Y. Lin, L.F. Yu, Using fuzzy inference system to improve neural network for predicting hospital wastewater treatment plant effluent, Comput. Chem. Eng. 33 (2009) 1272–1278.
[46] O. Nelles, Neural networks, in: Nonlinear System Identification, Springer, Cham, 2020, https://doi.org/10.1007/978-3-030-47439-3_11.
[47] N. Karunanithi, W.J. Grenney, D. Whitley, K. Bovee, Neural networks for river flow prediction, J. Comput. Civil Eng. 8 (1994) 201–220.
[48] S. Tengeleng, N. Armand, Performance of using cascade forward back propagation neural networks for estimating rain parameters with rain drop size distribution, Atmosphere 5 (2014) 454–472.
[49] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780.
[50] P. Oliveira, B. Fernandes, F. Aguiar, M.A. Pereira, C. Analide, P. Novais, A deep learning approach to forecast the influent flow in wastewater treatment plants, in: C. Analide, P. Novais, D. Camacho, H. Yin (Eds.), Proceedings of the Intelligent Data Engineering and Automated Learning — IDEAL 2020, Springer International Publishing, Cham, 2020, pp. 362–373.
[51] HCWW (Holding Company for Water & Waste water), Reports of the General Administration of Central Laboratories and Quality Control in the Sanitation Company of Greater Cairo/El Berka WWPT, El Salam City, Cairo, Egypt, 2018.