Taylor & Francis
Taylor & Francis Group

Journal of Applied Water Engineering and Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjaw20

Application of cascade feed forward neural network to predict coagulant dose

Dnyaneshwar Vasant Wadkar, Rahul Subhash Karale & Manoj Pandurang Wagh

To cite this article: Dnyaneshwar Vasant Wadkar, Rahul Subhash Karale & Manoj Pandurang Wagh (2021): Application of cascade feed forward neural network to predict coagulant dose, Journal of Applied Water Engineering and Research, DOI: 10.1080/23249676.2021.1927210

To link to this article: https://doi.org/10.1080/23249676.2021.1927210

Journal of Applied Water Engineering and Research, 2021
https://doi.org/10.1080/23249676.2021.1927210

Application of cascade feed forward neural network to predict coagulant dose

Dnyaneshwar Vasant Wadkar [a], Rahul Subhash Karale[b] and Manoj Pandurang Wagh [c*]

[a]*Department of Civil Engineering, AISSM'S College of Engineering, Pune, India* [b]*Department of Civil Engineering, TSSMs Bhivarabai Sawant College of Engineering and Research, Narhe, Pune, India* [c]*Department of Civil Engineering, Dr. Vithalrao Vikhe Patil College of Engineering, Ahmednagar, India*

(*Received 15 November 2020; accepted 9 April 2021*)

Inlet water quality fluctuations affect mainly coagulant dose, and outlet water quality of the water treatment plant (WTP). Many complex physical and chemical processes are involved in WTP and water distribution networks (WDN). These technologies show non-linear behavior, which is challenging to be described by linear mathematical models. Thus, there is a need to develop prediction models for coagulation dose. The present study involves the application of cascade feed-forward neural networks (CFFNN) to predict coagulant dose. CFFNN Model was developed by using the Levenberg-Marquardt Training Algorithm and Bayesian Regularization Training Algorithm to predict coagulant dose. During the development of these models, hidden nodes are varied from 15 to 60, and R is found between 0.914 and 0.947. The best results were obtained by the CFFNN model using the Bayesian Regularization Training Algorithm (CFNNCD2) with hidden node 40, where R = 0.945 for training and 0.947 for testing.

Keywords: Water quality; water treatment plant; residual chlorine concentration; coagulant dose; chlorine dose

1. Introduction

The routine WTP consists of screening, coagulation-flocculation (CF), sedimentation, filtration, and disinfection units. In India, most WTP operators apply approximate coagulant dose due to lack of automation (Kote and Wadkar 2019). Coagulation is a significant process for drinking water treatment, is deliberated to withdraw colloidal and fine particles (Bello et al. 2014). Coagulation dose is essential to minimize the turbidity of drinking water. Turbidity is present in the drinking water due to the presence of colloidal and fine suspended matter and microscopic organisms (Djeddou et al. 2019). The coagulant is typically determined through laboratory analysis, which requires a long experimental time in WTP (Loc et al. 2020). Thus identified operational challenges such as approximations of dosing and extended period of experimentation can be handled by developing models for coagulant dose (Amali et al. 2018). The main component of the WSS is a WDN which preserves the quality and quantity of water as well as maintains sufficient pressures in the distribution of water (Bekkari and Zeddouri 2019). The physical, operational, and environmental conditions for WTP and WDN should be analyzed as a crucial part of a WSS (Abba and Elkiran 2017). The uncertainty level and risk of time-dependent factors increase with age, which thereby creates very challenging situations for managers to plan future short-term and long-term operational and maintenance protocols for WTP (Asnaashari et al. 2013; Ayvaz and Kentel 2015). An effective evaluation of WTP requires extensive involvement of specialized personnel, reliable databases, substantial time, application of cutting-edge technology, and equipment (Ayvaz and Kentel 2015). Over the last few decades, modeling in WSS has been shifted from process-based technique to artificial intelligent (AI) techniques (Bowden et al. 2006; Bobadilla et al. 2019; Santos et al. 2019). Because of the growing accessibility of information in the water industry, data-driven modeling methods are becoming more popular (Chandwani V. et al. 2016). Out of the several AI techniques, artificial neural network (ANN) is the most popular techniques in WSS whereas fuzzy logic, genetic programming, and model trees (MT) are comparatively new to this system (Constans et al. 2003; Wu and Lo 2008; Zeinolabedini and Najafzadeh 2019). ANNs have been successfully used in many water supply studies and this was a motivating factor for its application to the present study (Heddam et al. 2012; Cuesta and Tau 2014). Drinking water quality tends to degrade as it passes through distribution systems owing to: (1) intrusion of contaminated groundwater through cracks in underground pipes, (2) possibility of microbial regeneration in stagnant water and formation of pipe-wall biofilm (Erickson et al. 2017). The adoption of hygienic processes from treatment to distribution system is necessary to prevent waterborne epidemics (Hamed et al. 2004; Gibbs et al. 2006; Hanbay et al. 2008; Hebati et al. 2017).

*Corresponding author. Email: profmpwagh@gmail.com

Furthermore, WTP and WDN include many complex physical and chemical processes. These processes are non-linear, which are difficult to be defined by linear mathematical models (Haghiri et al. 2018). Thus, there is a need to develop prediction models for water quality, coagulant dose. This evaluation is too expensive for most municipalities. Thus, there is a need for cost-effective soft computing model for performance evaluation of WTP using inlet and outlet water quality parameters (Heddam et al. 2012). This model will help owners/operators to plan effectively short-term and long-term activities. Therefore, the WTP operators and related professionals will benefit from this model in assessing, managing, planning, and budgeting for the WTP. In the present paper, the Cascade Feed Forward Neural Network Water Quality Model was developed by using Levenberg-Marquardt Training Algorithm and Bayesian Regularization Training Algorithm to monitor the water treatment plant. In the present paper, coagulant dose and performance evaluation of WTP in the farthest zones in WDN namely Indrayani Nagar WTP, Pimpri Chinchwad Municipal Corporation (PCMC), Maharashtra, India is investigated. The present study also involves the application of ANN model. The methodology adopted for ANN modeling with modified MATLAB code is explored in predicting water quality, coagulant dose, with cascade feed-forward neural networks (CFNN).

2. Material and methods

The data is collected from WTP PCMC for the development of the models. It includes mainly 27 water quality parameters *viz*., pH, Ca, Cl^-, alkalinity, K, TDS, total hardness, turbidity, conductivity, DO, color, Fe, Cr, Al, Mn, F^-, Cu, nitrate, BOD, TSS, phosphate, Ni, Co, detergent, Mo, B & MPN and coagulant dose, and major water supply pipelines of Indrayani Nagar (Figure 1). Indrayani Nagar having a population of 46,242 and has three ESR with capacity of 0.3, 0.6, and 2.5 ML. The distance of ESR from the first valve is 9.95 km and detention time of ESR is 3 h.

3. Experimental plan

3.1. *Coagulant dose neural network model*

Traditionally jar test is used to determine the optimum dose of coagulant that requires more time (Jayaweera and Aziz 2018; Kim et al. 2014; Swetland et al. 2013). In India, coagulant dose in a WTP remains constant during certain periods due to delay in jar testing, which may lead to the formation of under-dosing and over-dosing of coagulant dose (Kumpel and Nelson 2013; Krishnaiah et al. 2007; Kennedy et al. 2015; Koleva 2017). From literature, it is found that many ANN and fuzzy approaches are available for the prediction of coagulant dose for a particular WTP (Librantz et al. 2018; Lee et al. 2004; Muharemi et al. 2019). However, the same approach cannot give the same prediction efficiency for other WTPs. Therefore, in this study, CFNN models are developed and applied for prediction of coagulant dose at WTP of PCMC, Maharashtra (India).

3.1.1. *Methodology for coagulant dose neural network model*

Input parameters such as inlet and outlet water turbidity and output parameter as coagulant dose are identified for the coagulant dose neural network (CDNN) model. The CDNN model is established with inlet and outlet water turbidity in the input layer whereas the output layer predicts

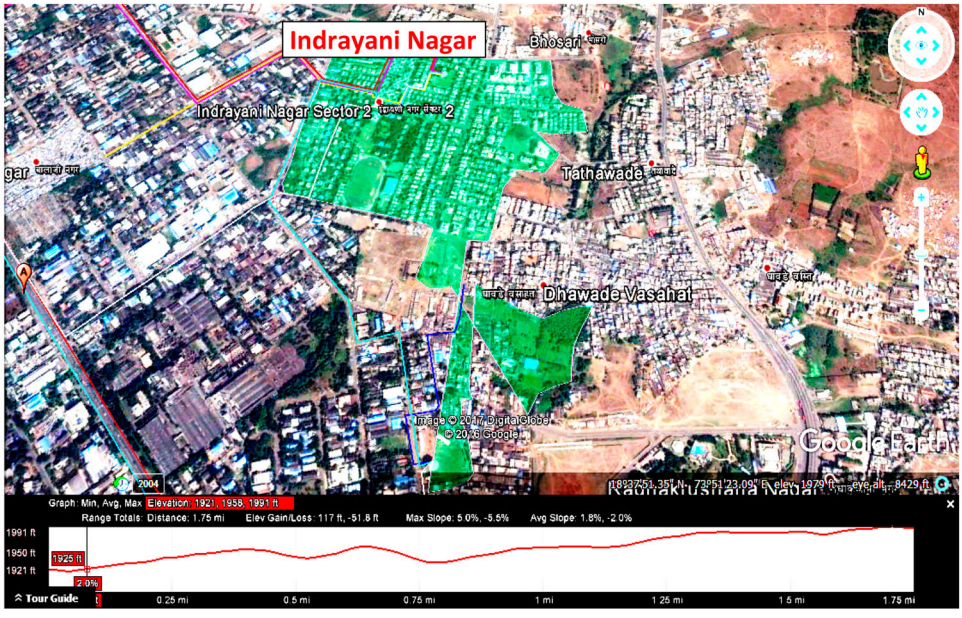


Figure 1. Satellite image of Indrayani Nagar, Pune.

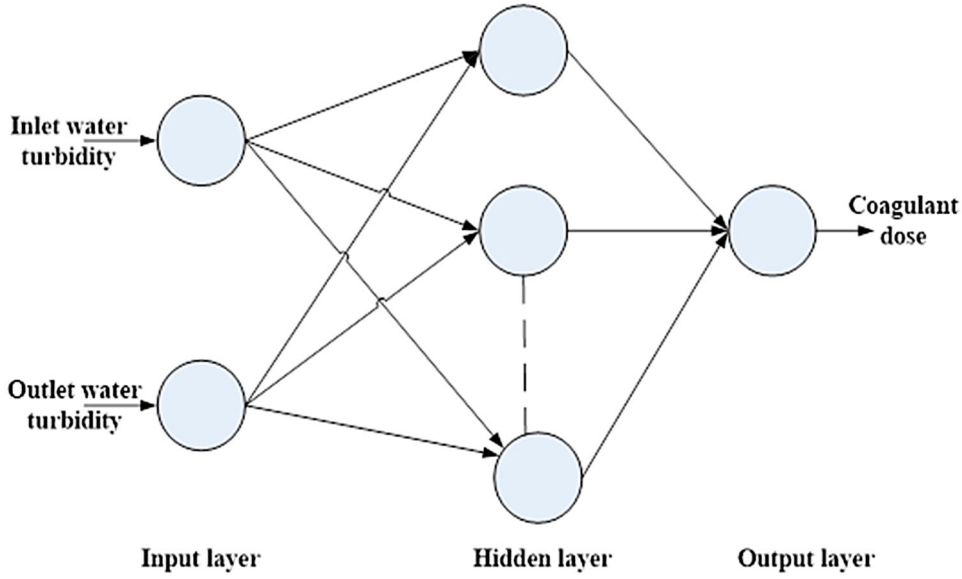


Figure 2. ANN Model for coagulation dose.

coagulant dose as shown in Figure 2. Daily data of input and output parameters spanning from 1 January 2012 to 31 December 2015 are obtained from the plant laboratory (Heddam et al. 2011). The database of input and output parameters required for the ANN modeling consists of 11,688 data points spanning eight data points per day. The data interval is three hours starting from 7 AM current day to 7 AM of the next day.

3.2. *Analysis of coagulant–turbidity*

Coagulation is an essential process of water treatment. Determination of optimal coagulant dose is vital, as insufficient dosing will result in a high-value of turbidity in treated water (Zhang et al., 2013). On the other hand, doses that are too high can result in high cost and health problems related to high levels of residual aluminum (if alum is used as the coagulant). Thus, the turbidity of water is an assessment parameter for the coagulation process. Coagulant–turbidity dataset is used for the prediction of optimum coagulant dose in WTP. This dataset has 11688 data points (1 January 2012 – 31 December 2015) which include eight readings per day at an interval of three hours namely inlet water turbidity, outlet water turbidity, and coagulant dose. As we know determination of coagulant dose in the laboratory requires nearly three hours, the inlet water turbidity, outlet water turbidity, and coagulant dose observed at the interval of three hours were collected.

The standard statistics of 3 hourly inlet water turbidity, outlet water turbidity, and coagulant dose are given in Table 1. The highest and lowest values of turbidity observed are 208 and 4 mg/L, respectively. And, the highest and lowest values of coagulant dose observed are 379.5 and 5.52 mg/L, respectively. It is seen that the value of ɣ1 of inlet water turbidity and coagulant dose are near to each other, which implies that the pattern of data distribution of both closely matches. The close match pattern of data distribution indicates applied coagulant dose nearly proportional to inlet water turbidity. Also, inlet water turbidity, outlet water turbidity, and coagulant dose show high skewness indicating that the data is asymmetrical whereas kurtosis has the leptokurtic distribution of data with long and fat tail and higher and sharper central peak.

Table 1. Standard statistics of inlet water turbidity, outlet water turbidity, and coagulant dose.

Standard statistics	3 hourly Inlet water turbidity (mg/L)	3 hourly outlet water turbidity (mg/L)	3 hourly Coagulant dose (mg/L)
Mean ($\bar{x}$)	13.967	1.635	20.157
Standard deviation (σ)	23.212	0.394	32.486
Skewness (ɣ1)	5.332	2.562	5.284
Kurtosis (ɣ2)	32.213	14.812	31.601

The observed inlet water turbidity and coagulant dose at WTP for four years is shown in Figure 3(a,b). It is seen that the pattern of inlet water turbidity and coagulant dose is nearly matching because coagulant dose at the WTP is provided as per variation in the raw water turbidity Kim and Parnichkunet 2017). Higher the turbidity, the applicable coagulant dose is also higher. However, it can also be seen that more peak values are observed for inlet water turbidity during the year 2015 which may be due to high turbidity present in the Pawana river. This high turbidity in the Pawana River may be due to rapid development along the bank of the river, high rainfall, industrial activities, and many more other activities.

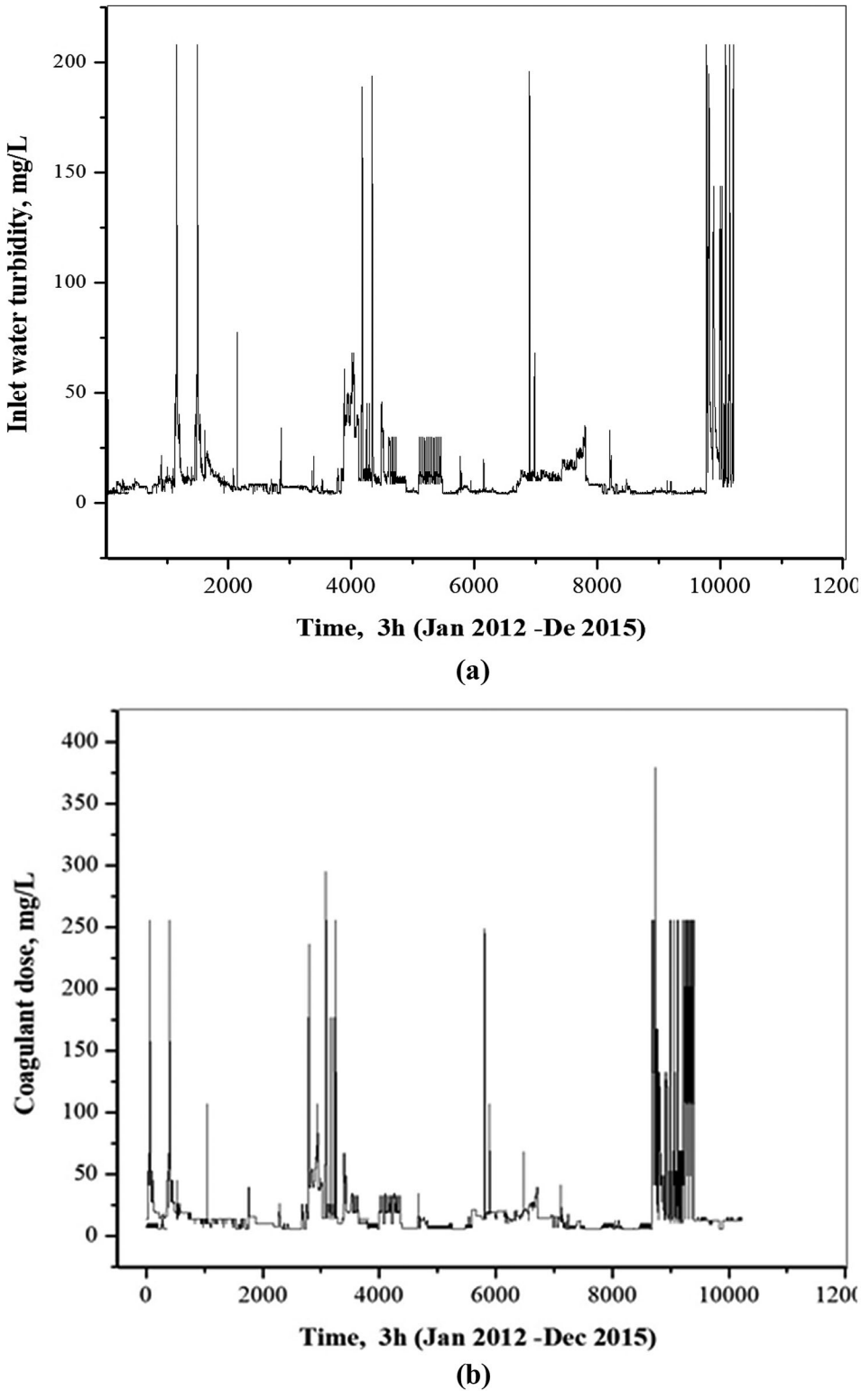


Figure 3. (a) Time series plot of inlet water turbidity. (b) Time series plot of coagulant dose.

3.3. Input and training algorithm identification

Many researchers used input parameters *viz.* turbidity of inlet and outlet water, observed coagulant dose, pH of water for ANN modeling of coagulant dose. Input parameters such as inlet and outlet water turbidity and output parameter such as coagulant dose are identified for ANN modeling (Pitta and Babu 2010; McCoy and VanBriesen 2012; Gamboa-Medina and Reis

Luisa 2017; O'Reilly et al. 2018; Saberi-Movahed et al. 2020). Daily data of input and output parameters spanning 4 years are obtained from the plant laboratory. The database of input and output parameters required for the ANN modeling consisted of 11,688 data points. Diversified training algorithms such as Bayesian regularization (BR), Levenberg-Marquardt (LM), Randomized polynomial (RP), BFGS (Broyden–Fletcher–Goldfarb–Shanno), one step secant (OSS), Conjugate gradient back-propagation with Fletcher-Powell (CGBF), Conjugate gradient back-propagation (CGB), Variable learning rate gradient descent (VLRGD), and gradient descent (GD) are used for the development of ANN models. It is observed that BR and LM training algorithms produced excellent predictions as compared to other training algorithms. Further best performed LM and BR training algorithms are used for the development of the best FFNN model and CFNN model by a varying number of hidden layers, and hidden nodes and epochs.

3.4. Feed forward neural network coagulant dose model

The number of training algorithms is analyzed for the prediction of coagulant dose. Summary of the performance of training algorithms for model development is shown in Figure 4. It is noticed that most training algorithms showed good performance while some model showed a negative correlation thereby indicating in-capabilities in modeling (Raduly and Gernaey 2007; Heddam and Dechemi 2015; Saha et al. 2017; Najafzadeh and Saberi-Movahed 2018, 2019). From Figure 4 it observed that BFG, NRP (Neonatal Resuscitation Algorithm), CGB, CGF, OSS, GDX, GDM (gradient descent with momentum), and GD required only one epoch and 15 hidden nodes; whereas for LM and BR training algorithm epoch is varied from 12 to 7228 and hidden node varied from 15 to 60. Still, FFNNCD model with LM ($R = 0.944$) and BR ($R = 0.945$) training algorithms are more effective than other training algorithms.

3.5. Feed forward neural network coagulant dose model using Levenberg-Marquardt training algorithm

FFNN coagulant dose model using LM training algorithm (FFNNCD1) is developed with one hidden layer in MATLAB software. Model was established with LM that updates weight and bias values conferring to LM optimization. LM is frequently the fastest back-propagation algorithm and is extremely endorsed as a primary-choice supervised algorithm, although it does require more memory than other algorithms. This model gave maximum 'R' with 60 numbers of hidden node and properties are results in Table 2.

Authentication vectors were used to stop training initially if the network performance on the validation vectors fails to recover or remains the same for determiningepochs in a row. Test vectors are used as a further check that the network is generalizing well but no effect on training. The time series and scatter plot of FFNNCD1 (2-60-1)

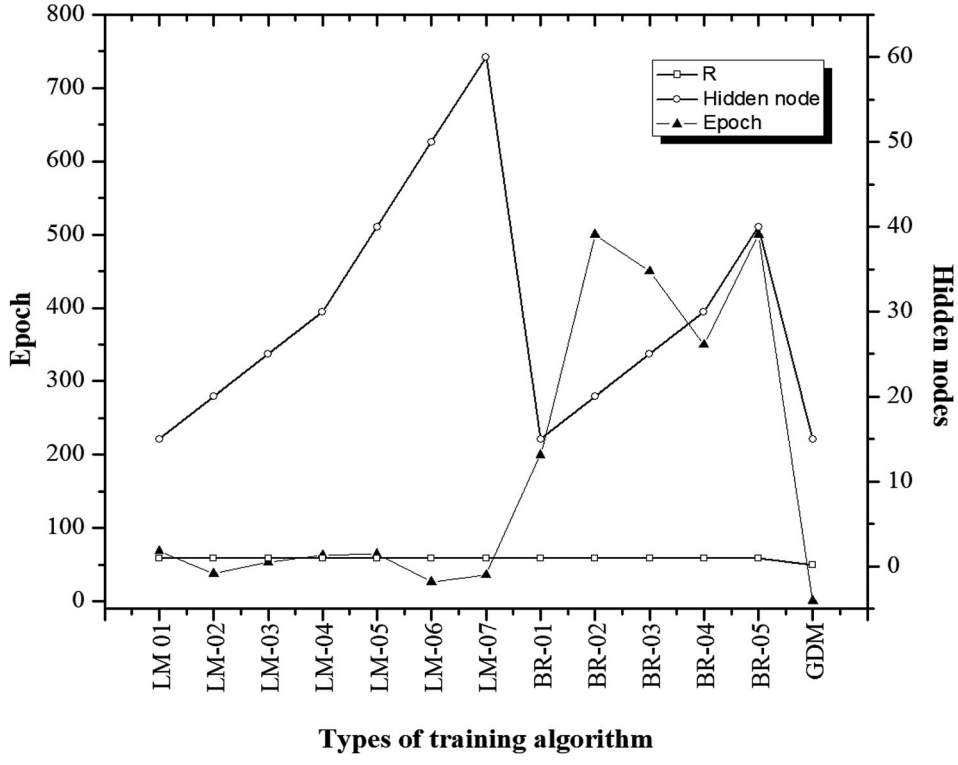
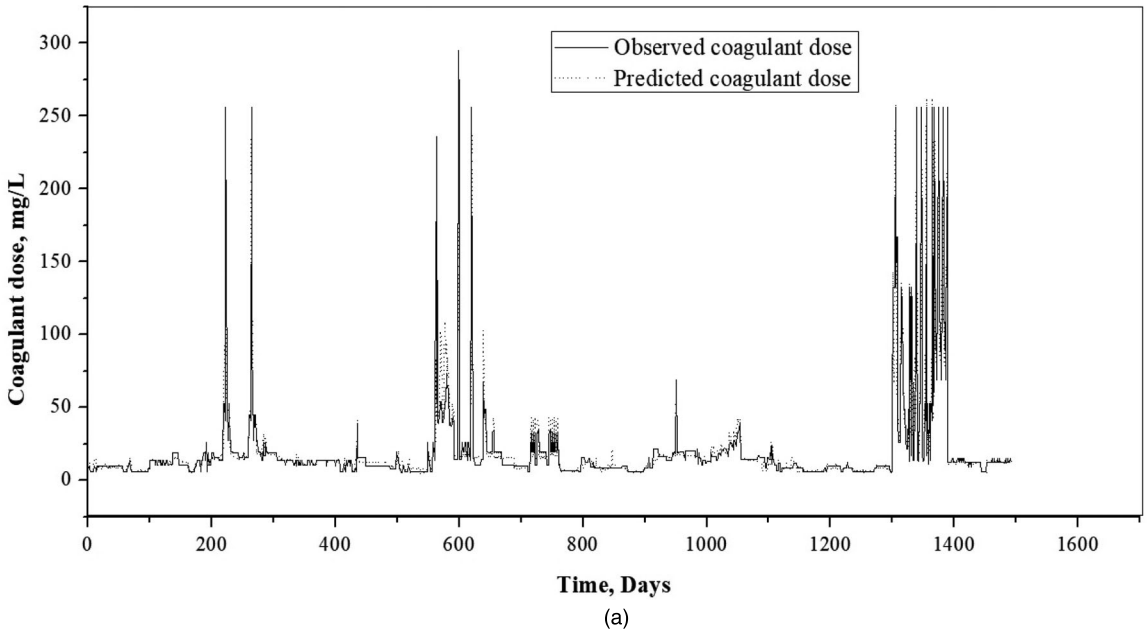
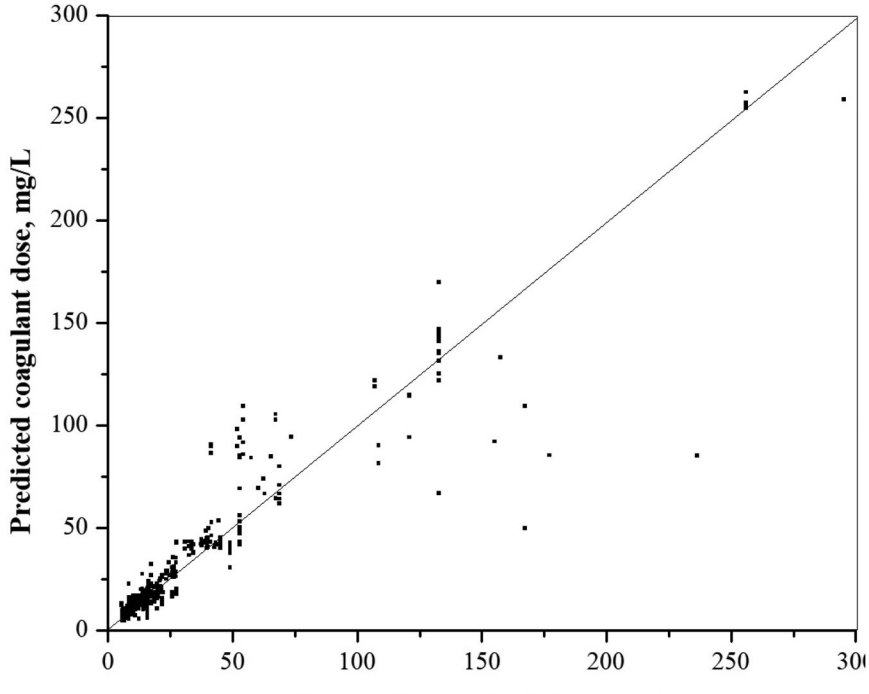


Figure 4. Performance of training algorithms for FFNNCD model.



(a)



(b)

Figure 5. (a) Time series plot of coagulant dose for FFNNCD1 (2-60-1) model. (b) Scatter plot of observed verus predicted coagulant dose for FFNNCD1 (2-60-1) model during training and testing period.

Table 2. Properties of best FFNNCD1 model.

Type of network	Training algorithm	No of Epoch	No of hidden node	R	MSE
FFNN	LM	26	60	0.944	185.09

model during the training and testing period shown in Figure 5(a,b) indicate a good fit. Peak values of coagulant dose are well captured. It could be due to the most peak values included in dataset training. Thus, rigorous training has captured this trend but failed in the testing period. This result shows that if the data series are not following the normal distribution, but changing the training algorithm

Table 3. Standard statistics of FFNNCD1 (2-60-1) model.

ANN model	Training period				Testing period			
	$\bar{x}$	σ	ɣ1	ɣ2	$\bar{x}$	σ	ɣ1	ɣ2
Observed values	20.157	32.486	5.284	31.601	8.899	1.493	1.233	0.171
FFNNCD1 (2-60-1)	20.109	30.711	5.361	33.919	9.067	2.280	1.212	1.851

Table 4. Standard statistics of FFNNCD2 (2-50-1) model.

ANN model	Training period				Testing period			
	$\bar{x}$	σ	ɣ1	ɣ2	$\bar{x}$	σ	ɣ1	ɣ2
Observed values	20.157	32.486	5.284	31.601	8.899	1.493	1.233	0.171
FFNNCD2 (2-50-1) model	20.175	30.811	5.314	33.240	9.051	2.302	1.200	1.612

and by rigorous training, slightly better results only could be achieved, even though low and average values are well captured by rigorous training of network.

The standard statistics of FFNNCD1 (2-60-1) model is shown in Table 3, where observed and predicted $\bar{x}$ and ɣ1 are closely matching during training and testing periods. The percentage change between observed and predicted σ values had 5.47% and 52.75% during training and testing period, respectively. Similarly, the percentage change between observed and predicted ɣ2 values had 7.33% and 980.51% during training and testing period, respectively.

3.6. Feed forward neural network coagulant dose MODEL using bayesian regularization training algorithm

FFNN water quality model using BR training algorithm (FFNNCD2) model is created with one hidden layer in MATLAB software. By varying values of the hidden node after ending epochs, different R and MSE values are obtained. It is distinguished that hidden nodes are increased from 5 to 15 where R and MSE values were changed accordingly. Though, after changing hidden nodes from 20 to 60, R remains identical. Similarly, it is observed that the best prediction obtained against hidden node 50 with R for training = 0.943 and R for testing = 0.945. In this case, the network response is satisfactory, and validation can be used for entering new inputs. The Bayesian methods are able to efficiently address the issue of over fitting and penalizes complicated models. The Bayesian methods include probably distribution of network weights in comparison to standard network training in which optimum weights are selected by minimizing an error algorithm. This algorithm utilizes the Jacobean matrix to calculate the output as a medium or total of square error (Zhang et al. 2011; Wang et al. 2017; Najafzadeh and Zeinolabedini et al. 2019). The BR algorithm feature can train any network as long as its derivative features include weight, net input and transfer algorithms.

The time series and scatter plot of FFNNCD2 (2-50-1) model during testing period shown in Figure 6(a,b) indicates average fit. Peak values are poorly captured because most of the time on WTP observed coagulant dose is kept constant. Thus, rigorous training has captured this trend but failed in the testing period, even though low and normal values are well captured by rigorous training of network (Wu and Lo 2010). The standard statistics of FFNNCD2 (2-50-1) model is shown in Table 4, where observed and predicted $\bar{x}$ and ɣ1 are close during the training and testing period. The percentage change between observed and predicted value of σ = 5.16 and 54.21% during the training and testing period, respectively. Similarly, the % change among observed and predicted ɣ2 values had 5.19% and 841.39% during the training and testing period correspondingly.

3.7. Cascade feed forward neural network coagulant dose model

The CFNN are similar to feed forward networks, but include a weight connection from the input to each layer, and from each layer to the successive layers. For example, a three-layer network has connections from layer 1 to layer 2, layer 2 to layer 3, and layer 1 to layer 3. The 3-layer network has connections from the input to all 3 layers. The supplementary networks might progress the speed at which the network acquires the desired relationship. Summary of the performance of training algorithms is shown in Figure 7. It is noticed that many training algorithms showed good performance while some models have negative correlation thereby indicating the in-capabilities in modeling. From Figure 7, it is observed that BFG, NRP, CGB, CGF, OSS, GDX, GDM, and GD required only one epoch and 15 hidden nodes; whereas LM and BR training algorithm required epoch 27–700 epochs and 15–60 hidden nodes.

However, FFNNAD model with LM algorithm (R = 0.943) and BR algorithm (R = 0.947) is more effective than other training algorithms.

3.8. Cascade feed forward neural network coagulant dose model using Levenberg-Marquardt training algorithm

CFNN coagulant dose model (CFNNCD1) developed with LM training algorithm that updates weight and bias values according to LM optimization. Gauss–Newton algorithm and steepest descent method are building block for LM training algorithm. Gauss–Newton algorithm provide speed and steepest descent method establish stability

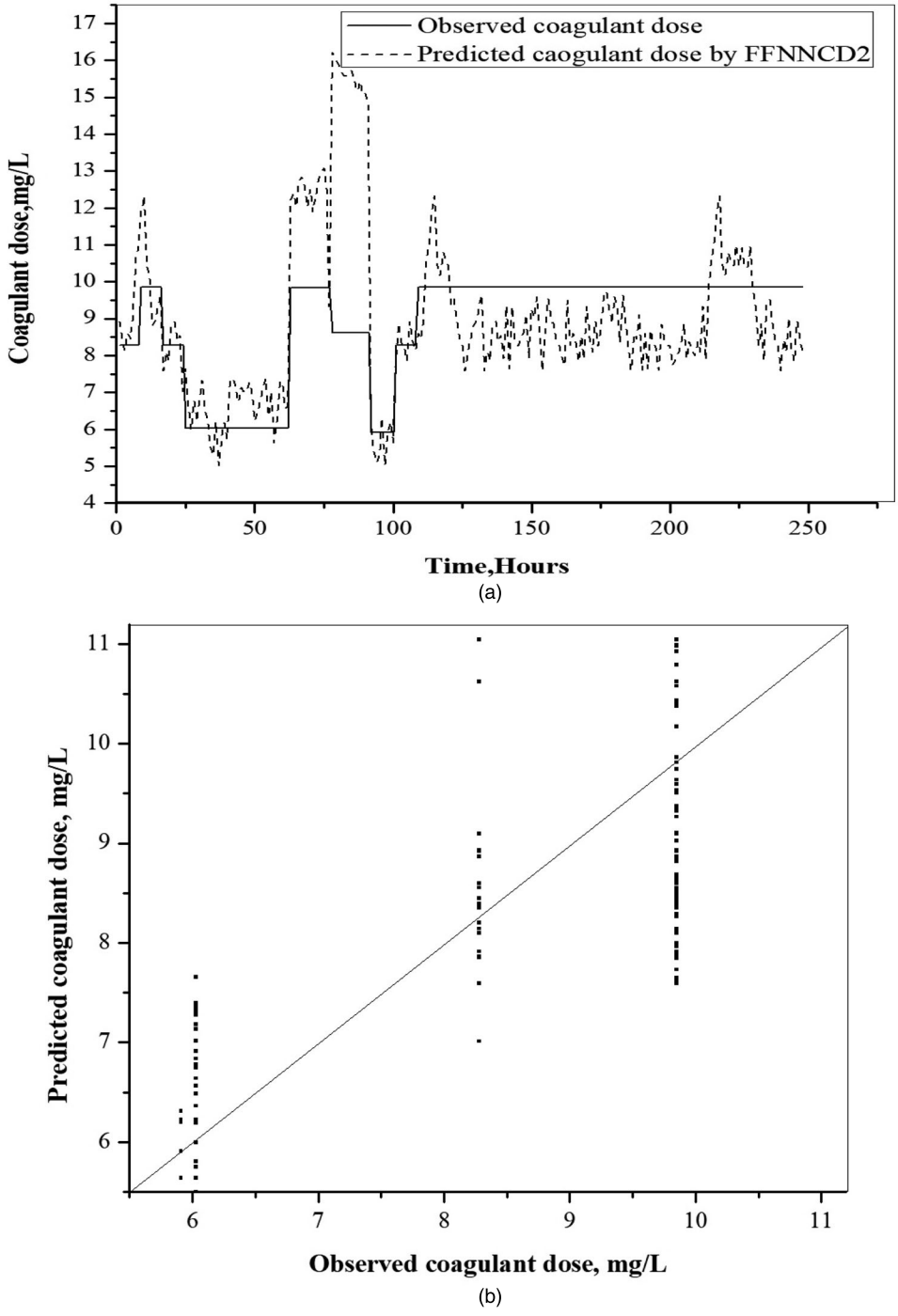


Figure 6. (a) Time series plot of coagulant dose for FFNNCD2 (2-50-1) model. (b) Scatter plot of observed verus predicted coagulant dose for FFNNCD2 (2-50-1) model during training and testing period.

in LM training algorithm. It is stronger than the Gauss–Newton algorithm as it can converge well, in many cases even if the error surface is far more complex than the quadratic situation. This model gave maximum 'R' with 60 numbers of hidden node and properties are mentioned in Table 5.

Table 5. Properties of best CFNNCD1model.

Type of network	Training algorithm	No of Epoch	No of hidden node	R	MSE
FFNN	LM	36	60	0.943	59.22

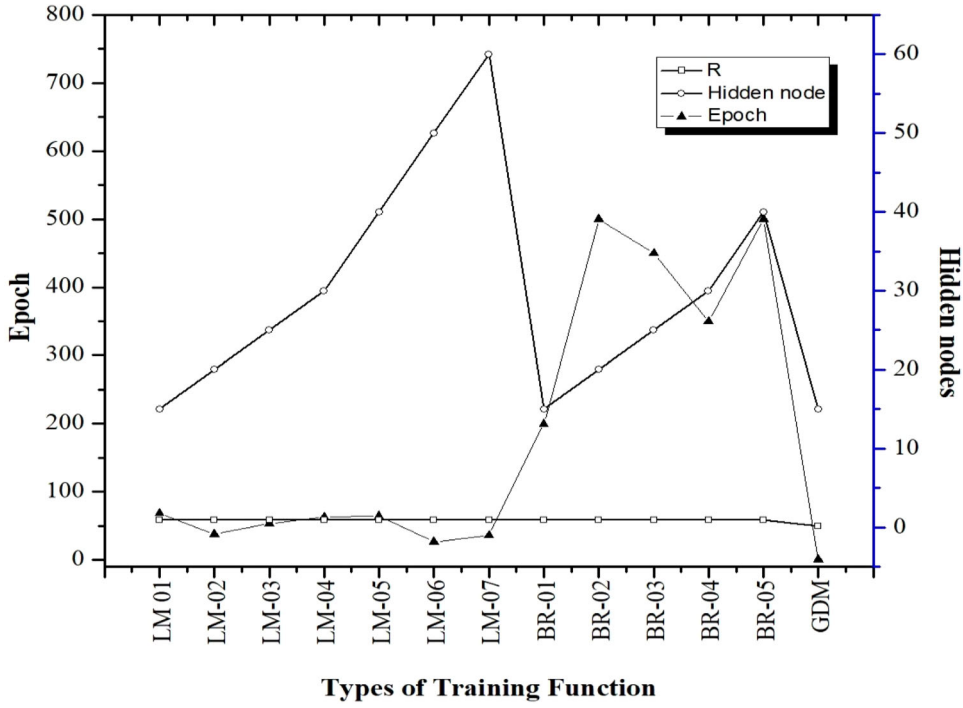


Figure 7. Performance of training algorithms for CFNNCD model.

During the development of this model hidden nodes are provided with 15, 20, 25, 30, 40, 50 and 60, it is found that *R* value ranged from 0.927 to 0.943 and MSE value ranged from 59.22 to 130.94. The best results were obtained with hidden node = 60 where $R = 0.943$ in training and 0.943 for testing. The time series and scatter plot of CFNNCD1 (2-60-1) model during the testing period shown in Figure 8(a,b) indicates average fit. Peak values are poorly captured because most of the time on WTP observed coagulant dose is kept constant.

In this way, thorough training has identified this pattern however failed in testing period, despite the fact that low and normal values are all around caught by thorough training of network. The standard statistics of CFNNCD1 (2-60-1) model is shown in Table 6, where observed and predicted $\bar{x}$ and $\gamma 1$ are close during the training and testing period. The percentage change between observed and predicted σ values had 5.16% and 54.21% during the training and testing period, respectively. Similarly, the percentage change between observed and predicted $\gamma 2$ values had 5.19% and 841.39% during the training and testing period, respectively.

Table 6. Standard statistics of CFNNCD1 (2-60-1) model.

ANN model	Training period				Testing period			
	$\bar{x}$	σ	$\gamma 1$	$\gamma 2$	$\bar{x}$	σ	$\gamma 1$	$\gamma 2$
Observed values	20.157	32.486	5.284	31.601	8.899	1.493	1.233	0.171
CFNNCD1 (2-60-1) model	20.223	30.694	5.339	33.711	9.150	1.888	1.983	4.348

3.9. *Cascade feed forward neural network coagulant dose model using Bayesian regularization training algorithm*

CFNN water quality model using BR training algorithm (CFNNCD2) is created with one input layer, one hidden layer and one output layer in MATLAB software that offers a platform for the simulation application. During the development of this model hidden nodes are provided with 15, 20, 25, 30, 40, 50 and 60, it is found that the best results were obtained with hidden node 40 where R is for training (0.945), and for testing (0.947). The time series and scatter plot of CFNNCD2 (2-60-1) model during the testing period are shown in Figure 9(a,b).

The prediction of coagulant dose by the developed ANN models during the testing period are carried out with 248 data points. In Figure 9, the predicted values of coagulant dose are not following the pattern of observed coagulant dose from 60 to 90 data points during the year 2016. These changes are obtained due to a large range of values of inlet water turbidity and coagulant dose during training of neural networks. In spite of the large variations, the mean value of observed coagulant dose and predicted coagulant dose is very close to each other (Liu et al. 2018).

The standard statistics of CFNNCD2 (2-40-1) model is shown in Table 7, where observed and predicted $\bar{x}$ and $\gamma 1$ are close during training and testing period. The % change

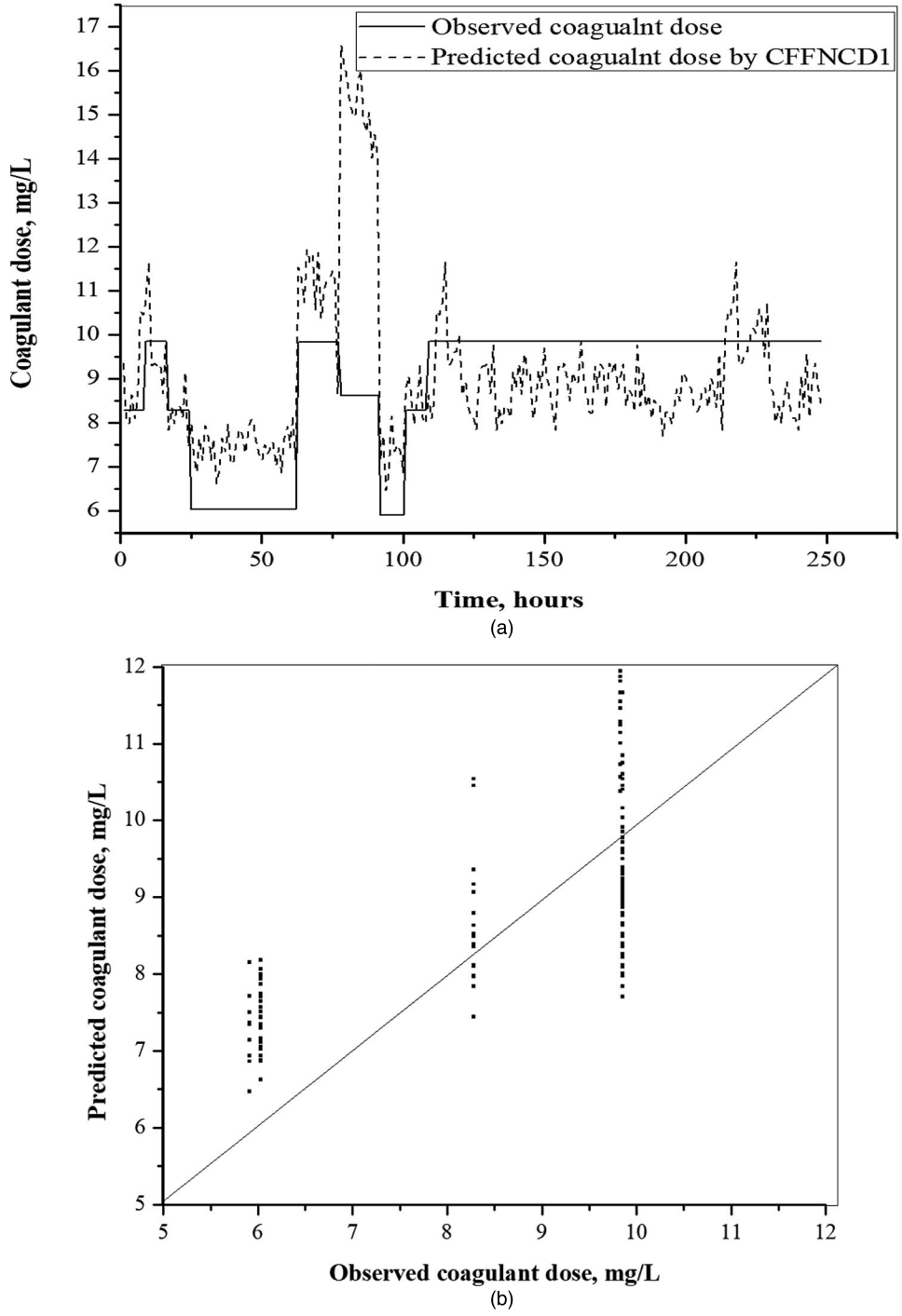


Figure 8. (a) Time series plot of coagulant dose for CFNNCD1 (2-60-1) model. (b) Scatter plot of observed verus predicted coagulant dose for CFNNCD1 (2-60-1) model during training and testing period.

between observed and predicted σ values had 5.10% and 53.76% during training and testing period, respectively. Similarly, the % change between observed and predicted $\gamma 2$ values had 7.37% and 574.24% during training and testing period respectively.

4. Conclusion

1 For the development of FFNN and CFNN models, the best performed LM and BR training algorithms are used among all models. During the development of these models hidden nodes are varied

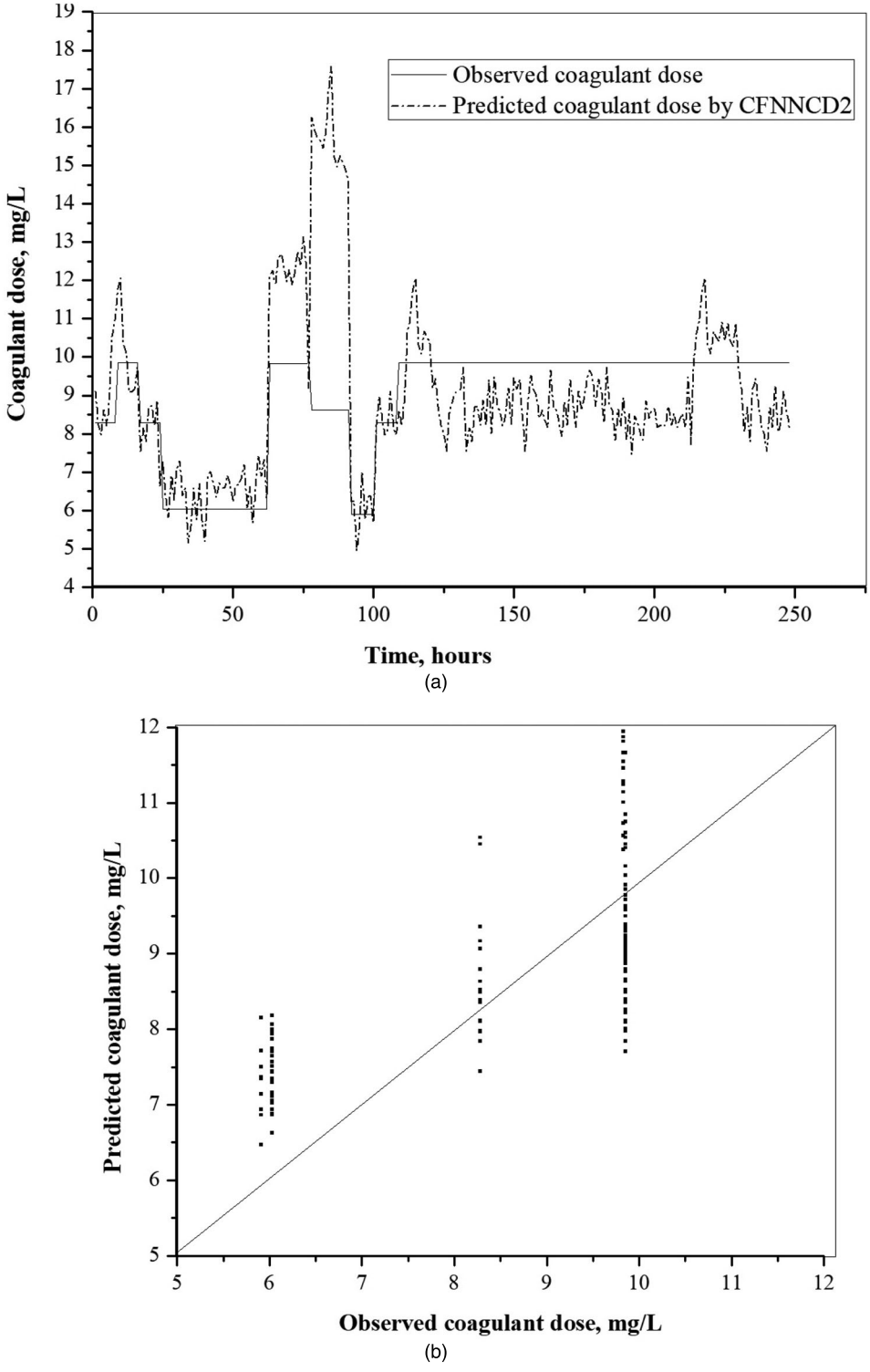


Figure 9. (a) Time series plot of coagulant dose for CFNNCD2 (2-40-1) model. (b) Scatter plot of observed verus predicted coagulant dose for CFNNCD2 (2-40-1) model during training and testing period.

from 15 to 60 and R is found between 0.914 and 0.947. It is observed that the best results are obtained by CFNNCD2 model with hidden node 40, where $R = 0.945$ for training and 0.947 for testing.

2 Standard statistics of CFNNCD2 (2-40-1) model is quite closer to the observed series as compared to other models. This showed the trend and pattern of coagulant dose by CFNNCD2 model were mapped closely with the observed series as compared to

Table 7. Standard statistics of CFNNCD2 (2-40-1) model.

ANN model	Training period				Testing period			
	$\bar{x}$	σ	ɣ1	ɣ2	$\bar{x}$	σ	ɣ1	ɣ2
Observed values	20.157	32.486	5.284	31.601	8.899	1.493	1.233	0.171
CFNNCD1 (2-40-1) model	20.121	30.830	5.373	33.930	9.063	2.295	1.312	1.155

other ANN models. The standard statistics of the CFNNCD2 model exhibited wasfound that values of ɣ1 (Skewness) and ɣ2 (Kurtosis) of predicted series is closely associated with observed series of coagulant dose.

3 The performance CFNNCD2 (2-40-1) model could be improved because both the hidden and the output layer had a weighted connection with input layer. Similarly, the BR training algorithm provides a crucial benchmark for completing the training step and counteracts the network's overtraining. This potential of the BR training algorithm produces good prediction with adaptive and convergent network.

4 Peak values are poorly captured because most of the time on WTP, the observed coagulant dose is kept constant due to lack of automation. In spite of the large variations, the mean value of the observed coagulant dose and predicted coagulant dose are very close to each other. It is seen that the value of ɣ1 of inlet water turbidity and coagulant dose are near to each other, which implies that the pattern of data distribution of both closely matches. The close match pattern of data distribution indicates applied coagulant dose nearly proportional to inlet water turbidity.

The results of the best ANN models are shown in Table 8 during the testing period. The predictions of the CFNN model values are better than that of the FFNN. There is a better consensus among the results of the CFNN model than the FFNN model. In the test period, the MSE reduction of the CFNN model amounted to almost 46.85%. In addition, there is a small improvement over the FFNN model in the forecast results from the CFNN model for the coagulant dose value during the testing period. The results using the CFNN model are very near to the observed values. Therefore, the CFNN model is more capable and precise in the modeling of the coagulation process.

Table 8. Comparison of performance of best ANN models during testing period.

ANN models	Training algorithm	*R*	MSE
FFNNCD1 (2-60-1)	LM	0.944	185.09
FFNNCD2 (2-50-1)	BR	0.945	113.13
CFNNCD1 (2-60-1)	LM	0.943	59.22
CFNNCD2 (2-40-1)	BR	0.947	99.28

Nomenclature

R:	Coefficient of correlation
R^2:	Coefficient of determination
$\bar{x}$:	Mean
σ:	Standard deviation
ɣ1:	Skewness
ɣ2:	Kurtosis
ANN:	Artificial neural networks
BR:	Bayesian regularization
LM:	Levenberg-Marquardt
BOD:	Bio-chemical oxygen demand
COD:	Chemical oxygen demand
CFNN:	Cascade feed forward neural network
CDNN:	Coagulant dose neural network
CCDNN:	Coagulant and chlorine dose neural network
CFNNCD1:	Cascade feed forward neural network coagulant dose model using Levenberg-Marquardt training algorithm
CFNNCD2:	Cascade feed forward neural network coagulant dose model using Bayesian regularization training algorithm
FFNNCD1:	Feed forward neural network coagulant dose model using Levenberg-Marquardt training algorithm
FFNNCD2:	Feed forward neural network coagulant dose model using Bayesian regularization training algorithm

Notes on contributors

Dr. Dnyaneshwar Vasant Wadkar completed his Ph.D. (Civil engineering) in Feb 2020. He is presently working as an assistant professor at AISSMS College of Engineering, Pune, Maharashtra, India. Having 19 years of teaching experience. Published 11 papers in various journals out of which 6 are in the Scopus and web of science and others are in referred journals indexed by UGC and other agencies. 4 research articles were published in reputed newspapers like Times of India, Indian Express, and DNA. He received a research grant of Rs 2 lakh from the Board of College and University (BCUD) Savitribai Phule University Pune. Completed 5 faculty enrichment courses which include NPTEL and Coursera. He published 3 books at the university level and attended 42 FDP, workshop, seminar, conference, and organized 36 interactions with the outside world in terms of the webinar, conference, seminar, and workshops.

Dr. Rahul Subhash Karale Professor in Civil Engineering Department, TSSMs Bhivarabai Sawant College of Engineering and Research, Narhe, Pune, Maharashtra, India.

Dr. Manoj Pandurang Wagh Presently working as Dean (Academics) and Associate professor in Civil Engineering Department, Dr. Vithalrao Vikhe Patil College of Engineering Ahmednagar, Maharashtra, India. He Having 19 years of teaching experience and published 33 papers in science citation index expanded, Scopus, web of science, and UGC referred journals. He received a research grant of Rs 1.90 lakh from the Board of College and University (BCUD) Savitribai Phule University Pune. He is a peer reviewer of many journals such as Indian Chemical Engineer (Taylor and Francis Journal), International Journal of Environmental Analytical Chemistry, (Taylor and Francis Journal), Clean Soil Air water (Wiley journal), Walailak Journal of Science and Technology. He is a reviewer of many international conferences such as Water Resource and Environment (WRE 2019) Macao, China, New Energy, and Future Energy Systems (NEFES 2019, Macao China. Advances in Civil and Ecological Engineering Research (ACEER 2020), Beijing, China. 6th International Conference on Water Resource and Environment (WRE 2020). Tokyo, Japan. He published a patent titled "Waste Management for Ayurveda and Urban Hospitals using IoT" He had published 6 books at the national and international levels.

ORCID

Dnyaneshwar Vasant Wadkar http://orcid.org/0000-0001-6444-3415
Manoj Pandurang Wagh http://orcid.org/0000-0002-3654-0194

References

Abba SI, Elkiran G. 2017. Effluent prediction of chemical oxygen demand from the astewater treatment plant using artificial neural network application. Procedia Comput Sci. 120:156–163.

Amali S, EL Faddouli NE, Boutoulout A. 2018. Machine learning and graph theory to optimize drinking water. Procedia Comput Sci. 127:310–319.

Asnaashari A, McBean EA, Bahram GB, Tutt D. 2013. Forecasting watermain failure using artificial neural network modelling. Can Water Resour J. 38(1):24–33.

Ayvaz MT, Kentel E. 2015. Identification of the Best Booster Station Network for a Water Distribution System. J Water Resour Plan Manag. 141(5):04014076.

Bekkari N, Zeddouri A. 2019. Using artificial neural network for predicting and controlling the effluent chemical oxygen demand in wastewater treatment plant. Manag Environ Qual: Int J. 30(3):593–608.

Bello O, Hamam Y, Djouani K. 2014. Coagulation process control in water treatment plants using multiple model predictive control. Alex Eng J. 53:420–435.

Bobadilla MC, Lorza RL, Garcia RE, Gomez FS, Gozalez EV. 2019. Coagulation: determination of key operating parameters by multi-response surface methodology using desirability functions. Water. 11:1–21.

Bowden GJ, Nixon JB, Dandy GC. 2006. Forecasting chlorine residuals in a water distribution system using a general regression neural network. Math Comput Model. 44:469–484.

Chandwani V. G, K N, Nagar R, Agrawal V, Jethoo A. 2016. Artificial neural networks aided conceptual stage design of water harvesting structures. Perspect Sci. 6:51–155.

Constans S, Bremond B, Morel P. 2003. Simulation and control of chlorine levels in water distribution networks. J Water Resour Plan Manag. 129:135–145.

Cuesta CL, Tau MT. 2014. Using artificial neural network models to assess water quality in water distribution networks. Procedia Eng. 70:399–408.

Djeddou M, Aouatef H, Loukam M. 2019. Wastewater treatment plant performances modelling using artificial neural networks. Eco Technol Wastewater Treat. 1–10.

Erickson JJ, Smith CD, Goodridge A, Nelson KL. 2017. Water quality effects of intermittent water supply in Arraiján, Panama. Water Res. 114:338–350.

Gamboa-Medina MM, Reis Luisa FR. 2017. Sampling design for leak detection in water distribution networks. Procedia Eng. 186:460–469.

Gibbs MS, Morgan N, Maier HR, Dandy GC, Nixon JB, Holmes M. 2006. Investigation into the relationship between chlorine decay and water distribution parameters using data driven methods. Math Comput Model. 44:485–498.

Haghiri S, Sina MA, Daghighi A. 2018. Optimum coagulant forecasting with modeling the Jar test experiments using ANN. J Drink, Water Eng Sci. 11:1–8.

Hamed MM, Mona GK, Ezzat AH. 2004. Prediction of wastewater treatment plant performance using artificial neural networks. Environ Model Softw. 21:465–475.

Hanbay D, Turkoglu I, Demir Y. 2008. Prediction of wastewater treatment plant performance based on wavelet packet decomposition and neural networks. Expert Syst Appl. 34:1038–1043.

Hebati M, Stedmon CA, Murphy K. 2017. Assessment of drinking water quality at the tap using fluorescence spectroscopy. Water Res. 125:1–10.

Heddam S, Bermad A, Dechemi N. 2011. ANFIS-based modelling for coagulant dosage in drinking water treatment plant: a case study. Environ Monit Assess. 184(4):1953–1971. DOI:10.1007/s10661-011-2091-x.

Heddam S, Bermad A, Dechemi N. 2012. Applications of radial-basis function and generalized regression neural networks for modeling of coagulant dosage in a drinking water-treatment plant: comparative study. J Environ Eng. 137:1209–1214.

Heddam S, Dechemi N. 2015. A new approach based on the dynamic evolving neural-fuzzy inference system (DENFIS) for modelling coagulant dosage (Dos): case study of water treatment plant of Algeria. Desalin Water Treat. 53:1045–1053.

Jayaweera CD, Aziz N. 2018. Development and comparison of extreme learning machine and multi-layer perceptron neural network models for predicting optimum coagulant dosage for water treatment. J Phys: Conf Ser. 11:1–8.

Kennedy MJ, Gandomia AH, Miller CM. 2015. Coagulation modeling using artificial neural networks to predict both turbidity and DOM-PARAFAC component removal. J Environ Chem Eng. 3(4):2829–2838.

Kim H, Kim S, Kim J. 2014. Prediction of chlorine concentration in various hydraulic conditions for a Pilot Scale Water Distribution System. Procedia Eng. 70:934–942.

Kim CM, Parnichkun M. 2017. Prediction of settled water turbidity and optimal coagulant dosage in drinking water treatment plant using a hybrid model of k-means clustering and adaptive neuro-fuzzy inference system. Appl Water Sci. DOI:10.1007/s13201-017-0541-5.

Koleva M,N. 2017. Optimisation approaches for the synthesis of water treatment plants. Comput Chem Eng. 106:849–871.

Kote AS, Wadkar DV. 2019. Application of feed forward neural network for prediction of optimum coagulant dose in water treatment plant. Int J Innov Technol Explor Eng. 8(12):1853–1856.

Krishnaiah D, Siva K, Isidore KM, Sarbatly R. 2007. Prediction of clarified water turbidity of Moyog Water Treatment Plant using artificial neural network. J Appl Sci. 7(15):2006–2010.

Kumpel E, Nelson KL. 2013. Comparing microbial water quality in an intermittent and continuous piped water supply. Water Res. 47(14):5176–5188.

Lee L, Lu C, Kung SL. 2004. Spatial diversity of chlorine residual in a drinking water distribution system. J Environ Eng. 130:1263–1268.

Librantz AF, Santos FC, Dias CG. 2018. Artificial neural networks to control chlorine dosing in a water treatment plant, Acta Scientiarum. Technology (Elmsford, N Y). 40:1–9.

Liu W, Ratnaweera H, Kvaal K. 2018. Model–based measurement error detection of a coagulant dosage control system. Int J Environ Sci Technol. DOI:10.1007/s13762-018-1895-4.

Loc HH, Do QH, Cokro AA, Irvine KN. 2020. Deep neural network analyses of water quality time series associated with water sensitive urban design (WSUD) features. J Appl Water Eng Res. 1–20. DOI:10.1080/23249676.2020.1831976.

McCoy ST, VanBriesen JM. 2012. Temporal variability of bacterial diversity in a chlorinated drinking water distribution system. J Environ Eng. 138:786–795.

Muharemi F, Leon F, Logofatu D. 2019. Machine learning approaches for anomaly detection of water quality on a real-world data set. J Inf Telecommun. DOI:10.1080/24751839.2019.1565653.

Najafzadeh M, Zeinolabedini M. 2018. Derivation of optimal equations for prediction of sewage sludge quantity using wavelet conjunction models: an environmental assessment. Environ Sci Pollut Res. 25:22931–22943. DOI:10.1007/s11356-018-1975-5.

Najafzadeh M, Saberi-Movahed F. 2019. GMDH-GEP to predict free span expansion rates below pipelines under waves. Mar Georesources Geotechnol. 37:375–392. DOI:10.1080/1064119X.2018.1443355.

Najafzadeh M, Zeinolabedini M. 2019. Prognostication of waste water treatment plant performance using efficient soft computing models: an environmental evaluation. Measurement. 138:690–701. DOI:10.1016/j.measurement.2019.02.014.

O'Reilly G, Bezuidenhout CC, Bezuidenhout JJ. 2018. Artificial neural networks: applications in the drinking water sector. Water Supply. 18(6):1869–1887.

Pitta KS, Babu JR. 2010. Performance evaluation of waste water treatment plant. Int J Eng Sci Technol. 2(12):7785–7796.

Raduly B, Gernaey KV. 2007. Artificial neural networks for rapid WWTP performance evaluation: methodology and case study. Environ Model Softw. 22:1208–1216.

Saberi-Movahed F, Najafzadeh M, Mehrpooya A. 2020. Receiving more accurate predictions for longitudinal dispersion coefficients in water pipelines: training group method of data handling using extreme learning machine conceptions. Water Resour Manag. 34:529–561. DOI:10.1007/s11269-019-02463-w.

Saha AK, Choudhury S, Majumder M. 2017. Performance efficiency analysis of water treatment plants by using MCDM and neural network model. Int J Sci Technol. 3(1):27–35.

Santos S, Marina K, Celeste AB, El-Shafie A. 2019. ANNs and inflow forecast to aid stochastic optimization of reservoir operation. J Appl Water Eng Res. 1–10. DOI:10.1080/23249676.2019.1687017.

Swetland KA, Weber-Shirk ML, Lion LW. 2013. Gravity-powered chemical dose controller for sustainable, municipal-scale drinking water treatment. J Environ Eng. 139:1023–1034.

Wang X, Ratnaweera H, Holam JA, Olsbu V. 2017. Statistical monitoring and dynamic simulation of a wastewater treatment plant: a combined approach to achieve model predictive control. J Environ Manag. 193:1–7.

Wu G-D, Lo S-L. 2010. Effects of data normalization and inherent-factor on decision of optimal coagulant dosage in water treatment by artificial neural network. Expert Syst Appl. 37(7):4974–4983.

Wu GD, Lo SL. 2008. Predicting real-time coagulant dosage in water treatment by artificial neural networks and adaptive network-based fuzzy inference system. Eng Appl Artif Intell. 21:1189–1195.

Zeinolabedini M, Najafzadeh M. 2019. Comparative study of different wavelet-based neural network models to predict sewage sludge quantity in wastewater treatment plant. Environ Monit Assess. 191:163. DOI:10.1007/s10661-019-7196-7.

Zhang K, Achari G, Li H, Zargar A, Sadiq R. 2013. Machine learning approaches to predict coagulant dosage in water treatment plants. Int J Syst Assur Eng Manag. 4(2):205–214.

Zhang JM, Lee HP, Khoo BC, Teo CP, Haja N, Peng KQ. 2011. Modeling and simulations of flow pattern, chlorine concentration, and mean age distributions in potable water service reservoir of Singapore. J Environ Eng. 137:575–584.