

Comparative study of different wavelet-based neural network models to predict sewage sludge quantity in wastewater treatment plant

Maryam Zeinolabedini  · Mohammad Najafzadeh

Received: 13 March 2018 / Accepted: 3 January 2019 / Published online: 16 February 2019
© Springer Nature Switzerland AG 2019

Abstract In this study, artificial neural networks (ANNs) including feed forward back propagation neural network (FFBP-NN) and the radial basis function neural network (RBF-NN) were applied to predict daily sewage sludge quantity in wastewater treatment plant (WWTP). Daily datasets of sewage sludge have been used to develop the artificial intelligence models. Six mother wavelet (W) functions were employed as a pre-processor in order to increase accuracy level of ANNs. In this way, a 4-day lags were considered as input variables to conduct training and testing stages for the proposed W-ANNs. To compare performance of W-ANNs with traditional ANNs, coefficient of correlation (R), root mean square error (RMSE), mean absolute error (MAE), and Nash-Sutcliffe efficiency coefficient (NSE) were considered. In the case of all wavelet functions, it was found that W-FFBP-NN ($R = 0.99$ and $MAE = 5.78$) and W-RBF-NN ($R = 0.99$ and $MAE = 6.69$) models had superiority to the FFBP-NN ($R = 0.9$ and $MAE = 21.41$) and RBF-NN ($R = 0.9$ and $MAE = 20.1$) models. Furthermore, the use of DMeyer function to improve ANNs indicated that W-FFBP-NN (RMSE =

7.76 and $NSE = 0.98$) and W-RBF-NN (RMSE = 9.35 and $NSE = 0.98$) approaches stood at the highest level of precision in comparison with other mother wavelet functions used to develop the FFBP-NN and RBF-NN approaches. Overall, this study proved that application of various mother wavelet functions into architecture of ANNs led to increasing accuracy of artificial neural networks for estimation of sewage sludge volume in the WWTP.

Keywords Sewage sludge quantity · Wastewater treatment plant · Artificial neural networks · Wavelet functions

Introduction

Sludge is chiefly composed of water with lower quantity of solid materials eliminated from liquid sewage. Sewage sludge has two main sources including industrial and domestic wastewaters (Danish et al. 2016). With emergence of modern communities, huge quantity of sewage sludge, which is emanated from human activities, is one of the most challenging environmental issues in industrialized countries (e.g., Belanche et al. 1999; Hanbay et al. 2008). Compared to the past decades, ferocious demands for contemporary wastewater treatment plant (WWTP) in order to increase level of water quality for masses in the community have enhanced the quantity of sewage sludge generation. Fundamentally, sludge production is inextricably bound up with the

M. Zeinolabedini · M. Najafzadeh (✉)
Department of Water Engineering, Faculty of Civil and Surveying Engineering, Graduate University of Advanced Technology, Kerman, Iran
e-mail: m.najafzadeh@kgut.ac.ir
e-mail: moha.najafzadeh@gmail.com

M. Zeinolabedini
e-mail: mary.zeinolabedini@gmail.com

wastewater treatment process. This means that an increase of wastewater treatment can produce huge quantity of sewage sludge. On account of putrefaction of the organic components of the sludge and difficulties of its disposal process, gradual gathering of huge quantity of the sewage sludge from various sources such as domestic and industrial wastewaters has deteriorative effects on environment (e.g., Llorens et al. 2008; Pai et al. 2011; Huang and Chen 2015).

With rapid rise of the industrial revolution and dwelling in megacities, recycling process of the wastewater plays a substantial role in management of water resources. WWTP units are one of important places for conducting recycling processes. In fact, treatment of sewage sludge is performed due to decrease in quantity of sludge physical properties such as mass and humidity in order to cut down high costs of disposal process. There is no denying the fact that an accurate estimation of sewage sludge volume can have significant influences on performance of WWTP (e.g., Belanche et al. 1999; Hamoda et al. 1999; Wan et al. 2011; Kaira and Christian 2013; Filipović et al. 2013; Li et al. 2014; Danish et al. 2016; Fernandez de Canete et al. 2016; Kolečka et al. 2017). In the case of designing the sludge treatment units (STU), pieces of information related to the quantities, sources, and weight-volume relationship of sludge are essentially needed to consider as well as providing disposal facilities (Mirabi et al. 2014). This is highly recommended to predict the volume of sludge for designing sludge treatment facilities in the wastewater treatment systems (Kaira and Christian 2013). Generally, it can be said that there is no relationship to predict sewage sludge quantity owing to the fact that sewage sludge treatment is a complicated process with high level of non-linearity, uncertainty, and large time-delay. For instance, during storm events, upsets to the various physical and biological processes occasionally occur in a WWTP, and thus, possibility of forecasting the hydraulic loads to treatment facilities during storm is highly parsimonious for optimization of WWTP operation. Most of the hydrological and hydraulic models which are capable to describe sewage collection systems are grouped into deterministic techniques (El-Din and Smith 2002). On the other hand, such numerical models need in-depth information about system and occasionally rely on the large number of variables which some of them are uncertain or not easy-to-define. In this way, WWTP units are the systems with degree of complexity due to existence of physical and chemical processes.

Nowadays, artificial intelligence (AI) approaches are used to solve problems of a complex system. In accordance with potential of AI models, for more than two decades, soft computing techniques such as artificial neural networks (ANNs), adaptive neuro-fuzzy inference system (ANFIS), support vector machine (SVM), random forest (RF), multivariable adaptive regression splines (MARS), gene-expression programming (GEP), evolutionary polynomial regression (EPR), and model tree (MT) have been applying to assess operation of WWTP (e.g., Boger and Guterman 1990; Cohen et al. 1997; Belanche et al. 1999; Gontarski et al. 2000; El-Din and Smith 2002; Çinar 2005; Llorens et al. 2008; Moral et al. 2008; Noori et al. 2010; Li et al. 2014; Honggui et al. 2014; Safavi et al. 2015; Guo et al. 2015; Han et al. 2016; Najafzadeh and Zeinolabedini 2018; Cong and Yu 2018). Applications of AI models showed that physical and chemical processes of WWTP can be well evaluated. On the basis of previous investigations, the most predominant AI approach used to assess performance of WWTP was ANN models as, multilayer perceptron (MLP) and radial basis function neural networks (RBF-NN). These two common ANNs have remarkable advantages. The MLP, known as feed forward neural network (FF-NN), is the most common supervised training model of ANN which is essentially applied for regression and classification. This approach has high rate of convergence to find non-linear relationships between input variables and output variables. Moreover, RBF-NN is more flexible than other structure of the ANNs due to the fact that any non-linear function can be used in the structures of RBF-NN as an activation function. Hence, structure of the RBF-NN will be capable of improving accuracy level of model performance. Furthermore, RBF-NN models are easy-to-design because of having just three layers. In this way, volume of computations is intrinsically low in comparison with other types of ANNs (Oja 1994). Sewage sludge values, known as stationary signals, occasionally have high level of irregularly, and additionally, this problem can have repercussion on accuracy level of ANN models. To get rid of this problem, researchers have applied wavelet transformation (WT), as a preprocessor, to smooth irregularity of sewage sludge signals. In fact, on the most useful advantage of WT is de-noising of the non-stationary signals in different decomposition levels. In this way, wavelet function has the capability to reduce irregularity of signals which are out-of-band (Najafzadeh and Zeinolabedini 2018).

In this research, two types ANNs including feed forward back propagation neural network (FFBP-NN) and RBF-NN are used to predict the volume of sludge produced in WWTP, located in Kerman City, Iran. Furthermore, various mother wavelet functions, known as model preprocessors, such as Daubechies (db), Coiflet (coif), DMeyer (dmey), Haar, Symlets (Sym), and BiorSplines (bior) are used to increase precision level of FFBP-NN and RBF-NN techniques. Performances of the proposed W-ANNs models are investigated by means of various statistical criteria.

Intelligence predictive models: a brief review

Since the middle of 1980s, AI techniques have been applied to assess physical and chemical processes in WWTP. A survey of literature indicated that Rumelhart and McClelland (1986) were probably the first researchers who have used FFBP-NN model to predict removal of total organic carbon (TOC) in the treatment plant. Boger and Guterman (1990) have applied successfully a three-layer ANN to predict weekly $\text{NH}_4\text{-N}$ concentration using a two-year data for Shafdan WWTP, Israel. Côté et al. (1995) have proposed a two-step model to improve accuracy level of activated sludge process (ASP) prediction. In the first step, optimal selection of parameters was conducted using downhill simplex approach and consequently five effective variables as effluent suspended solid (SS_{ef}), effluent COD, NH_4 , dissolved oxygen (DO) in the mixed liquor, and volatile suspended solids (VSS) in the returned activated sludge (AS) were considered. In the second phase, feed forward network was successfully utilized to estimate the ASP. Cohen et al. (1997) used neuro-fuzzy technique to prosperously evaluate load entering the WWTP and biochemical conversion through the sequencing batch reactors (SBR) process. Belanche et al. (1999) have predicted behavioral patterns of WWTP, located in Catalonia, Spain, by means of fuzzy heterogeneous neural network (FH-NN). They found that performance of FH-NN to estimate BOD, COD, and total suspended solid (TSS) was better than other ANNs.

Gontarski et al. (2000) have employed ANN models to predict environmental characteristics of output flow from WWTP at Rhodiaco Ltda, Brazil. From their study, statistical results indicated that ANNs estimated accurately elimination of TOC. El-Din and Smith (2002) have proposed FFBP-NN approach to predict

wastewater inflow rate at the entrance of plant. They found that FFBP-NN could simulate efficiently physical-chemical processes occurring within storm events in WWTP located in the Edmonton area, Canada. Onkal-Engin et al. (2005) have presented a reasonable relationship between sewage odor and BOD of WWTP by means of ANNs. Mjalli et al. (2007) have applied ANN model to appraise WWTP at Doha city, Qatar. They considered COD, BOD, and TSS as input variables and found that ANN had good performance. Hanbay et al. (2008) collected TSS data from WWTP in Malatya, Turkey and found that W-ANN had an acceptable level of efficiency in assessment of WWTP performance. Moreover, Jalili Ghazizade and Noor (2008) have used ANN model to estimate mass of municipal solid waste production in Mashhad, Iran. They concluded that ANN approach had lower computational error in comparison with traditional models. On the basis of mass balance modeling, Llorens et al. (2008) predicted outflow of WWTP into Tet River, south of France. They used Kohonen Self-Organizing Maps (KSOMs) for retrieval of missing datasets and found that KSOMs-ANN was an appropriate tool for system evaluation. Moral et al. (2008) have predicted ASP using ANN model. They considered nine various chemical variables to obtain an acceptable level of accuracy. In Noori et al. (2009a) research, principle component analysis (PCA) and SVM used to estimate solid waste management system at Mashhad city, Iran. They found that application of PCA could increase precision level of SVM technique. Furthermore, Noori et al. (2009b) have employed PCA-ANN to predict solid waste production in Tehran, Iran. They concluded that performance of the PCA-ANN stood at the higher level of linear regression model. Fernandez et al. (2009) have used successfully neurofuzzy model to precisely forecast flow rate of WWTP. Pai et al. (2011) used ANFIS models to obtain precise estimate effluent suspended solids (SS_{ef}), effluent COD, and effluent pH from WWTP in Taiwan. Nasr et al. (2012) have applied FFBP-NN models to evaluate operation of WWTP in El-Agamy, Egypt. They considered BOD, COD, and TSS as input variables to develop FFBP-NN model. From their research, they concluded that ANN is a robust tool to assess the performance of WWTP. Boniecki et al. (2012) employed successfully ANNs to predict ammonia emissions sent out from sewage sludge composting. Ongen et al. (2013) indicated that ANN had the capability of providing an accurate prediction for calorific value of synthetic gas generated

by dewatering sludge in WWTP. Verma et al. (2013) used different AI models such as ANN, k-nearest neighbor, MARS, SVM, and RF to predict TSS in WWTP located at Des Moines, USA. They considered effluent flow rate and the influent carbonaceous biochemical oxygen demand (CBOD) as input variables. They concluded that FFBP model was the best approach among other AI techniques. In Zare-Abyaneh (2014) study, concentration of water quality indices for WWTP was estimated by multivariate linear regression (MLR) and ANN models. Also, temperature, pH, TSS, and TS were taken into account as input variables. Statistical results demonstrated that ANN had better performance than MLR model. Honggui et al. (2014) used a fuzzy neural network (FNN) to precisely simulate fault detection patterns of WWTP. From their research, DO, pH, COD, and total nutrients (NT) were considered as input parameters. Guo et al. (2015) predicted total Nitrogen concentration using ANN and SVM for WWTP in Ulsan, Korea. They found that SVM model had better performance in comparison with ANN approach. Moreover, Huang and Chen (2015) have employed two types of ANN model including FFBP-NN and generalized regression neural network (GRNN) to estimate the thin layer pattern in municipal sewage sludge (MSS). They found that both ANN models could provide an accurate prediction of average temperature and moisture content in MSS. Han et al. (2016) concluded that recurrent self-organizing neural network (RSONN) provided precise estimations of SVI for WWTP in the Beijing, China. Cong and Yu (2018) found that W-ANN could produce more accurate prediction of water quality indices for WWTP than ANN and SVM techniques. In Han et al. (2018) study, a fuzzy neural network (FNN) based on PCA technique was used to predict concentrations of effluent total phosphorous (TP) and ammonia nitrogen (NH₄-N). They tested PCA-FNN for various WWTP in China and, as a result, robustness of the proposed AI model was found. Recently, Najafzadeh and Zeinolabedini (2018) have used three conjunction models of W-EPR, W-MT, and W-GEP to precisely predict rate of sewage sludge in the WWTP located in Kerman city, Iran.

Overview of the study region

Kerman city, located in the southeast of Iran, has stood at the 14th most populated city in Iran. In terms of

geographical properties, Kerman has altitude of 1755 m above sea level, latitude 30° 17', and longitude 57° 04'. In this city, the WWTP was designed in eight modules and for population about 75,000. The plant, located in the last 6 km (approximately 3.5 miles) of northwestern part of Kerman city, has area of 45 ha and design capacity of 259,200 m³/day. Location of Kerman WWTP was illustrated in Fig. 1. It is worth mentioning that the proposed method applied for refining process in this WWTP unit is conventional activated sludge (CAS) (Forster and Foundation 2003). In fact, CAS system basically comprises a tank for aeration process, which is applied for biological degradation. Additionally, there is a secondary tank, known as a clarifier, for sedimentation. In the clarifier, separation of sewage sludge from treated wastewater is performed. In fact, CAS has some advantages. Not only is CAS the low-cost project, but also it needs small land. Furthermore, it has effluent with good quality and relatively low head losses.

In the present research, 1736 datasets of daily sewage sludge values have been collected from Kerman WWTP for a period of 57 weeks beginning at 2011 to develop the proposed models. Figure 2 illustrates variations of sewage sludge values versus time. Generally, the datasets (23rd October 2011 until 21st July 2016) was divided randomly into two parts including training (calibration) and testing (or validation). In other words, exactly 70% of the datasets were dedicated to training purpose while the remaining 30% used for testing. Statistical characteristics of sewage sludge were presented in Table 1.

Methodology

Wavelet transformation

Wavelet function has the capability to analyze stationary (or periodic) and non-stationary datasets by means of various types of wavelet mother functions. This technique can transform datasets to both time and frequency information with a higher resolution, which is not provided by conventional transformation (e.g., Hamed et al. 2004; Noori et al. 2010; Kisi and Shiri 2011; Shoaib et al. 2015; Shiri 2018). Basically, wavelet transformation is defined in form of two important functions as, continuous wavelet transformation



Fig. 1 (a) Study area location; (b) general scheme of Kerman WWTP; and (c) sludge drying beds

(CWT) and discrete wavelet transformation (DWT). The CWT of a time series $f(t)$ is presented as follows (Kisi and Shiri 2011):

$$w_{a,b(t)} = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{a}} \psi^* \left(\frac{t-b}{a} \right) dt \tag{1}$$

Where * is indicative of the complex conjugate function, $\psi(t)$ is the wavelet function (or mother

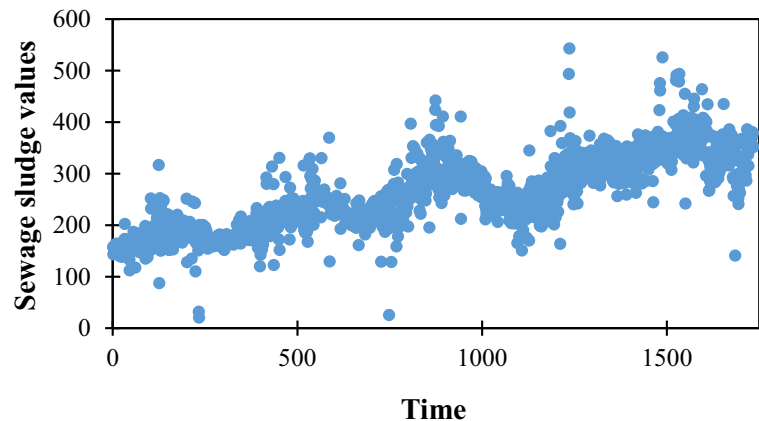
wavelet), a and b are the frequency factors and time factor, respectively.

In the case of a discrete time series $[f(t)]$, the DWT can be defined as follows:

$$w(a,b)_D = 2^{-j/2} \sum_{k=1}^{j=J} \psi^* \left(2^{-j/2-k} \right) f(t) dt \tag{2}$$

in which j and k are the integer values. As seen in Fig. 3, the DWT of a specific time series is computed by sending it through a series of filters,

Fig. 2 Variations of sewage sludge versus time for Kerman WWTP



known as low-pass and high-pass filters. This process is known as Mallat algorithm (Mallat 1989). The time series (or signal) is decomposed into two main parts including the approximation and detail. In the current investigation, coefficients of detail (d) and approximation (a) related to the decomposition of a non-periodic signal are provided by means of Eq. (2).

Selection of wavelet families

One of the most challenging issues in Wavelet transform is to select the most appropriate mother wavelet for analyzing the time series. In fact, a wide range of mother wavelets employed on the time series can provide various results. In this study, wavelet functions, originated from six various wavelet families with different orders of subclass, were presented in Table 2. Other descriptions about principle of mother wavelet functions were mentioned in Seo et al. (2015). All the wavelet families are used for DWT phase of hybrid models (Maheswaran and Khosa 2012; Joseph and Anto 2012). In the present investigation, 32 selected wavelet functions are employed to evaluate performance of hybrid W-ANN models.

Table 1 Statistical properties of under study sewage sludge

| Maximum | Minimum | Mean | Standard deviation |
|---------|---------|--------|--------------------|
| 543.16 | 20.36 | 259.28 | 71.38 |

Definition of decomposition level

In addition to selection of typical mother wavelet functions, determining decomposition level can influence on accuracy of sewage sludge time series analysis. Different researches considered various decomposition levels to analyze non-stationary signals in their investigations (Partal and Kisi 2007; Kisi and Shiri 2011; Catalão et al. 2011; Adamowski and Sun 2010; Seo et al. 2015; Hu and Wang 2015; Kasiviswanathan et al. 2016; Najafzadeh and Zeinolabedini 2018). In Aussem et al. (1998) research, the decomposition level was determined by the following equation as:

$$L = \text{int}(\log(N)) \quad (3)$$

in which, L = the decomposition level, N = the number of time series datasets, and $\text{int}[\cdot]$ = the integer-part function. L value is equal to 3 due to 1736 datasets. Therefore, for a given wavelet function, three decomposition levels (d_1, d_2, d_3) and approximation (a_3) were generated. In this study, 128 (4×32) sub-time series components (a_3, d_1, d_2, d_3) were generated by means of six mother wavelet functions. For instance, Fig. 4 shows variations of Haar wavelet function components with time.

Artificial neural networks

Feed forward neural network

The multilayer perceptron (MLP), known as feed forward neural network, is one of the most common and practical architecture of ANN including three

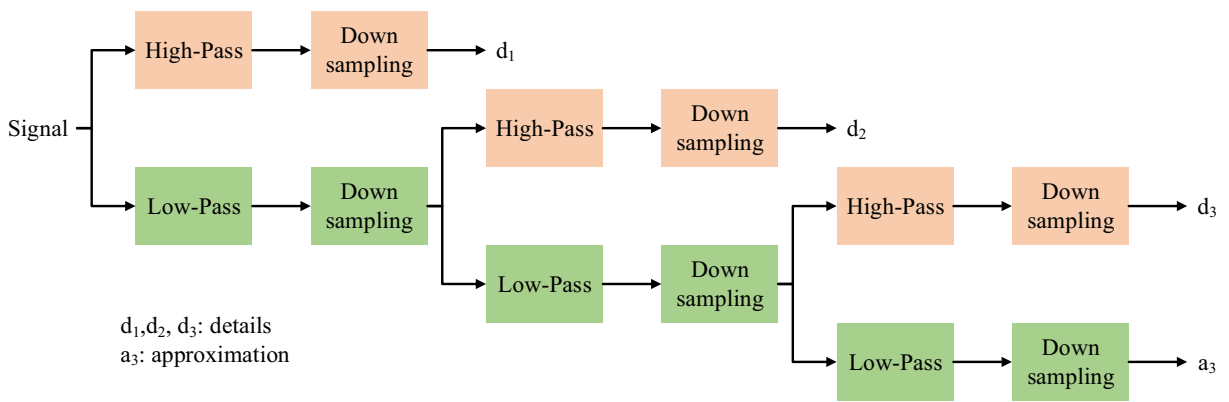


Fig. 3 Mallat algorithm for three-level decomposition of a signal

typical layers: an input layer, a hidden layer(s), and an output layer. The volume of datasets according to input and output variables plays a substantial role in determining number of neurons in each layer and number of hidden layers. Basically, the FF-NN is considered as a model with m number of neurons in the input layer $\{x_1, x_2, \dots, x_m\}$, r number of neurons in the hidden layer $\{h_1, h_2, \dots, h_r\}$ and p number of neurons in the output layer $\{z_1, z_2, \dots, z_p\}$:

$$h_j = f\left(\sum_{i=1}^m x_i w_{ij} + \alpha_j\right) \tag{4}$$

$$z_k = g\left(\sum_{j=1}^r h_j \theta_{kj} + \beta_k\right) \tag{5}$$

Where, $f(\cdot)$ and $g(\cdot)$ are the activation functions, α_j is the bias for j th neuron in hidden layer and β_k is the bias for k th neuron in output layer. Moreover, w_{ij} is the weight of the connection from neuron x_i to neuron h_j , and θ_{ij} is the weight of the connection from neuron h_j to z_k . Eventually, $i, j,$ and k denote input, hidden, and output layers, respectively. In-depth information about FF-NN can be found in literature (Altunkaynak and Nigussie 2015; Heidari et al. 2016).

Radial basis function neural network

Radial basis function neural network (RBF-NN) is capable to present solutions to the various problems such as approximation (or prediction) of linear and non-linear functions, pattern recognition of signal, and data classification. In this regard, radial basis functions were employed as activation functions in the structure of ANNs. The final output of the RBF-NN is a linear combination of radial basis functions of the inputs and neuron parameters. RBF is three-layer neural networks which there are input vectors constructing the input layer in the first layer. In the second layer, introduced as a hidden layer, there are non-linear activation functions which are applied in the hidden layer. Final neuron in output layer is a linear function. Furthermore, basic descriptions of RBF-NN model can be found in literature (e.g., Zounemat-Kermani et al. 2009; Bateni et al. 2007).

Models development

In the similar investigations for Kerman WWTP, Najafzadeh and Zeinolabedini (2018) concluded that use of a 4-day lag has produced more accurate

Table 2 Results of ANNs performance without wavelet transformation

| Model | R | | RMSE (m ³ /day) | | MAE (m ³ /day) | | NSE | |
|---------|----------|---------|----------------------------|---------|---------------------------|---------|----------|---------|
| | Training | Testing | Training | Testing | Training | Testing | Training | Testing |
| FFBP-NN | 0.89 | 0.90 | 32.1 | 30.8 | 21.60 | 21.41 | 0.79 | 0.81 |
| RBF-NN | 0.89 | 0.90 | 31.7 | 30.8 | 21.41 | 20.10 | 0.80 | 0.81 |

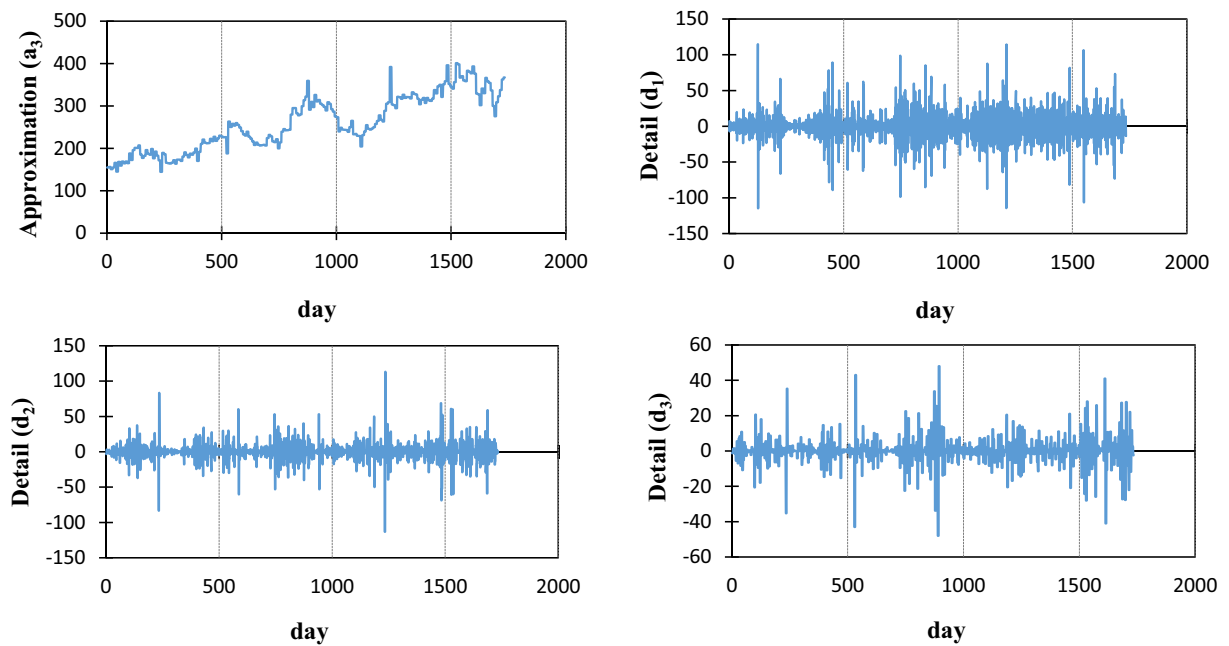


Fig. 4 Decomposed Haar wavelet sub-time series components (Ds) of sewage sludge datasets

prediction of $S(t)$ than other lags. Therefore, in this study, prediction of sewage sludge quantity $[S(t)]$, as an output variable, is conducted using a 4-day lag for $S(t)$ $[S(t - 1), S(t - 2), S(t - 3), \text{ and } S(t - 4)]$ as input variables.

To model FFBP-NN, the Levenberg Marquardt (LM) algorithm was applied and additionally, the logsig and pureline functions were chosen as activation functions for the hidden layers and output layer, respectively. Learning rate of 0.01 and epoch number of 100 were used in training phase. A diagram of the FFBP-NN model including a network with two hidden layers of six neurons was illustrated in Fig. 5. To gain the lowest computational error of training stage for RBF-NN model, 30 neurons in hidden layers were considered. Thus, proposed structure of RBF-NN for predicting $S(t)$ rates has been shown in Fig. 6.

In order to obtain the temporal and spectral information of the time series sewage sludge data, input variables (lagging from 1 to 4 days) are completely decomposed into sub-series of approximation and details using the DWT. As mentioned in Table 2, 32 selected mother wavelet functions were employed to analyze the effects of the DWT on the precision level of ANNs in comparison with simple FFBP-NN and RBF-NN models. Each lagged signal (or time series) was decomposed into three levels. In W-ANN models, four input variables were transformed into

16 sub-time series. Training of W-FFBP-N model was performed using 32 mother wavelet functions. All the W-FFBP-NN models have two hidden layers with 12 neurons. Learning rate value of 0.01 and epoch of 100 were considered to obtain minimum computations errors in training stages. Similarly, performance of training stages for 32 W-RBF-NN indicated that the models had the same structure with 30 hidden neurons. Implementation of the wavelet function with FFBP-NN and RBF-NN techniques has conceptually been depicted in Fig. 7.

Results and discussion

Training and testing for both ANNs and W-ANNs were evaluated using various statistical parameters as, the correlation coefficient (R), the root mean square error (RMSE), the mean absolute error (MAE), and the Nash-Sutcliffe efficiency coefficient (NSE). These parameters were expressed as,

$$R = \frac{\sum_{i=1}^n (S(t)_{\text{Obs}}^i - \overline{S(t)_{\text{Obs}}}) (S(t)_{\text{Pre}}^i - \overline{S(t)_{\text{Pre}}})}{\sqrt{\sum_{i=1}^n (S(t)_{\text{Obs}}^i - \overline{S(t)_{\text{Obs}}})^2} \sqrt{\sum_{i=1}^n (S(t)_{\text{Pre}}^i - \overline{S(t)_{\text{Pre}}})^2}} \quad (6)$$

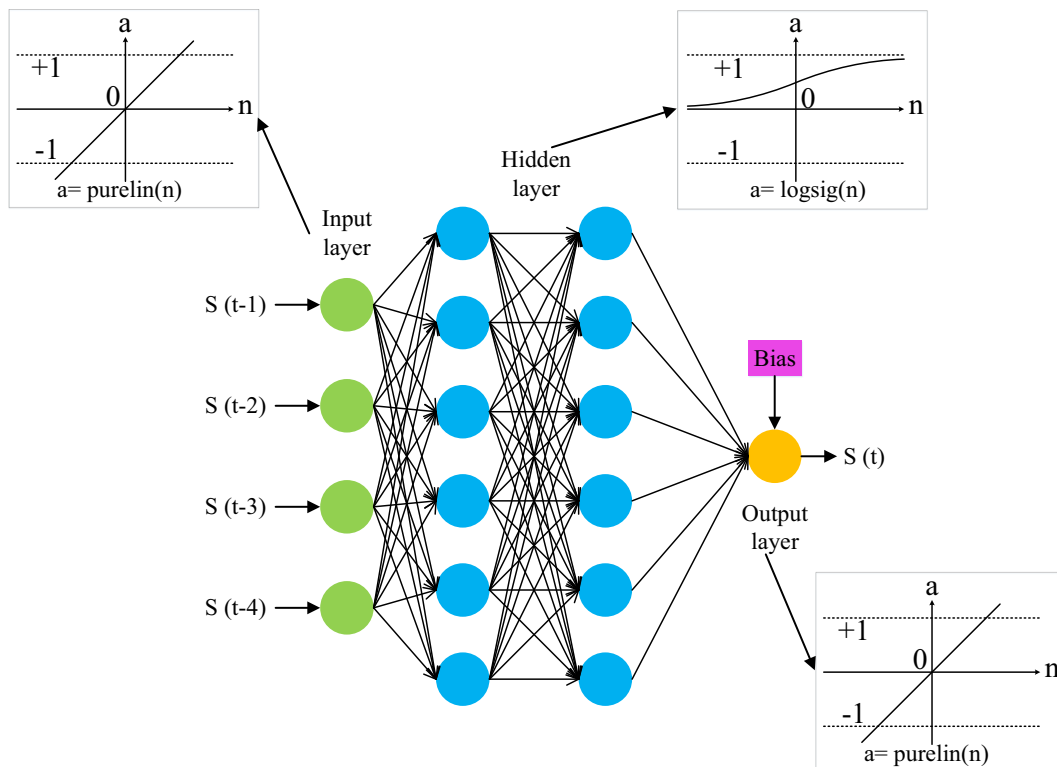


Fig. 5 Proposed architecture of FFBP-NN for sewage sludge prediction

$$MAE = \frac{1}{n} \sum_{i=1}^n |S(t)_{Obs}^i - S(t)_{Pre}^i| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (S(t)_{Obs}^i - S(t)_{Pre}^i)^2} \tag{8}$$

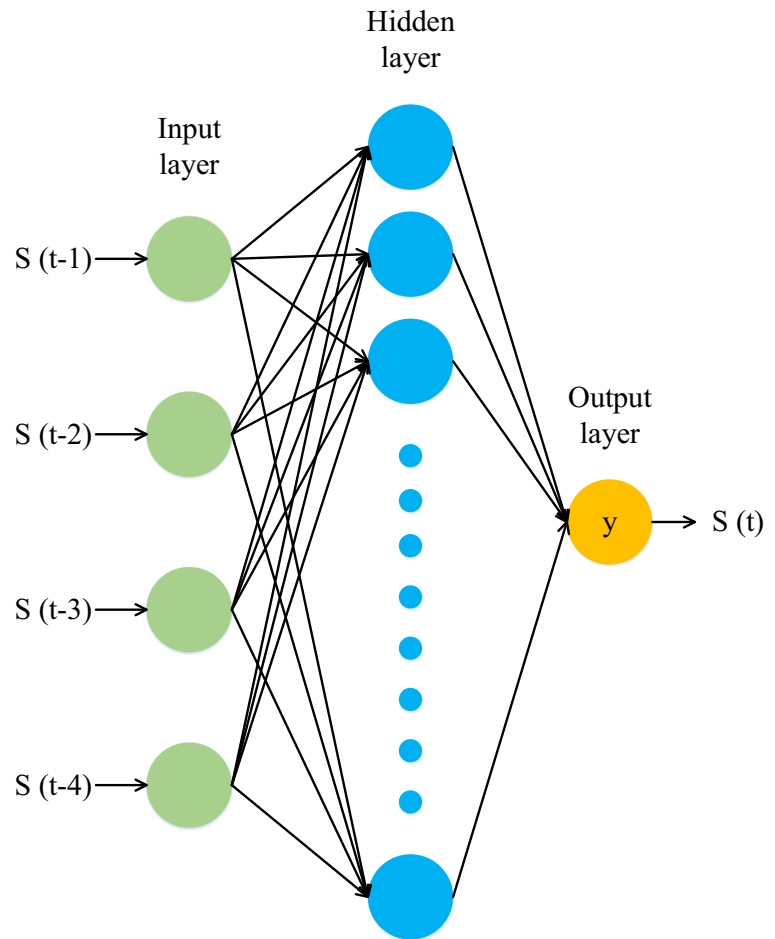
$$NSE = 1 - \frac{\sum_{i=1}^n (S(t)_{Obs}^i - S(t)_{Pre}^i)^2}{\sum_{i=1}^n (S(t)_{Obs}^i - \overline{S(t)_{Obs}})^2} \tag{9}$$

where $S(t)_{Obs}$ is the observed values, $\overline{S(t)_{Obs}}$ is the mean of $S(t)_{Obs}$, $S(t)_{Pre}$ is the estimated (predicted) values, $\overline{S(t)_{Pre}}$ is the mean of $S(t)_{Pre}$, and n is the number of data sets.

Quantitative results of RBF-NN and FFBP-NN techniques were presented in Table 5. In training stage, Table 3 indicated that efficiency of RBF-NN model stood at the relatively higher level of accuracy (RMSE = 31.7 and MAE = 21.41) than FFBP-NN (RMSE = 32.1 and MAE = 21.6). For testing phase, statistical parameters of R (0.9), RMSE (30.8), and

NSE (0.81) showed that both RBF-NN and FFBP-NN approaches have the relatively same performance to predict $S(t)$. Furthermore, statistical performance of W-FFBP-NN for training and testing phases were presented in Table 4. In training stage, Table 4 showed that Daubechies function with db10 sub-class has produced lower computational error ($R = 0.99$ and $RMSE = 9.52$) than other db sub-classes. Similarly, values of MAE and NSE showed this trend. For the Coiflets function, the use of fifth sub-class provided the most accurate estimation of sewage sludge (MAE = 6.63 and NSE = 0.98). Moreover, Symlets function with sub-class of sym8 predicted $S(t)$ with the highest level of accuracy (RMSE = 10.68 and MAE = 7.55) in comparison with other sub-classes. Statistical parameters of RMSE (10.6) and NSE (0.97) indicated that BioSplines function has stood at the highest level of precision in the prediction of sewage sludge compared with other bio sub-classes. According to Table 4, it should be generally mentioned that DMeyer and Haar functions have the lowest and highest computational error in

Fig. 6 Proposed architecture of RBF-NN for prediction of sewage sludge



prediction of $S(t)$, respectively. In testing stage, superiority of mother wavelet functions had as same manner as training phases. For instance, DMeyer function had the best performance (RMSE = 7.76 and MAE = 5.73) than other mother wavelet functions.

Table 5 presents quantitative results for both training and testing stages of W-RBF-NN models. In training phase, DMeyer function with RMSE of 9.56 and MAE of 6.81 has provided the most accurate prediction than other mother wavelet function, while Haar function resulted in the worst performance (RMSE = 20.82 and MAE = 13.02). The Coiflet function with *coif5* sub-class has produced more precise prediction of $S(t)$ (RMSE = 10.52 and MAE = 7.64) than *sym8* sub-class (RMSE = 11.68 and MAE = 8.24), *bior6.8* sub-class (RMSE = 11.21 and MAE = 7.92), *db9* sub-class (RMSE = 11.33 and MAE = 8.07). Furthermore, values of R and NSE indicated superiority of *coif5* sub-class to

the *sym8*, *bior6.8*, and *db9* sub-classes. For testing stages of W-RBF-NN models, DMeyer function wavelet had the best efficiency (RMSE = 9.35 and MAE = 6.69) than *db10* (RMSE = 10.45 and MAE = 7.64), *coif5* (RMSE = 10.17 and MAE = 7.64), *sym8* (RMSE = 11.68 and MAE = 7.77), *bior6.8* (RMSE = 11.04 and MAE = 0.97), and Haar (RMSE = 20.91 and MAE = 12.8). Values of R and NSE were indicative of superiority of DMeyer-RBF-NN to the other mother wavelet functions with various sub-classes.

In order to evaluate the efficiency of hybrid wavelet models, the percentage of increase (or decrease) in the statistical indices was compared with those obtained using simple ANN models. Figure 8a–d illustrated the percentage of variations in the R, RMSE, MAE, and NSE values for 32 W-FFBP-NN models, respectively. As seen in Fig. 8, W-FFBP-NN model developed by the *dmey* wavelet function had the best efficiency and the corresponding increase in the R and NSE values were

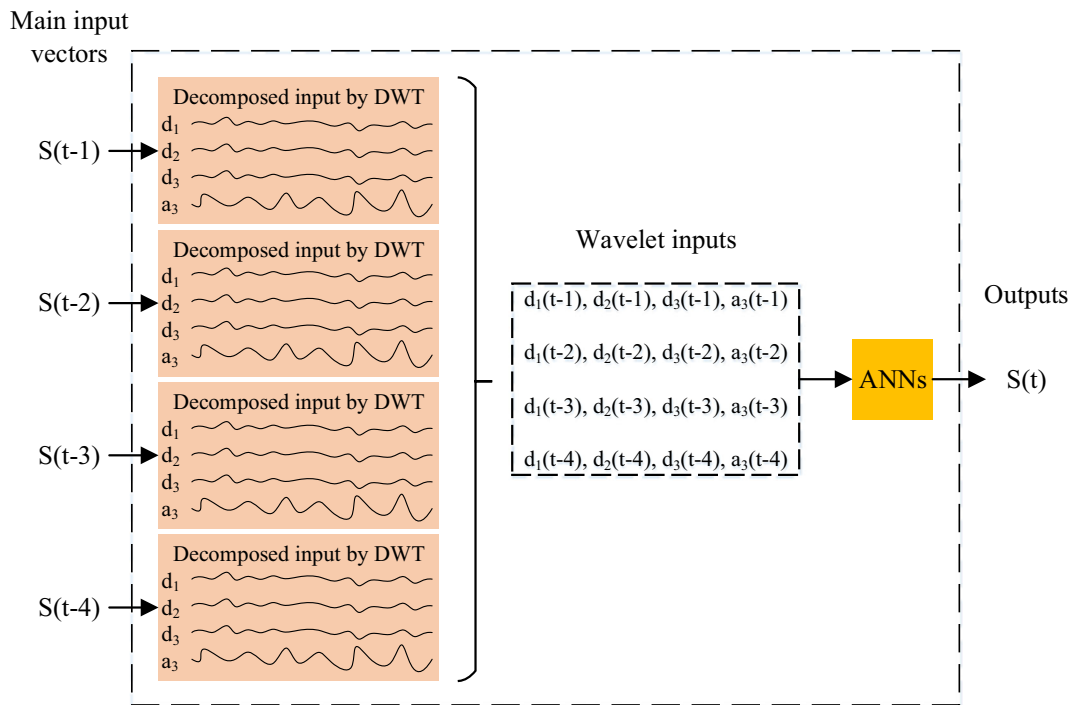


Fig. 7 Flow chart of W-ANNs

about 10.3% and 22.5%, respectively whereas RMSE and MAE values decreased to 75.3% and 73.2% respectively. With respect to R and NSE values, the efficiency of W-FFBP-NN model improved by the db10 and coif5 wavelet functions was similar to the W-FFBP-NN developed using dmey wavelet function, while R and NSE values for all functions have increased about 10.3% and 22.5%, respectively. In accordance with MAE value, the use of Haar and db3 wavelet functions has gone through similar trend. Applying Haar wavelet function had the highest level of error for the sewage sludge quantity prediction (RMSE = 22.96 and MAE = 15.48) in comparison with those obtained using other wavelet functions. For the Haar-FFBP-NN model, the increase in the R and NSE values reached to 5.8% and 11.8% respectively in comparison with the simple FFBP-NN model. What is more, the decrease in the RMSE and MAE value were roughly 26.9% and 27.6%, respectively.

Figure 9a–d presented the percentage of changes in the R, RMSE, MAE, and NSE values for the W-RBF-NN models in comparison with the simple RBF-NN model, respectively. For an improved W-RBF-NN by dmey function, R and NSE values have increased to roughly 10.1% and 21.7%, respectively whereas decrease in the RMSE and MAE values were about

69.7% and 67.4%, respectively. The efficiency level of W-RBF-NN model developed by Haar function was almost similar to the bior1.1 function. Compared with the simple RBF-NN model, an increase in the R and NSE values obtained by the Haar-RBF-NN model was roughly 6.4% and 13.3%, respectively whereas decrease in the RMSE and MAE value was about 34.7% and 38%, respectively. From Tables 4 and 5, values of statistical parameters confirmed that an increase of the number of sub-class order in each wavelet family led to an increase in precise level of W-NN models. In this study, application of dmey and haar wavelet functions in structure of ANN models have provided the best and worst accuracy level, respectively. Figures 10 and 11 showed qualitative performance of hybrid W-FFBP-NN models for training and testing stages, respectively. Also, for the best mother wavelet function, Figs. 12 and 13 illustrated performance of W-RBF-NN models for training and testing phases, respectively.

Results of DR analysis

In this study, the crucial aim of using statistical parameters is to select the AI approach which has the most

Table 3 W-FFBP-NN models results with wavelet transformation

| Mother wavelet | Sub class | R | | RMSE (m ³ /day) | | MAE (m ³ /day) | | NSE | |
|----------------|-------------|----------|---------|----------------------------|---------|---------------------------|---------|----------|---------|
| | | Training | Testing | Training | Testing | Training | Testing | Training | Testing |
| Haar | haar | 0.95 | 0.95 | 22.96 | 22.96 | 15.65 | 15.48 | 0.89 | 0.90 |
| Daubechies | db2 | 0.97 | 0.97 | 15.79 | 15.74 | 11.13 | 11.63 | 0.95 | 0.94 |
| | db3 | 0.97 | 0.96 | 20.96 | 20.52 | 16.42 | 15.54 | 0.91 | 0.91 |
| | db4 | 0.98 | 0.98 | 13.21 | 13.19 | 9.03 | 9.21 | 0.96 | 0.96 |
| | db5 | 0.98 | 0.98 | 12.28 | 11.87 | 8.45 | 8.49 | 0.97 | 0.97 |
| | db6 | 0.98 | 0.98 | 10.90 | 10.75 | 7.50 | 7.47 | 0.97 | 0.97 |
| | db7 | 0.98 | 0.98 | 11.07 | 10.94 | 7.73 | 7.71 | 0.97 | 0.97 |
| | db8 | 0.98 | 0.98 | 11.49 | 10.56 | 7.72 | 7.43 | 0.97 | 0.97 |
| | db9 | 0.98 | 0.99 | 10.58 | 10.29 | 7.18 | 7.20 | 0.97 | 0.98 |
| | db10* | 0.99* | 0.99* | 9.52* | 9.51* | 6.75* | 6.80* | 0.98* | 0.98* |
| | Coiflets | coif1 | 0.97 | 0.96 | 17.36 | 17.26 | 11.59 | 11.41 | 0.94 |
| coif2 | | 0.98 | 0.98 | 14.53 | 14.46 | 11.10 | 11.18 | 0.95 | 0.96 |
| coif3 | | 0.98 | 0.98 | 11.2 | 11.13 | 7.75 | 7.97 | 0.97 | 0.97 |
| coif4 | | 0.98 | 0.98 | 11.02 | 10.61 | 7.8 | 7.6 | 0.97 | 0.97 |
| coif5* | | 0.99* | 0.99* | 9.49* | 9.42* | 6.63* | 6.64* | 0.98* | 0.98* |
| Symlets | sym2 | 0.97 | 0.96 | 17.78 | 17.57 | 12.58 | 11.92 | 0.93 | 0.93 |
| | sym3 | 0.97 | 0.97 | 14.94 | 14.92 | 10.19 | 10.07 | 0.95 | 0.95 |
| | sym4 | 0.98 | 0.98 | 12.99 | 12.68 | 9.27 | 8.92 | 0.96 | 0.96 |
| | sym5 | 0.98 | 0.98 | 12.11 | 11.71 | 8.29 | 8.29 | 0.97 | 0.97 |
| | sym6 | 0.98 | 0.98 | 11.83 | 11.67 | 8.4 | 8.19 | 0.97 | 0.97 |
| | sym7 | 0.98 | 0.98 | 10.83 | 10.71 | 7.67 | 7.42 | 0.97 | 0.97 |
| | sym8* | 0.98* | 0.98* | 10.68* | 10.16* | 7.55* | 7.20 | 0.97* | 0.97* |
| | BiorSplines | bior1.1 | 0.95 | 0.95 | 21.55 | 20.9 | 13.79 | 13.76 | 0.9 |
| bior1.5 | 0.96 | 0.96 | 18.35 | 16.66 | 11.53 | 11.35 | 0.93 | 0.93 | |
| bior2.4 | 0.98 | 0.98 | 13.58 | 12.79 | 9.40 | 9.25 | 0.96 | 0.96 | |
| bior2.8 | 0.98 | 0.98 | 12.32 | 12.21 | 8.67 | 8.41 | 0.97 | 0.97 | |
| bior3.3 | 0.97 | 0.97 | 16.40 | 16.15 | 10.54 | 10.43 | 0.94 | 0.94 | |
| bior3.7 | 0.98 | 0.98 | 12.75 | 12.52 | 8.70 | 8.86 | 0.96 | 0.96 | |
| bior4.4 | 0.98 | 0.98 | 12.84 | 12.85 | 8.98 | 8.98 | 0.96 | 0.96 | |
| bior6.8* | 0.98* | 0.98* | 10.6* | 10.32* | 7.45* | 7.35* | 0.97* | 0.97* | |
| DMeyer | dmey* | 0.99* | 0.99* | 7.76* | 7.76* | 5.73* | 5.79* | 0.98* | 0.98* |

*The proposed wavelet-conjunction models with the best performance

accurate performance. In other words, these error indices may have no capability of presenting valuable information about error distribution. In this way, performance of the W-ANNs was evaluated using discrepancy ratio (DR) to present more robust mother wavelet function. Fundamentally, mathematical shape of DR was expressed as (White et al. 1973),

$$DR = \log \left(\frac{Q(t)_{(Predicted)}}{Q(t)_{(Observed)}} \right) \quad (10)$$

DR values can be evaluated in three different ways. When DR is equal to 1, the proposed model has the best performance can be obtained. Furthermore, for $DR > 1$, the proposed technique had over prediction for S(t)

Table 4 W-RBF-NN models results with wavelet transformation

| Mother wavelet | Sub class | R | | RMSE (m ³ /day) | | MAE (m ³ /day) | | NSE | |
|----------------|-------------|----------|---------|----------------------------|---------|---------------------------|---------|----------|---------|
| | | Training | Testing | Training | Testing | Training | Testing | Training | Testing |
| Haar | haar | 0.95 | 0.95 | 20.82 | 20.91 | 13.02 | 12.8 | 0.91 | 0.91 |
| Daubechies | db2 | 0.97 | 0.96 | 17.14 | 16.37 | 11.86 | 11.70 | 0.94 | 0.94 |
| | db3 | 0.97 | 0.97 | 15.20 | 14.63 | 10.11 | 9.82 | 0.95 | 0.95 |
| | db4 | 0.98 | 0.98 | 13.7 | 13.71 | 9.54 | 9.28 | 0.96 | 0.96 |
| | db5 | 0.98 | 0.98 | 12.61 | 12.40 | 8.89 | 8.9 | 0.96 | 0.96 |
| | db6 | 0.98 | 0.98 | 12.22 | 11.34 | 8.5 | 8.27 | 0.97 | 0.97 |
| | db7 | 0.98 | 0.98 | 11.76 | 11.62 | 8.3 | 7.86 | 0.97 | 0.97 |
| | db8 | 0.98 | 0.98 | 12.65 | 11.75 | 8.85 | 8.17 | 0.96 | 0.97 |
| | db9 | 0.98 | 0.98 | 11.43 | 11.33 | 7.84 | 7.68 | 0.97 | 0.97 |
| | db10* | 0.98* | 0.98* | 11.33* | 10.45* | 8.07* | 7.64* | 0.97* | 0.97* |
| | Coiflets | coif1 | 0.97 | 0.97 | 16.44 | 14.86 | 11.29 | 10.02 | 0.94 |
| coif2 | | 0.98 | 0.98 | 13.18 | 12.87 | 9.14 | 8.94 | 0.96 | 0.96 |
| coif3 | | 0.98 | 0.98 | 12.37 | 11.17 | 8.5 | 7.8 | 0.96 | 0.97 |
| coif4 | | 0.98 | 0.98 | 11.42 | 10.11 | 8.07 | 7.37 | 0.97 | 0.97 |
| coif5* | | 0.98* | 0.98* | 10.52* | 10.17* | 7.64* | 7.17* | 0.97* | 0.97* |
| Symlets | sym2 | 0.96 | 0.97 | 17.65 | 16.43 | 12.18 | 11.48 | 0.93 | 0.94 |
| | sym3 | 0.97 | 0.97 | 15.48 | 14.33 | 10.46 | 9.83 | 0.95 | 0.95 |
| | sym4 | 0.98 | 0.98 | 12.9 | 12.84 | 8.79 | 8.67 | 0.96 | 0.96 |
| | sym5 | 0.98 | 0.98 | 13.29 | 13.22 | 9.36 | 9.15 | 0.96 | 0.96 |
| | sym6 | 0.98 | 0.98 | 12.05 | 11.83 | 8.39 | 8.19 | 0.97 | 0.97 |
| | sym7 | 0.98 | 0.98 | 12.48 | 12.38 | 8.68 | 8.03 | 0.96 | 0.96 |
| | sym8* | 0.98* | 0.98* | 11.68* | 11.14* | 8.24* | 7.77* | 0.97* | 0.97* |
| | BiorSplines | bior1.1 | 0.95 | 0.96 | 22.12 | 17.9 | 13.8 | 11.92 | 0.9 |
| | bior1.5 | 0.97 | 0.97 | 17.29 | 16.61 | 11.28 | 10.97 | 0.94 | 0.94 |
| | bior2.4 | 0.98 | 0.98 | 14.09 | 13.47 | 9.8 | 9.55 | 0.96 | 0.96 |
| | bior2.8 | 0.98 | 0.98 | 12.71 | 12.45 | 8.84 | 8.44 | 0.96 | 0.96 |
| | bior3.3 | 0.97 | 0.97 | 17.16 | 16.26 | 11.18 | 10.71 | 0.94 | 0.94 |
| | bior3.7 | 0.98 | 0.98 | 13.27 | 12.82 | 9.19 | 9.02 | 0.96 | 0.96 |
| | bior4.4 | 0.97 | 0.98 | 14.19 | 13.45 | 9.99 | 9.32 | 0.95 | 0.96 |
| | bior6.8* | 0.98* | 0.98* | 11.21* | 11.04* | 7.92* | 7.86* | 0.97* | 0.97* |
| DMeyer | dmey* | 0.99* | 0.99* | 9.56* | 9.35* | 6.81* | 6.69* | 0.98* | 0.98* |

*The proposed wavelet-conjunction models with the best performance

Table 5 Results of DR analysis for W-FFBP-NN model

| Mother wavelet function | Sub-class | Maximum DR | Minimum DR | Average DR |
|-------------------------|-----------|------------|------------|------------|
| Haar | haar | 0.757 | -0.154 | 0.016 |
| Daubechies | db10 | 0.083 | -0.116 | -0.0004 |
| Coiflets | coif5 | 0.600 | -0.066 | 0.001 |
| Symlets | sym8 | 0.088 | -0.086 | -0.00001 |
| BiorSplines | bior6.8 | 0.151 | -0.054 | 0.002 |
| DMeyer | dmey | 0.057 | -0.056 | 0.002 |

values; otherwise, for DR < 1, possibility of under prediction state is inevitable.

In this study, DR analysis is conducted using testing dataset. Minimum, maximum, and average of DR values for the testing stage of the W-ANN models have been computed in Tables 6 and 7. From Table 6, an improved FFBP-NN model by dmey function had the minimum range of DR values [-0.056, 0.057] and average of 0.002,

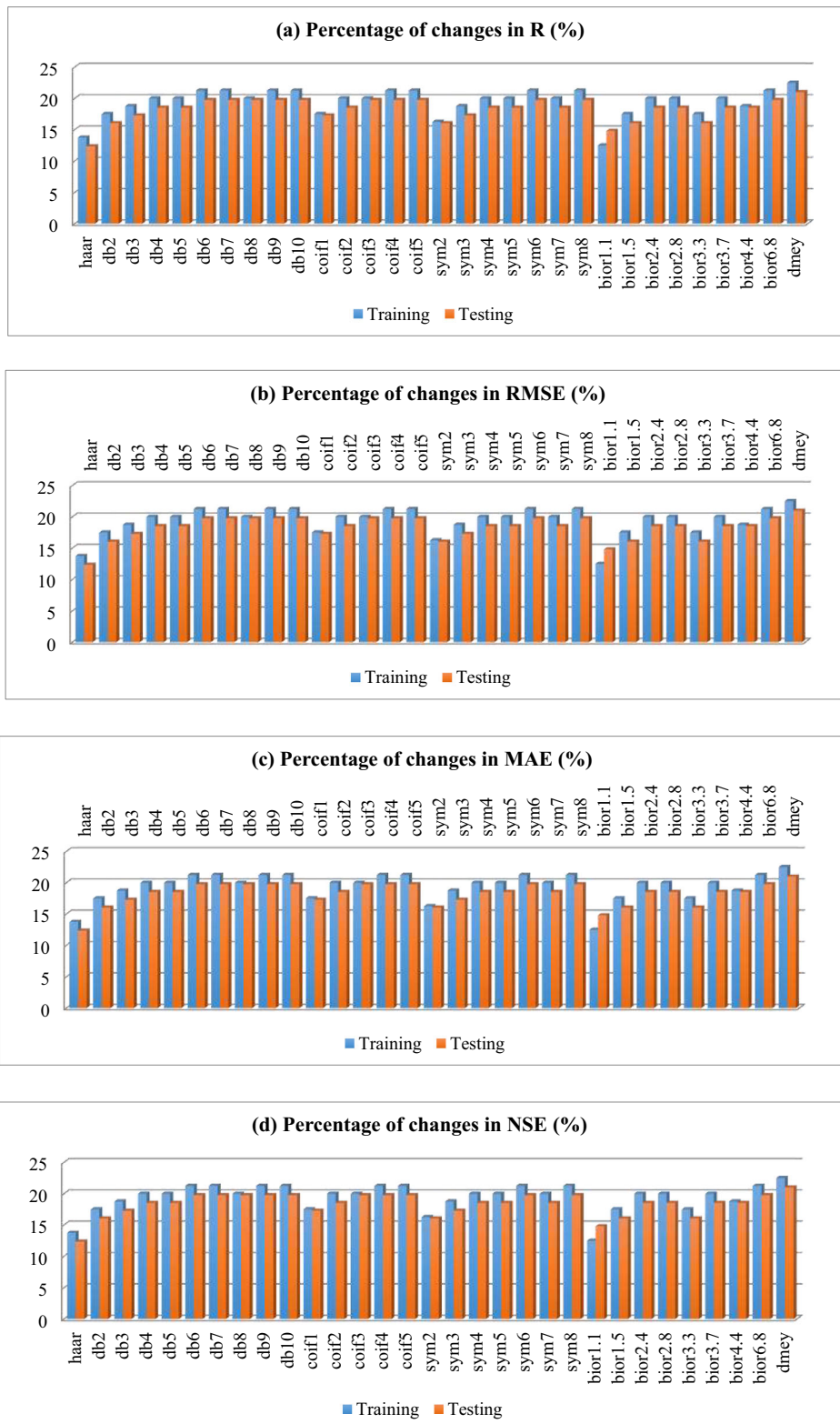


Fig. 8 Percentage of changes in (a) R, (b) RMSE, (c) MAE, and (d) NSE values obtained by W-FFBP-NN



Fig. 9 Percentage of changes in (a) R, (b) RMSE, (c) MAE, and (d) NSE values obtained by W-RBF-NN

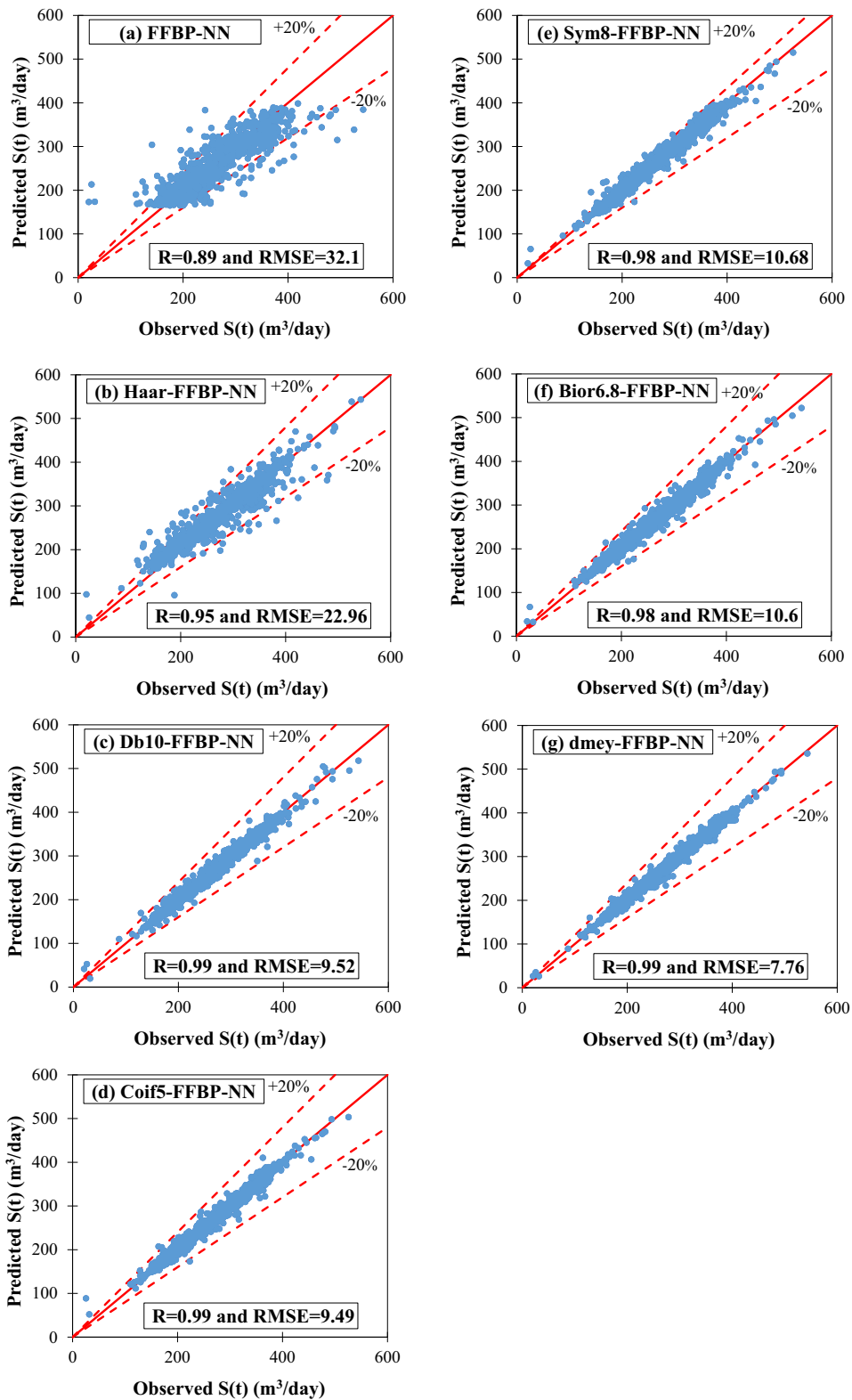


Fig. 10 Scatter plots between observed and predicted sewage sludge quantity for training stages using (a) FFBP-NN; (b) haar-FFBP-NN; (c) db10-FFBP-NN; (d) coif5-FFBP-NN; (e) sym8-FFBP-NN; (f) bior6.8-FFBP-NN; and (g) dmey-FFBP-NN

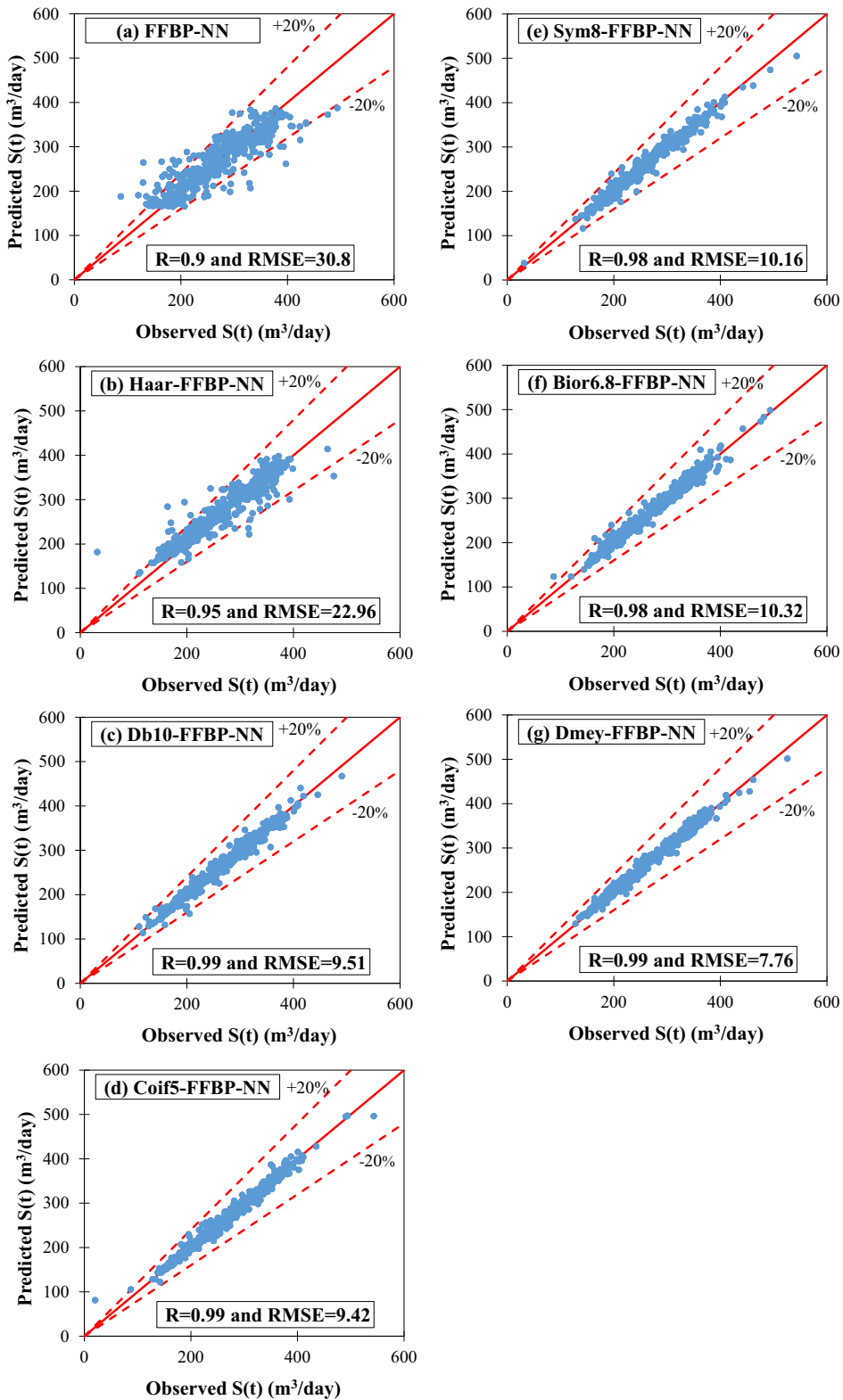


Fig. 11 Scatter plots between observed and predicted sewage sludge quantity for testing stages using (a) FFBP-NN; (b) haar-FFBP-NN; (c) db10-FFBP-NN; (d) coif5-FFBP-NN; (e) sym8-FFBP-NN; (f) bior6.8-FFBP-NN; and (g) dmey-FFBP-NN

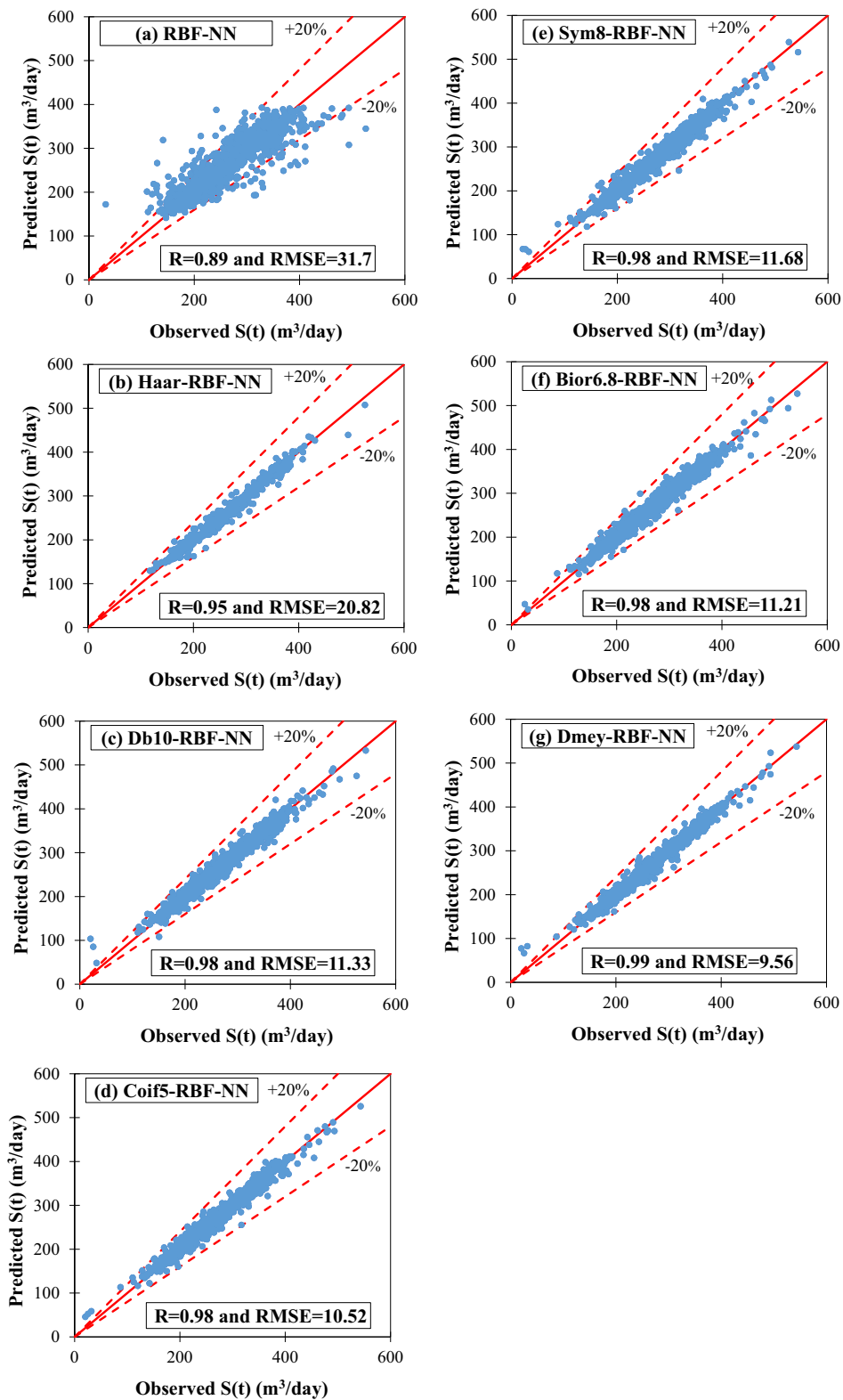


Fig. 12 Scatter plots between observed and predicted sewage sludge quantity for training stages using (a) RBF-NN; (b) haar- RBF-NN; (c) db10- RBF-NN; (d) coif5- RBF-NN; (e) sym8-RBF-NN; (f) bior6.8-RBF-NN; and (g) dmey-RBF-NN

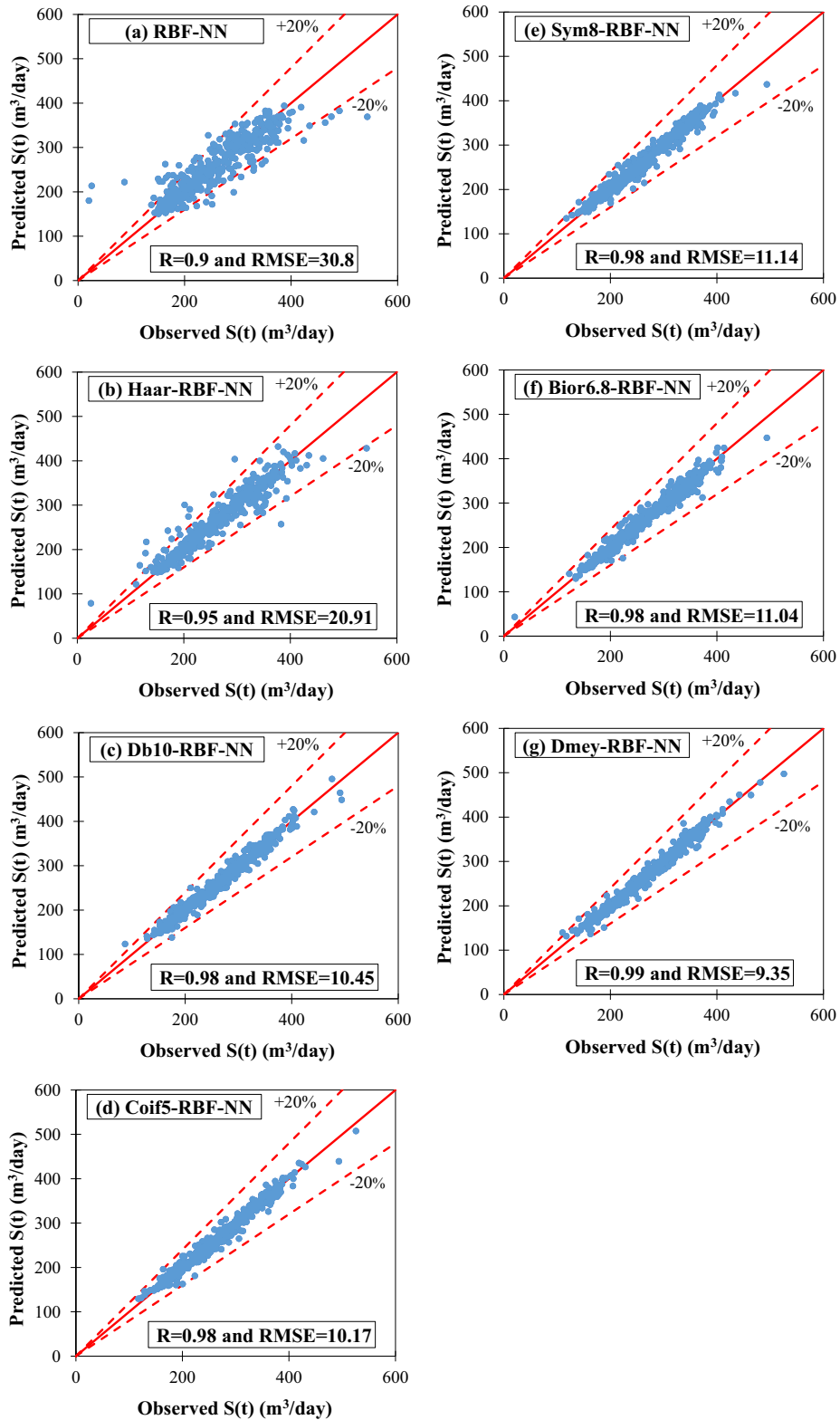


Fig. 13 Scatter plots between observed and predicted sewage sludge quantity for testing stages using (a) RBF-NN; (b) haar- RBF-NN; (c) db10- RBF-NN; (d) coif5- RBF-NN; (e) sym8-RBF-NN; (f) bior6.8-RBF-NN; and (g) dmey-RBF-NN

Table 6 Results of DR analysis for W-RBF-NN model

| Mother wavelet function | Sub-class | Maximum DR | Minimum DR | Average DR |
|-------------------------|-----------|------------|------------|------------|
| Haar | haar | 0.488 | -0.172 | 0.004 |
| Daubechies | db10 | 0.151 | -0.101 | 0.0004 |
| Coiflets | coif5 | 0.078 | -0.091 | -0.0004 |
| Symlets | sym8 | 0.085 | -0.088 | 0.001 |
| BiorSplines | bior6.8 | 0.328 | -0.103 | -0.0009 |
| DMeyer | dmey | 0.103 | -0.094 | -0.0002 |

resulting the best performance in comparison with applying other mother wavelet functions. Moreover, mother wavelet of sym8 has stood at the second rank in terms of accuracy for prediction of sewage sludge. Average of DR value obtained by haar-FFBP-NN model (average DR = 0.016) showed the largest computational error than other W-FFBP-NN models. Variations of DR values versus number of data points for the W-FFBP-NN models were illustrated in Fig. 14. In Table 7, dmey-RBF-NN model with average DR of -0.0002 had the best performance in prediction of sewage sludge. Furthermore, haar function with DR between -0.17 and 0.48 produced the lowest accuracy for estimation of sewage sludge in comparison with W-RBF-NN models. Variations of DR values with number of data samples for the W-RBF-NN approaches have been shown in Fig. 15.

Conclusions

In this research, the effect of various mother wavelet functions on the efficiency of W-NN models was investigated to predict sewage sludge quantity for Kerman WWTP, Iran. In this way, two common types of ANNs including RBF-NN and FFBP-NN models have been used to assess quantity of

sewage sludge. To obtain more accurate performance of ANNs, six wavelet families including Haar, Daubechies, Coiflets, Symlets, BiorSplines, and DMeyer were employed to reduce irregularity of non-periodical time series of sewage sludge. With respect to the mother wavelet functions of sub-class order, 32 wavelet functions were used as preprocessors of input variables for the FFBP-NN and RBF-NN models. Statistical parameters obtained through training stage indicated that FFBP-NN has provided more precise prediction than RBF-NN model. Moreover, 32 W-FFBP-NN approaches had superiority to simple FFBP-NN model. Similarly, 32 W-RBF-NN models decreased remarkably computational error of sewage sludge quantity in comparison with the conventional RBF-NN network. Results of testing stages had the same trend as the training phases. In training and testing stages, application of dmey and haar functions had the best and worst performance among other wavelet functions, respectively.

From a logical point of view, performance of proposed W-ANN models showed that precision level was inextricably bound up with type of mother wavelets. Finally, results of the proposed hybrid models indicated that the conjunction of wavelet decomposition with ANNs could be considered as a robust tool for sewage sludge quantity prediction. In the current investigation, one of the main shortcomings is availability of dataset for a

Table 7 Results of DR analysis for W-RBF-NN model

| Mother wavelet function | Sub-class | Maximum DR | Minimum DR | Average DR |
|-------------------------|-----------|------------|------------|------------|
| Haar | haar | 0.488 | -0.172 | 0.004 |
| Daubechies | db10 | 0.151 | -0.101 | 0.0004 |
| Coiflets | coif5 | 0.078 | -0.091 | -0.0004 |
| Symlets | sym8 | 0.085 | -0.088 | 0.001 |
| BiorSplines | bior6.8 | 0.328 | -0.103 | -0.0009 |
| DMeyer | dmey | 0.103 | -0.094 | -0.0002 |

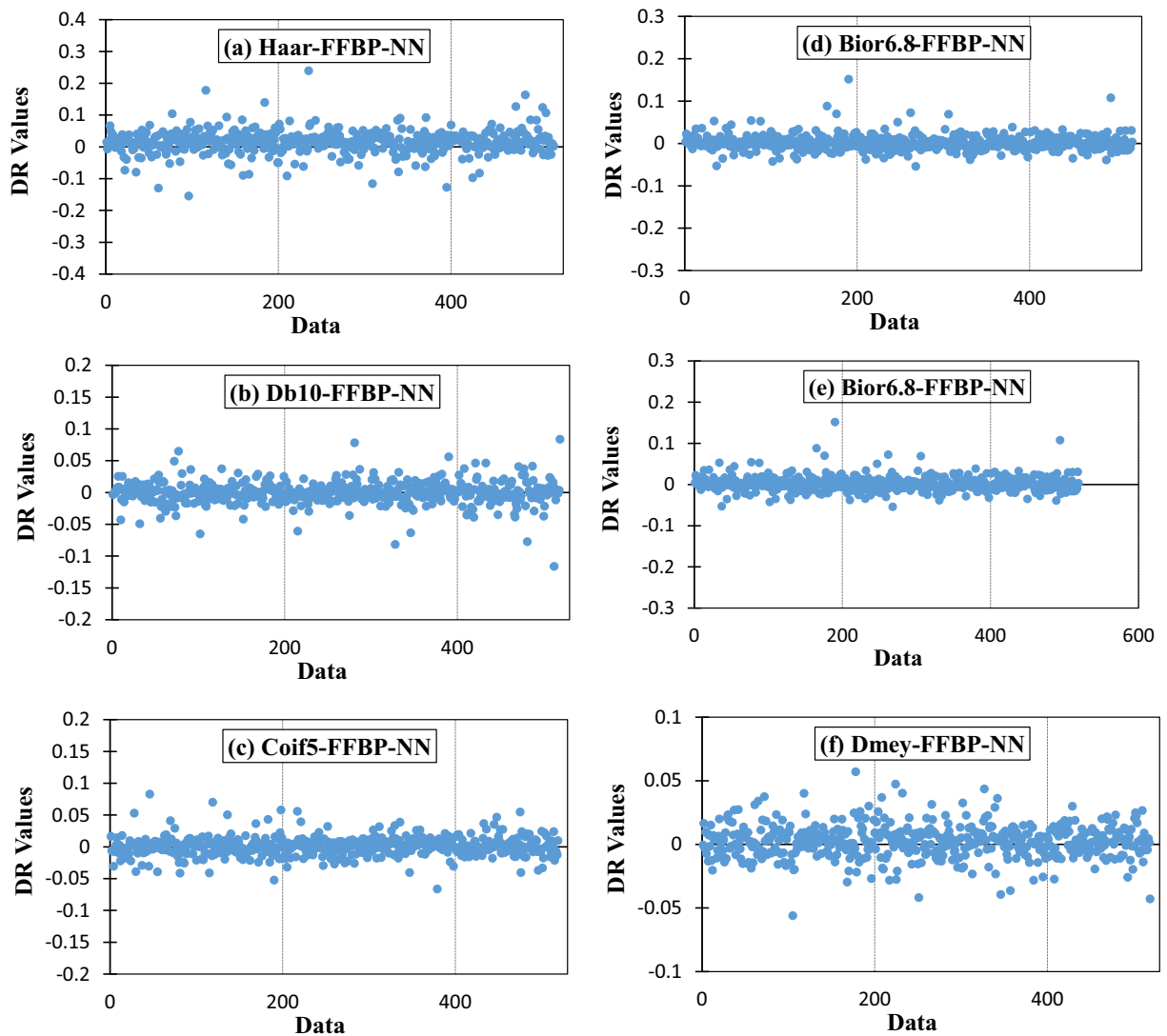


Fig. 14 Variations of the *DR* values versus time for the W-FFBP-NN model using various mother wavelet functions (a) haar; (b) db10; (c) coif5; (d) sym8; (e) bior6.8; and (f) dmey

limited time period. There is no denying the fact that applying signals of sewage sludge with a longer time period can provide a reliable prediction. Kerman WWPT was newly constructed and consequently, time duration of daily measured sewage sludge is intrinsically limited. As mentioned in this study, LM technique, as an iterative algorithm, was used to train FFBP-NN and W-FFBP-NN models. In order to improve accuracy level of W-ANN techniques, there is a possibility to employ evolutionary algorithms (EAs) such as

particle swarm optimization (PSO), genetic algorithm (GA), and gravitational search algorithm (GSA) in order to compute weighting coefficients and biases related to general structure of ANNs. In fact, statistical results of W-ANNs based EAs can be compared with those obtained by the W-ANNs. To conduct more in-depth research in future, MARS and RF, as powerful data-mining approaches, can be applied instead of ANNs. Furthermore, results of W-MARS and W-RF models will be comparable with the current study.

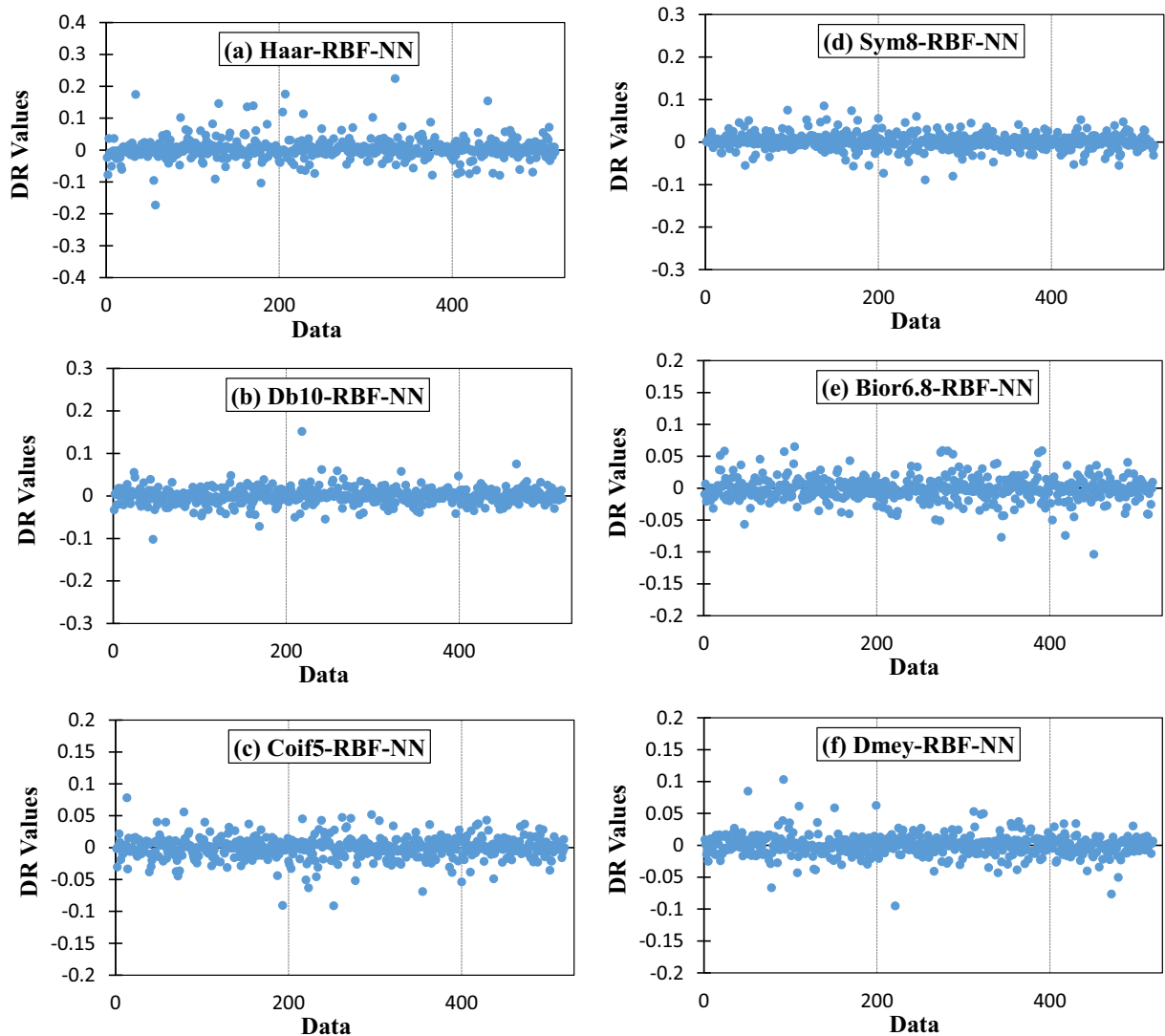


Fig. 15 Variations of the *DR* values versus time for the W-RBF-NN model using various mother wavelet functions (a) haar; (b) db10; (c) coif5; (d) sym8; (e) bior6.8; and (f) dmey

Acknowledgements We would like to thank the manager of Kerman province wastewater treatment plant for providing us with the field data used in this study.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Adamowski, J., & Sun, K. (2010). Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. *Journal of Hydrology*, *390*, 85–91.
- Altunkaynak, A., & Nigussie, T. A. (2015). Prediction of daily rainfall by a hybrid wavelet-season-neuro technique. *Journal of Hydrology*, *529*, 287–301. <https://doi.org/10.1016/j.jhydrol.2015.07.046>.
- Aussem, A., Campbell, J., & Murtagh, F. (1998). Wavelet-based feature extraction and decomposition strategies for financial forecasting. *Journal of Computational Intelligence in Finance*, *6*, 5–12.

- Bateni, S. M., Jeng, D.-S., & Melville, B. W. (2007). Bayesian neural networks for prediction of equilibrium and time-dependent scour depth around bridge piers. *Advances in Engineering Software*, 38(2), 102–111. <https://doi.org/10.1016/j.advengsoft.2006.08.004>.
- Belanche, L. s. A., Valdés, J. J., Comas, J., Roda, I. R., & Poch, M. (1999). Towards a model of input-output behaviour of wastewater treatment plants using soft computing techniques. *Environmental Modelling & Software*, 14(5), 409–419. [https://doi.org/10.1016/S1364-8152\(98\)00102-9](https://doi.org/10.1016/S1364-8152(98)00102-9).
- Boger, Z., & Guterman, H. (1990). Knowledge representation of wastewater treatment plant operating history and current state. Proceeding of the 5th International Association on Water Pollution Research and Control Workshop, Yokohama and Kyoto, Japan. Pergamon Press.
- Boniecki, P., Dach, J., Pilarski, K., & Piekarska-Boniecka, H. (2012). Artificial neural networks for modeling ammonia emissions released from sewage sludge composting. *Atmospheric Environment*, 57, 49–54.
- Catalão, J., Pousinho, H. M. I., & Mendes, V. M. F. (2011). Short-term wind power forecasting in Portugal by neural networks and wavelet transform. *Renewable Energy*, 36(4), 1245–1251.
- Çinar, Ö. (2005). New tool for evaluation of performance of wastewater treatment plant: artificial neural network. *Process Biochemistry*, 40(9), 2980–2984. <https://doi.org/10.1016/j.procbio.2005.01.012>.
- Cohen, A., Janssen, G., Brewster, S. D., Seeley, R., Boogert, A. A., Graham, A. A., Mardani, M. R., Clarke, N., & Kasabov, N. K. (1997). Application of computational intelligence for on-line control of a sequencing batch reactor (SBR) at Morrinsville sewage treatment plant. *Water Science and Technology*, 35(10), 63–71.
- Cong, Q., & Yu, W. (2018). Integrated soft sensor with wavelet neural network and adaptive weighted fusion for water quality estimation in wastewater treatment process. *Measurement*, 124, 436–446.
- Côté, M., Grandjean, B. P. A., Lessard, P., & Thibault, J. (1995). Dynamic modelling of the activated sludge process: improving prediction using neural networks. *Water Research*, 29(4), 995–1004. [https://doi.org/10.1016/0043-1354\(95\)93250-W](https://doi.org/10.1016/0043-1354(95)93250-W).
- Danish, M., Jing, H., Pin, Z., Ziyang, L., & Pansheng, Q. (2016). A new drying kinetic model for sewage sludge drying in presence of CaO and NaClO. *Applied Thermal Engineering*, 106, 141–152. <https://doi.org/10.1016/j.applthermaleng.2016.05.191>.
- El-Din, A. G., & Smith, D. W. (2002). A neural network model to predict the wastewater inflow incorporating rainfall events. *Water Research*, 36, 1115–1126.
- Fernandez de Canete, J., Del Saz-Orozco, P., Baratti, R., Mulas, M., Ruano, A., & Garcia-Cerezo, A. (2016). Soft-sensing estimation of plant effluent concentrations in a biological wastewater treatment plant using an optimal neural network. *Expert Systems with Applications*, 63, 8–19. <https://doi.org/10.1016/j.eswa.2016.06.028>.
- Fernandez, F. J., Seco, J. A., Ferrer, J., & Rodrigo, M. A. (2009). Use of neurofuzzy networks to improve wastewater flow-rate forecasting. *Environmental Modelling & Software*, 24, 686–693.
- Filipović, J., Grčić, I., Beranec, V., & Kniewald, G. (2013). Monitoring of total metal concentration in sludge samples: case study for the mechanical–biological wastewater treatment plant in Velika Gorica, Croatia. *Science of the Total Environment*, 447, 17–24. <https://doi.org/10.1016/j.scitotenv.2012.12.078>.
- Forster, C. F., & Foundation, W. E. R. (2003). *Wastewater treatment and technology*. Thomas Telford.
- Gontarski, C. A., Rodrigues, P. R., Mori, M., & Prenem, L. F. (2000). Simulation of an industrial wastewater treatment plant using artificial neural networks. *Computers & Chemical Engineering*, 24, 1719–1723.
- Guo, H., Jeong, K., Lim, J., Jo, J., Kim, Y. M., Park, J.-p., Kim, J. H., & Cho, K. H. (2015). Prediction of effluent concentration in a wastewater treatment plant using machine learning models. *Journal of Environmental Sciences*, 32, 90–101. <https://doi.org/10.1016/j.jes.2015.01.007>.
- Hamed, M. M., Khalafallah, M. G., & Hassanien, E. A. (2004). Prediction of wastewater treatment plant performance using artificial neural networks. *Environmental Modelling & Software*, 19(10), 919–928. <https://doi.org/10.1016/j.envsoft.2003.10.005>.
- Hamoda, M. F., Al-Ghusain, I. A., & Hassan, A. H. (1999). Integrated wastewater treatment plant performance evaluation using artificial neural networks. *Water Science and Technology*, 40(7), 55–65. [https://doi.org/10.1016/S0273-1223\(99\)00584-3](https://doi.org/10.1016/S0273-1223(99)00584-3).
- Han, H.-G., Li, Y., Guo, Y.-N., & Qiao, J.-F. (2016). A soft computing method to predict sludge volume index based on a recurrent self-organizing neural network. *Applied Soft Computing*, 38, 477–486. <https://doi.org/10.1016/j.asoc.2015.09.051>.
- Han, H., Zhu, S., Qiao, J., & Guo, M. (2018). Data-driven intelligent monitoring system for key variables in wastewater treatment process. *Chinese Journal of Chemical Engineering*, 26(10), 2093–2101.
- Hanbay, D., Turkoglu, I., & Demir, Y. (2008). Prediction of wastewater treatment plant performance based on wavelet packet decomposition and neural networks. *Expert Systems with Applications*, 34(2), 1038–1043. <https://doi.org/10.1016/j.eswa.2006.10.030>.
- Heidari, E., Sobati, M. A., & Movahedirad, S. (2016). Accurate prediction of nanofluid viscosity using a multilayer perceptron artificial neural network (MLP-ANN). *Chemometrics and Intelligent Laboratory Systems*, 155, 73–85. <https://doi.org/10.1016/j.chemolab.2016.03.031>.
- Honggui, H., Ying, L., & Junfei, Q. (2014). A fuzzy neural network approach for online fault detection in waste water treatment process. *Computers and Electrical Engineering*, 40(7), 2216–2226.
- Hu, J., & Wang, J. (2015). Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression. *Energy*, 93, 1456–1466. <https://doi.org/10.1016/j.energy.2015.10.041>.
- Huang, Y. W., & Chen, M. Q. (2015). Artificial neural network modeling of thin layer drying behavior of municipal sewage sludge. *Measurement*, 73, 640–648.
- Jalili Ghazizade, M., & Noor, A. R. (2008). Prediction of municipal solid waste generation by use of artificial neural network: a case study of Mashhad. *International Journal of Environmental Research*, 2(1), 13–22.

- Joseph, S. M., & Anto, B. P. (2012). Speech coding based on orthogonal and biorthogonal wavelet. *Procedia Technology*, 6, 397–404.
- Kaira, G. L., & Christian, R. A. (2013). Wastewater treatment: concepts and design approach: PHI learning.
- Kasiviswanathan, K. S., He, J., Sudheer, K. P., & Tay, J.-H. (2016). Potential application of wavelet neural network ensemble to forecast streamflow for flood management. *Journal of Hydrology*, 536, 161–173. <https://doi.org/10.1016/j.jhydrol.2016.02.044>.
- Kisi, O., & Shiri, J. (2011). Precipitation forecasting using wavelet-genetic programming and wavelet-neuro-fuzzy conjunction models. *Water Resources Management*, 25(13), 3135–3152.
- Kołecka, K., Gajewska, M., Obarska-Pempkowiak, H., & Rohde, D. (2017). Integrated dewatering and stabilization system as an environmentally friendly technology in sewage sludge management in Poland. *Ecological Engineering*, 98, 346–353. <https://doi.org/10.1016/j.ecoleng.2016.08.011>.
- Li, S., Li, Y., Lu, Q., Zhu, J., Yao, Y., & Bao, S. (2014). Integrated drying and incineration of wet sewage sludge in combined bubbling and circulating fluidized bed units. *Waste Management*, 34(12), 2561–2566. <https://doi.org/10.1016/j.wasman.2014.08.018>.
- Llorens, E., Thiery, F., Grieu, S., & Polit, M. (2008). Evaluation of WWTP discharges into a Mediterranean river using KSOM neural networks and mass balance modelling. *Chemical Engineering Journal*, 142, 135–146.
- Maheswaran, R., & Khosa, R. (2012). Comparative study of different wavelets for hydrologic forecasting. *Computers & Geosciences*, 46, 284–295. <https://doi.org/10.1016/j.cageo.2011.12.015>.
- Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674–693. <https://doi.org/10.1109/34.192463>.
- Mirabi, M., Mianabadi, H., Zarghami, M., Sharifi, M. B., & Mostert, E. (2014). Risk-based evaluation of wastewater treatment projects: a case study in Niasar city, Iran. *Resources, Conservation and Recycling*, 93, 168–177. <https://doi.org/10.1016/j.resconrec.2014.10.002>.
- Mjalli, F. S., Al-Asheh, S., & Alfadala, H. E. (2007). Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance. *Journal of Environmental Management*, 83, 329–338.
- Moral, H., Aksoy, A., & Gokcay, C. F. (2008). Modeling of the activated sludge process by using artificial neural networks with automated architecture screening. *Computers & Chemical Engineering*, 32(10), 2471–2478. <https://doi.org/10.1016/j.compchemeng.2008.01.008>.
- Najafzadeh, M., & Zeinolabedini, M. (2018). Derivation of optimal equations for prediction of sewage sludge quantity using wavelet conjunction models: an environmental assessment. *Environmental Science and Pollution Research*, 25, 22931–22943.
- Nasr, M. S., Moustafa, M. A. E., Seif, H. A. E., & Kobrosy, G. E. (2012). Application of artificial neural network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT. *Alexandria Engineering Journal*, 51, 37–43.
- Noori, R., Abdoli, M. A., Ameri Ghasrodashti, A., & Jalili Ghazizade, M. (2009a). Prediction of municipal solid waste generation with combination of support vector machine and principal component analysis: a case study of Mashhad. *Environmental Progress & Sustainable Energy: An Official Publication of the American Institute of Chemical Engineers*, 28(2), 249–258.
- Noori, R., Abdoli, M. A., Farokhnia, A., & Abbasi, M. (2009b). Retracted: results uncertainty of solid waste generation forecasting by hybrid of wavelet transform-ANFIS and wavelet transform-neural network. *Expert Systems with Applications*, 36(6), 9991–9999. <https://doi.org/10.1016/j.eswa.2008.12.035>.
- Noori, R., Karbassi, A., & Sabahi, M. S. (2010). Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste prediction. *Journal of Environmental Management*, 91(3), 767–771.
- Oja, E. (1994). Neural networks—advantages and applications. *Machine Intelligence and Pattern Recognition*, 16, 359–365.
- Ongen, A., Ozcan, H. K., & Arayıc, S. (2013). An evaluation of tannery industry wastewater treatment sludge gasification by artificial neural network modeling. *Journal of Hazardous Materials*, 263, 361–366.
- Onkal-Engin, G., Demir, I., & Engin, S. N. (2005). Determination of the relationship between sewage odour and BOD by neural networks. *Environmental Modelling & Software*, 20, 843–850.
- Pai, T. Y., Yang, P. Y., Wang, S. C., Lo, M. H., Chiang, C. F., Kuo, J. L., Chu, H. H., Su, H. C., Yu, L. F., Hu, H. C., & Chang, Y. H. (2011). Predicting effluent from the wastewater treatment plant of industrial park based on fuzzy network and influent quality. *Applied Mathematical Modelling*, 35(8), 3674–3684. <https://doi.org/10.1016/j.apm.2011.01.019>.
- Partal, T., & Kisi, Ö. (2007). Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. *Journal of Hydrology*, 342, 199–212.
- Rumelhart, D. E., & McClelland. (1986). *Parallel distributed processing: explorations in the microstructure of cognition*. Cambridge: MIT Press.
- Safavi, S., Bateni, S. M., & Xu, T. (2015). Predicting the amount of municipal solid waste via hybrid principal component analysis-artificial neural network approach. *Applied Mechanics and Materials*, 768, 722–727.
- Seo, Y., Kim, S., Kisi, O., & Singh, V. P. (2015). Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. *Journal of Hydrology*, 520, 224–243. <https://doi.org/10.1016/j.jhydrol.2014.11.050>.
- Shiri, J. (2018). Improving the performance of the mass transfer-based reference evapotranspiration estimation approaches through a coupled wavelet-random forest methodology. *Journal of Hydrology*, 561, 737–750.
- Shoab, M., Shamseldin, A., Melville, B., & Khan, M. (2015). Runoff forecasting using hybrid wavelet gene expression programming (WGEP) Approach (Vol. 527).
- Verma, A., Wei, X., & Kusiak, A. (2013). Predicting the total suspended solids in wastewater: a data-mining approach. *Engineering Applications of Artificial Intelligence*, 26, 1366–1372.
- Wan, J., Huang, M., Ma, Y., Guo, W., Wang, Y., Zhang, H., Li, W., & Sun, X. (2011). Prediction of effluent quality of a paper mill wastewater treatment using an adaptive network-based

- fuzzy inference system. *Applied Soft Computing*, 11(3), 3238–3246. <https://doi.org/10.1016/j.asoc.2010.12.026>.
- White, W. R., Crabbe, A. D., Mill, H., & Station, H. R. (1973). *Sediment Transport: An Appraisal of Available Methods*, Hydraulics Research Station.
- Zare-Abyaneh, H. (2014). Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters. *Journal of Environmental Health Science & Engineering*, 12(1), 1–8.
- Zounemat-Kermani, M., Beheshti, A.-A., Ataie-Ashtiani, B., & Sabbagh-Yazdi, S.-R. (2009). Estimation of current-induced scour depth around pile groups using neural network and adaptive neuro-fuzzy inference system. *Applied Soft Computing*, 9(2), 746–755. <https://doi.org/10.1016/j.asoc.2008.09.006>.