Desalination and Water Treatment
www.deswater.com
doi: 10.1080/19443994.2013.878669

53 (2015) 1045–1053
January

Taylor & Francis
Taylor & Francis Group

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

Salim Heddam[a,*], Noureddine Dechemi[b]

[a]*Faculty of Science, Agronomy Department, Hydraulics Division University 20 Août 1955, Route El Hadaik, BP 26, Skikda, Algeria Tel. +213 06 61 74 51 22; email: heddamsalim@yahoo.fr*
[b]*Laboratory Construction et Environnement, Polytechnical National School, 10, avenue Hassen Badi, B.P. 182 El Harrach, Algiers, Algeria*

Received 23 May 2013; Accepted 20 December 2013

ABSTRACT

This study was developed using data from a drinking water treatment plant located at Boudouaou, Algeria, it is located at about 7 km from the Keddara dam which supplies potable water to Algiers capital of Algeria. The treatment consists essentially of preliminary disinfection, coagulation–flocculation, settling, filtration and final disinfection. Traditionally, optimum coagulant dosages are determined using jar tests. However, jar tests are relatively expensive and time consuming. In this study, we present a new Artificial Intelligence Techniques model called dynamic evolving neural-fuzzy inference system (DENFIS) based on an evolving clustering method, for modelling coagulant dosage rate used in the coagulation stage. Six online variables of raw water quality including turbidity, conductivity, temperature, apparent colour, ultraviolet absorbance, water pH and alum dosage were used to build the coagulant dosage model. Two DENFIS-based evolving neural-fuzzy inference system are presented and compared. The two DENFIS systems are: (1) Offline-based system named DENFIS-OF and (2) Online-based system, named DENFIS-ON. The performances of the models are evaluated using root-mean square errors (RMSE), mean absolute error and correlation coefficient (CC) statistics. The low RMSE and high CC values were obtained with DENFIS-ON method.

Keywords: Water treatment plant; Coagulant dosage; DENFIS; Modelling

1. Introduction

Water treatment plants (WTP) are expected to provide safe and aesthetically acceptable water to consumers at a reasonable cost. Conventionally, management and operation of a WTP is based on monitoring finished water quality parameters and then comparing them to the regulatory requirements [1]. Coagulation–flocculation followed by sedimentation and filtration is the most commonly used water treatment process, in which turbidity or particle removal is strongly dependent on proper coagulant

*Corresponding author.

Presented at the International Joint CB-WR-MED Conference / 2nd AOP'Tunisia Conference for Sustainable Water Management, 24–27 April 2013, Tunis, Tunisia

dosage, flocculation mixing time, mixing intensity and effective size of filter media [2]. Coagulation involves three basic steps: coagulant addition and mixing, colloid particle destabilisation and floc formation. Coagulation describes the initial colloid destabilisation, principally by charge neutralisation after adding the coagulant. In practice, there is a little distinction between coagulation and the early phase of flocculation because it occurs very rapidly. Hence, the term either "coagulation" or "flocculation" could be used to describe the overall treatment process [3]. The coagulation effectiveness depends on various factors, including coagulant types and dosage, pH and particle properties [4]. Numerous other operational parameters can affect coagulation including initial coagulant, applied shear stresses by agitation and hydraulic retention time for coagulation and floc formation [5].

One of the factors influencing coagulation behaviour is the coagulant which plays a key role in water treatment process [6]. Commonly used metal coagulants are generally based on aluminium and iron and have been widely used in water treatment since the early twentieth century [7]. The coagulants used in this study were aluminium sulphate also called "alum" ($Al_2SO_4{\cdot}18H_2O$). Improper coagulant and dosages may lead to an inefficient operation and may increase the quantity of chemical sludge [8]. Traditionally, optimum alum doses are determined using jar tests [9,10]. However, jar tests are relatively expensive and take a long time to conduct. The limitations of using jar tests for determining optimum alum doses can be overcome using models [11]. Models that have the ability to capture underlying relationships using examples of the desired input-output mapping are very suitable [12]. Until now, the majority of coagulant dosage models have been statistically based, including polynomial equations, artificial neural networks (ANN) and neuro fuzzy models.

During the last few decades, many authors have carried out modelling work on coagulant dosage concentration. Bazer-Bachi et al. developed two separate models based on polynomial equations used to determine coagulant feed rates for the Clairfont WTP in France [13]. Gagnon et al. developed an ANN model for predicting the optimal alum dose for the SteFoy WTP in Quebec, Canada [14]. Joo et al. developed a similar model for the Chungju WTP in Korea [9] and Van Leeuwen et al. developed an ANN model for the prediction of optimal alum doses based on jar tests conducted on surface waters collected in southern Australia [15]. Maier et al. used the same database as Van Leeuwen et al. to predict optimal alum dosage and treated water quality parameters [11]. Heddam et al. developed and compared two ANNs namely, generalised regression neural network and radial basis function neural network for modelling the nonlinear complex process of coagulant dosage using raw water data collected at Boudouaou WTP in Algiers, located in the northern zone of Algeria [16].

Another approach to modelling coagulant dosage is a neuro-fuzzy-based approach. Heddam et al. developed and compared two ANFIS models namely (i) grid partition-based fuzzy inference system (FIS), named ANFIS-GRID and (ii) subtractive clustering-based inference system (FIS), named ANFIS-SUB, for the Boudouaou WTP [17]. Wu and Lo developed and compared the performance of ANNs and Adaptive Neuro-Fuzzy Inference System (ANFIS) approaches for coagulant dosage modelling. The study was conducted at the WTP in Taipei County, Taiwan, and the models were used to model poly aluminium chloride dosing [18].

These models can be differentiated from each other based on the type of raw water data used as input. Bazer-Bachi et al. [13] used four variables (turbidity, resistivity, temperature and organic material) as input data with a correlation coefficient (CC) between observed and calculated coagulant dose of 0.94. The model of Gagnon et al. [14] is based on the pH, turbidity, temperature and conductivity, the mean absolute error (MAE) was approximately (2.69 mg/L) in the validation phase. The parameters used in the model developed by Joo et al. [9] are temperature, pH, turbidity and alkalinity. The root-mean-square error (RMSE) obtained ranged over two orders of magnitude with minimum and maximum values of 17.31 and 24.04. Van Leeuwen et al. [15] and Maier et al. [11] selected the following parameters: dissolved organic carbon, absorbance at 254 nm, turbidity and alkalinity. The standard deviation of the difference between the actual coagulant dose levels and those obtained using the ANN model developed by van Leeuwen et al. [15] was (6.2 mg/L), and the MAEs obtained using the model developed by Maier et al. [11] was (3.2 mg/L). Heddam et al. [16,17] used six variables (turbidity, conductivity, pH, temperature, dissolved oxygen and absorbance at 254 nm); the CC obtained ranged over two orders of magnitude, with minimum and maximum values of 0.90 and 0.93 in the validation phase. Finally, Wu and Lo [18] used turbidity, pH, colour and temperature as input to the developed model. The CC between observed and calculated coagulant dosage was relatively strong (CC = 0.860). Hence, we can conclude that the performances of these models vary depending on the type of data. Thus, it is difficult to tell which model will be more suitable for a particular application.

Recently, dynamic evolving neural-fuzzy inference system (DENFIS) has been used in many fields of civil engineering applications. Nevertheless, to the author's knowledge, coagulant dosage modelling and predicting with DENFIS has been not reported in the literature. The present study investigates the use of DENFIS system in the development of robust model for coagulant dosage modelling.

2. Materials and methods

2.1. WTPs and raw water data

The raw water used in this study was taken from the WTPs of Boudouaou City Boudouaou water treatment plant (BWTP), located at Boudouaou, Algeria. It is located at about 7 km from the Keddara dam which supplies potable water to Algiers, capital of Algeria. BWTP is an important drinking WTP in Algeria and operates with a production capacity of 540,000 CMD (cubic meters per day), and provides drinking water to more than four million inhabitants. This drinking water plant uses conventional treatment as the main treatment steps. The plant uses alum as the primary coagulant at an average dosage of 23.87 mg/L, as $Al_2(SO4)_318H_20$. The raw water quality parameters and coagulant dosage of BWTP drinking water are presented in Table 1. In this research, raw water pH, temperature (TE), conductivity (CO), turbidity (TU), apparent colour (AC) and UV_{254} were measured at the inlet of the treatment plant. A jar test was used to simulate the coagulation and sedimentation processes, providing the necessary target values (the optimal dosing rates to be estimated from raw water quality data (Dos)). The raw water database consisted of 808 samples of six input variables (TU, CO, TE, AC, UV_{254} and the pH of raw water) were used as input parameters to model the coagulant dosage and sampled between January 2009 and June 2012. Nevertheless, the data collected at the BWTP were incomplete; and therefore, days with inadequate data were removed from the patterns. In addition to the above-mentioned input parameters, the optimum coagulant dosage rate was determined through Jar-tests (Dos) and were related to the above reported six raw water variables, which was used as an output parameter of the developed model [16,17,19].

Table 1
The statistical summary of raw water data

Variables		Unit	X_{max}	X_{min}	X_{mean}	S_x	C_v
Temperature (TE)	Training	℃	19.40	10.20	15.15	2.05	0.13
	Validation		19.50	11.00	15.10	2.03	0.13
	All data		19.50	10.20	15.14	2.05	0.13
pH	Training	/	8.33	7.48	7.96	0.18	0.02
	Validation		8.24	7.56	7.97	0.17	0.02
	All data		8.33	7.48	7.97	0.18	0.02
Conductivity (CO)	Training	uS/cm	1312.00	691.00	982.57	100.08	0.10
	Validation		1274.00	709.00	995.98	106.03	0.10
	All data		1312.00	691.00	985.22	101.36	0.10
Turbidity (TU)	Training	NTU	62.60	1.12	5.66	5.88	1.03
	Validation		38.95	1.14	5.16	4.23	0.81
	All data		62.60	1.12	5.56	5.595	1.00
Apparent colour (AC)	Training	UH(1)	70.00	5.00	16.09	7.04	0.43
	Validation		50.00	5.00	16.58	6.96	0.41
	All data		70.00	5.00	16.19	7.02	0.43
UV_{254} (UV)	Training	cm^{-1}	1.11	0.01	0.11	0.04	0.40
	Validation		0.25	0.04	0.11	0.02	0.25
	All data		1.11	0.01	0.11	0.04	0.38
Coagulant dosage (Dos)	Training	mg/L	50.00	2.00	24.06	12.61	0.52
	Validation		50.00	2.00	23.09	12.23	0.52
	All data		50.00	2.00	23.87	12.53	0.52

(1) Hanzen unit = (mg Pt-Co L^{-1}), Cv = S_x/X_{mean}.

2.2. Dynamic evolving neural-fuzzy inference system

The DENFISs was introduced by Kasabov and Song [20]. DENFIS evolves through incremental, hybrid (supervised and unsupervised) learning, and accommodates new input data, including new features, new classes, etc. through local element tuning. New fuzzy rules are created and updated during the operation of the system. At each time moment, the output of DENFIS is calculated through a FIS based on the *m*-most activated fuzzy rules, which are dynamically chosen from a fuzzy rule set. A set of fuzzy rules can be inserted into DENFIS before or during its learning process. Fuzzy rules can also be extracted during or after the learning process [20,21]. An evolving clustering method (ECM) is used in DENFIS models to partition the input space for creating the fuzzy rules [20]. DENFIS is used for online (DENFIS-ON) and offline (DENFIS-OF) learning. In the online model of DENFIS, the linear functions in the consequent parts are created and updated through learning from data using least square estimator [20]. In the online mode, the fuzzy rules in the rule set can also be updated as new training data appear in the system [22]. DENFIS offline was proposed together with the online version of DENFIS, which sacrifices the dynamic evolving aspect of the DENFIS algorithm and replaces it with more sophisticated learning algorithm aimed at providing higher accuracy. It has shown improvement in prediction accuracy; however, more optimisation can be applied to further improve its accuracy [23]. The Takagi–Sugeno fuzzy inference engine is used in both online and offline modes of DENFIS. The difference between them is that for forming a dynamic inference engine, only first-order Takagi–Sugeno fuzzy rules are employed in DENFIS online mode, and both first-order Takagi–Sugeno fuzzy rules and expanded high-order Takagi–Sugeno fuzzy rules are used in DENFIS offline modes. To build such a fuzzy inference engine, several fuzzy rules are dynamically chosen from the existing fuzzy rule set depending on the position of current input vector in the input space [24].

2.2.1. General principles

The DENFIS [20], both online and offline models, use Takagi–Sugeno type fuzzy inference engine [25,26]. The inference in DENFIS is performed on m fuzzy rules indicated as follows.

$$\left\{\begin{array}{l} \text{if } x_1 \text{is } R_{11} \text{ and } x_2 \text{ is } R_{12} \text{ and} \ldots \text{and } x_q \text{ is } R_{1q}, \text{ then } y \text{ is } f_1\ (x_1, x_2, \ldots, x_q) \\ \text{if } x_1 \text{is } R_{21} \text{ and } x_2 \text{ is } R_{22} \text{ and} \ldots \text{and } x_q \text{ is } R_{2q}, \text{ then } y \text{ is } f_2\ (x_1, x_2, \ldots, x_q) \\ \text{if } x_1 \text{is } R_{m1} \text{ and } x_2 \text{ is } R_{m2} \text{ and} \ldots \text{and } x_q \text{ is } R_{mq}, \text{ then } y \text{ is } f_m\ (x_1, x_2, \ldots, x_q) \end{array}\right\} (1)$$

where "x_j is R_{ij}", $i = 1, 2 \ldots, m$ and $j = 1, 2, \ldots, q$ are $m \times q$ fuzzy propositions that form m antecedents for m fuzzy rules, respectively; x_j, $j = 1, 2, \ldots, q$, are antecedent variables defined over universes of discourse X_j, $j = 1, 2, \ldots, q$ and R_{ij}, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, q$ are fuzzy sets defined by their fuzzy membership functions μR_{ij}: $X_j \rightarrow [0,1]$, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, q$. In the consequent parts of the fuzzy rules, y is the consequent variable, and crisp functions fi, $i = 1, 2 \ldots, m$ are employed.

In both DENFIS online and offline models, all fuzzy membership functions are triangular type functions which depend on three parameters as given by the following equation:

$$\mu(x) = mf(x, a, b, c) = \begin{cases} 0, & x \leqslant a \\ \frac{x-a}{b-a}, & a \leqslant x \leqslant b \\ \frac{c-x}{c-b}, & b \leq x \leq c \\ 0. & c \leq x \end{cases} \quad (2)$$

where b is the value of the cluster center on the x dimension, $a = b - d \times Dthr$ and $c = b + d \times Dthr$, d ranged between 1.2 and 2; the threshold value, $Dthr$, is a clustering parameter.

If the consequent functions are crisp constants, i.e. $f_i\ (x_1, x_2, \ldots, x_q) = C_i$, $i = 1, 2, \ldots, m$, we call such system a zero-order Takagi–Sugeno type FIS. The system is called a first-order Takagi–Sugeno type FIS, if $f_i\ (x_1, x_2, \ldots, x_q)$, $i = 1, 2, \ldots, m$, are linear functions. If these functions are nonlinear functions, it is called high-order Takagi–Sugeno FIS [20].

For an input vector $x^0 = [x_1^0\ x_2^0\ \ldots,\ x_q^0]$, the result of inference, y^0 (the output of the system) is the weighted average of each rule's output indicated as follows:

$$y^0 = \frac{\sum_{i=1}^{m} w_i f_i(x_1^0, x_2^0, \ldots, x_q^0)}{\sum_{i=1}^{m} w_i} \quad (3)$$

where

$$w_i = \prod_{j}^{q} \mu\ R_{ij}(x_j^0); \quad i = 1,2,\ldots,m, j = 1,2,\ldots,q. \quad (4)$$

2.3. Ranges of water quality data

The daily statistical parameters of the water quality data and coagulant dosage are given in Table 1, in which the X_{mean}, X_{max}, X_{min}, S_x and Cv denote the mean, maximum, minimum, standard deviation and variation coefficient, respectively. Because the seven variables above had different dimensions, and there

were major differences among values, it was considered necessary to standardise the primary data in order to enhance the training speed and the precision of the models. This was done to correct the differing units used to measure each variable. As can be seen from Table 1, the data ranges for the input and output variables vary significantly. The inputs and outputs of the data-sets were normalised to improve the performance of the model [16,17,19]. The normalisation applied was as follows:

$$xn_{i,k} = \frac{x_{i,k} - m_k}{SD_K} \quad (5)$$

where $xn_{i,k}$ is the normalised input k or target data at $i = 1,2, \ldots, N$, the index number of the data value, $x_{i,k}$ the original data, and m_k and SD_k are the mean value and standard deviation of input k or target data. All the input and output variables were normalised to have zero mean and unit variance.

2.4. Division of data

In this study, the DENFIS-OF and DENFIS-ON models for coagulant dosage were developed and compared. The 808 daily samples of six input variables (raw water quality) and one output variable (Dos) were used for the models. In both models, 80% of the data-set was used for training (calibration), which corresponded to 648 input–output pairs while 20% was used for testing the performance of the model predictions (validation), which corresponded to 160 input–output pairs. Note that both the DENFIS-OF and DENFIS-ON models employ the same training and validation data-sets for an appropriate performance comparison. The training data (1–648) were the first part of the data-set and the validation (1–160) data were the second.

A brief data summary are presented herein to highlight background raw water quality, concentration ranges and notable differences between distribution of the variables at the training and validation phases of the data set. It can be seen from Table 1, concentrations of coagulant dosage ranged over three orders of magnitude, with minimum and maximum values of 2.0 and 50 mg/L. The mean of all observations was nearly 24.00 mg/L (23.87 mg/L). The distribution of coagulant dosage in the training phase was very similar to that for the validation phase, with mean, min and maximum values of 24.06, 2.0 and 50.0 mg/L. The mean value was slightly higher than the value at the validation phase. Conductivity ranged from 691 to 1,312 uS/cm, with a mean value of 982.57 uS/cm (Table 1), in the training phase. At the validation phase, the minimum value of conductivity (709 uS/cm) was slightly higher than the value at the training phase. The maximum for conductivity was slightly less than the value at the training phase (1,274 uS/cm) and the mean of all observations was nearly 996 uS/cm over one-tenth the mean value at the training phase. The distribution of water temperature and pH shows that the values at the training phase were very similar to that for the validation phase, with no significant difference.

As seen from Table 1, the mean value of water turbidity at the training phase was higher than the value in the validation phase, and was the same value in comparison with the all data. However, there is no shift in the min and mean values, as seen from Table 1. The mean observation for apparent colour was 5 UH; concentrations ranged from 16.09 to 70 UH in the training phase. At the validation phase, the mean value of apparent colour, 16.58 UH, was slightly higher than the value at the training phase. The maximum for apparent colour was 50 UH, and the mean of all observations was 5 UH. UV_{254} ranged from 0.01 to 1.1 (cm^{-1}), with a mean value of 0.11 in the training phase. The min value at the validation phase, 0.04 (cm^{-1}), was higher the min value in the training phase; the maximum observation (0.25) was less than the value in the training phase. Obviously, training data-set must cover all the characters of the problem, and it is desirable that training data includes all the maximal and minimal values. Moreover, in the process of segregation, the average and standard deviation of three independent subsets were computed to ensure the data-sets were divided equally among two subsets.

Pearson CC were calculated to identify the statistically significant correlation between the variables. Pearson CCs among the variables showed a number of strong, moderate and weak (positive and negative) associations. The correlation between these variables is shown in Table 2. As shown in this table, there is no significant correlations between these variables, except the very strong positive CC (0.754) between the apparent colour and turbidity and the strong negative CC (–0.66) between the pH and water temperature. The value and sign of CC demonstrate the relative contribution and direction of influence of each variable. As shown in Table 2, the most evident features are the strong positive correlations between apparent colour and coagulant dosage (0.533), turbidity and coagulant dosage (0.428), implying that any model built using these two variables will certainly be able to compute the coagulant dosage concentrations satisfactorily, these two variables are

Table 2
Pearson CCs between and among physical raw water variables, and coagulant dosage concentration

	TE (°C)	pH	CO (uS/cm)	TU (NTU)	AC (UH)	UV (cm^{-1})	Dos (mg/L)	Dos-of (mg/L)	Dos-on (mg/L)
TE	1.000								
PH	−0.665	1.000							
CO	0.035	−0.194	1.000						
TU	−0.268	0.217	0.239	1.000					
AC	−0.507	0.359	0.136	0.754	1.000				
UV	−0.107	0.122	−0.131	0.334	0.248	1.000			
Dos	−0.302	0.268	−0.022	0.428	0.533	0.143	1.000		
Dos-of*	−0.070	0.007	−0.039	0.002	0.049	0.045	0.710	1.000	
Dos-on**	−0.104	0.044	−0.009	0.011	0.054	0.015	0.820	0.750	1.000

*Dos-of coagulant dosage calculated using DENFIS-OF model. **Dos-on coagulant dosage calculated using DENFIS-ON model.

highly interrelated with the coagulant dosage, the positive CC indicates that as one variable increases, the other increases and vice versa. Coagulant dosage was weakly positively correlated with both pH and UV_{254} (CC = 0.268 and 0.143, respectively). Coagulant dosage and water temperature show moderate negative correlation with CC of (−0.302). Also, it can be seen from Table 2, coagulant dosage presented negligible correlation with conductivity. It is observed from Table 2, the parameters such as AC and TE exhibit strong negative relationship with CC of –0.507. TU, CO and UV_{254} are correlated with TE with coefficients of −0.268, 0.035 and −0.107, respectively. pH and AC show moderate CC of 0.359. CO, TU and UV_{254} show weak correlation with pH and with Pearson CCs of −0.194, 0.217 and 0.122, respectively. TU, AC and UV_{254} are correlated with CO with coefficients of 0.239, 0.136 and –0.131, respectively.

2.5. Performance indices

To assess the fitting and predictive accuracy of the models, the data-sets were mathematically evaluated by calculating the following evaluation criteria: coefficient of correlation (CC), RMSE and MAE.

$$\mathrm{CC} = \frac{\frac{1}{N}\sum (O_i - O_m)(P_i - P_m)}{\sqrt{\frac{1}{N}\sum_{i=1}^{n} (O_i - O_m)^2}\sqrt{\frac{1}{N}\sum_{i=1}^{n} (P_i - P_m)^2}} \quad (6)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (O_i - P_i)^2} \quad (7)$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} |O_i - P_i| \quad (8)$$

where N is the number of data points, O_i is some measured value and P_i is the corresponding model prediction. O_m and P_m are the average values of O_i and P_i.

3. Results and discussion

The modelling approaches above presented in section (2.2) were applied to modelling coagulant dosage (Dos) in WTP of Boudouaou. In this section, we present the numerical results and their evaluation. Table 3 presents the numerical results obtained, with comparisons in terms of accuracy and reliability. The DENFIS-ON and DENFIS-OF models were performed using program codes written in Matlab language available from the Knowledge Engineering and Discovery Research Institute (KEDRI) (http://www.aut.ac.nz/research/research-institutes/kedri/books). In order to compare the performance of both DENFIS-ON and DENFIS-OF models, in addition to CCs and RMSE, MAE was also considered as model performance

Table 3
Performances of the two DENFIS models in different phases

	Training			Validation		
Model	CC	RMSE	MAE	CC	RMSE	MAE
DENFIS-ON	0.824	7.435	4.837	0.804	7.858	5.531
DENFIS-OF	0.721	8.744	7.030	0.666	9.279	7.433

evaluation criteria. All the two models are trained and tested with the same data. Results for the DENFIS-ON and DENFIS-OF are tabulated in Table 3. In addition, the performance of DENFIS depends mostly on the choice of distance threshold value (*Dthr*), the two models are trained with different distance threshold value (*Dthr*) and its effects on the prediction performance are analysed.

Table 3 shows the results of different DENFIS-ON architectures, in terms of RMSE, MAE and CC statistics, respectively. The DENFIS model, on the other hand, needs only one parameter to be tuned up; the distance threshold value (*Dthr)*. *Dthr* parameter was changed from 0.09 to 0.01, with 0.01 steps in order to seek the most suitable value to achieve the best prediction performance. The best result was obtained when *Dthr* was 0.01. As shown in Table 3, the DENFIS-ON predictions for the coagulant dosage yield a MAE of 4.837, a RMSE of 7.435 and a CC of 0.824 in the training phase. Table 3 indicates that the DENFIS-ON has the smallest MAE (5.531) and RMSE (7.858), and the highest CC (0.804) in the validation phase. These values show that the DENFIS-ON predicts coagulant dosage very well. Figs. 1(a)–(b) and 2(a)–(b) shows the plots between observed and model calculated, and the Scatterplots of observed vs. calculated values of coagulant dosage concentration in training and validation, for the DENFIS-ON model, respectively. Generally, very good prediction accuracy is attained with the DENFIS-ON model. The

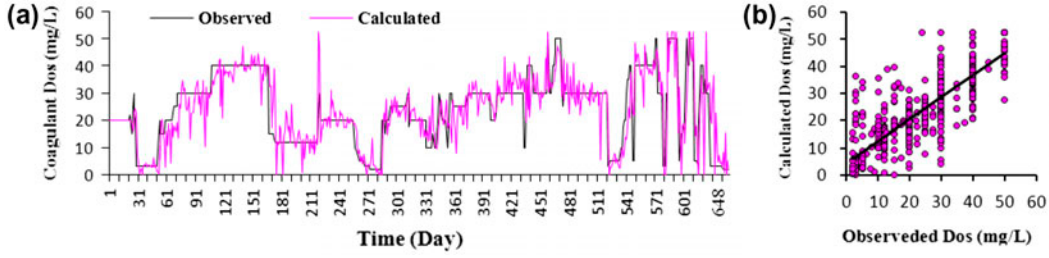

Fig. 1. Results obtained with DENFIS-ON model for coagulant dosage modelling in the training phase: (a) comparison of observed and simulated series of Dos and (b) scatter plots of observed and calculated Dos.

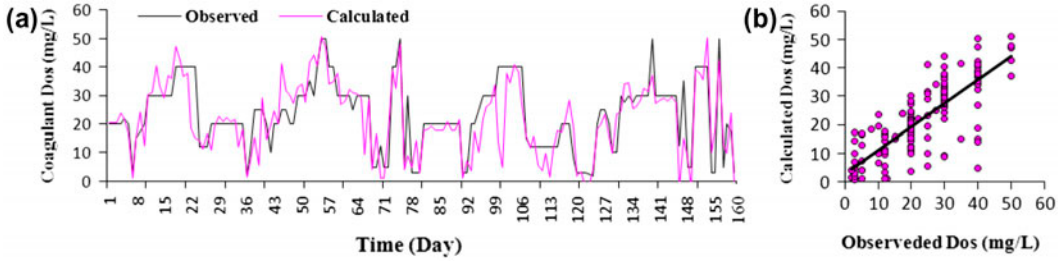

Fig. 2. Results obtained with DENFIS-ON model for coagulant dosage modelling in the validation phase: (a) comparison of observed and simulated series of Dos and (b) scatter plots of observed and calculated Dos.

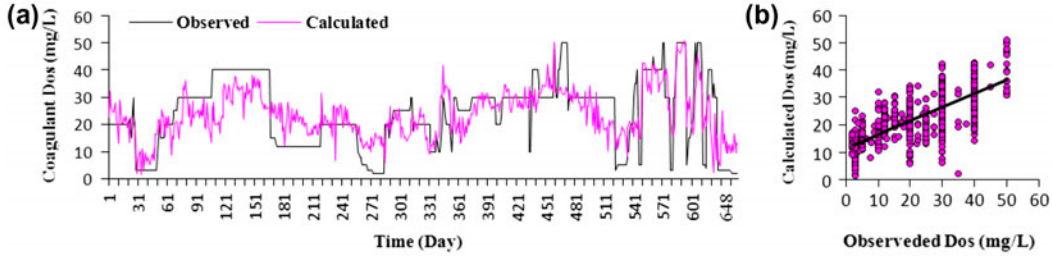

Fig. 3. Results obtained with DENFIS-OF model for coagulant dosage modelling in the training phase: (a) comparison of observed and simulated series of Dos and (b) scatter plots of observed and calculated Dos.

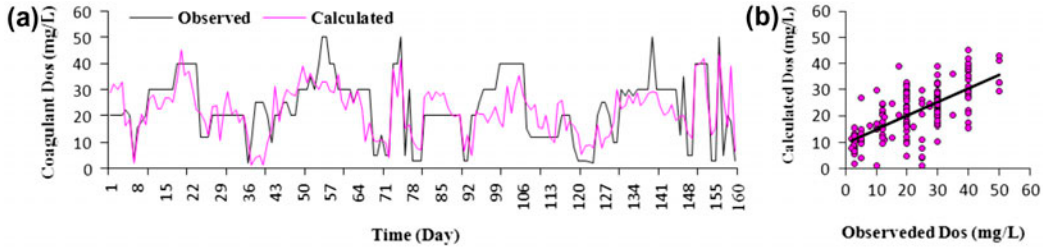

Fig. 4. Results obtained with DENFIS-OF model for coagulant dosage modelling in the validation phase: (a) comparison of observed and simulated series of Dos and (b) scatter plots of observed and calculated Dos.

comparison of results showed that DENFIS-ON predicted coagulant dosage concentration was found to be closer with the measured values.

Table 3 shows that the DENFIS-ON performed better during training and validation, and it outperforms the DENFIS-OF in terms of all the standard statistical measures. In the training phase, the DENFIS-ON improved the DENFIS-OF forecast of about 14.97 and 31.19% reduction in RMSE and MAE values, respectively. In addition, improvements of the forecast results regarding the CC value during the training phase were approximately 10.30%. In addition, in the validation phase as seen in Table 3, the values with the DENFIS-ON prediction were able to produce a good forecast, as compared to those with DENFIS-OF prediction. In the validation phase, the DENFIS-ON improved the DENFIS-OF M6 forecast of about 15.31 and 25.58% reduction in RMSE and MAE values, respectively. In addition, improvements of the forecast results regarding the CC value during the validation phase were approximately 13.80%. Overall, the performance of the DENFIS-ON model is very good. The results demonstrate that the DENFIS-ON can be successfully applied to establish the predicting model that could provide accurate and reliable coagulant dosage concentration (Dos), raw water quality data. It is understood that the developed DENFIS-ON model has been successful in the prediction of coagulant dosage. Fig. 3(a)–(b) and Fig. 4(a)–(b) shows the plots between observed and model calculated, and the Scatterplots of observed vs. calculated values of coagulant dosage concentration in training and validation, for the DENFIS-OF model, respectively.

4. Conclusion

Feasibility of applying DENFIS to predict coagulant dosage concentration in WTP of Boudouaou, Algeria was studied. Results show that the trained model can be considered as very satisfactory. Six raw water quality variables, such as water pH, temperature, conductivity, turbidity, apparent colour, and UV_{254}, were used as input parameters and the DENFIS model results showed a very good agreement with the measured values (CC = 0.824 − 0.666). Two types of DENFIS networks were considered regarding to the structure identification method. It has been concluded that DENFIS-ON could be used as a powerful and simple alternative technique for prediction of coagulant dosage in WTP.

Acknowledgements

The authors would like to thank the editor and anonymous reviewers for their valuable comments and suggestions, which were helpful in improving the paper. Also, we would like to thank the staff of The Water Laboratory of Boudouaou water treatment plant (BWTP) for providing the data that makes this research possible.

References

[1] K. Zhang, G. Achari, R. Sadiq, C.H. Langford, M.H.I. Dore, An integrated performance assessment framework for water treatment plants, Water Res. 46 (2012) 1673–1683.
[2] C.N. Lei, I.C. Lou, H.U. Song, P. Sun, Turbidity removal improvement for Yangtze River raw water, Desalin. Water Treat. 45(1–3) (2012) 215–221.
[3] J.Q. Jiang, Development of coagulation theory and pre-polymerized coagulants for water treatment, Sep. Purif. Rev. 30(1) (2001) 127–141.
[4] I. Lou, S. Gong, X. Huang, Y. Liu, K.M. Mok, Coagulation optimization using ferric and aluminum salts for treating high algae and high alkalinity source water in a typical North-China plant, Desalin. Water Treat. 51 (16–18) (2013) 3361–3370.
[5] M. Dixon, C. Staaks, R. Fabris, V. Vimonses, C.W.K. Chow, S. Panglisch, J.A. van Leeuwen, M. Drikas, The impact of optimised coagulation on membrane fouling for coagulation/ultrafiltration process, Desalin. Water Treat. 51(13–15) (2013) 2718–2725.
[6] F. Ying, Investigation on flocculation process of composite poly-Si-Fe coagulant, Desalin. Water Treat. 30 (1–3) (2011) 122–133.

[7] N.A. Awang, H. Aziz, M.J.K. Bashir, M. Umar, Comparative removal of suspended solids from landfill leachate by *Hibiscus rosa-sinensis* leaf extract and alum, Desalin. Water Treat. 51(10–12) (2013) 2005–2013.
[8] W.P. Cheng, R.F. Yu, Y.J. Hsieh, S.Y. Wu, Y.W. Huang, Optimizing coagulant demand by nephelometric turbidimeter monitoring system (NTMS), Desalin. Water Treat. 16(1–3) (2010) 95–100.
[9] D.S. Joo, D.J. Choi, H. Park, The effects of data preprocessing in the determination of coagulant dosing rate, Water Res. 34 (2000) 3295–3302.
[10] R.F. Yu, S.F. Kang, S.L. Liaw, M.C. Chen, Application of artificial neural network to control the coagulant dosing in water treatment plant, Water Sci. Technol. 42(3–4) (2000) 403–408.
[11] H.R. Maier, N. Morgan, C.W.K. Chow, Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters, Environ. Model. Softw. 19 (2004) 485–494.
[12] H.R. Tashaouie, G.B. Gholikandi, H. Hazrati, Artificial neural network modeling for predict performance of pressure filters in a water treatment plant, Desalin. Water Treat. 39(1–3) (2012) 192–198.
[13] A. Bazer-Bachi, E. Puech-Coste, R. Ben Aim, J.L. Probst, Mathematical modelling of optimal coagulant dose in water treatment plant, Revue des sciences de l'eau 3 (1990) 377–397.
[14] C. Gagnon, B. Grandjean, Modelling of coagulant dosage in a water treatment plant, Artif. Intell. Eng. 11(1997) 401–404.
[15] J.A. Van Leeuwen, C.W.K. Chow, D. Bursill, M. Drikas, Empirical mathematical models and artificial neural networks for the determination of alum doses for treatment of southern Australian surface waters, J. Water Sci. Res. Technol. AQUA 48(3) (1999) 115–127.
[16] S. Heddam, A. Bermad, N. Dechemi, 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(12) (2011) 1209–1214.
[17] S. Heddam, A. Bermad, N. Dechemi, ANFIS-based modelling for coagulant dosage in drinking water treatment plant: A case study, Environ. Monit. Assess. 184 (2012) 1953–1971.
[18] G.D. Wu, S.L. Lo, Predicting real-time coagulant dosage in water treatment by artificial neural networks and adaptive network-based fuzzy inference system, Eng. Appl. Artif. Intell. 21(8) (2008) 1189–1195.
[19] S. Heddam, Modeling water quality using artificial neural networks approaches, neuro fuzzy systems and the classification techniques: Case study of the drinking water treatment plant of Boudouaou-Algeria, Doctorate thesis, Superior National School of Agronomics (ENSA), Algiers, Algeria, 2012. 249 p.
[20] N. Kasabov, Q. Song, DENFIS: Dynamic, evolving neural-fuzzy inference system and its application for time-series prediction, IEEE Trans. Fuzzy Syst. 10 (2002) 144–154.
[21] N. Kasabov, Evolving Connectionist Systems & the Knowledge Engineering Approach, 2nd ed., Springer-Verlag, London, 2007, 465. ISBN 978-1-84628-345-1.
[22] N. Kasabov, B. Woodford, Rule insertion and rule extraction from evolving fuzzy neural networks: Algorithms and applications for building adaptive, intelligent expert systems, in: Proceedings of IEEE International Conference Fuzzy Systems Conference, Seoul, Korea, 1999, pp. 1406–1411.
[23] Y.C. Hwang, Q. Song, Dynamic neural fuzzy inference system, in: Proceedings of the International Conference on Advances in Neuro-Information Processing ICONIP. Lecture Notes in Computer Science, Vol. 5506/2009, Springer-Verlag, Berlin Heidelberg, 2009, pp. 1245–1250.
[24] N. Kasabov, Q. Song, M.M. Tian, Fuzzy-neuro systems for local and personalized modelling. Forging new frontiers: Fuzzy pioneers II, Stud. Fuzz. Soft Comput. 218/2008(2008) 175–197.
[25] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, IEEE Trans. Syst., Man Cybern SMC 15(1) (1985) 116–132.
[26] M. Sugeno, G. Kang, Structure identification of fuzzy model, Fuzzy Sets Syst. 26(1) (1988) 15–33.