Journal of Advances in Environmental Health Research

Application of artificial neural network (ANN) for the prediction of water treatment plant influent characteristics

Mehri Solaimany-Aminabad[1], Afshin Maleki[1], Mahdi Hadi[2]

1 Kurdistan Environmental Health Research Center, School of Health, Kurdistan University of Medical Sciences, Sanandaj, Iran
2 Center for Water Quality Research (CWQR), Institute for Environmental Research (IER), Tehran University of Medical Sciences, Tehran, Iran

Original Article

Abstract

Application of a reliable forecasting model for any water treatment plant (WTP) is essential in order to provide a tool for predicting influent water quality and to form a basis for controlling the operation of the process. This would minimize the operation and analysis costs, and assess the stability of WTP performances. This paper focuses on applying an artificial neural network (ANN) approach with a feed-forward back-propagation non-linear autoregressive neural network to predict the influent water quality of Sanandaj WTP. Influent water quality data gathered over a 2-year period were used to building the prediction model. The study signifies that the ANN can predict the influent water quality parameters with a correlation coefficient (R) between the observed and predicted output variables reaching up to 0.93. The prediction models developed in this work for Alkalinity, pH, calcium, carbon dioxide, temperature, total hardness, turbidity, total dissolved solids, and electrical conductivity have an acceptable generalization capability and accuracy with coefficient of determination (R^2) ranging from 0.86 for alkalinity to 0.54 for electrical conductivity. The predicting ANN model provides an effective analyzing and diagnosing tool to understand and simulate the non-linear behavior of the influent water characteristics. The developed predicting models can be used by WTP operators and decision makers.

KEYWORDS: Neural Network, Time Series, Influent Water Characteristics, Forecasting

Date of submission: 7 June 2013, ***Date of acceptance:*** 9 September 2013

Citation: Solaimany-Aminabad M, Maleki A, Hadi M. **Application of artificial neural network (ANN) for the prediction of water treatment plant influent characteristics.** J Adv Environ Health Res 2013; 1(2): 89-100.

Introduction

To maintain a stable performance in a water treatment plant (WTP), it is desirable to know in advance the influent water characteristics of the WTP. Water characteristics such as turbidity, total suspended solids, and pH are important water quality parameters. There is a significant relationship between these parameters and the amounts of coagulants and flocculants used in treatment processes. Prediction of the influent water characteristics is helpful in the optimal scheduling of the coagulation and flocculation process. In practice, the influent water characteristics are usually estimated by the operators based on experience and or using online sensors. Such estimations, however, are not accurate enough to manage WTPs, especially for operators that want to manage the WTP performance for the next day. The precipitation may cause large variability of the influent water characteristics, thus reducing the efficiency of WTPs. Moreover, heavy rainfall overwhelms the water treatment system, causing spills and overflows. Thus, prediction models for water

Corresponding Author:
Afshin Maleki
Email: maleki43@yahoo.com

quality characteristics, based on their registered historical data, can be built by the data mining approach.[1] Data mining is a promising approach for building prediction models. It is the process of finding patterns from data by algorithms versed on the crossroads of statistics and computational intelligence.[2]

Artificial neural networks (ANNs) are one of the most accurate and widely used data mining processes and forecasting models. It has been shown that a network can approximate any continuous function to any desired accuracy. ANNs are nonlinear and non-parametric methods, and unlike traditional approaches, such as the Box–Jenkins or ARIMA, do not assume that the time series under study are generated from linear processes. However, they may be inappropriate if the underlying mechanism is nonlinear. In fact, real world systems are often nonlinear.[3] Artificial neural networks have been found to be a viable contender to various traditional time series models.[4,5] Lapedes has reported the first attempt to model nonlinear time series with artificial neural networks.[6] Imrie et al. have reported the application of ANN for the river flow prediction.[7] Wu and Lo used the ANN to model the nonlinear relationship between accumulated input and output numerical data for the coagulation processes in water treatment.[8] Melessea et al. have presented the application of a multilayer perceptron (MLP) ANN with an error back propagation algorithm for the prediction of suspended sediment load of river systems.[9] An ANN data driven modeling approach was used by Huo et al. to predict the water quality indicators of Lake Fuxian, the deepest lake of southwest China.[10] Patil et al. have presented a study of predicting sea surface temperature with nonlinear autoregressive neural networks.[11]

In this study, we used an ANN approach to predict daily influent water characteristic to Sanandaj water treatment plant. This paper presents a data-mining approach to predict influent water characteristic in a WTP for a short-term period (one day ahead). In this work, the proposed approach is based on the classical nonlinear autoregressive time series using time-lagged feed-forward networks, in which the data from the daily time series are used to forecast the next day. In this study the prediction models are developed for alkalinity (Alk), pH, calcium (Ca), carbon dioxide (CO_2), temperature (T), total hardness (TH), turbidity (Tur), total dissolved solids (TDS), electrical conductivity (EC), and chloride (Cl) as the influent water characteristics. The models output is evaluated using statistical indices and observed water quality data.

Materials and Methods

ANN model was developed to predict the characteristic parameters of influent water of Sanandaj water treatment plant. This plant is one of the oldest water treatment plants in Iran. It is located in the northeast of the city of Sanandaj at an altitude of 1510 meters above sea level and near Nanaleh village road. Nominal design capacity of the treatment plant is 0.7 cubic meters per second, and can increased up to 1.5 cubic meters per second when needed. The raw water is supplied from Gheslagh dam. The water is transferred through a concrete and steel transmission line, with the length of 8 kilometers, by gravity force. The treated water, after disinfection and storage, is pumped by a steel transmission pipeline with the length of 2.2 km to Faizabad storage tank and then the distribution network. Registered daily historical data of the influent water quality parameters including carbon dioxide, total hardness, chloride, total calcium, total dissolved solids, total alkalinity, electrical conductivity, pH, turbidity, and temperature were used to conduct the study. The data was provided by the urban Water and Wastewater Company of Kurdistan and collected over a 2-year period. This period was satisfactory as it covered all probable seasonal variations in the studied variables.

The numbers of data points for plant data used for the training and test data sets together are 707 points. The description of the variables, units of measure, range of the data, together with the mean and standard deviation of the plant's raw data are presented in table 1.

Artificial Neural Network is an information processing tool that is inspired by systems such as biological nervous systems. The objective of a neural network is to compute output values from input values by some internal calculations.[12]

Neural network is trained to construct the specific black box function by adjusting the values of the connections (weights) between layers of elements based on a comparison of the output and the target until the network output matches the target.[13]

Figure 1 illustrates neural network training structure. There are many different types of training algorithms. One of the most common classes of training algorithms for feed forward neural networks (FFNNs) is called back propagation (BP).[14]

The basic component of a neural network is the neuron, also called node. Figure 2 illustrates a single node of a neural network. The inputs are represented by a_1, a_2, and a_n, and the output by O_j. Several signals can be inserted into the node. The node manipulates these inputs in such a way to give a single output signal. The values W_{1j}, W_{2j}, and W_{nj}, are weight factors associated with each of the inputs to the node. The other input to the node, b_j, is the node's internal threshold, also called bias. This is a randomly chosen value that governs the node's net input through the following equation:[15]

Table 1. Raw influent water characteristics data of Sanandaj, Iran water treatment plant (WTP)

	Mean (µ)	SD	Min	Max	µ-4SD	µ+4SD
CO_2	3.24	10.89	0.10	199.00	-40.32	46.80
TH	153.51	11.79	122.00	205.00	106.35	200.66
Cl	9.40	7.63	5.50	160.40	-21.12	39.92
Ca	47.14	7.01	32.40	146.30	19.12	75.17
TDS	209.87	19.16	157.00	252.00	133.22	286.53
Alk	157.95	16.57	0.00	193.00	91.68	224.23
EC	332.56	138.65	0.00	3337.00	-222.03	887.15
pH	8.47	7.18	7.16	175.90	-20.24	37.17
Tur	3.71	5.26	0.50	65.00	-17.33	24.76
T	11.42	6.31	2.00	90.00	-13.81	36.64

SD: Standard deviation; TH: Total hardness; TDS: Total dissolved solids; Alk: Alkalinity; EC: Electrical conductivity; Tur: Turbidity; T: Temperature

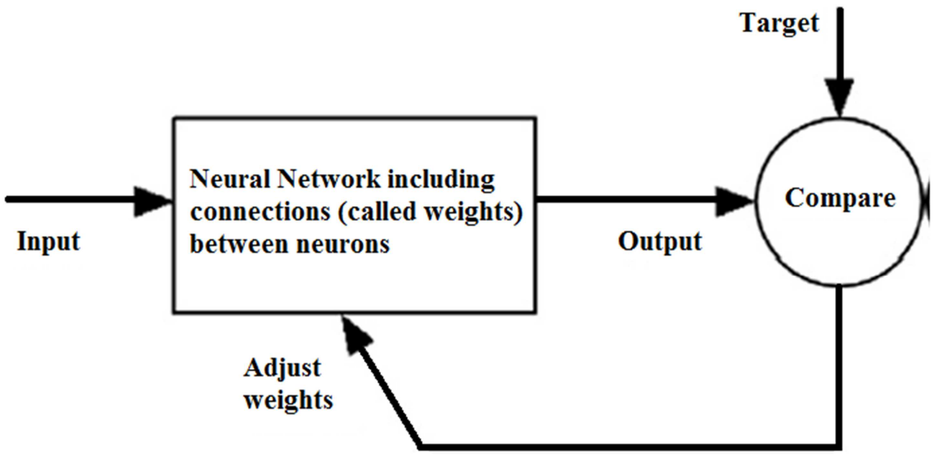


Figure 1. Neural network training structure

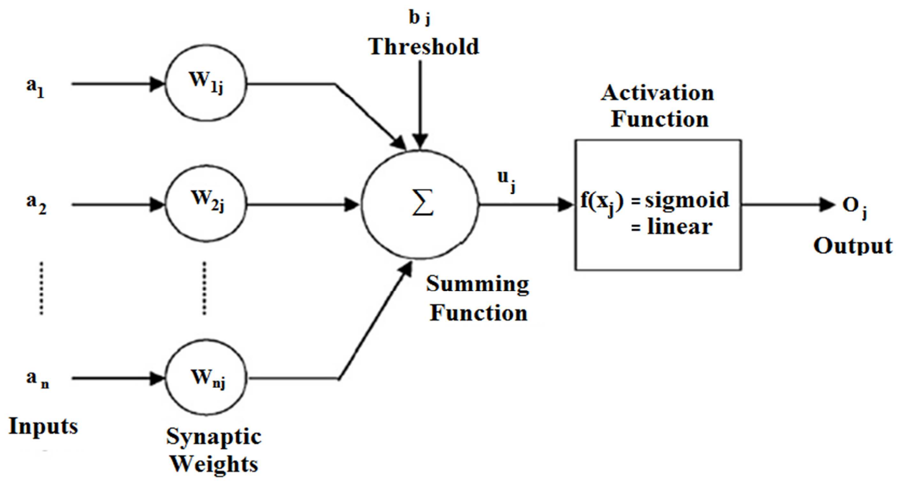


Figure 2. Single node anatomy

$$u_j = \sum_{i=1}^{n} (W_{ij} \times a_i) + b_j$$

The transfer function can transform the node's net input in a linear or non-linear manner. Commonly used transfer functions in hidden layer are sigmoid transfer function and hyperbolic tangent transfer function as follows:[15]

- Sigmoid transfer function

$$f(x) = \frac{1}{1+e^{-x}} \qquad 0 \le f(x) \le 1$$

- Hyperbolic tangent transfer function

$$f(x) = \tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \qquad -1 \le f(x) \le 1$$

The neuron's output O_j is found by performing linear function on the hidden layer neuron's outputs as follows:[15]

- Linear transfer function

$$f(x) = x \qquad -\infty < f(x) < +\infty$$

Neural networks very rarely operate directly on the raw data, although this is possible. The disadvantage of using raw data values is that the training time for the neural network would be significantly longer as the various variables have very different ranges. Data pre-processing can have a significant effect on the generalization performance of a supervised neural network.[16]

Plant data is, most often, not very reliable and many problems can occur which can affect the reliability or integrity of the data. Missing data is a common problem in statistical analysis.[17]

Tarassenko proposes some strategies to deal with missing data.[18] One method consists of replacing the missing data value by its mean or median across the training set.[19] The other method is to estimate the missing value for an n-dimensional input vector from knowledge of the other n-1 input variables. The last method uses either a linear model or a NN network to predict the nth value given the set (n-1)-dimensional vectors as inputs. The approach eventually adopted was to use a linear interpolation method to replace the missing data values in the Sanandaj WTP plant data set. In a few instances, the missing data points were consecutive, but this did not extend to more than 5 consecutive missing points.

Data that appear to be very distant from the normal data distribution may be classified as being outliers. In certain instances however, this outlying value may be correct and is a natural product of the variables distribution.[20] One approach for data rejection is to plot the histogram of the data distribution and then carefully scrutinize the data which appear as outliers. The standard deviation based outlier analysis is also a mechanism for revealing values that are numerically distant from the rest of the data. In this study, we took a normal distribution with cutoff 4 standard deviations from the mean to detect the outliers. Thus, the data that were more extreme than μ ± 4SD were considered as outliers.

Neural networks can be trained by using raw data as inputs, but the training time will be considerably longer. However, if scaled data is presented to the network, the weights can remain in small, similar predictable ranges. Box-cox transforms non-normally distributed data to a set of data that has approximately normal distribution. The Box-Cox transformation is a family of power transformations. The values of λ parameter for studied variables are shown in table 2. These values were not zero for all water characteristic parameters and the transformation of data was performed according to the following relationships:

If λ is not = 0, then $$data(\lambda) = \frac{data^{\lambda} - 1}{\lambda}$$

If λ is = 0, then $$data(\lambda) = \log(data)$$

After pre-processing the raw data, the neural network model was created in MATLAB software (version 7.12; Mathworks Inc., USA) that offers a platform for the simulation application. A nonlinear autoregressive (NAR) time series neural network was used and trained to predict the variable for the next day from that series' past values. The NAR is a network with feedback arrangement as shown in figure 3. In NAR network, there is only one series involved. The future values of a time series y (t) are predicted only from past values of that series. This form of prediction is called nonlinear autoregressive and can be written as follows:

$$y(t+1) = f(y(t), y(t-1),..., y(t-d))$$

The network was trained using the common algorithm of Levenberg-Marquardt. The network had non-linear sigmoid transfer function for the hidden layer and a linear transfer function for the output layer neurons. The number of feedback delays was determined by depicting partial autocorrelation function (PACF). The numbers of delays in PACF chart with a significant correlation coefficient were considered as the numbers of feedback delays. The numbers of feedback delays are shown in table 2. The other network properties are as follows:

- Network type: Feed-Forward Back-Propagation
- Training function: TRAINLM
- Performance function: MSE
- Number of hidden layers: 1

Hidden layer size: A single hidden layer with different count of neurons (i.e. 1 to 20) has been assessed for this study. As shown in figure 4 the performance of the network decreased almost by increasing the hidden layer size. These decreases are more considerable almost after ten sizes for neurons in all models. Thus, in this study 10 neurons were considered as the maximum possible size for the hidden layer for all models.

Table 2. Pre-processed influent water characteristics data and number of feedback delays

	Description and unit of measure	Min.	Median	Mean	Max.	SD	λ	Feedback Delays*
CO_2	Carbon dioxide (mg/l)	0.10	2.3	2.6	8.2	1.5	0.41	6
TH	Total Hardness (mg/l)	122.00	155.3	154.0	197.2	10.1	1.86	4
Cl	Chloride (mg/l)	5.50	9.0	8.9	12.5	1.2	-0.20	8
Ca	Calcium (mg/l)	32.40	48.0	47.2	59.8	4.0	2.43	6
TDS	Total dissolved solids (mg/l)	157.00	214.0	211.0	252.0	12.6	5.05	7
Alk	Total alkalinity (mg/li)	120.60	160.2	158.2	193.0	14.0	1.76	5
EC	Electrical conductivity (μ.mohs/cm)	260.00	333.5	330.0	393.0	18.5	4.59	11
pH	-	7.16	8.2	8.2	8.9	0.3	-0.02	5
Tur	Turbidity (NTU)	0.50	2.0	3.3	24.0	3.5	-0.20	8
T	Temperature (°C)	2.00	6.0	11.0	11.4	5.3	0.72	6

* Feedback delays: The number of auto-regressive series lags as inputs of the NAR neural network; SD: Standard deviation; TH: Total hardness; TDS: Total dissolved solids; Alk: Alkalinity; EC: Electrical conductivity; Tur: Turbidity; T: Temperature

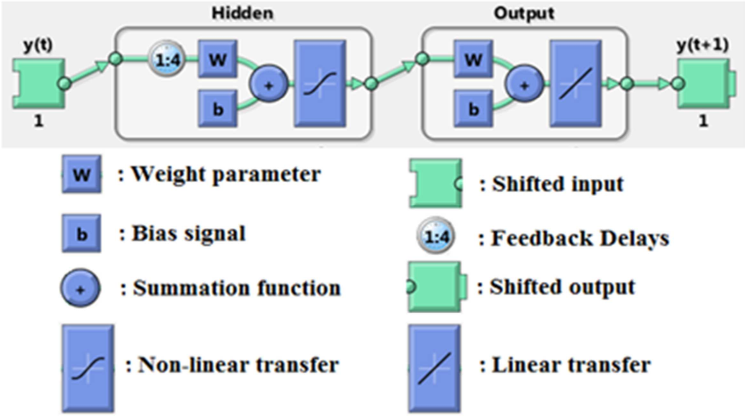


Figure 3. Neural network predicting structure with a hidden layer

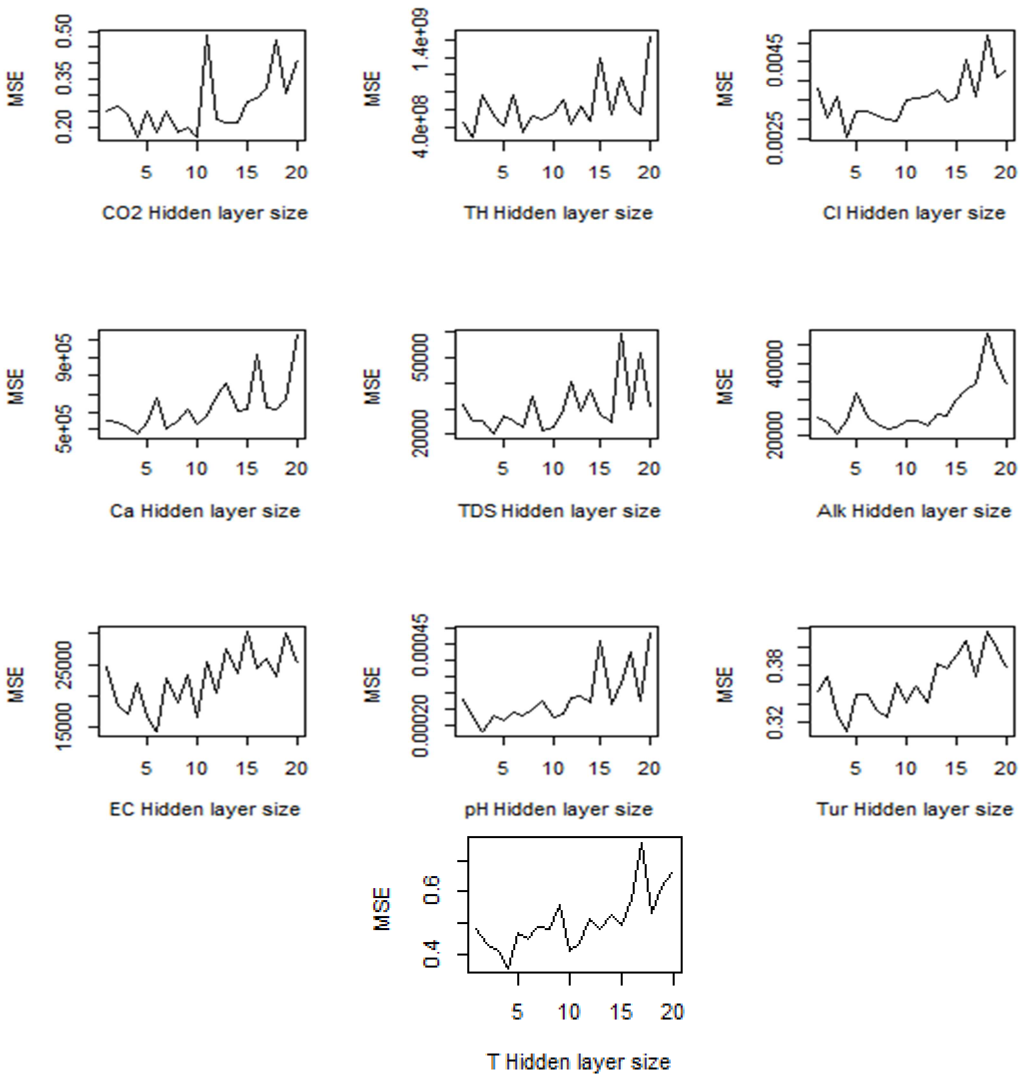


Figure 4. Hidden layer size versus network performance

The MATLAB routine *trainlm* (training with Levenberg-Marquardt algorithem) was used for the optimization. This algorithm attains fast learning speed and high performance relative to other optimization algorithms and the details of this algorithm are reported by Hagan et al.[15] The performance function used for training is based on the mean square errors (MSE) between actual WTP influent water characteristics and network predictions. Based on the selected network structure, the training process was activated to achieve a performance target of 1×10^{-3} for a maximum training of 1000 epochs. The learning rate was chosen to be 0.01. The value of this parameter was obtained after performing several trial and error runs. It was found that this value insures stable fast learning.

In order to study the relative performance of the network, the correlation coefficient (R) and root of mean square error (RMSE) were worked out. The underlying expressions as well as the strengths and weaknesses of these parameters are given as below.[11]

Correlation coefficient (R):

$$R = \frac{\sum_{i=1}^{n}(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X-\bar{X})^2(Y-\bar{Y})^2}}$$

where X = observed y_t , $\bar{X}$ = mean of X, Y = predicted y_t, $\bar{Y}$ = mean of Y, and n = number of observations.

The correlation coefficient (R) shows the extent of the linear association and similarity of trends between the target and the realized outcome. It is a number between 0 and 1; such that the higher the correlation coefficients the better the model fit. It, however, is heavily affected by the extreme values.

Root of mean square error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X-Y)^2}{n}}$$

RMSD is a good measure of accuracy, but only to compare forecasting errors of different models for a particular variable and not between variables, as it is scale-dependent.[21]

Results and Discussion

The most common training algorithm used in the ANN literature is called back propagation (BP). Back propagation was developed and popularized by Rumelhart et al. and it is the most widely implemented of all neural network algorithms.[22] It is based on a multi-layered feed forward topology with supervised learning. The network uses the default Levenberg–Marquardt algorithm for training. The input vectors and target vectors are randomly divided into three sets as follows: 70% are used for training, 10% are used to validate that the network is generalizing and to stop training before over-fitting, the last 20% are used as a completely independent test of network generalization. The number of networks to fit with different random starting weights was 20 times. These are then averaged when producing predicts.

Figure 4 shows the results of regression between network outputs and data sets of validation, and training and test targets. It is observed that the output tracks the targets well. Data from table 3 shows R and RMSE of each ANN for validation, training, and test steps. The correlation coefficient (R) measure the correlation between outputs and targets. An R value of 1 means a close relationship and 0 a random relationship while the RMSE is the root of mean squared difference between outputs and targets. The lower the values are the better.

The coefficient R for the validation phase upon application of the test set, ranges from 0.61 for Cl to 0.93 for Alk, and the coefficient of determination R^2 ranges from 0.37 for Cl to 0.86 for Alk. These figures indicate that 37% of the variation in the Cl variable can be explained by the variable time delays. The remaining 64% can be attributed to unknown, lurking variables, or inherent variability. Neural network model for Cl may, therefore, not able to solve this particular input-output mapping problem well.

The results for Alk in table 3 are interesting as the R correlation coefficient is 0.93 (R^2 = 0.86) for the validation phase. This indicates that this

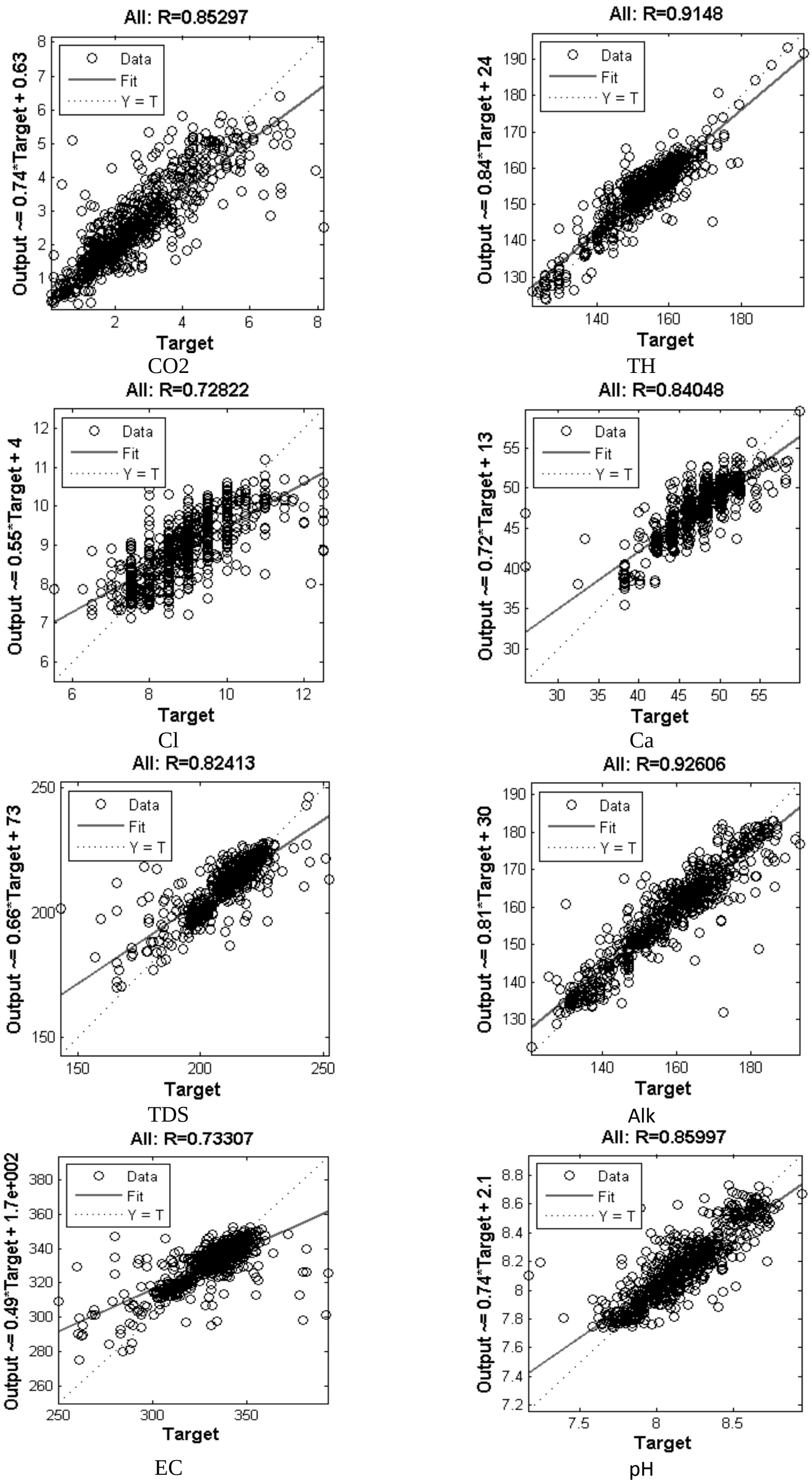


Figure 4. Measured and predicted output variables

TH: Total hardness; TDS: Total dissolved solids; Alk: Alkalinity; EC: Electrical conductivity model has the best result throughout. However, all other results are also acceptable. The regression coefficient ranges from 0.74 for EC (R^2 = 0.54) to 0.89 (R^2 = 0.79) for pH. Thus, for these variables more than 50% of the variation in them can be explained by their time delays. Time series models for these variables, therefore, are able to solve the input-output mapping problem well.

For all studied water influent characteristics, the simulation results of influent parameters are presented in figure 5 by plotting the measured and predicted output variables. The network response is satisfactory, and simulation can be used for entering new inputs. Given the above, it can be conclude that a feed-forward neural network based nonlinear autoregressive (NAR) model can be used for forecasting time series values well. The results of this study indicated high correlation coefficient between the measured and predicted output variables, reaching up to 0.93. Therefore, the prediction models developed in this work for Alk, pH, Ca, CO_2, T, TH, Tur, TDS, and EC have an acceptable generalization capability and accuracy with coefficient of determination ranging from 0.86 for Alk to 0.54 for EC. As a result, the neural network modeling could effectively simulate and predict these influent water quality parameters of Sanandaj WTP. In the case of Cl, neural network modeling may not able to predict this influent water characteristic well, at least in this study. Thus, it is necessary to conduct more studies to make its behavior clear. The testing step of the models also provided similar results to validation step results. The correlation coefficient ranges from 0.65 (R^2 = 0.42) for Cl to 0.88 (R^2 = 0.77) for TH and T. for Alkalinity, the test phase R is 0.85 (R^2 = 0.72). Thus, these results confirm the validation step.

The application of predictive models for wastewater influent characteristics has been reported in several studies.[23-25] Neural network modeling has rarely been used in water treatment plant influent forecasting. Zhang and Stanley, in their two studies, used ANN models to predict treated water turbidity and color at the Rossdale water treatment plant in Edmonton, Alberta, Canada.[26,27] Lamrini et al. adapted the Levenberg–Marquardt method in ANN to predict the coagulant dosage for the raw water with high turbidity.[28]

Conclusion

This study presented a detailed methodology for developing successful ANN models for modeling influent water characteristics. The utility and applicability of this methodology is demonstrated through a case study in which some successful NAR models to predict influent water characteristics were developed. It is concluded that nonlinear autoregressive or NAR neural network provides an effective analyzing and diagnosing tool to understand and simulate the non-linear behavior of influent water characteristics of the water treatment plant. Moreover, it is a valuable predicting tool for

Table 3. Performance of MLP networks

Parameter	Training phase		Validation phase		All phases		Testing phase	
	RMSE	R	RMSE	R	RMSE	R	RMSE	R
Cl	0.68	0.59	0.89	0.37	0.74	0.53	0.87	0.42
EC	12.81	0.55	12.41	0.55	12.69	0.53	12.43	0.49
TDS	6.70	0.72	6.96	0.61	7.20	0.67	9.34	0.55
Tur	1.76	0.72	1.97	0.69	1.91	0.71	2.41	0.69
TH	3.61	0.86	4.77	0.71	4.07	0.83	5.19	0.77
T	1.20	0.86	2.19	0.71	1.53	0.83	2.01	0.77
CO_2	0.72	0.76	0.77	0.71	0.76	0.72	0.92	0.64
Ca	2.27	0.71	2.03	0.72	2.18	0.71	1.87	0.74
pH	0.14	0.76	0.14	0.79	0.14	0.74	0.17	0.62
Alk	4.79	0.88	4.75	0.86	5.33	0.86	7.74	0.72

RMSE: Root of mean square error; TH: Total hardness; TDS: Total dissolved solids; Alk: Alkalinity; EC: Electrical conductivity; Tur: Turbidity; T: Temperature

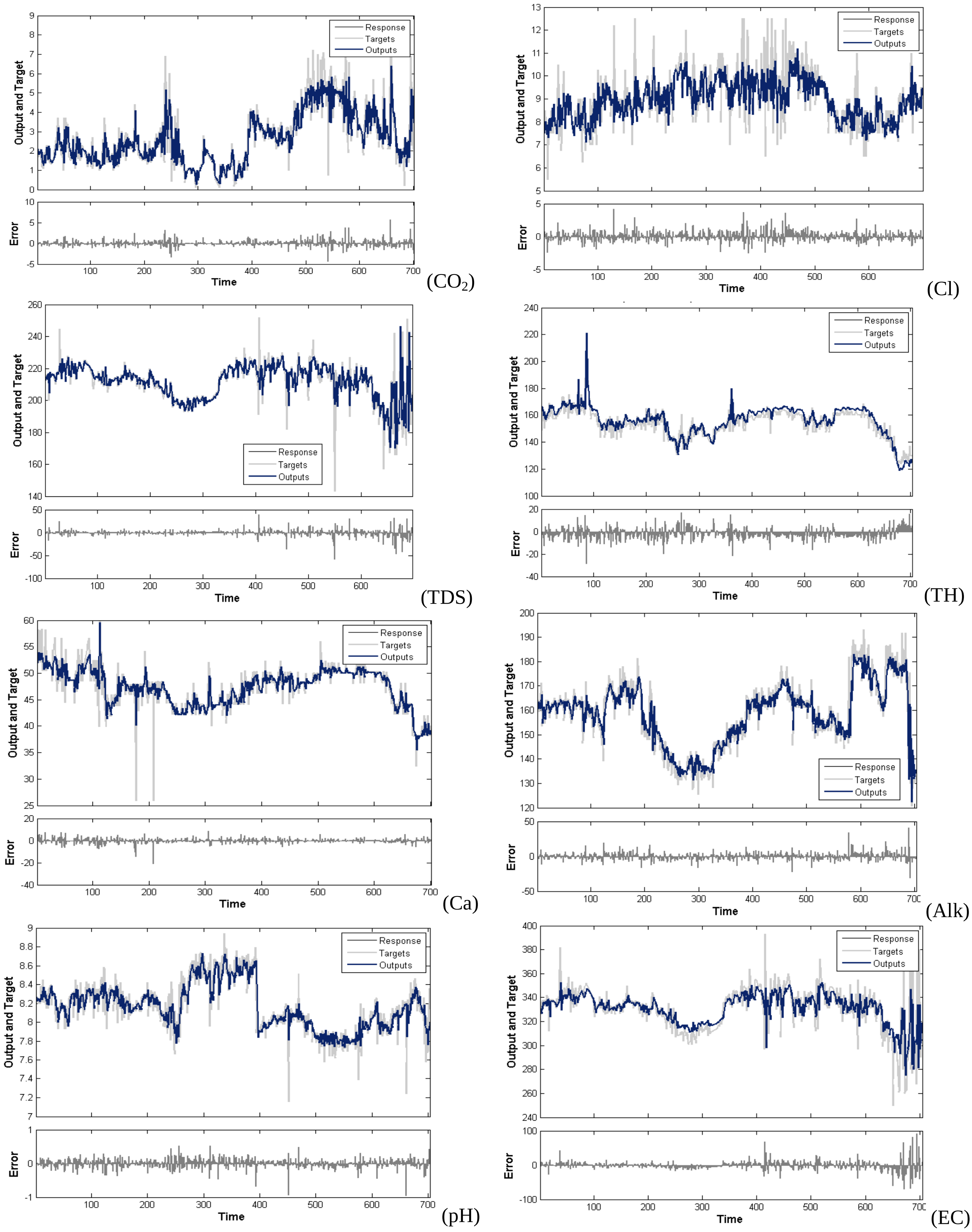
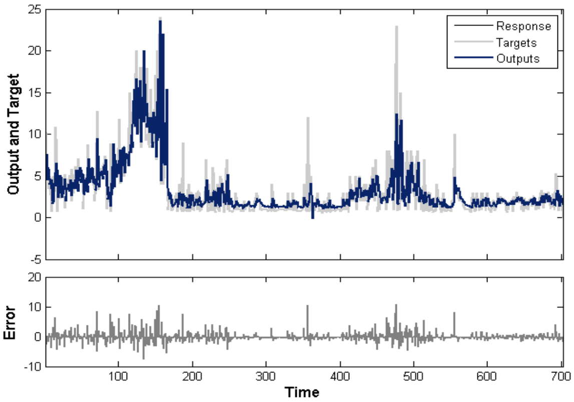


Figure 5. Networks responses and errors
TH: Total hardness; TDS: Total dissolved solids; Alk: Alkalinity; EC: Electrical conductivity; Tur: Turbidity; T: Temperature



(Tur)

(T)

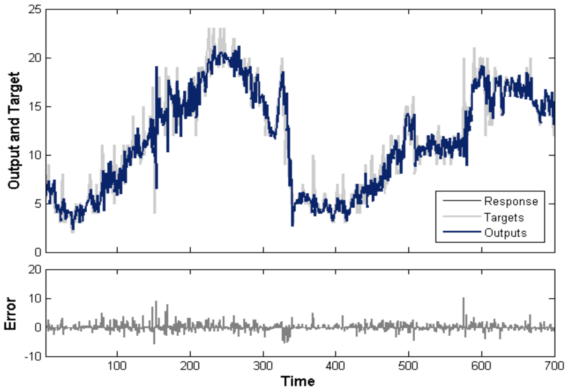


Figure 5. Networks responses and errors (Continue)
TH: Total hardness; TDS: Total dissolved solids; Alk: Alkalinity; EC: Electrical conductivity; Tur: Turbidity; T: Temperature

plant operators and decision makers. The NAR models are robust artificial intelligence models that can be proposed as a useful tool to understand the complex and dynamic nature of influent water characteristic.

Conflict of Interests

Authors have no conflict of interests.

Acknowledgements

The authors thank the Kurdistan University of Medical Sciences for their financial support. The authors would also like to thank the Kurdistan Urban Water and Wastewater Company for the provision of the two years period results of chemical analysis of the influent water of Sanandaj WTP.

References

1. Maest AS, Kuipers JR, Travers CL, Atkins DA. Predicting water quality at hardrock mines: methods and models, uncertainties, and state-of-the-art. Washington, DC: Earthworks; 2005.
2. Witten IH, Frank E, Hall MA. Data mining: practical machine learning tools and techniques. 2nd ed. Burlington, MA: Morgan Kaufmann; 2005.
3. Zhang G, Eddy Patuwo B, Hu MY. Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting 2013; 14(1): 35-62.
4. Chen Y, Yang B, Dong J, Abraham A. Time-series forecasting using flexible neural tree model. Information Sciences 2013; 174(3-4): 219-35.
5. Giordano F, La Rocca M, Perna C. Forecasting nonlinear time series with neural network sieve bootstrap. Computational Statistics & Data Analysis 2007; 51(8): 3871-84.
6. Lapedes A. Nonlinear signal processing using neural networks: Prediction and system modelling. Los Alamos National Laboratory; 1987.
7. Imrie CE, Durucan S, Korre A. River flow prediction using artificial neural networks: generalisation beyond the calibration range. Journal of Hydrology 2000; 233(1-4): 138-53.
8. Wu GD, Lo ShL. Effects of data normalization and inherent-factor on decision of optimal coagulant dosage in water treatment by artificial neural network. Expert Systems with Applications 2010; 37(7): 4974-83.
9. Melessea AM, Ahmad S, McClaina ME, Wangc X, Limd YH. Suspended sediment load prediction of river systems: An artificial neural network approach. Agricultural Water Management 2011; 98(5): 855-66.
10. Huo Sh, He Z, Su J, Xi B, Zhu Ch. Using artificial neural network models for eutrophication prediction. Procedia Environmental Sciences 2013; 18: 310-6.
11. Patil K, Deo MC, Ghosh S, Ravichandran M. Predicting sea surface temperatures in the North Indian Ocean with nonlinear autoregressive neural networks. International Journal of Oceanography 2013; 2013.
12. Delgrangea N, Cabassuda C, Cabassudb N, Durand-Bourlierc L, Lainec JM. Neural networks for prediction of ultrafiltration transmembrane pressure – application to drinking water production. Journal of Membrane Science 1998; 150(1): 111-23.
13. Looney CG. Pattern recognition using neural networks: theory and algorithms for engineers and scientists. New York, NY: Oxford University Press Inc; 1997.
14. Demuth H, Beale M, Hagan M. Neural network toolbox™ 6. user guide. Natick, MA: Math Works; 2008.

15. Hagan MT, Demuth HB, Beale MH. Neural network design. Stamford, CT: Thomson Learning; 1996.
16. Teng CM. Correcting noisy data. Proceedings of 16th International Conference on Machine Learning; 1999 Jun 27-30; San Francisco, USA.
17. Arthur RM. Application of on-line analytical instrumentation to process control. Proceedings of the 1st Annual Conference on Activated Sludge Process Control; 1982;. Chicago, Ill, USA.
18. Tarassenko LA. Guide to neural computing applications. Oxford, UK: Butterworth-Heinemann; 1998.
19. Olssen G, Newell B. Wastewater treatment systems, modelling, diagnosis and control. London, UK: IWA Publishing; 1999.
20. Masters T. Practical neural network recipes in C++. San Diego, CA: Academic Press; 1993.
21. Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. International Journal of Forecasting 2006; 22(4): 679-88.
22. Rumelhart DE, Hinton GE, Williams RJ. Learning internal representations by error propagation. In: Rumelhart DE, McClelland JL, PDP Research Group. Parallel distributed processing: explorations in the microstructure of cognition. Cambridge, MA: A Bradford Book; 1987.
23. Zhu J, Zurcher J, Rao M, Meng MQH. An on-line wastewater quality predication system based on a time-delay neural network. Engineering Applications of Artificial Intelligence, 1998; 11(6): 747-58.
24. Pai TY, Tsai YP, Lo HM, Tsai CH, Lin.CY. Grey and neural network prediction of suspended solids and chemical oxygen demand in hospital wastewater treatment plant effluent. Computers & Chemical Engineering, 2007; 31(10): 1272-81.
25. El-Din AG, Smith DW. A neural network model to predict the wastewater inflow incorporating rainfall events. Water Res 2002; 36(5): 1115-26.
26. Zhang, Q, Stanley SJ. Real-time water treatment process control with artificial neural networks. Journal of Environmental Engineering 1999; 125(2): 153-60.
27. Zhang, Q, Stanley SJ. Forecasting raw-water quality parameters for the North Saskatchewan River by neural network modeling. Water Res 1997; 31(9): 2340-50.
28. Lamrini B, Benhammou A, Le Lann MV, Karama A. A neural software sensor for online prediction of coagulant dosage in a drinking water treatment plant. Transactions of the Institute of Measurement and Control 2005; 27(3): 195-213.