International Journal of Civil Engineering and Technology (IJCIET)
Volume 8, Issue 8, August 2017, pp. 1264–1270, Article ID: IJCIET_08_08_136
Available online at http://http://www.iaeme.com/ijciet/issues.asp?JType=IJCIET&VType=8&IType=8
ISSN Print: 0976-6308 and ISSN Online: 0976-6316

PREDICTION OF RESIDUAL CHLORINE IN A WATER TREATMENT PLANT USING GENERALIZED REGRESSION NEURAL NETWORK

D. V. Wadkar
Research Scholar, DYPIET, and Asst. Prof., AISSMS COE, Pune, India

Dr. A. S. Kote
Professor, DYCOE, Akurdi, Pune, India

ABSTRACT

Generally, in India in water treatment plant (WTP) chlorine dose is decided by plant operators based on their experience, which may result in under-chlorination or over-chlorination. Thus, there is a need to develop predictive models for residual chlorine in a WTP. This research work focuses on applying artificial neural network (ANN) approach to predict residual chlorine in a WTP. Weekly water quality data spanning 4 years was obtained from plant laboratory for modeling with ANN.

Thirty-two ANN models were developed to predict residual chlorine with four combinations of input variables. The ANN models for prediction of residual chlorine were established by several steps of training and testing using feed forward neural network, cascade feed forward neural network, pattern recognition neural network, radial basis neural network and generalized regression neural network. The ANN models were evaluated with error statistics viz coefficient of determination (R^2), mean square error (MSE), mean absolute error (MAE) and standard statistics viz mean($\bar{X}$), standard deviation (σ), skewness ($\gamma 1$) and kurtosis ($\gamma 2$).

Performance of the ANN models increased with the increasing number of input variables such as turbidity, pH and chlorine dose. The best performing model was found to be GRNN with MSE = 0.001 mg/lit, MAE = 0.019 mg/lit and R^2= 0.979. Thus, ANN provides a valuable performance assessment tool for plant operators and decision makers to predict residual chlorine.

Key words: artificial neural network, residual chlorine, turbidity, water treatment plant.

Cite this Article: D. V. Wadkar and Dr. A. S. Kote, Prediction of Residual Chlorine in a Water Treatment Plant Using Generalized Regression Neural Network. *International Journal of Civil Engineering and Technology,* 8(8), 2017, pp. 1264–1270
http://www.iaeme.com/IJCIET/issues.asp?JType=IJCIET&VType=8&IType=8

1. INTRODUCTION

In a water treatment plant (WTP), there are various treatment stages but most important treatment stage is disinfection because it provides assurance of safe water. Generally, chlorine is the most commonly used disinfectant due to its ease of application and monitoring, its low cost and its effectiveness in killing bacteria. The effectiveness of chlorination process mainly depends upon three major parameters namely turbidity of water, pH of water and applied chlorine dose.

Turbidity increases chlorine demand and presents a challenge to maintaining a microbiocidal free available chlorine residue in highly turbid waters (Le Chevallier *et al*., 2004), and the chemical characteristics and effects of turbidity may vary widely across settings. Due to particle association, higher turbidity may also shield microbes from inactivation by chlorine. (Crump *et al*., 2004). Relationship of turbidity with chlorination exhibit non-linear behavior, which is difficult to describe by linear mathematical model. Thus there is a need to develop prediction models for residual chlorine using artificial neural networks (ANN).

Several studies show prediction of residual chorine in water distribution network with various soft computing techniques. Wu *et al*. (2014) developed first-order back-propagation ANN models for prediction of total chlorine and free ammonia in water distribution system with Nash-Sutcliffe efficiencies .Cuesta *et al*. (2014) used the Monte-Carlo calculations in combination with ANN for residual chlorine prediction. Soyupak *et al*. (2012) developed single-input single-output time series ANN model for free residual chlorine forecasting in water distribution networks. Gibbs *et al*.(2006) investigated chlorine decay in water distribution network using multi layer preceptor (MLP) and GRNN. Similarly Bowden *et al*. (2006) forecasted chlorine residuals in a water distribution system using a GRNN and multiple linear regression (MLR). The GRNN achieved a significantly lower error for the training, testing and validation sets as compared to the MLR.

However, literature on prediction of residual chorine at WTP is not available. Generally, in India, chlorine dose in a WTP is determined from operator’s knowledge, which may result in low residual chorine producing bacterial re-growth or excess residual chorine producing objectionable taste and odour. Thus, there is a need to develop predictive models for residual chlorine using ANNs. The developed ANN model will be useful for operators of WTP. This paper presents modeling by ANN for prediction of residual chlorine in a WTP in Pimpri-Chinchwad Municipal Corporation (PCMC), Pune,India.

2. MATERIAL AND METHODS

2.1. Materials

The twin town of Pimpri Chinchwad has population of 1,729,320 and average water supply to the city is 170 lpcd with 59 elevated storage reservoirs and 1, 17,936 water connections. The WTP under study supplies 428 MLD water to an area of 177 km^2. The treatment process includes aeration, coagulation, sedimentation, flocculation, filtration, disinfection, and distribution.

An attempt has been made to predict residual chorine considering water quality parameters: pH, turbidity, chorine dose as input variables in ANN modeling. One of the most important step in the ANN development process is the determination of significant input variables. Usually, not all of the input variables collected will be equally informative, since some may be correlated, noisy, or have no significant relationship with the target variable being modeled. The inclusion of extraneous inputs only serves to increase the computational complexity of the model and make the training process more difficult (Bowden *et al.,* 2006). Therefore, there are obvious advantages in using analytical procedures to select an appropriate set of inputs for ANN models.

Total number of 1521 data samples of inlet and outlet water from WTP *viz* pH, turbidity, chorine dose, residual chorine were considered in the study as these parameters were closely associated with chlorination process. The input variables to different networks were varied from 1 to 4 as listed below:

Group I: One input variables (inlet water turbidity) and one target variable (residual chlorine)

Group II: Two input variables (pH, inlet water turbidity) and one target variable (residual chlorine)

Group III: Three input variables (pH, inlet water turbidity, outlet turbidity) and one target variable (residual chlorine)

Group IV: Four input variables (pH, inlet water turbidity, outlet turbidity, chlorine dose) and one target variable (residual chlorine)

2.2. Methods

Artificial neural networks (ANNs) are biologically inspired systems composed of many simple interconnected elements called neurons, which are organized in input layer, hidden layer and output layer. These layers are connected with weighted connections corresponding to human brain synapses (Haykin., 2009). The neural networks adapt the weights of their hidden neurons built on the input and target data. Inclusion of defective data in the training set will change the mean errors, which are propagated back for weight optimization resulting in inaccurate predictions. Pre-processing of data points was done to ensure the reliability of the input and target data. The development of ANN model requires the partitioning of the parent database into statistically similar subsets in order to calibrate and then validate the model. The data was divided into three sets *namely*, training, validation and testing as 70%, 15% and 15% respectively.

ANN models were developed using modified MATLAB code. Levenberg Marquardt (LM) and Bayesian regularization (BR) training functions were analyzed in various networks, *namely* feed forward neural network (FFNN), cascade feed forward neural network (CFNN), pattern recognition neural network (PRNN) for variations number of epochs, hidden layers and hidden nodes during training of networks. Similarly, radial basis neural network (RBNN) and generalized regression neural network (GRNN) models were analyzed by varying the values of spread factor. In RBNN and GRNN models, spread factor plays an important role in establishing a good ANN regression model with high prediction accuracy and stability. The spread factor values providing the best testing performance of the RBNN and GRNN were equal to 1 and 0.1, respectively (Heddam *et al*., 2011).In this study RBNN and GRNN models were trained by varying value of spread factor (S) from 0.01 to 25.

Further detailed analysis of best performed LM and BR training function was carried out for FFNN, CFNN and PRNN. Determining the optimal number of hidden nodes has always been challenging task in neural network applications and there is no direct or precise way to determine the optimal number of nodes in each hidden layer. Several guidelines have been developed by researchers for approximately determining the required number of hidden nodes in a hidden layer from knowledge of the number of nodes in both the input and output layers (Najjar *et al.,* 1997; Salchenberger *et al.*, 1992; Hajela *et al.*, 1991; Hecht *et al.*, 1989 and Caudill, 1988).

The performance predictions of all ANN models were evaluated for the different ANN configurations with one hidden layer. The number of neurons in the hidden layer was varied from 5 to 100 for FFNN, CFNN and PRNN models for group I, II, III and IV. Several runs were performed for each network structure to prevent wrong selection of initial weights. All the trained neural networks were validated using the validation data set. The model with least validation error was selected and further tested using test data set for computing the final network error.

2.3. Performance Evaluation

Three performance evaluation statistics and four standard statistics were used to test the effectiveness of each model, which included coefficient of determination (R2), mean square error (MSE), mean absolute error (MAE) and standard statistics viz mean($\overline{X}$), standard deviation (σ), skewness (ɣ1) and kurtosis (ɣ2). ANN models with different combinations of input variables for Group I to IV were tested in order to develop a best performing network. Best combination of training function, number of neurons in the hidden layer and the number of epochs for highest R^2 and lowest MSE were determined. Monitoring R2 value is one of the most common methods for comparing the obtained results and a perfect prediction, where R2= 1 is considered as 100% accuracy. On the other hand, MSE and MAE describe a quantitative measure of the average prediction error over entire database.

3. RESULTS AND DISCUSSION

3.1. Analysis of ANN Models

In this study, thirty two residual chlorine neural network models were developed with four combinations of input variables. Each model was trained 10 times and the best performance was evaluated.

3.2. Model Comparison

It can be seen from Figure 1 that, MSE and MAE for all the models is marginally same (MSE = 0.001 to 0.017 mg/lit & MAE= 0.031 to 0.106 mg/lit) whereas R^2 varies from 0.023 to 0.979. However, there is a distinct superiority in prediction with GRNN models. Also, it shows that, models with 4 inputs gives better performance as compared to other input variation to the networks. It was observed that FFNN(4-50-1) with LM training function, PRNN(4-15-1) with BR training function ,CFNN(4-60-1) with BR training function displayed good performance compared with other FFNN , PRNN and CFNN models. Therefore, performance of best GRNN models under Group I, II, III and IV during training and testing phase are shown in Table 1

Table 1 Summary of best ANN models from Group I, II, III and IV during training and testing phase

ANN models	Spread factor	Training Phase			Testing Phase		
		MSE (mg/lit)	MAE (mg/lit)	R^2	MSE (mg/lit)	MAE (mg/lit)	R^2
GRNN 1	0.01	0.022	0.100	0.431	0.009	0.074	0.479
GRNN 2	0.01	0.015	0.082	0.607	0.007	0.068	0.553
GRNN 3	0.01	0.001	0.014	0.969	0.003	0.045	0.964
GRNN 4	0.1	0.001	0.023	0.972	0.001	0.019	0.979

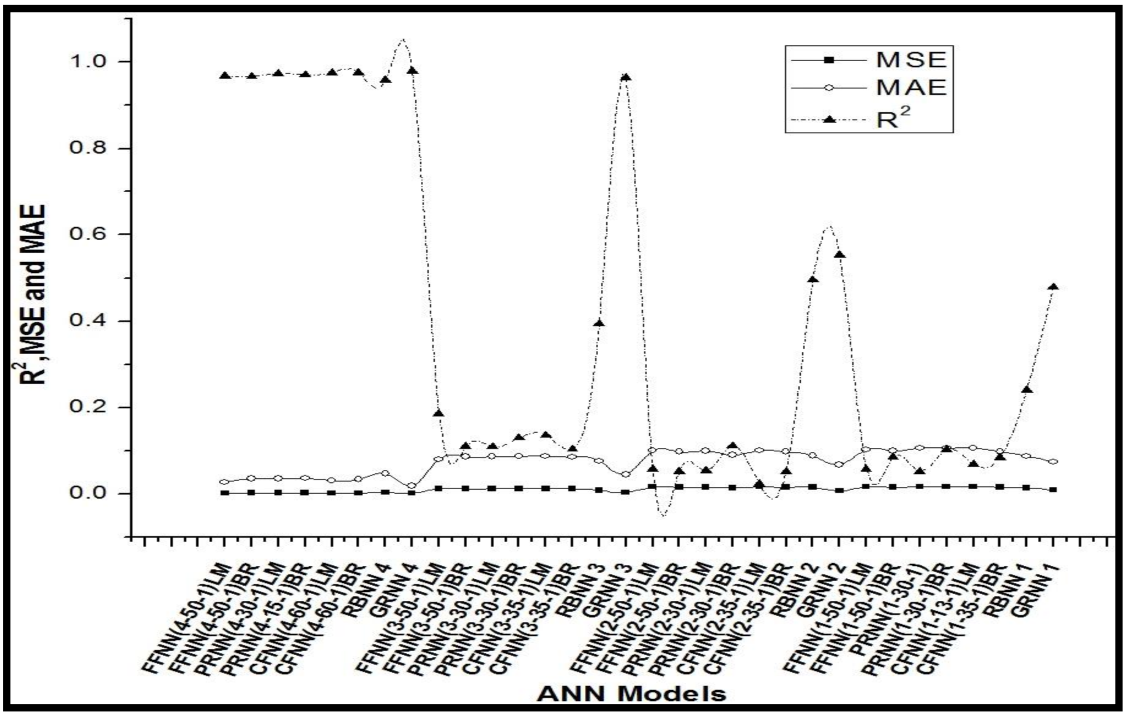

Figure 1 Performance of ANN models of Group I, II, III and IV during testing phase

Amongst all the thirty two models, GRNN 4 model gave the best performance with MSE=0.001 mg/lit, MAE=0.019 mg/lit and R2=0.979 during testing as compared to other ANN models of Group I, II, III and IV. And , it was observed that performance of RBFN models and GRNN models increased as the value of spread factor increased up to 25 and decreases up to 0.01 respectively.

Standard statistics such as $\bar{X}$,σ, ɣ1 and ɣ2 were evaluated for all ANN models and displayed in Table 2 during training and testing. GRNN 4 model showed lowest σ and ɣ2 during training phase and testing phase. It was observed that best ANN models produced positive kurtosis where heavier tails are accompanied by a higher peak for provided common data points. GRNN 4 model with low standard deviation indicates that the data points tend to be close to the expected value of the set, while GRNN 1 had high standard deviation indicating that the data points are spread out over a wider range of values.

Table 2 Summary of standard statistics of GRNN models during training and testing phase

ANN Models	Training				Testing			
	$\bar{X}$	σ	ɣ1	ɣ2	$\bar{X}$	σ	ɣ1	ɣ2
Statistics of observed values	1.701	0.194	1.206	7.862	1.758	0.117	-1.017	4.664
GRNN1	1.701	0.147	-0.992	8.571	1.747	0.096	1.329	4.327
GRNN 2	1.702	0.122	-0.558	7.432	1.757	0.086	0.964	3.835
GRNN 3	1.702	0.034	-0.212	18.202	1.761	0.047	2.409	16.133
GRNN 4	1.728	0.034	0.451	5.939	1.702	0.037	-0.032	2.167

The scatter and time series plot of observed and predicted residual chlorine by GRNN4 model during testing phase is shown in Figure 2 and 3 respectively and it can be seen that the predicted values are very close to observed values which implies best performance of the model

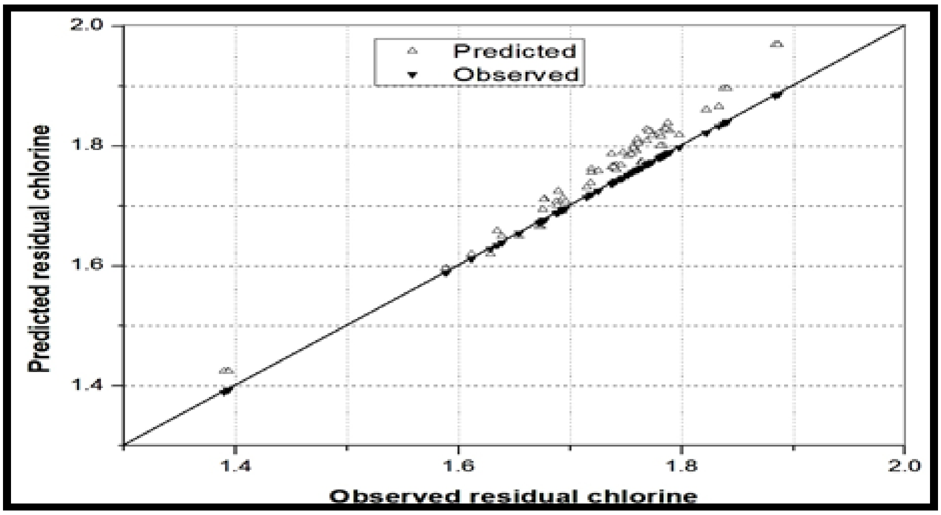

Figure 2 Scatter plot of GRNN 4 model during testing phase

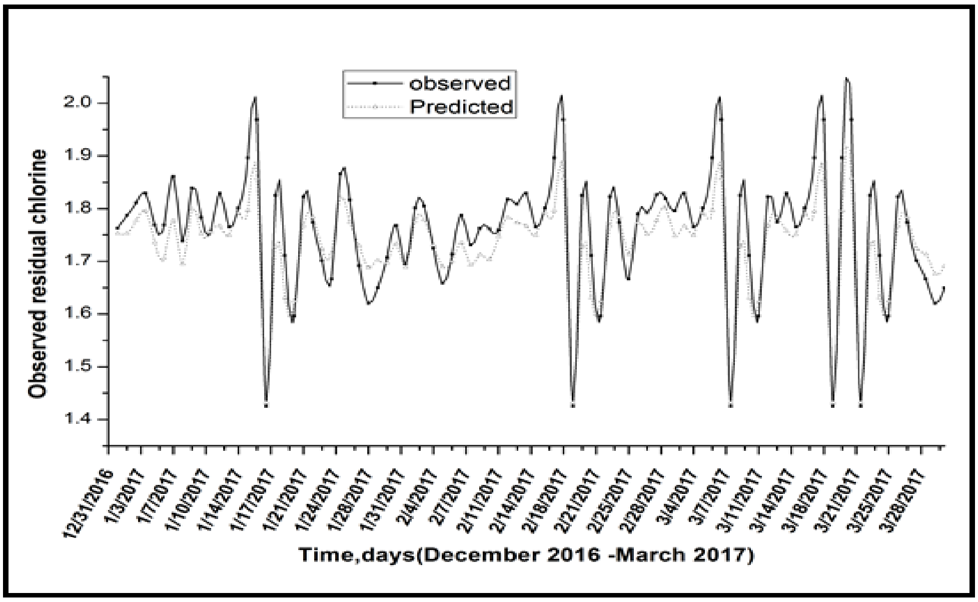

Figure 3 Times series plot of GRNN 4 model during testing phase

4. CONCLUSIONS

This study explores thirty two ANN models for prediction of residual chlorine in a WTP. It is seen that GRNN models showed good prediction for various combinations of input and target variables. All models were developed using measured data, displaying the ability to make accurate predictions using existing data. By comparing performance of ANN models, it was observed GRNN 4 model had highest R^2=0.979 and lowest MSE= 0.001 mg/lit and MAE=0.019 mg/lit. Similarly, it was found that performance ANN models increases as the number of input increases. As we know chlorine dose need to change according to turbidity of water but in actual practice in most of the WTP in India, chlorine dose is kept constant. Therefore, for real time estimation of residual chlorine and chlorine dose in a WTP, GRNN 4 model was more useful as compare to other ANN models. The developed ANN model can be used as a valuable performance assessment tool for decision makers and operations of a WTP.

REFERENCES

[1] Bowden, G. J., Nixon, J.B., and Dandy, G. C. Forecasting chlorine residuals in a water distribution system using a general regression neural network, *Mathematical and Computer Modelling*, **44,** 2006, pp. 469–484.

[2] Caudill, M. 1988. *Neural networks primer, Part III*. AI Expert 3,**6**: 53–59.

[3] Crump, J.A., Okoth, G.O., Slutsker, L., Ogaja, D.O., Keswick,B.H.& Luby, S. P. Effect of point-of-use disinfection, flocculation and combined flocculation –disinfection on drinking water quality in western Kenya., *J. Appl. Microbiol*., **97,** 2004, pp. 225–231.

[4] Cuesta Cordoba, L. and Tuhovka, M. T. Using artificial neural network models to assess water quality in water distribution networks, *Procedia Engineering*, **70**, 2014, pp. 399-408.

[5] Gibbs, M. S., Morgan ,N. and Maier, H.R. Investigation into the relationship between chlorine decay and water distribution parameters using data driven methods , *Mathematical and Computer Modelling*, **44,** 2006, pp. 485–498

[6] Hajela, P., and Berke, L., Neurobiological computational modes in structural analysis and design. *Computers & Structures*, **41**(4),1991, pp. 657–667

[7] Hecht-Nielsen, R. Theory of the back-propagation neural network. *Proceedings of the International Joint Conference on Neural Networks*, **1**, 1989, pp. 593–606.

[8] Heddam Salim , Abdelmalek B. and Noureddine D. 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**, 2011, pp. 1209-1214

[9] Haykin, S.,NeuralNetworksandLearningMachines,2009,3rded. Prentice Hall.

[10] LeChevallier, M. W., & Au, K .K. Water Treatment and Pathogen Control: Process Efficiency in Achieving Safe Drinking Water”. *IWA Publishing*, 2004, London, UK.

[11] Najjar, Y. M., Basheer, I. A., and Hajmeer, M. N. Computational neural networks for predictive microbiology: I. methodology, *International Journal of Food Microbiology*, **34** (1),1997, pp. 27–49

[12] Salchenberger, L. M., Cinar, E. M. and Lash, N. A. Neural networks: A new tool for predicting thrift failures, *Journal of Decision Sciences*, **23**, 1992, pp. 899–916

[13] Soyupak, S., Kilic, H., Karadirek, I. and Muhammetoglu, H. Single Input Single Output Time Series Artificial Neural Network Models for Free Residual Chlorine Forecasting in Water Distribution Networks, *Proceedings of the 4th International Joint Conference on Computational Intelligence*, 2012, pp. 588-593

[14] Wu,W., G. C. Dandy, M.ASCE; and H. R. Maier, Optimal Control of Total Chlorine and Free Ammonia Levels in a Water Transmission Pipeline Using Artificial Neural Networks and Genetic Algorithms, J. Water Resour. Plann. Manage,(2014) DOI: 10.1061/(ASCE) WR.1943-5452.0000486

[15] Dr. Pankaj Singh, Shashwat Sharda and Subhra Singh Cauhan, Domestic Waste Water treatment by fly and wood ash along with additives materials , International Journal of Civil Engineering and Technology, 7(2), 2016, pp. 67 – 75 .

[16] Ali H. Al-Aboodi, Husham T.Ibrahim and Sarmad A.Ab bas, Performance of Sequencing Batch Reactor (SBR) for Domestic Wastewater Treatment Under Low Temperature in Basrah City (South Of Iraq), International Journal of Civil Engineering and Technology , 8(6), 2017, pp. 787–796.

[17] V.Y Katte, J.F.N Seukep, A Moundom, A.S.L Wouatong and K.B. V Kamgang The Effect of Partial Replacement of Waste Water Treatment Sludge on The Properties of Burnt Clay Brick. International Journal of Civil Engineering and Technology , 8(6), 2017, pp. 567–583.