

# Modeling of Chlorine Decay using Artificial Neural Network



S.Vanitha, C.Sivapragasam

**Abstract:** In this paper an attempt is made to model chlorine decay using Artificial Neural Network (ANN). Initial chlorine concentration, fast and slow reacting organic and nitrogenous compounds and reaction rate constants of the compounds are used as inputs to the ANN model and the chlorine decay at different points in the decay curve are evaluated. ANN is trained by two different methods namely single output model and multi output models. Predicted data are compared with observed using correlation coefficient. Result indicates multi output model able to model more accurately than single output model.

**Keywords:** Organic Compounds and Inorganic Compounds, Chlorine decay, Modeling, Artificial Neural Network.

## I. INTRODUCTION

The main objective of water supply system is to supply qualitatively safe water to consumers. The major challenging task for water supply engineers is to maintain water quality till delivery end. However, within the distribution system water is contaminated again due to presence of micro organisms in the distribution system. The quality of water should be maintained using disinfection process to kill pathogenic bacteria and to oxidize the organic matter present in water. The residual or free chlorine should be maintained to protect from future contamination [1].

Many disinfectant methods are available such as chlorination, ozonation, ultraviolet rays etc. to treat the contaminated water [1]. Of all methods, chlorination is most widely used. However, the concentration of chlorine decreases as water travels through distribution system due to bulk decay and wall decay [2]. This is known as chlorine decay. To avoid chlorine decay, sufficient amount of chlorine should be added with water in the distribution system. If the dosing rate of chlorine is too high it produces disinfection by products (DBP) [1]. At the same time, if it is low there may not be sufficient residual to protect the water from future contamination.

A model is required to describe the changes in the quality of water that occurred in the distribution system [2]. The model can be used to ensure that sufficient amount of residual chlorine is maintained in the distribution system.

Traditionally process based models are used for modeling of chlorine decay. The advantage of process based model is that it involves understanding of the process or behavior of chlorine decay process with underlying principles. Process based models require many factors such as flow in the pipes, water temperature, initial chlorine concentration, source water quality and presence of bacterial film. Determining these parameters is the main challenging task in chlorine decay.

In recent past, there is a shift towards modeling using data driven statistical models. In case of limited availability of data, statistical model helps to understand the relationship between independent and dependent variable. The main advantage of this model is that it does not require prior knowledge of the laws of chemistry and mathematics governing the behavior of residual chlorine or the distribution system being studied. However the knowledge of factors that will influence the chlorine decay can help to identify which data are relevant for the analysis. In this study, data driven model namely Artificial Neural Network (ANN) is used to predict the chlorine decay rate for specific initial conditions. The main objective of this study is to find the decay rate of chlorine by having initial free chlorine, fast & slow reacting organic and nitrogenous compounds and the rate reactions of above compounds. ANN is already used widely in various applications and its ability to model complex process is well established [3,4].

### Tools Used

Basically ANN is based on simulation of functioning of brain system to which inputs are fed and outputs are estimated through the number of experimental data. The number of inputs to the neural network depends upon the number of variables assumed to affect the process. A sufficient large data set is required to effectively train neural network architecture. The simplest neural network consists of one input layer, one hidden layer and one output layer. These layers are characterized by number of neurons which are connected through interdependent weights. Each input vector consisting of number of selected input variables is rooted through the ANN architecture and modulated by an activation function to yield the output. These outputs are compared with actual output and error is redistributed through the network system to update the weights appropriately to minimize the error measure. Typically the total data set is divided into three portions namely training, testing and validation set. The number of training set varies from 60-70%. However exact data set is determined by trial and error. Typically ANN is stopped when the error in the testing set data is minimum. Correlation Coefficient (CC) is used to verify the error measurement between predicted and observed data.

Manuscript published on 30 December 2019.

\* Correspondence Author (s)

**Dr.S.Vanitha**, department of Civil Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, India. Email:svanithacivil@gmail.com.

**Dr.C.Sivapragasam\***, Center of water technology, department of Civil Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Tamilnadu, India..Email: sivapragasam@klu.ac.in

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

II. METHODOLOGY

In this study the following inputs are taken

1. Time of initial concentration (p1x)
2. Initial concentration of chlorine (p1y)
3. Turning point of time (p2x)
4. Turning point of chlorine concentration (p2y)
5. Constant value of time (p3x)
6. Constant value of chlorine concentration (p3y).

Figure 1 shows an example of chlorine decay curve. It is desired to model this curve only with three points as indicated in the figure. The point p1x indicates the time of initial chlorine concentration which can be taken as 0. The p1y is the initial concentration of residual chlorine in the distribution system. The input variables selected to model this curve are the initial concentrations of organic and inorganic species and the associated reaction constants. These inputs are detailed in Table 1.

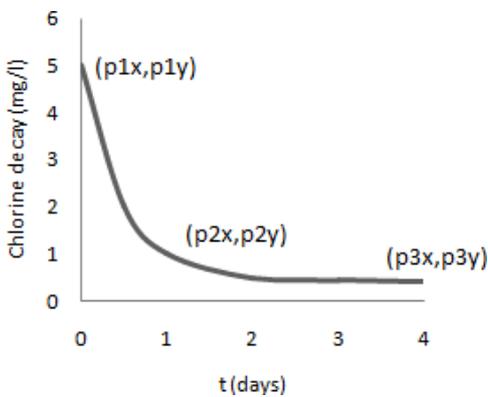


Fig.1 Chlorine decay curve

The input vectors consists of the following variables

Table- I: Inputs and outputs

Input	Output
Initial concentration of free chlorine( $cl_i$ )	Initial concentration of chlorine (point 1y)
Initial concentration of fast reacting nitrogenous compound( $Nf_i$ )	Turning point of time (point 2x)
Initial concentration of slow reacting nitrogenous compound( $NS_i$ )	Turning point of concentration of chlorine (point 2y)
Initial concentration of fast reacting organic compound( $Of_i$ )	Constant value of time (point 3x)
Initial concentration of slow reacting organic compound( $Os_i$ )	Constant of value chlorine concentration (point 3y)
Reaction rate of combined chlorine( $CC_k$ )	
Reaction rate of fast reacting nitrogenous compound( $Nf_k$ )	
Reaction rate of slow reacting nitrogenous compound( $Ns_k$ )	
Reaction rate of fast reacting organic compound( $Of_k$ )	
Reaction rate of slow reacting organic compound( $Os_k$ )	

Out of a total of 278 data, 211 data are taken as training set, 39 data are taken as test set and the remaining 28 data are taken as validation set. Two different ANN models are considered.

Multi output model: In this model, all the output variables are given at once and a single ANN architecture is trained. This is represented in functional form as follows

$$(p_{1y}, p_{2x}, p_{2y}, p_{3x}, p_{3y}) = f(cl_i, Nf_i, NS_i, Of_i, Os_i, CC_k, Nf_k, Ns_k, Of_k, Os_k)$$

Single output model: In this model, the output variables are given one at a time, and a new architecture is designed for each of the output variable which can be functionally represented as follows

$$(p_{2x}) = f(cl_i, Nf_i, NS_i, Of_i, Os_i, CC_k, Nf_k, Ns_k, Of_k, Os_k)$$

$$(p_{2y}) = f(cl_i, Nf_i, NS_i, Of_i, Os_i, CC_k, Nf_k, Ns_k, Of_k, Os_k)$$

$$(p_{3x}) = f(cl_i, Nf_i, NS_i, Of_i, Os_i, CC_k, Nf_k, Ns_k, Of_k, Os_k)$$

$$(p_{3y}) = f(cl_i, Nf_i, NS_i, Of_i, Os_i, CC_k, Nf_k, Ns_k, Of_k, Os_k)$$

III. RESULTS AND DISCUSSIONS

A. Optimal ANN Architecture

For the multi output model, the optimal number of hidden neurons is found to be 22 with tanh as the activation function. The network is trained with initial weight as 0.3. The training is stopped when number of epochs since minimum average error exceeded 200000. However, for single output model the optimal number of hidden neurons is 17 with activation function as logistic function. The training is stopped when number of epochs since minimum average error exceeded 200000.

B. Prediction Results

Correlation coefficient value for Training, Testing and Validation set is shown in Table 2 and 3 for multi output model and single output model respectively. It is seen that the results from multi-output model are better than single output model. Figure 2,3,4,5 and 6 shows the result obtained from multi output model

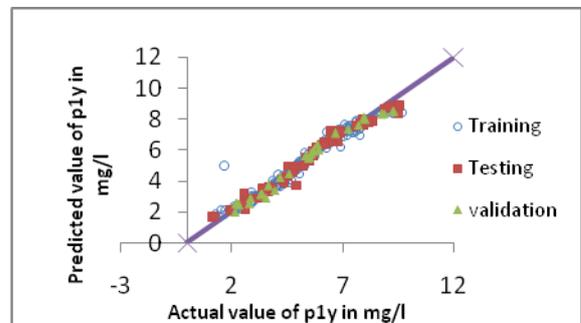


Fig.2 Scatter plot for p1y

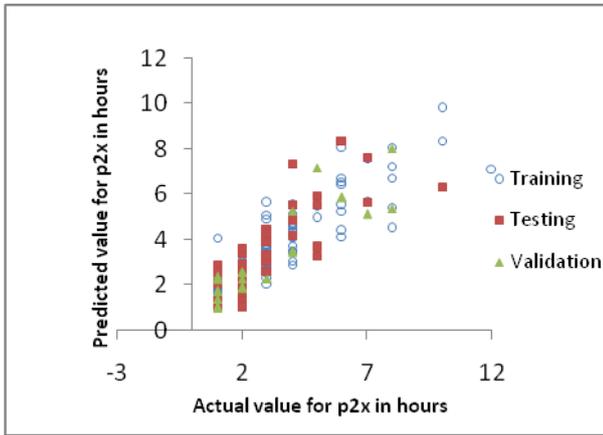


Fig.3 Scatter plot for p2x

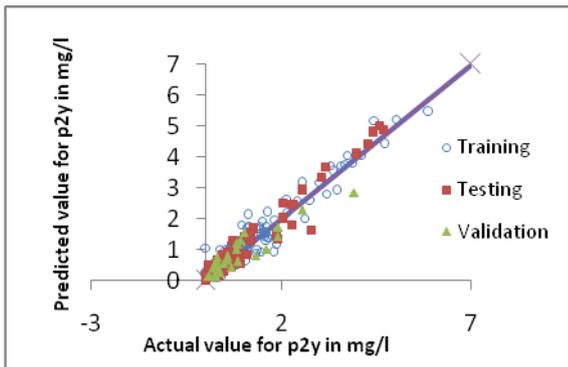


Fig.4 Scatter plot for p2y

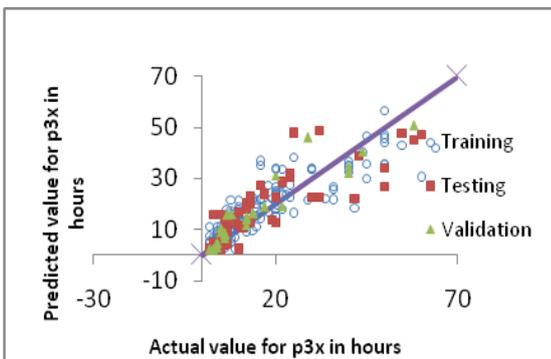


Fig.5 Scatter plot for p3x

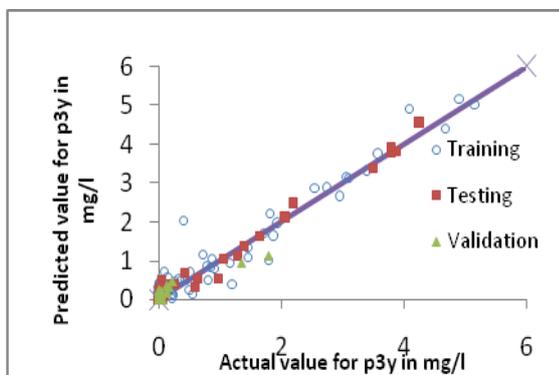


Fig.6 Scatter plot for p3y

Table- II: Multi-output model

Output	Training Correlation value	Testing Correlation value	Validation Correlation value
p1y	0.982	0.979	0.988
p2x	0.900	0.841	0.904
p2y	0.971	0.943	0.914
p3x	0.890	0.855	0.927
p3y	0.980	0.965	0.920

Table- III: Multi-output model

Output	Training Correlation value	Testing Correlation value	Validation Correlation value
p2x	0.908	0.787	0.904
p2y	0.976	0.941	0.895
p3x	0.891	0.901	0.925
p3y	0.985	0.976	0.938

#### IV. CONCLUSION

C The following conclusions can be arrived

- (i) ANN is effective tool for modeling of chlorine decay.
- (ii) In this study only three points are considered for chlorine decay. However better results can be obtained if more points can be taken.
- (iii) It is recommended to model chlorine decay process using additional parameters for getting better insight of the process.

#### ACKNOWLEDGMENT

The authors sincerely thank Nithya. N, Priya. R and Sheema Devi.P for help rendered for doing the analysis.

#### REFERENCES

1. Metclaf and Eddy, "Wastewater Engineering Treatment and Reuse", 4<sup>th</sup> edition, Mc Graw Hill Publishers, 2003
2. S.Palani, S.Liong, P.Tkalich, "ANN application for water quality forecasting", Marine Pollution Bulletin, 56, (2008), 1586-1597.
3. A.M. Kalteh, P. Hjorth, R. Berndtsson Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application, Environmental Modelling & Software, 23, (2008), 835-845.
4. L.Monteiro, D.Figueiredo, S.Dias, R.Freitas, D.Covas, J.Menaia, S.T.Coelho, "Modeling of chlorine decay in drinking water supply systems using EPANET MSX" 12th International Conference on Computing and Control for the Water Industry, CCWI2013, Procedia Engineering, 70, (2014), 1192-1200.

#### AUTHORS PROFILE



**Dr. S.Vanitha** has completed her graduation and post-graduation from Government of College of engineering, Tirunelveli and Arulmigu Kalasalingam College of Engineering. She has completed her PhD in Kalasalingam Academy of Research And Education. She has published more than 20 papers in journals and conferences.



**Dr. C.Sivapragasam** has completed his graduation and post-graduation from IIT-Roorkee and IIT-Delhi respectively. He completed his PhD from NUS, Singapore. He has 2 years of industry experience and more than 17 years of teaching experience. He has completed 3 sponsored research projects from various funding agencies and published more than 70 papers in peer reviewed journals and conferences.