



# Journal of Applied Sciences

ISSN 1812-5654

**science**  
alert

**ANSI***net*  
an open access publisher  
<http://ansinet.com>

## Prediction of Clarified Water Turbidity of Moyog Water Treatment Plant Using Artificial Neural Network

<sup>1</sup>Duduku Krishnaiah, <sup>1</sup>Siva Kumar Kumaresan, <sup>2</sup>Matthew Isidore and <sup>1</sup>Rosalam Sarbatly

<sup>1</sup>School of Engineering and Information Technology, Universiti Malaysia Sabah,  
Locked Bag 2073, 88999 Kota Kinabalu, Sabah, Malaysia

<sup>2</sup>Corporate Dynamics Sdn Bhd, Lot B6.2 Bangunan KWSP, 88000 Kota Kinabalu, Sabah, Malaysia

---

**Abstract:** This study outlines the artificial neural networks application to improve the prediction capability by investigating the effect of data sampling, network type and configuration as well as the inclusion of past data at the neural network input. Multi layered perception and Elman network were used. Validation results using input data based on 5 min and 1 h sampling was compared. It was found that the 1 h sampling yielded better prediction. Different network configurations were also compared and it was observed that although the larger network showed better prediction capability during the training phase, it was the smaller network that demonstrated better prediction in the validation stage. The inclusion of past data into the neural network was also studied. The generalisation degraded as more past data were included.

**Key words:** Coagulation control, artificial neural network, water quality prediction, network validation

---

### INTRODUCTION

Coagulation refers to the agglomeration of suspended particles in raw water through the neutralisation of negative charges inherently present in suspended particles. The coagulant used in the Moyog water treatment plant (Sabah, Malaysia) was in the form of crystallised  $Al_2(SO_4)_3$ . To date, the complexity of the dynamics of coagulation is not fully understood. Preliminary observation of plant data suggests that it is a multivariable and non-linear physicochemical process depended on a range of operating parameters as well as water characteristics. It is also a time-dependent process where the resultant is a function of previous parameters. Clarified water turbidity is a parameter that can be used for immediate on line assessment on the success of coagulation (Holger *et al.*, 2004). The ability to successfully predict clarified water turbidity allows for the adjustment of the plant operating variables to counteract any anticipated quality deterioration.

Current methods to achieve the correct alum dosage for coagulation includes jar tests (Baxter *et al.*, 1999; Joo *et al.*, 200; Yu *et al.*, 2000) look-up tables or empirical equations and operator's experience. All these methods are reactive, based control where deterioration in water quality would have to be detected first before the calculation of the correct chemical dosage can be initiated. This can lead to poor water quality at the plant effluent.

Predicting Clarified Water turbidity (CW Turbidity) allows the detection of potential water quality deterioration before it happens. Operators are then able to manipulate certain variables, specifically chemical dosages (alum, lime and polymer) at the input of neural network and determine whether these values yield desirable future clarified water turbidity. Successful prediction of clarified water turbidity depends on water quality at required regulation standards, reduction of expensive waste coagulation chemicals and to reduce manpower required to control the process. Clarified Water turbidity (CW turbidity) can be used as an immediate indicator on the success of the coagulation process. The main objective of this study was to develop an artificial neural network to predict the clarified water turbidity at the effluent of the pulsator type clarifier at Moyog water treatment plant.

### MOYOG WATER TREATMENT PLANT

This study is based upon the operation of the Moyog water treatment plant, situated 15 km from Kota Kinabalu, Sabah as shown in Fig. 1. The Raw Water (RW) from the Babagon dam is channeled to the aerator for the removal of taste and odour as well as oxidising dissolved manganese. After aeration, the water is added with alum, lime and polymer before going to the pulsator where clarification takes place. The pulsator is the critical unit in

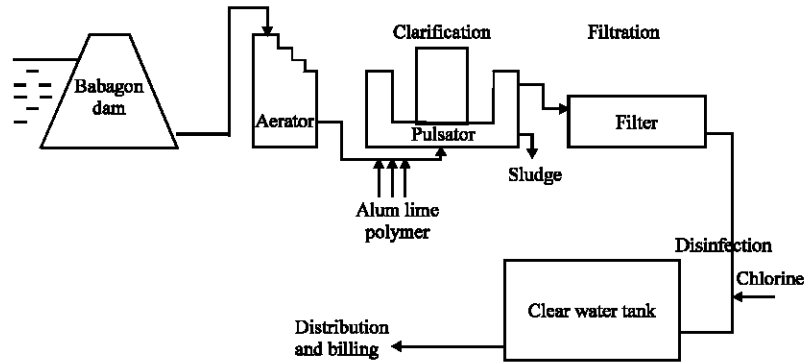


Fig. 1: Moyog plant schematic process diagram

the clarification process as it removes all particulates that have precipitated due to coagulation and flocculation by the addition of alum. The water from the pulsator is known as dosed water (DW). After leaving the pulsator, water is filtered and disinfected before distribution.

### BACKGROUND

Isidore *et al.* (2000) utilised multi layered perception (MLP) for the prediction of CW turbidity. Inputs to the network include:

- RW flow, RW turbidity, RW turbidity (lag 1 h), RW turbidity (lag 2 h)
- RW pH, RW pH (lag 1 h)
- CW turbidity, CW turbidity (lag 0.5 h), CW turbidity (lag 1 h), CW turbidity (lag 2 h)

All of which were sampled at 5 min interval.

The most successful network in predicting CW turbidity in the validation phase had a (10-10-1) architecture with a Mean Absolute Error (MAE) of 4.5% and a correlation coefficient,  $R^2$  of 0.15. The poor  $R^2$  was mainly contributed by the 1 h time lag in the prediction. When the prediction was time shifted forward by 1 h, the MAE and the  $R^2$  improved to 1.9 and 0.84%, respectively.

The next stage of the research was focused on the removal of the time lag. Literature review shows that MLP's with 1 or 2 hidden layers are frequently used for the water treatment application. Performance of networks having 3 or more hidden layers is no different when compared to networks having 1 or 2 hidden (de Villiers and Barnard, 1992). Besides MLP's, hybrid networks used (Evans *et al.*, 1998) where Kohonen self organising map was combined with MLP. Combination of fuzzy logic and neural networks like Adaptive Neural Fuzzy Inference System (ANFIS) (Evan *et al.*, 1998; Jang, 1993; Han *et al.*, 1997), or a hybrid combination of simultaneous fuzzy and neural control were also utilised.

Data chosen as network inputs and required past data inputs were also investigated (2, 3, 5, 6, 7, 10). Inputs included RW turbidity, RW flow, RW temperature, RW conductivity, RW alkalinity, RW apparent colour, RW real colour, RW UV absorbitivity, polymer dosing and alum dosing. It was reported that the temperature is irrelevant to the performance of the network in predicting water colour (Fletcher *et al.*, 2001). Sensitivity analysis was done (Mirsepassi *et al.*, 1995) to determine the minimum critical number of past data required for prediction. Through sensitivity analysis, it was found that the raw water temperature and turbidity were the main determinants in forecasting the Clarified Water Turbidity (CWT) followed by alum dosage.

### METHODOLOGY

Further enhancement for development of an artificial neural network to predict CW turbidity is the main objective of this paper. In particular, special attention was focused on the removal of the one hour time lag in the prediction encountered (Isidore *et al.*, 2000). Therefore, the objectives of the research project were as follows:

- To determine the effect of sampling on neural network predictions.
- To explore the different types of neural networks.
- To optimise the number of past data as network inputs.
- To determine the effectiveness of the network in predicting the difference between the present CWT and the future CWT (i.e.,  $\Delta CWT$ ).

Two main neural networks, a Multiple Layer Perception (MLP) and an Elman network were chosen to predict the clarified water turbidity. The MLP is a feed forward network which was chosen for its ability to predict the model. The Elman network is a recursive

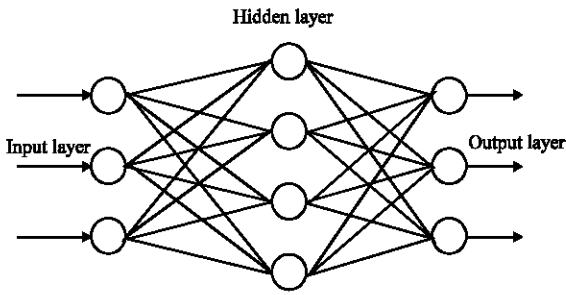


Fig. 2: Multiple layer perception architecture

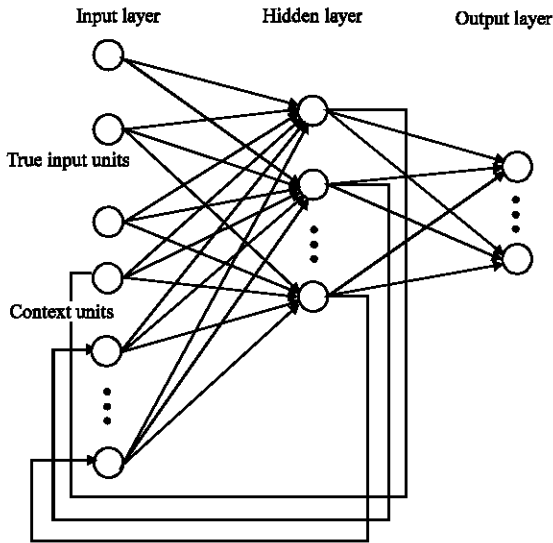


Fig. 3: Elman network architecture

network which was chosen for its ability to due to its conflicting patterns (Skapura, 1996). Figure 2 and 3 show the MLP and Elman network schematics, respectively.

Data from the Moyog water treatment plant database were extracted at 1 h interval for effectiveness in prediction comparison with data sampled every 5 min.

Lastly, the effects of the past number of data included in the inputs on the prediction error were compared.

**RESULTS AND DISCUSSION**

**Effect of sampling period:** Both MLP and Elman network were configured to (4 10 1) with inputs of RW pH, RW flow, RW turbidity and CW turbidity. The network was trained to predict CW turbidity 1 h ahead to 500 epochs with the Levenberg Marquardt (LM) training algorithm. The results of the training and validation with 5 min sampled data and 1 h sampled data is shown in Table 1.

From Table 1, it can be seen that data sampled at 1 h intervals improves the networks performance for the

Table 1: Effects of sampling on network prediction

Network/ sampling	Training MSE	Training R <sup>2</sup>	Validation MSE	Validation R <sup>2</sup>
MLP (5 min)	0.017337	0.828546	0.064081	0.576920
Elman (5 min)	0.017967	0.821630	0.039340	0.703456
MLP (1 h)	0.014976	0.847784	0.061514	0.721815
Elman (1 h)	0.008884	0.912777	0.034331	0.719274

Table 2: Effects of architecture on network prediction

Network/ architecture	Training MSE	Training R <sup>2</sup>	Validation MSE	Validation R <sup>2</sup>
MLP (5 10 1)	0.033593	0.813318	0.045722	0.740173
Elman (5 10 1)	0.033657	0.812924	0.038386	0.783863
MLP (5 20 1)	0.029683	0.837187	0.058717	0.680086
Elman (5 20 1)	0.030321	0.833340	0.077029	0.611138
MLP (5 10 5 1)	0.030028	0.835111	0.049119	0.733349
Elman (5 10 5)	0.029746	0.836812	0.074830	0.639445

training and validation phase. In addition it was observed that the Elman network performed slightly better than the MLP.

**Effect of network type and architecture:** Both MLP and Elman networks were configured with (5-10-1), (5-20-1) and (5-10-5-1) architectures. The networks were trained for the prediction of CW turbidity 2 h ahead using data collected at 2 h intervals. The training was limited to 500 epochs since the training errors did not show any improvements beyond that epoch. LM training algorithm were utilised due to its rapid convergence during training. The network inputs were RW pH, RW flow, CW turbidity, RW turbidity and time (to correlate to temperature). Training and validation results are shown in Table 2.

From Table 2, it can be seen that although the more complex networks have better training, it is the simplest network that provided the best prediction thus confirming to the earlier findings (de Villiers and Barnard, 1992). The difference in performance between the MLP and Elman network is also marginal with the MLP performing better with the simpler network. The time prediction lag was also still present. A typical validation plot with time lags is shown in Fig. 4.

**Effect of past data:** With 3 or 9 sets of past inputs were included into the input of both the MLP and the Elman network as mentioned above. The architecture was set to a single hidden layer with 10 nodes.

From Table 3, it can be seen that more past data does not result in better predictions, although training improved, validation reduced for both networks. This is most likely due to the fact that the retention time between the chemical injection points to the pulsator's effluent was approximately 1.5 h. Parameters earlier than that would have insignificant effect. Besides this, it was also noticed that there was little difference between both network performances.

Table 3: Effects of past data on network prediction

Network/ past inputs	Training MSE	Training R <sup>2</sup>	Validation MSE	Validation R <sup>2</sup>
MLP (17 10 1)				
3 past inputs	0.028303	0.845457	0.057262	0.684916
Elman (17 10 1)				
3 past inputs	0.027001	0.853181	0.067661	0.660984
MLP (41 10 1)				
9 past inputs	0.022505	0.879334	0.058059	0.674098
Elman (41 10 1)				
9 past inputs	0.019321	0.897387	0.09899	0.545906

Table 4: Effectiveness in prediction of ΔCWT(2 h)

Network/ architecture	Training MSE	Training R <sup>2</sup>	Validation MSE	Validation R <sup>2</sup>
MLP (8 10 1)	0.051857	0.820835	0.101206	0.554470
Elman (8 10 1)	0.051637	0.805474	0.222293	0.489172
MLP (8 10 5 1)	0.055552	0.810375	0.101204	0.619331
Elman (8 10 5 1)	0.054793	0.812459	0.086852	0.630233

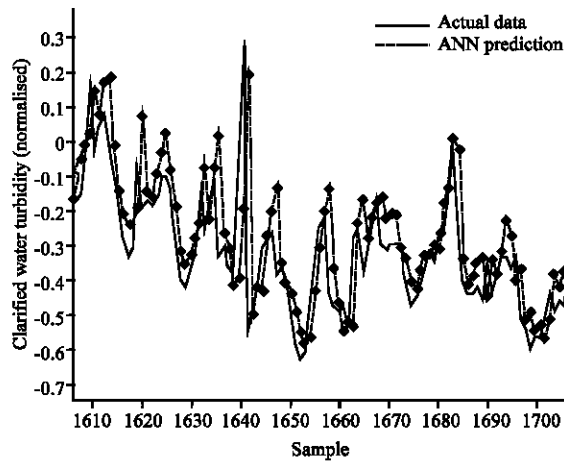


Fig. 4: Validation plot MLP (10 5 1)

**Predicting ΔCWT:** MLP and Elman type network were utilised to predict ΔCWT (2 h) where,

$$\Delta CWT(2\text{ h}) = CWT(2\text{ h}) - CWT(0\text{ h}) \quad (1)$$

Inputs to the network also includes the present differences of RW pH and turbidity and ΔCWT (0 h) that are defined as,

$$\Delta X(0\text{ h}) = X(0\text{ h}) - X(-2\text{ h}) \quad (2)$$

The output of the network would then be used to calculate CWT by unnormalising ΔCWT (2 h) and adding it to CWT (0h). Then (8 10 1) and (8 10 5 1) architectures were used. Training and validation results are shown in Table 4. It can be seen that double hidden layered ANN (MLP and Elman) generally performs well compared to their single hidden layered counterpart. The time lag however was still observed. Over all MLP and Elman networks performed similarly although the Elman network required longer calculations.

## CONCLUSIONS

The complexity of the process dynamics made artificial neural networks highly suitable to predict clarified water turbidity, a resultant of the coagulation process. The Elman network showed no significant difference when compared to MLP. In addition, Elman network required more calculations due to the inclusion of the context units. It was also demonstrated that the past inputs played no major role in improving the networks performances. Thus, future works would only make use of data at n-th hour to predict CW turbidity at (n+τ)-th hour where n, τ are positive integers.

## ACKNOWLEDGMENT

The authors wish to acknowledge the staff at Corporate Dynamics Sdn Bhd Moyog plant, Sabah, Malaysia.

## REFERENCES

Baxter, C.W., S.J. Stanley and Q. Zhang, 1999. Development of a full scale artificial neural network model for the removal of natural organic matter by enhanced coagulation. *J. Water SRT-Aqua.*, 48: 12-136.

de Villiers, J. and E. Barnard, 1992. Back propagation neural nets with one or two hidden layers. *IEEE Trans. Neural Networks*, 4: 136-141.

Evans, J., C. Enoch, M. Johnson and P. Williams, 1998. Intelligent based auto-coagulation control applied to a water treatment works. *UKACC International Conference on Control*, 98: 141-145.

Fletcher, I., A. Adgar, C.S. Cox and T.J. Boehme, 2001. Neural network applications in the water industry. *A DERA/IEE Workshop on Intelligent Sensor Processing*, pp: 1611-1616.

Han, T., E. Nahm, K. Woo, Kim and C. Ryu, 1997. Optimization of coagulant dosing process in water purification system. *Proceedings of the 36th SICE Annual Conference International Session Papers*, pp: 1105-1109.

Holger, R.M., N. Morgan and C.W.K. Chow, 2004. Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters. *Environ. Modeling Software*, 19: 485-494.

Isidore, J.M., S. Kumaresan and Abdul Noor, 2000. Enhanced coagulation performance through artificial neural networks. *Proceedings of the Symposium of Malaysian Chemical Engineers, Putra Jaya, Malaysia*, pp: 453-460.

- Jang, J., 1993. ANFIS: Adaptive network based fuzzy inference systems. *IEEE Trans. Systems, Man Cybernetics*, 23: 665-685.
- Joo, D.S., D.J. Choi and H. Park, 2000. The effects of data preprocessing in the determination of coagulant dosing rate. *Water Res.*, 34: 3295-3302.
- Mirsepasi, A., B. Cathers and H.B. Dharmappa, 1995. Application of artificial neural networks to the real time operation of water treatment plants. *IEEE Proceedings on Neural Networks*, 1: 516-521.
- Skapura, D.M., 1996. *Building Neural Networks*. ACM Press, New York.
- Yu, R.F., S.F. Kang, S.L. Liaw and M.C. Chen, 2000. Application of artificial neural network to control the coagulant dosing in water treatment plant. *Water Sci. Technol.*, 42: 403-408.