African Journal of Biotechnology Vol. 10(81), pp. 18755-18762, 16 December, 2011
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.2748
ISSN 1684–5315 ©2011 Academic Journals

Full Length Research Paper

Simulation of ammoniacal nitrogen effluent using feedforward multilayer neural networks

Mohammed S. Jami, Mustapha Mujeli* and Nassereldeen A. Kabbashi

Bioenvironmental Research Group (BERU), Department of Biotechnology Engineering, International Islamic University Malaysia, 50728, Kuala Lumpur, Gombak, Malaysia.

Accepted 7 November, 2011

Ammoniacal nitrogen in domestic wastewater treatment plants has recently been added as the monitoring parameter by the Department of Environment, Malaysia. It is necessary to obtain a suitable model for the simulation of ammonical nitrogen in the effluent stream of sewage treatment plant in order to meet the new environmental laws. Therefore, this study explores the robust capability of artificial neural network in solving complex problems, which are similar to physical, chemical and biological conditions of wastewater treatment plant. Data obtained from Bandar Tun Razak Sewage Treatment plant was used for the model design. The simulation of ammoniacal nitrogen in the effluent stream by model shows a satisfactory result because the mean square error and correlation coefficients were 0.1591 and 0.7980, respectively.

Key words: Ammoniacal nitrogen, wastewater treatment plants, artificial neural network.

INTRODUCTION

The amount of ammoniacal nitrogen (NH_3-N) discharged by domestic and industrial sewage treatment plants has a huge contribution in polluting rivers and subsequent impairment to ecological system of the water bodies. Eutrophication of lakes, rivers and estuaries caused by the discharge of nutrients such as nitrates and phosphates through fertilizers and/or sewages contributes to the growth of phytoplankton that depletes water bodies' oxygen; fish and other oxygen dependent organisms are suffocated and killed. In extreme cases, an anaerobic condition is favoured; it promotes the growth of microorganisms such as *Clostridium botulinum* that produces toxins deadly to birds and mammals. These rivers provide the main source of drinking water (about 98%) and will remain for a long time (Liew, U.S Department of Commerce, 2007).

Biological nitrogen removal in sewage treatment plants consists of nitrification and denitrification processes. Nitrification is a two-step process: Initially, *Nitrosomas* sp., bacteria oxidize ammonia to nitrite and then *Nitrobacter* sp., catalyze the nitrite to nitrate under aerobic conditions (Silver et al., 2002). In the denitrification process, heterotrophic bacteria convert the nitrate to molecular nitrogen under anoxic condition. A sound strategy is required for simultaneous removal of organic and nitrogen compounds from the treatment systems; such as appropriate placement of nitrification/ denitrification processes and proper control of reactor variables such as hydraulic retention time, temperature, waste activated sludge (WAS), return activated sludge (RAS) and organic-microorganisms ratio (F/M). Pollutants removal optimization is necessary in order to meet the current effluent permit and minimum operation cost.

The task of industrial and domestic treatment systems frequently expand in view of the emerging chemicals and pharmaceutical products that end up in the wastewater streams. These are the fundamental causes of incessant addition of new components to the common effluent requirements of present wastewater treatment plants (WWTPs). Recently, NH_3-N and nitrate-nitrogen (NO_3-N) were included as effluent quality parameters under the amended Environment Quality (Sewage) Regulations (2009) of Malaysian Environmental Quality Act 1974 as presented in Table 1 (Environmental Quality (Sewage) Regulations, 2009). To withstand the effluent standard

*Corresponding author. E-mail: mjlmustapha@yahoo.com. Tel: +601-7356-6083.

Table 1. Acceptable conditions for nitrogen compounds for effluents discharge of WWTP standard A and B (Malaysia).

Parameter	Standard (mg/L)	
	A	B
Ammoniacal nitrogen (enclosed water bodies)	5.5	5.0
Ammoniacal nitrogen (river)	10.0	20.0
Nitrate-nitrogen (river)	20.0	50.0
Nitrate-nitrogen (enclosed water body)	10.0	10.0

Malaysia Department of Environment {Environmental Quality (Sewage Regulations, 2009)}.

Table 2. Design influent and effluent quality for Bandar Tun Razak (STP).

Items	Influent concentration (mg/L)	Effluent concentration (mg/L)
BOD_5	220	20
Suspended solids (SS)	270	40
Total nitrogen (T-N)	40	10

BOD, Biochemical oxygen demand.

regulated by the environmental agencies is very challenging due to the complex nature of WWTP that integrates the physical, chemical and biological conditions. Maintaining effluent quality of any WWTP is possible by developing a model to simulate the plant performance using the important parameters of the system (Mjalli et al., 2007) for process control. Recently, many authors acknowledged the benefits of multilayer feedforward neural network (MFNN) for simulation of ecological systems, such as WWTP parameters (Hamoda et al., 1999; Hamed et al., 2004; Côté et al, 1995).

Models for prediction of nitrogen removal performance such as ammoniacal nitrogen were not given much attention because Department of Environment in the earlier years have not set an effluent discharge limit. Therefore, as nitrogen compounds were included in the effluent discharge limit, there is a need to have a model that represent the ideal situation of both nitrification and denitrification. This study explored the advantage of artificial neural network (ANN) to develop a model for the prediction of NH_3-N in the final effluent of sewage treatment plant located in Kuala Lumpur, Malaysia.

MATERIALS AND METHODS

A case study

Bandar Tun Razak Sewage Treatment Plant (STP) is among the treatment plant managed by Indah water konsotium (IWK), located at Jalan 11/118B, Desa Tun Razak, southeast of Federal Territory of Kuala Lumpur, Malaysia. The plant occupied an area of ten acres with about six acres reserved area. Initially, the plant was an oxidation pond with two pump stations that handle the population equivalent (P.E) of 35,000. Later, it was modified by equipping the ponds with surface aerators. Finally, the sequential batch reactor (SBR) plant was built to overcome the situation of raw sewage overflow. The current system was designed to serve a population equivalent of 200,000 and daily design influent of 50,000 m^3/day. The present average inflow of raw sewage to the plant is 14,500 m^3/day. Table 2 represents the design quality of influent and effluent parameters.

Sequential batch reactor (SBR)

As stated earlier, the plant is a mechanized sequential batch reactor for biological removal of carbonaceous and nitrogen compounds. The usage of SBR in wastewater treatments is receiving remarkable attention because of the additional benefits accompanying its application. SBR is a modified technology of activated sludge treatment process that operates in time series rather than space sequence. Apart from the preliminary treatment, all the remaining processes were taking place in a single tank reactor in contrast with the conventional activated sludge system. Feeding, reaction (aerobic, anoxic and anaerobic), sedimentation and discharge carried out in a single tank are illustrated in Figure 1.

The SBR has been widely accepted due to its ability to accommodate varying flow rates with optimum pollutants removal. According to de Sousa et al. (2008), an adequately designed SBR may reach higher removal of carbonaceous matter and suspended solids as well as better nitrification. Activated sludge system initially designed by Arden and Lockett (1914) was actually a variable-volume reactor (fill and draw batch reactor), but it was later modified to continuous stirred tank reactor (CSTR) for steady flow wastewater treatment (Mahvi et al., 2008; Wilderer et al., 2001). The variable-volume suspended growth activated sludge system was restored in 1970s when Irvine named his variable-volume reactor the SBR in 1967 (Wilderer et al., 2001).

Data collection

The data obtained from Bandar Tun Razak STP as earlier stated was used for the development of the model. About four years data was collected from the plant database, which comprises the influent/effluent parameters namely: BOD, COD, SS, NH_3-N, pH and influent flow rate (Q). There is no possibility of varying the sam-

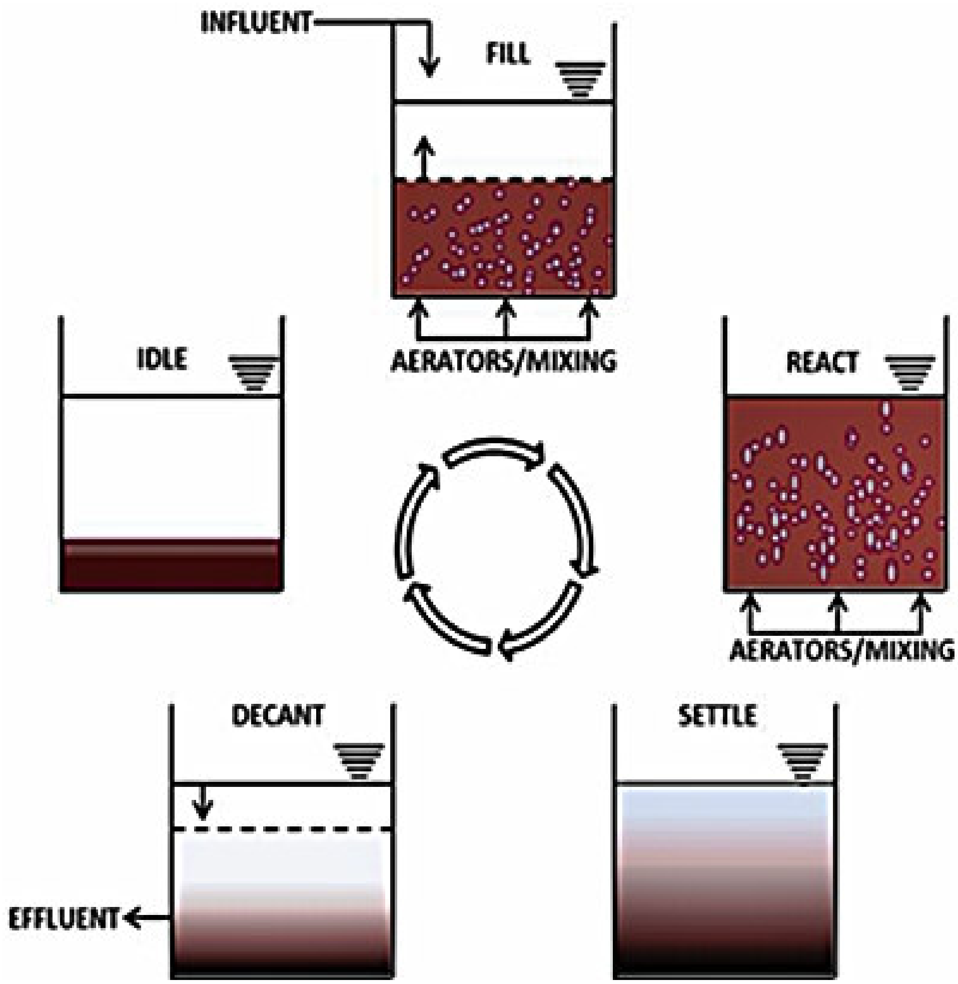

Figure 1. Typical phases of SBR operation.

ple locations because the plant sample points were authorized by the Department of Environment (DOE) on the weekly sampling for laboratory analysis to ensure that the plant meet the DOE requirement. Although, it might not affect the outcome of the model because Hamed et al. (2004) varied the sample locations and found that models with higher coefficient of determination were those in which the samples points were at the preliminary settling tank (PST) and final clarifier tank (FCT) as inputs and outputs parameters, respectively. The fortunate thing about this current research is the size of the data and the period in which they were collected when compared with the previous studies. The advantage of this study is the volume of data collected when compared with other previous researchers. They often complain on insufficient data and its consequences on decreasing model performance. The line plot of influent/effluent parameters: NH_3-N, BOD and SS for the operational period of the plant are illustrated in Figure 2. The figure demonstrates some seasonal pattern in the influent strength of all the parameters apart from a particular day where both BOD and SS concentrations were extremely high above the expected level. The removal rate of BOD and SS revealed in Figure 2 shows a cyclic behaviour throughout the plant operation time. The only exception was that of NH_3-N removal performance towards end of the data sequence with minimal or totally no removal of ammonical nitrogen. Inappropriate timing between anoxic and oxic stages in the biological reactor might be responsible for the instability on the removal rate of ammoniacal nitrogen.

Model development

Artificial neural network modelling requires trial and error procedure to obtain an optimum network structure. However, prior to training commencement, the data were pre-processed to remove any possible error.

Data pre-processing

The errors in the data may be in form of outliers or missing data points. A code was drafted in an M-file format of Matlab® and executed to remove sample points that were not within the range of ±3σ. In order for the network to able to learn fast, the data were standardized to have a zero mean and unit standard deviation using Equation 1:

$$y = \frac{x \quad \bar{x}}{\sigma} \tag{1}$$

Where, y is the scaled data point; x is the initial data point; $\bar{x}$ is the mean of the sample and σ is the standard deviation of the sample.

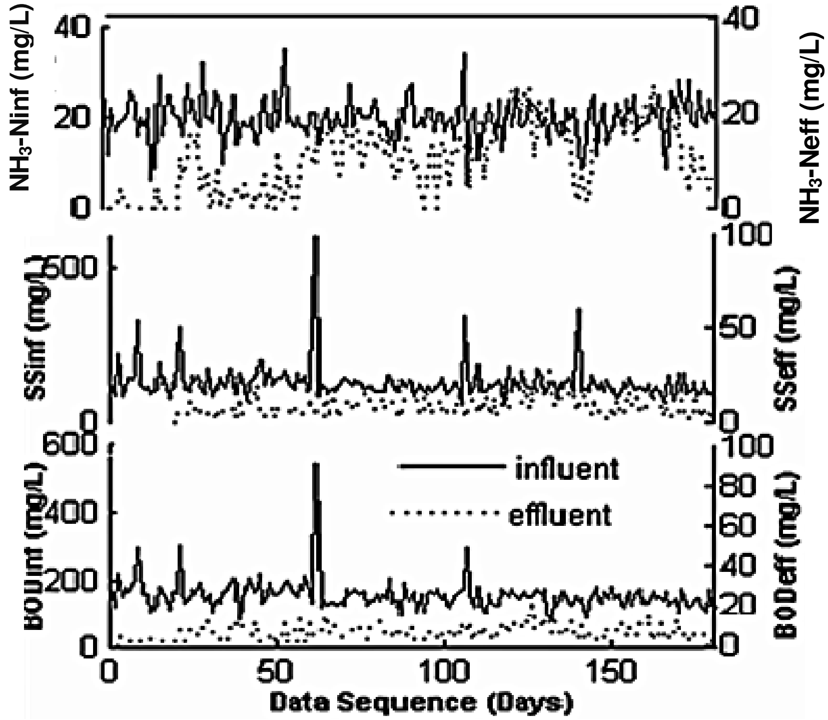

Figure 2. Raw data of influent and eflluent parameters of Bandar Tun Razak STP.

Network training

The topology of the network is the general multi-layer feedback neural networks (FBNN) because it is the most used structure. Therefore, the trial and error techniques were used to search for an optimum structure of the network architecture. Sensitivity analysis of the input parameters was the first study performed in order to optimize the structure of the network. It was performed using combination of the six input parameters until 41 combinations were trained, and each of the combination was tried ten times. The performance of the networks was measured using the mean square error and correlation coefficient for each training. The activation functions in the hidden and output layers of the network also affect the performance of the network. The sigmoidal functions such as logsig and tansig were the most common activation functions used in the hidden layer, while the linear activation function in the output layer (Haykin, 1999).

The number of neurons in the hidden layer also played an important role in determining the network performance in order to overcome overfitting and underfitting of the network. The former has been the condition of perfect training and poor generalization and may occur when the complex network try to fit the noise and the latter condition occurs when non-complex network fails to detect difficult situation in the data set (Mjalli et al., 2007). In this study, early stopping technique was utilized to prevent the occurrence of overfitting. Underfitting normally has been solved by providing large data for training and validation. Using the early stopping criteria, the data are divided into three sets: training, validation and testing. The training and validation data are used simultaneously during training, while the testing data is used for generalization. The training and validation error at the beginning of the training reduces but when the network start overfitting, the training error continues to decrease while the validation error begin*s* to rise. Therefore, few iterations were chosen in order to stop the training (default six) when the validation error starts increasing.

RESULTS AND DISCUSSION

The sensitivity analysis of six input variables: BOD_{inf} (influent BOD), COD_{inf} (influent COD), SS_{inf} (influent SS), NH_3-N_{inf} (influent NH_3-N), pH_{inf} (influent pH) and Q_{inf} (influent Q) established that only four parameters (BOD_{inf}, NH_3-$N_{inf,}$ pH_{inf} and Q_{inf}) have major significance on the output, that is, NH_3-N_{eff} (effluent NH_3-N). The combinations have the least mean square error (mse) and maximum coefficient of determination (R).

The results for the activation functions combination between the hidden layer and output layer of the network shows that both the sigmoidal (logsig) and the hyperbole tangent (tansig) in the hidden layer performed well when the linear (purelin) was the output transfer function. The plot of the optimization results of the transfer function combinations in the hidden and output layer are illustrated in Figure 3. The tansig-purelin and logsig-purelin has the best result of R = 0.9190, mse = 8.7 and R = 0.9221 and mse = 8.8, respectively. They are the two most common combinations used in neural network training (Mjalli et al., 2007; Hamed et al., 2004).

From Figure 4, it was clearly shown that 15 numbers of neurons in the hidden layer outperformed others for both the training and testing performance of the networks. The mean square error and correlation coefficient for the training is mse = 0.3532 and R = 0.7545, whereas it was mse = 0.3937 and R = 0.7212 for testing. The correlation coefficient of training and testing differ only by 4.41%, while that of mean square error is 10.29%.

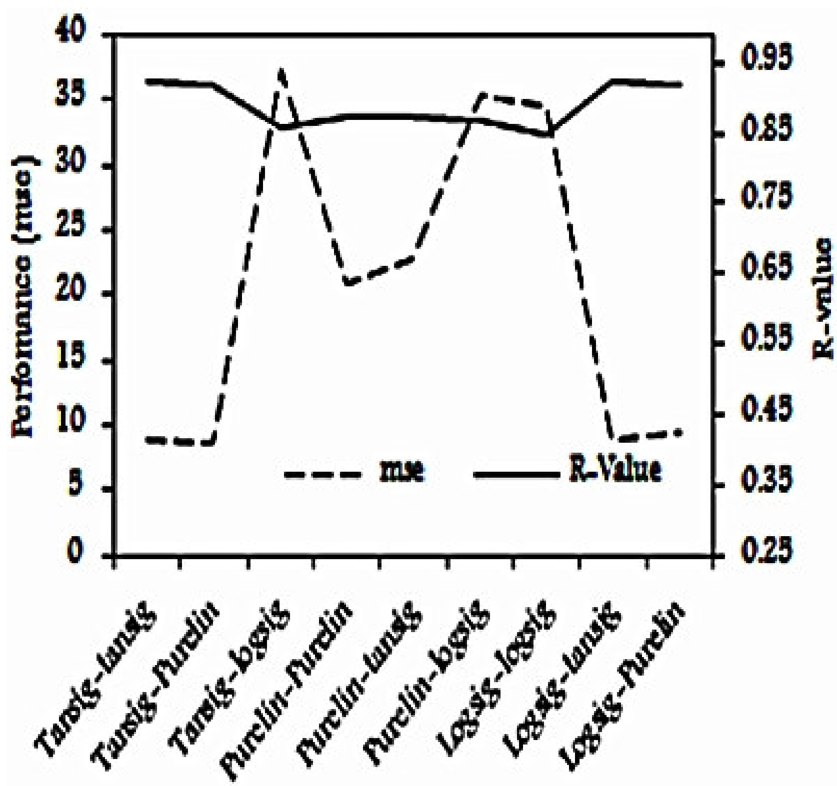

Figure 3. The optimized transfer functions in the hidden and output layer.

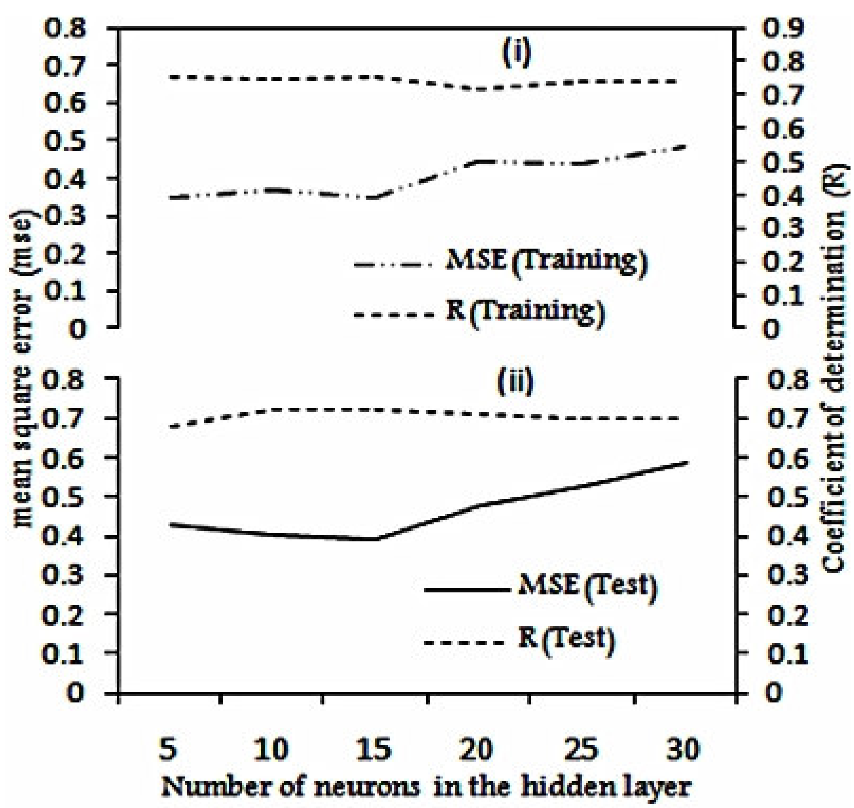

Figure 4. Plot of an optimum number of hidden neurons in the hidden layer.

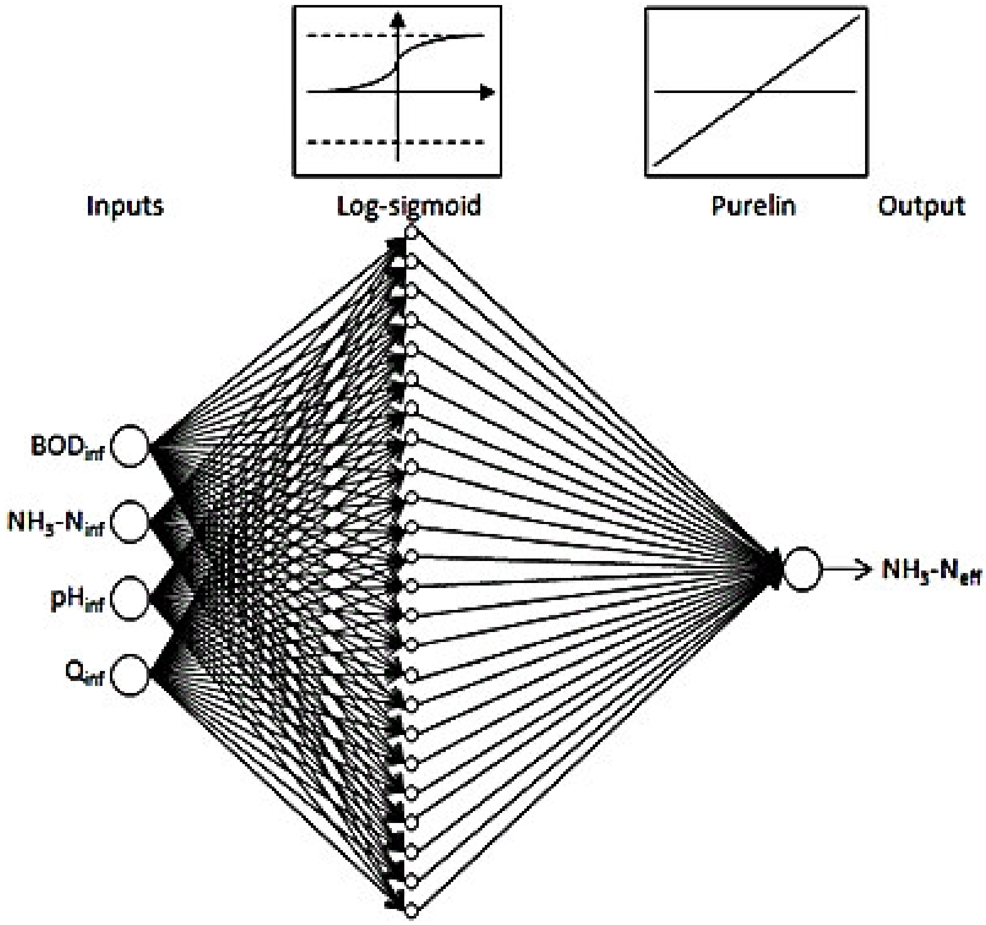

Figure 5. The structure of the best neural network.

After the entire critical functions of the neural network were optimized, the optimum functions of all the network components were assembled together for real model development. Figure 5 illustrates the structure of the best network structure that comprises the elements earlier optimized. The architecture consists of four input variables ($BOD_{inf,}$ NH_3-$N_{inf,}$ pH_{inf} and Q_{inf}) and one output variable (NH_3-N_{eff}). The hidden layer contains 15 neurons with log-sigmoid transfer function for each of the neurons. The transfer function for the output layer is purelin. Figure 6 shows the fitting result of the developed model, whereby the model output fit the target data very well for the training and testing data. The regression plot between the target and output for training and testing of the model is shown in Figures 7 and 8, respectively. The correlation coefficient (R) was 0.8283 and 0.7980 for the training and testing of the model. The mean square error of the trained model was 0.1274 and the model was simulated with a new set of data and the mean square error obtained was 0.1591. The model simulated about 79.80% of ammoniacal nitrogen characteristics in the sewage treatment plant and the result was satisfactorily due to the nature of the system complexity.

BOD has represented the amount of organic matter characterized by BOD, COD, and SS, thereby eliminating the redundancy within the input variables to the network. The influent ammoniacal nitrogen represents the nutrient required by the microorganisms to metabolize the organic matter. The pH and wastewater flow rate (Q) characterize the physical parameters that have high influence in varying other parameters strength, directly affect the symbiotic of the microbes and the system, and consequently disturb the performance of the wastewater treatment system. The proper balance among the numerous input variables make it essential for the success of the model simulation. The model result was excellent when compared to the work of Tezel and Sinan (2010). They found a correlation coefficient for NH_4-N between the simulated and actual output to be -0.02, and a mean square error of 10.65 and 12.48 for training and testing, respectively.

Conclusion

Wastewater treatment plant influent and effluent data were collected and analysed using multilayer feedforward artificial neural network. During the study time, many

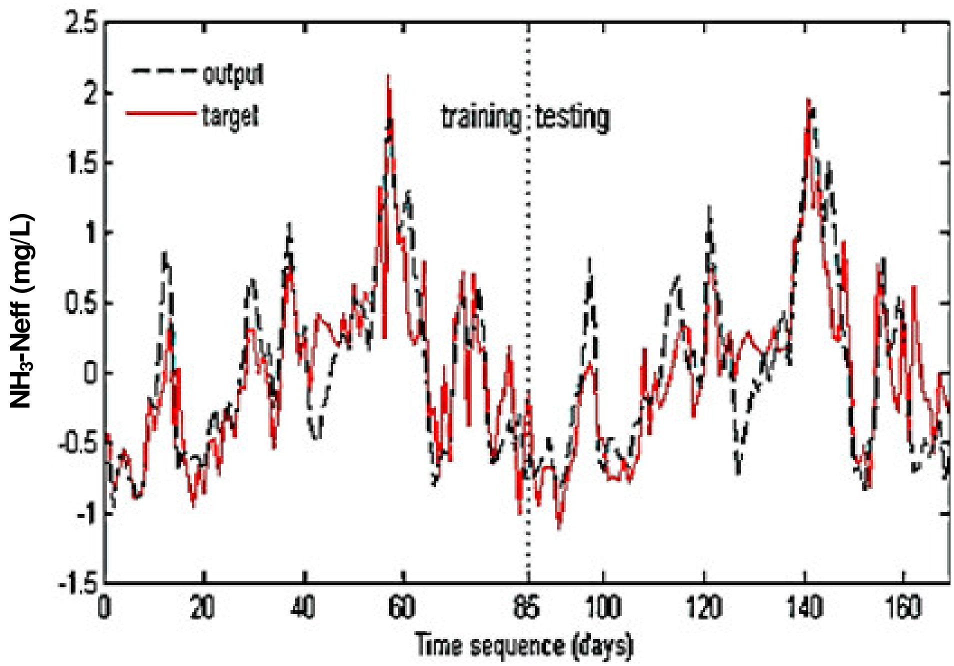

Figure 6. Plot of target and output for training and testing data samples.

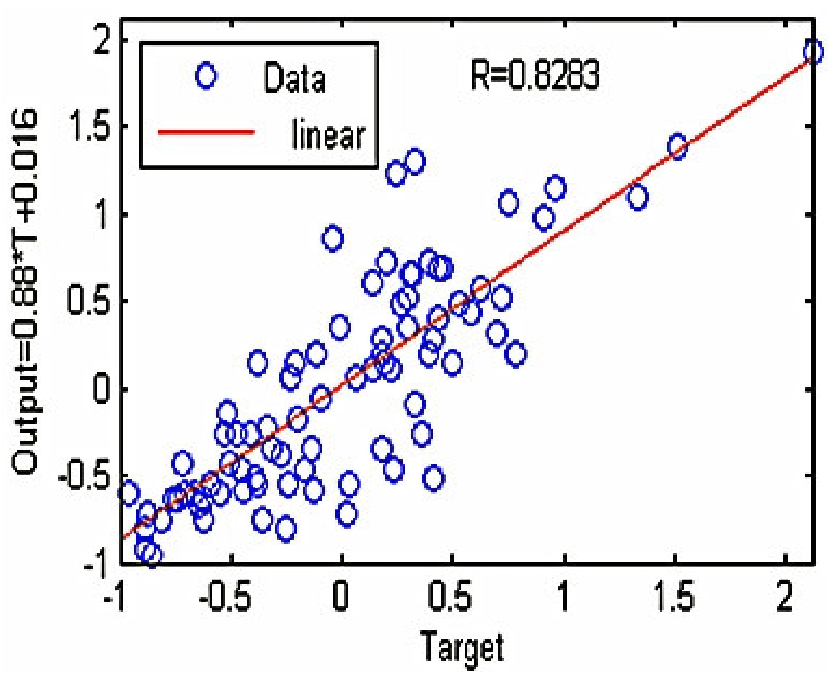

Figure 7. Regression plot between the target and output for network training.

important investigations were performed. The sensitivity analysis of input variables prior to network development was important for the final performance of the model. The sensitivity analysis was able to detect the response of the model based on the cause and effect of each parameter and its combinations with other parameters.

Similarly, many other factors contribute to the performance of neural network model and needs to be optimized before commencement of the actual network creation. These factors include: Transfer function in the hidden and output layers, number of neurons in the hidden layer, training function and the learning function.

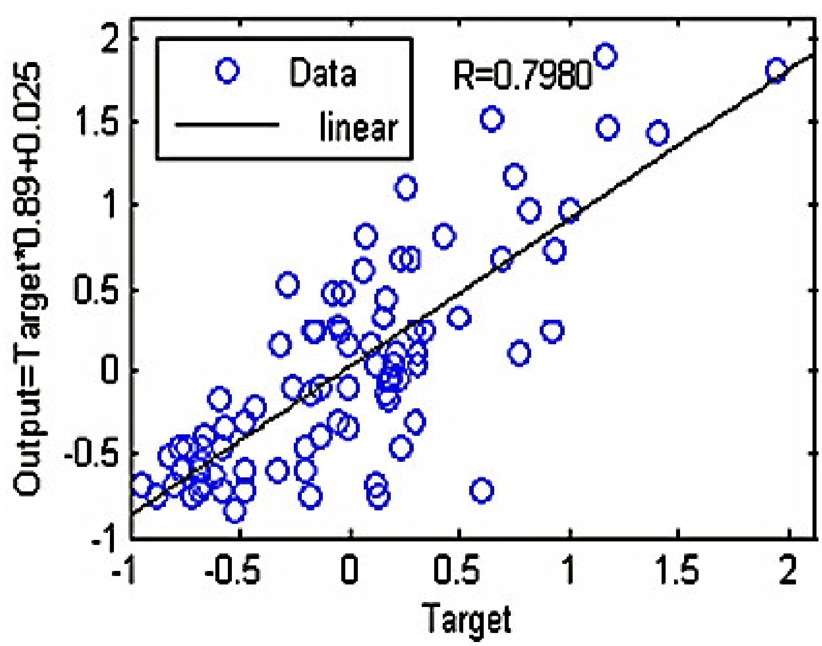

Figure 8. Regression plot between the target and output for the model testing.

Several iterations are essential in performing the optimization which requires patience and determination.

The model was built after all the necessary preliminary refining of the raw data such as, data pre-processing, standardization and optimum neural network structure evaluation. The model gave a brilliant simulation of ammoniacal nitrogen concentration of sewage treatment plant effluent. The model has the capability of explaining up to 79.80% of the wastewater treatment plant process for the purpose of simulation of NH_3-N performance in the effluent stream with only 0.1591 mean square error.

ACKNOWLEDGEMENTS

The authors would like to thank the staff of Bandar Tun Razak Sewage Treatment Plant (STP), Kuala Lumpur, for their cooperation during data collection. This work was partially supported by IIUM Endowment Fund (Type B) and Fundamental Research Grant Scheme (FRGS). The authors would like to acknowledge the financial support with sincere gratitude.

REFERENCES

Ardern E, Lockett WT (1914). Experiments on the oxidation of Sewage without the aid of filters. J. Soc. Chem. Ind. 33: 524.

Côté M, Grandjean BP, Lessard P, Thibault J (1995). Dynamic modeling of the activated sludge process: Improving prediction using Neural Networks. Water Res. 29(4): 995-1004.

De Sousa JT, Henrique IN, De Oliveira R, Lopes WS (2008). Nitrification in a submerged attached growth bioreactor using *Luffa Cylindrical* as solid substrate. Afr. J. Biotechnol. 7(15): 2702-2706.

Environmental Quality (Sewage) Regulations. (2009). Retrived January 24, 2011, from the DepartmentofEnvironmentMinistryofNaturalResourcesandEnvironment:http://www.doe.gov.my/portal/wpcontent/filesattachment/Legislation.Acts.Regulation.Order/Regulation/Environmental.quality.(sewage).regulations.2009.pdf.

Hamed MM, Khalafallah MG, Hassanien EA (2004). Prediction of wastewater treatment plant performance using artificial neural network. Environ. Modeling Software, 19: 919-928.

Hamoda MF, Al-Ghusain IA, Hassan AH (1999). Integrated wastewater treatment plant performance evaluation using artificial neural networks. Water Sci. Technol. 40(7): 55-65.

Haykin S (1999). Neural Networks: A Comprehensive foundation (2nd edn.), New Jersey: Prentice-Hall, Inc.

Liew R (2007). Country: Malaysian Water and Wastewater, U.S Commercial Service United State of America Department of Commerce.

Mahvi AH (2008). Sequencing batch reactor: A promising technology in wastewater treatment. Iran. J. Environ. Health Sci. Eng. 5(2): 79-90.

Mjalli FS, Al-Asheh S, Alfadala HE (2007). Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance. J. Environ. Manage. 83: 329-338.

Silver MR, Coelho MAZ, Araújo OQF (2002). Minimization of Phenol and ammoniacal nitrogen in refinery wastewater employing biological treatment. Engenharia Térmica, Edição Especial. pp. 33-37.

Tezel G, Yel E, Sinan RK (2010). Artificial neural network (Ann) model for domestic wastewater treatment plant control. The fourth scientific conference on water observation and information system for decision support (BALWOIS), Ohrid, Republic of Macedonia.

Wilderer PA, Irvine RL, Gronazy MC (2001). Sequencing batch reactor technology, IWA Scientific and Technical Report No. 10, IWA Puplishing (UK).