Analysis of Municipal Wastewater Treatment Plant Performance Using Artificial Neural Network Approach

N.A; Husain, Jami, M.S; Kabashi, , I.A; Abdullah, N.

Bioenvironmental Engineering Research Unit (BERU), Department of Biotechnology Engineering, International Islamic University Malaysia, P.O. Box 10, 50728 Kuala Lumpur, Malaysia
E-mail: saedi@iium.edu.my

*Artificial neural network **(ANN)** was used in this research as a statistical modeling tool for predicting the performance of wastewater treatment plant. A two years data of the waste water treatment plants' effluent and influent parameters was collected and applied in developing and training the ANN using the ANN toolbox in MATLAB. The data were obtained from Bandar Tun Razak Sewage Treatment Plant (BTR STP), that is managed by Indah Water Konsurtium (IWK), Malaysia's national sewerage company. The input and output parameters for the ANN were BOD, SS, and COD. It was found that the use of data screening is essential to come up with better ANNs model. Moreover, using multiple input-single output models was even a better model than single input-single output. The optimum number of hidden layer and neurons were determined which gave excellent results in predicting both the BOD and COD of the effluent which are required by the DOE. From the regression analysis, networks with one hidden layer and 20 nodes and BOD as input and COD as output were found to be the best one. The optimum number of hidden layers is 10 and the R value is improved by 30 %. The Mean Squared Error (MSE) is the lowest for the network. From the regression analysis, it is obvious that networks using screened data give better results in term of R values and MSE, and were selected for the subsequent modeling analysis in this study, that is prediction.*

1. Introduction

The application of artificial neural network (ANN) to solve real-life problem has only been developed in the last two decades. ANN provides a powerful tool for the analysis of the linear or non-linear relationships between variables. ANN have been trained to perform functions in various fields, including pattern recognition, identification, classification, control systems and many wastewater treatment applications i.e. predicting wastewater treatment plant operation, simulating and controlling of wastewater treatment plant, evaluating integrated wastewater treatment plant performance and effluent quality prediction [1]. ANN are different from the other traditional modeling approaches in the sense that they are designed and trained to learn solutions from previous data rather than modeling a specific problem [2].
In waste water treatment plants (WWTP), monitoring and controlling the influent parameters is a plant-dependent strategy that requires the understanding of the plant performance and the factors that affect the influent characteristics such as, time, season and the nature of the local life style. Moreover, the control of the waste water plant requires the monitoring of influent parameters such as biochemical oxygen demand (BOD), suspended solids (SS), ammonia nitrogen (NH_3N) and chemical oxygen demand (COD) [3].

In order to optimize the performance of the treatment plant and ensure the fulfillment of the department of environments (EOD) effluent requirements, both influent and effluent

parameters have to be monitored and controlled. However, the dynamic behavior of the WWTP complicates the process to generate relationships between input and output [4].
ANN can be trained to solve problems that are difficult for conventional computers or human beings. It actually works by learning from the trained data samples to understand the nature of the data [5]. In this work, wastewater treatment influent parameters; BOD, SS, and COD were analyzed using ANN tool available in MATLAB in order to find the relationships between the input and output values using the past data of Bandar Tun Razak Sewage Treatment Plant (BTR STP), one of the plants managed by Indah Water Konsurtium (IWK), Malaysia's national sewerage company.

2. Methodology

Data that were obtained from Bandar Tun Razak STP contains values of WWTP parameters measured over the span of almost two years' readings of each parameter.
An artificial neural net work can be defined as a distributed computational system composed of a number of individual processing elements operating largely in parallel, interconnected according to some specific topology (architecture), and having the capability to self-modify connection strengths during the processing of element parameters (learning) [1]. Figure 1 shows the topology of the feed-forward ANN model used in this work. The ANN basically consists of an input layer, one hidden or intermediate layer and an output layer. The nodes of one layer are connected to the nodes of another layer with connection weight, but they are not connected to nodes of the same layer. Each output value from a node, Y_i is weighted by the value W_i and the input to the next node, X_i is obtained by summing the product Y_iW_i. The output from this node is obtained by applying a non-linear activation function to X_i. The connection weights between nodes are optimized using the known input and target values through an iterative process and error-minimization technique, so that the network produces outputs close or equal to the known target values.

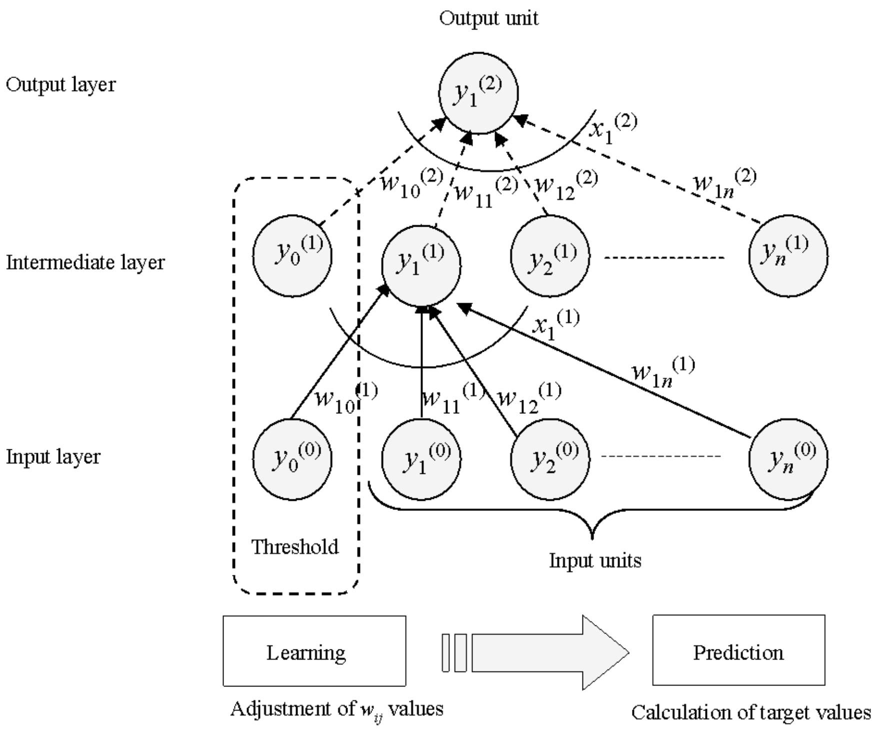

Fig. 1: Topology of the artificial neural net work model with an input layer, one hidden layer and an output layer

Data screening process is carried out on raw experimental data to eliminate all out-of-range values which might be due to transcription or transposition errors due to improper input of data, and experimental errors or human errors. The data screening was accomplished by removing values that are not within the range of $\pm 3\sigma$ and was completed using Microsoft Excel. Data screening has the advantages of accelerating the process, reducing training data and thus less forecasting time and homogenizing the data set leading to well-trained and accurate networks that are more constructive for analysis.

The components of input layers consist of three parameters; BOD influent (BOD_{in}), COD influent (COD_{in}) and SS influent (SS_{in}) while the outputs used in this modeling are BOD effluent (BOD_{out}), COD effluent (COD_{out}) and SS effluent (SS_{out}).

The data was subdivided into three groups; training, validating, testing, in the ratio of 4:2:1, respectively. Moreover, a single hidden layer with different count of nodes (i.e. 10, 20, or 30) was selected for this study. Back propagation network was used for the training of the network. The weights were first initialized to small random values within the range followed by give the input vectors and output vectors. After that, the output values were computed in a feed forward and the values computed were used by the final layer unit and the corresponding target value to compute the delta quantities. The weights were updated and the training was repeated until the iteration has been reached

The designed model was verified by comparing predicted values versus observed values. A good agreement between the observed and the predicted data confirms the validity of the methodology developed.

The previous steps were performed for the raw data and the screened data. Moreover, a single input-single output model was compared with multiple input-single output models.

3. Results and Discussion

The single input-single output modeling resulted with 54 networks for both the raw and screened data sets. Table 1 shows the R value and MSE for each network. From the table, it can be observed that the networks with higher number of nodes have higher R value and less MSE. The best fitting model concluded from this trainings is BODin-BODout with 30 nodes (R= 0.45) for the raw data whereas BODin-CODout with 20 nodes (R=0.58) for the screened data. The results also show that higher number of nodes results in a higher R value and better fitting model. Moreover, the regression analysis shows that networks using screened data give better result in term of R values and MSE as shown in Figures 4 and 5, and were selected for the subsequent modeling analysis in this study; prediction. The MSE resulted from the simulations show quite a significant deviation from the real values and predicted ones.
It can also be noticed from the regression results that the data that has been screened became more homogenous and thus more defined compared to non-screened data; consequently the correlation between parameters is more comprehensive and accurately linked, and lead to high R values.

Table 1: comparison of the training results for single input-single output networks

	R value & MSE					
NO of nodes / Network	Raw data			Screened data		
	10	20	30	10	20	30
BODin-BODout	2.18203e-1 8.70493e-0	5.62677e-2 14.78775e-0	4.53120e-1 6.96598e-0	3.29477e-1 3.70349e-0	3.55638e-1 5.37020e-0	4.01611e-1 3.94935e-0
BODin-CODout	2.31460e-1 99.04173e-0	2.2689e-1 112.32949e-0	-1.70396e-1 101069.43605 e-0	5.53426e-1 20.97258e-0	6.02204e-1 20.02581e-	5.39250e-1 23.00198e-0
BODin-SSout	2.37593e-1 14.30095e-0	2.55030e-1 14.12424e-0	1.11429e-1 33.77045e-0	4.75264e-1 4.58004e-0	3.76077e-1 5.20329e-0	1.67767e-1 11.03946e-0
CODin-CODout	5.87835e-2 116.29178e-0	-6.77928e-2 372.02702e-0	1.73716e-1 215.343888e-0	1.50919e-1 37.06769e-0	-8.87101e-2 128.61582e-0	8.45436e-2 58.78197e-0
CODin-BODout	1.65696e-1 9.42502e-0	1.47050e-1 10.89874e-0	4.44468e-1 7.34956e-0	3.05822e-1 4.39412e-0	2.22051e-1 5.42473e-0	1.25589e-1 11.43135e-0
CODin-SSout	2.43896e-1 14.16660e-0	3.05934e-1 14.08984e-0	9.6815e-2 23.84508e-0	1.67314e-1 14.32691e-0	1.82710e-1 6.51484e-0	4.94775e-1 4.59753e-0
SSin-SSout	2.43112e-1 23.76881e-0	1.88508e-1 140.97602e-0	2.33920e-1 22.41529e-0	3.87641e-1 4.88578e-0	-1.34887e-1 8.15365e-0	3.35314e-1 11.07111e-0
SSin-BODout	-5.10172e-2 10.13807e-0	1.75257e-1 67.30890e-0	5.10105e-2 16.85928e-0	2.79885e-1 3.68421e-0	1.92904e-1 7.46435e-0	-3.08416e-2 14.54059e-0

SSin- CODout	-1.76058e-1 61.03835e-0	1.31927e-1 117.21180e-0	1.93165e-1 306.05628e-0	-5.20707e-2 117.53354e-0	5.68348e-1 19.92609e-0	4.45934e-2 155.16622e-0

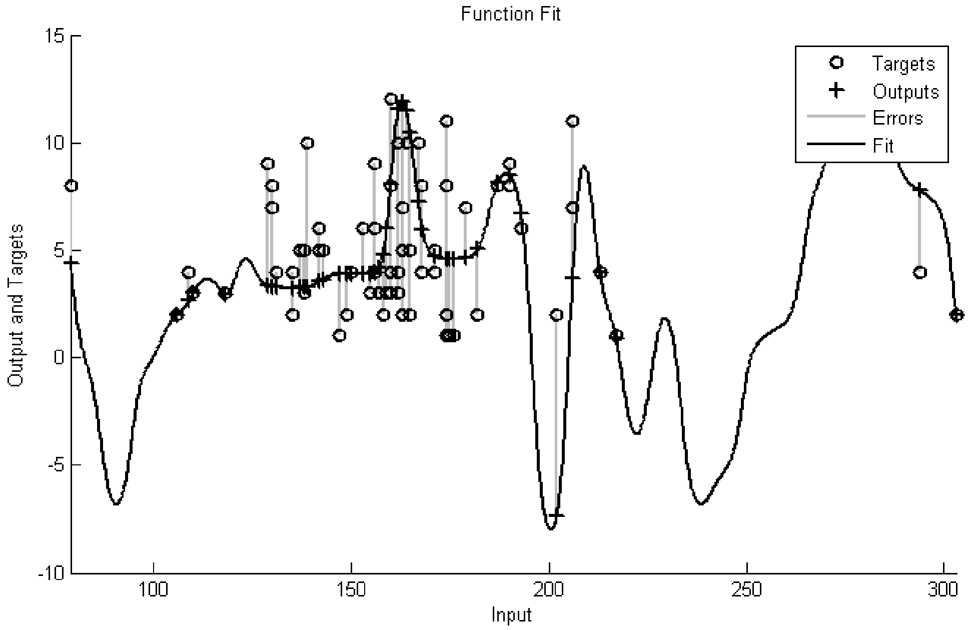

Fig. 2: Curve fitting for BODin-BODout obtained from raw data training with 30 nodes

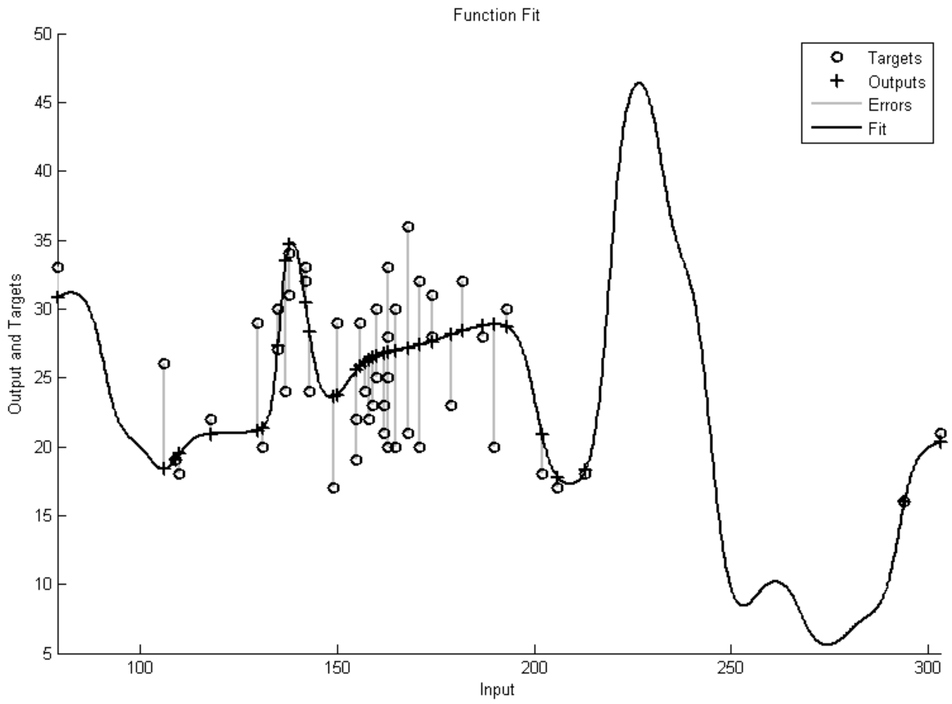

Fig. 3: Curve fitting for BODin-CODout obtained from screened data training with 20 nodes

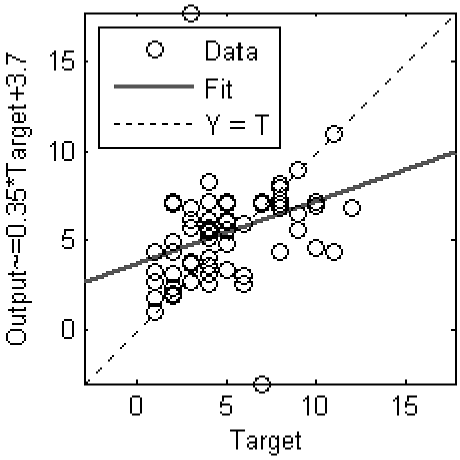

Fig. 4: linear regression of BODin-BODout of raw data at 30 nodes

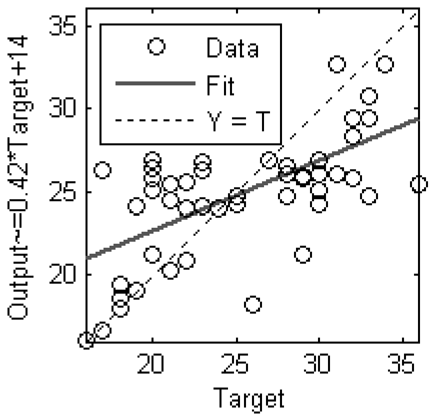

Fig. 5: linear regression of BODin-Codout of screened data at 20 neurons

However after screening, the number of set of data is limited (50 sets) in contrast to raw data (70 sets). This limitation of data caused inefficiency of the networks; consequently the MSE resulted from the prediction is relatively high for several networks. Not only is the quantity of the data important but also the quality. This includes the synchronization of data with time. The time gap between data used is not constant. Since, ANNs is dynamic itself, such biased in data could affect the overall performance of the networks. To achieve optimum result, especially in unveiling the interrelationship between parameters, it is important to train the networks as many times as possible. A better computerized system should be used, if ANN is to be implemented in real wastewater treatment plant, for better optimization procedures.

Networks with multiple inputs have been train as well. It can be concluded from the results shown in table 2 that the network with multiple inputs and BOD as an output and 10 nodes is the best model for the raw data. While the network with multiple inputs and BOD as output and 20 nodes in the best for the screened data. The R value for the raw data is higher than that for the screened data, this can be due to the reduction of the number of the data sets as a result of screening. Thus for a better model design and training a larger data set should be applied. Moreover, the data should be collected at a regular time interval for more reliable results.

Table 2: R value and MSE results for multiple inputs- single output networks

	R value & MSE					
NO of nodes Network	Raw data			Screened data		
	10	20	30	10	20	30
inputs-BODout	7.57295e-2 3.35977e-1	6.54542e-2 3.259032-1	6.75799e-2 4.92340e-1	1.11839e-1 5.29738e-1	1.00870e-1 4.42281e-1	3.46948e-1 4.00950e-2
inputs-CODout	5.12232e-2 4.3173e-1	5.10658e-2 3.63832e-1	5.09801e-2 3.32268e-1	6.46825e-2 3.48635e-1	2.52537e-1 3.35991e-1	5.26220e-2 6.17708e-1
inputs-SSout	4.15562e-2 2.85145e-1	4.19057e-2 3.81616e-1	1.04606e-1 1.32424e-1	6.10977e-2 3.96357e-1	9.96014e-2 1.36641e-1	1.58717e-1 1.95877e-1

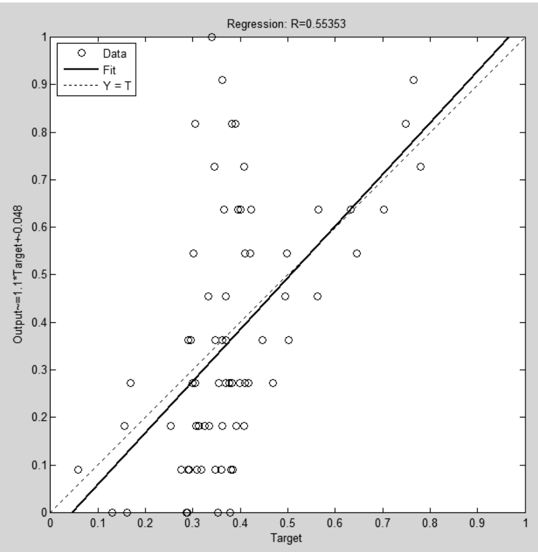

Figure 1: regression graph for the network inputs-BODout of raw data with 10 nodes

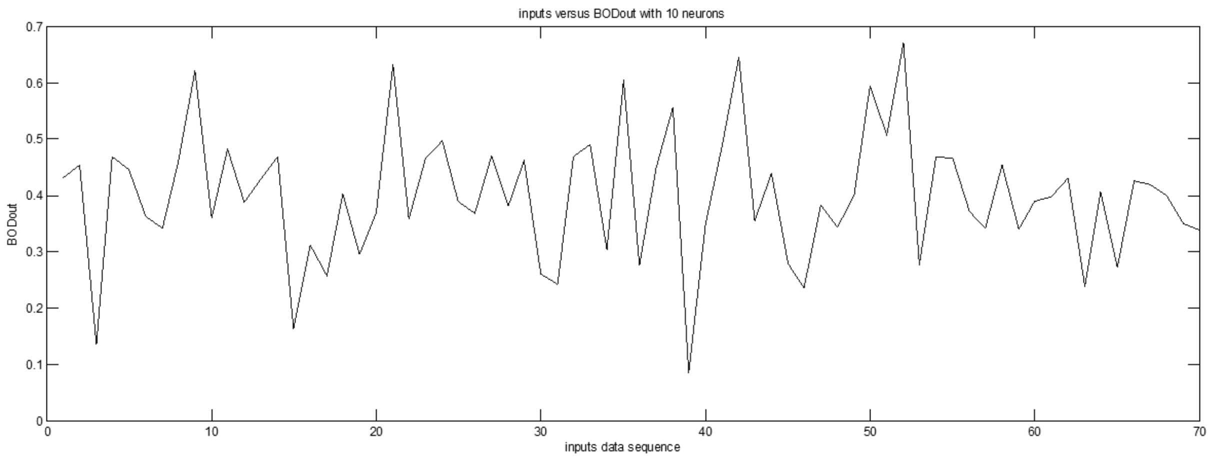

Figure 2: fitting graph for the input data with BOD as output for raw data

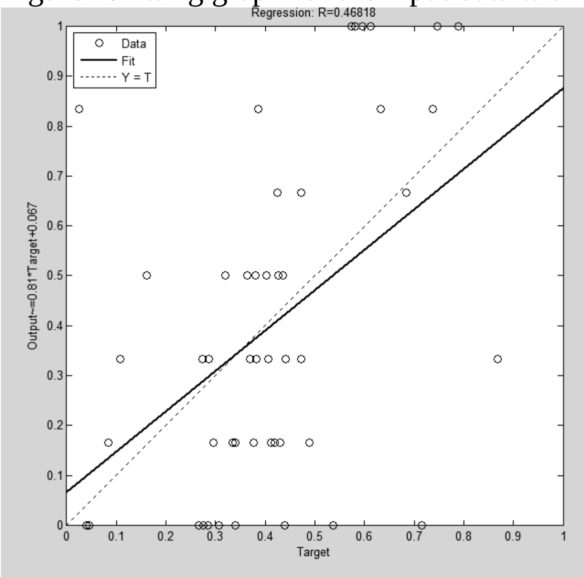

Figure 3: regression graph for the network inputs-BODout of screened data with 20 nodes

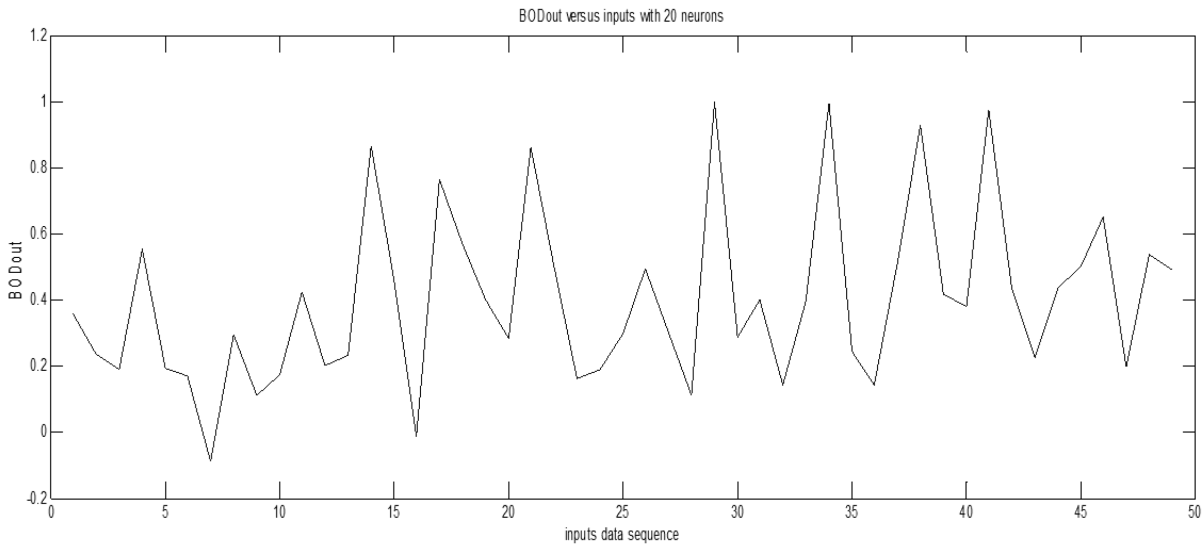

Figure 4: fitting graph for the input data with BOD as output for screened data

4. Conclusions

Several conclusions can be made from this work. In this study, the data from an existing WWTP are used to train ANNs. It was found that screening of the data is essential to come up with better ANNs model. ANNs developed in this study can be used to predict the performance of an existing WWTP thus solving the problem of variations in parameters of WWTP which contributes to difficulties in monitoring and processing activities. ANNs is found to be applicable to analyze complex, non-linear, and dynamic data. These make ANNs valid as a tool to study biological processes.

5. References

[1] Haykin, S. (1994), *Neural Networks: A Comprehensive Foundation*, Macmillan College Publishing Company, 696.

[2] Sencan,A.; Kalogirou, S. (2005), A new Approach Using Artificial Neural Networks for Determination of the Thermodynamic Properties of Fluid Couples, *Energy Conversion and Management*, 46, 2405–2418.

[3] Grishma, R. S.; Chellam, S. (2003), Predicting membrane fouling during municipal drinking water nanofiltration using artificial neural networks, *Journal of Membrane Science*, 217 (1-2), 69-86.

[4] Hanbay, D.; Turkoglu, I.; Demir, I. (2008), Prediction of wastewater treatment plant performance based on wavelet packet decomposition and neural networks. *Expert Systems with Applications: An International Journal*, 34 (2), 1038-1043.

[5] Hong, Y. S. T.; Rosen, M. R.; Bhamidimarri, R. (2003), Analysis of a municipal wastewater treatment plant using a neural network-based pattern analysis. *Water Research*, 37 (7), 1608-1618.

[6] Farouq S. Mjalli; S. Al-Asheh; H. E. Alfadala. (2007), Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance. *Journal of Environmental Management*, 83

Acknowledgements

The authors would like to thank Indah Water Konsortium at Bandar Tun Razak Sewage Treatment Plant for making available the data collected and IIUM Research and Innovation for financial support.