Research Article

Optimization of Anaerobic Treatment of Petroleum Refinery Wastewater Using Artificial Neural Networks

H.A. Gasim, S.R.M. Kutty, M. Hasnain Isa and L.T. Alemu

Department of Civil Engineering, and Department of Computer and Information Sciences, Universiti Teknologi PETRONAS (UTP), Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia

Abstract: Treatment of petroleum refinery wastewater using anaerobic treatment has many advantages over other biological method particularly when used to treat complex wastewater. In this study, accumulated data of Upflow Anaerobic Sludge Blanket (UASB) reactor treating petroleum refinery wastewater under six different volumetric organic loads (0.58, 1.21, 0.89, 2.34, 1.47 and 4.14 kg COD/m$^3$·d, respectively) were used for developing mathematical model that could simulate the process pattern. The data consist of 160 entries and were gathered over approximately 180 days from two UASB reactors that were continuously operating in parallel. Artificial neural network software was used to model the reactor behavior during different loads applied. Two transfer functions were compared and different number of neurons was tested to find the optimum model that predicts the reactor pattern. The tangent sigmoid transfer function (tansig) at hidden layer and a linear transfer function (purelin) at output layer with 12 neurons were selected as the optimum best model.

Keywords: Anaerobic digestion, artificial neural networks, chemical oxygen demand, petroleum refinery wastewater, UASB

INTRODUCTION

Petroleum refinery wastewater is hazardous to the environment as there are several organic and inorganic pollutants found in petroleum refinery wastewater of environmental concern including ammonia, oil, phenol, sulphur based contaminants and heavy metals (Vohra et al., 2006). One of the most important monitoring parameters for wastewater treatment plant is Chemical Oxygen Demand (COD) as it required relatively shorter time (Khan et al., 2006).

Anaerobic treatment of high-strength petroleum refinery wastewaters (generally>1000 mg COD/L) (Behling et al., 1997) has been shown to provide a very cost-effective alternative to aerobic processes with savings in energy, nutrient addition and reactor volume. Petroleum refinery wastewater is degraded under anaerobic conditions and many toxic and recalcitrant organic compounds that found in petroleum wastewater are serving as a growth substrate (Metcalf and Eddy, 2003). The Up flow Anaerobic Sludge Blanket (UASB) reactor is one of the most notable developments in anaerobic treatment process technology that commonly used for treating a wide range of industrial wastewater. UASB is a high rate system that can retain biomass with high treatment capacity and low site area requirement in addition to other advantages (Zinatizadeh et al., 2007).

The use of software to simulate existing historical experimental data and predict unknown data based on a model representing the process, help to minimize efforts and creates more data which doesn’t exist. The modeling and simulation of processes have been developed using ever more complex deterministic models, due to the recent evolution of personal computer (Gontarski et al., 2000).

Neural Networks (NN) or widely known as Artificial Neural Networks (ANN) is a mathematical modeling tool used to simulate complex relationships following a simplified level of the activity of the human brain through a large number of highly interconnected processing elements (neurons); and have been used in application of artificial intelligence that has shown quite a promise in engineering, pattern recognition and analysis (Hamed et al., 2004). Anaerobic digestion is a non-linear process which requires a non-linear control strategy; whereby artificial neural networks is the choice when a large amount of anaerobic digestion data are available but no reliable model and little knowledge of how the process works (Ward et al., 2008).

ANN has being used to model existing data and simulate for predicted behavior in many wastewater treatment processes to ease the operation activities. Artificial neural networks are claimed to have a

Corresponding Author: H.A. Gasim, Department of Civil Engineering, Universiti Teknologi PETRONAS (UTP), Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia, Tel.: +60 192292944

This work is licensed under a Creative Commons Attribution 4.0 International License (URL: http://creativecommons.org/licenses/by/4.0/).
distinctive advantage over some other nonlinear estimation methods used for bio-processes as they do not require any prior knowledge about the structure of the relationships that exist between important controlling variables (Holubar et al., 2002).

ANN has been used to simulate full working wastewater treatment plant using a model that was developed using laboratory data for ten months. Modeling of this wastewater treatment process used a configuration with tan sigmoid activation function for the input and hidden layers, while the linear activation function was used as the output activation function, resulted in $R^2$ values ranged from 0.63 to 0.81 for Biochemical Oxygen Demand (BOD) and from 0.45 to 0.65 for Suspended Solids (SS) (Hamed et al., 2004). Using the same mentioned configuration, Chemical Oxygen Demand (COD) removal was modeled using ANN in a wastewater treatment process for the prediction and simulation of degradation. The configuration of the back propagation neural network with 14 neurons and Levenberg-Marquardt back propagation training algorithm (TRAINLM) predicted the actual experimental results with correlation coefficient ($R^2$) of 0.997 and MSE of 0.000376 (Elmolla et al., 2010).

Anaerobic biological treatment of wastewater was modeled based on integrated fuzzy systems and neural networks for the simulation and control of complex anaerobic treatment systems (anaerobic fluidized bed reactor and up-flow anaerobic sludge blanket) (Tay and Zhang, 1998).

Several Feed-Forward Back Propagation neural networks (FFBP) were trained in order to model and subsequently control, methane production in four anaerobic continuous stirred tank reactors. The model was able to predict gas production and avoid shock loadings (Holubar et al., 2002).

Utilizing a neural network simulation, anaerobic wastewater treatment process has been modeled to occur during disturbances to an anaerobic digestion. The neural network was capable of rapid recognition of disturbances that in the form of an increase in influent COD concentration and by utilizing data from an online bicarbonate alkalinity sensor (Wilcox et al., 1995).

A high strength wastewater (7300 mg COD/L) batch from a local petroleum refinery was treated in UASB as part of a train of biological reactor; the COD removal was found to be 80% (Gasim et al., 2012a). Two parallel UASB reactors were used to evaluate the treatment efficiency of petroleum refinery wastewater under six organic volumetric loading rates (0.58, 0.89, 1.21, 1.47, 2.34 and 4.14 kg COD/m$^3$·d, respectively); the COD removals efficiencies were 78, 82, 83, 80, 81 and 75%, respectively as the load increased (Gasim et al., 2012b, c).

The present study follows from the previous investigation by modeling the anaerobic treatment of petroleum refinery wastewater considering the influent

- Minimum Mean Absolute Percentage Error
- and effluent COD concentration under different loads; the developed model was then used to simulate the reactor performance.

**MATERIALS AND METHODS**

**Experimental data:** The original raw data were adapted from previous work in which the data were representing two laboratory-scale Up-flow Anaerobic Sludge Blanket (UASB) reactors that were operated in parallel (A and B) at room temperature to treat petroleum refinery wastewater. The raw petroleum refinery wastewater was collected from a local petroleum refinery and fed to the two reactors in different concentration ranging on average from 982 to 6972 mg COD/L over approximately 180 days. Chemical Oxygen Demand (COD) was tested for influent and effluent samples following colorimetric method using a HACH DR 2000 spectrophotometer, other parameters were measured according to Standard Methods (APHA, 1980). The data that were gathered from this experiment were 160 entries for influent and effluent.

**ANN procedure:** The COD monitoring results during different loading were used for modeling. Artificial neural network was used as mathematical tool to simulate and predict the pattern of the reactor. Optimal generalization was targeted from this tool, therefore, the Levenberg-Marquardt algorithms was used as training function and batch gradient descent with momentum back propagation algorithms (TRAINGDM) as adaption learning function, Feed-Forward Back propagation network type was selected. The number of neurons has to be determined as it is related to the converging performance of the output error function during the training process. Increasing the number of neurons usually results in a better learning performance, as too few number of neurons limit the ability of the neural network to model the process, but too many number of neurons may results in losing the generalization and learning the noise present in the database used in training (Holubar et al., 2002).

Normalization of input data was performed by dividing all the input data with the maximum input; this resulted in the data to be in the range of 0 to 1. Output data were normalized by dividing all the output data with the maximum output; this resulted in the date to be in the range between 0 and 1.

Neurons were tested and varied the number of neurons in the range from 5 up to 35. For better initialization of the model, the model was run 100 times at every neuron tested. Optimum number of neuron was selected in this study based on:

- Minimum Root Mean Square Error (RMSE)
- Maximum Variance Accounted For (VAF)
- Maximum correlation coefficient ($R^2$)

(MAPE)
Neural Network in MATLAB (R2009a) software was used with back propagation neural network three layers in two configurations. First, with Log Sigmoid transfer function (LOGSIG) at hidden layer and a linear transfer function (PURELIN) at output layer. Second, with Tangent Sigmoid transfer function (TANSIG) at hidden layer and a linear transfer function (PURELIN) at output layer. The linear activation function (PURELIN) was used for both configurations for the output neuron since it is appropriate for continuous valued targets (Hamed et al., 2004).

RESULTS AND DISCUSSION

Modeling results: The raw data from anaerobic reactor was modeled using artificial neural networks software. Logsig4Purelin transfer function was compared to Tansig4Purelin transfer function to define the optimum model. The selected model was then used to predict the reactor performance. The simulation data were then used to find the optimum performance.

During testing and validation of data, number of neurons was tested ranging from 5 to 35. Table 1 shows the number of neurons tested and the score registered for RMSE, VAF, $R^2$ and MAPE during evaluation of Logsig-Purelin and Tansig-Purelin transfer functions.

Table 1: Number of neurons tested and the score for evaluation parameters

| Neurons | Logsig-purelin | Tansig-purelin |
|---------|----------------|----------------|
|         | RMSE   | VAF   | $R^2$ | MAPE | RMSE   | VAF   | $R^2$ | MAPE |
| 5       | 0.076  | 85.952 | 0.859 | 16.878 | 0.076  | 86.027 | 0.860 | 17.035 |
| 6       | 0.076  | 86.134 | 0.860 | 17.340 | 0.076  | 86.104 | 0.860 | 16.552 |
| 7       | 0.075  | 86.249 | 0.862 | 17.255 | 0.075  | 86.229 | 0.862 | 17.279 |
| 8       | 0.076  | 85.956 | 0.859 | 16.884 | 0.075  | 86.147 | 0.861 | 16.622 |
| 9       | 0.075  | 86.250 | 0.862 | 16.694 | 0.075  | 86.170 | 0.861 | 15.971 |
| 10      | 0.075  | 86.500 | 0.864 | 16.810 | 0.075  | 86.154 | 0.861 | 16.576 |
| 11      | 0.075  | 86.516 | 0.864 | 16.977 | 0.075  | 86.388 | 0.864 | 16.750 |
| 12      | 0.074  | 86.741 | 0.867 | 17.385 | 0.074  | 86.638 | 0.866 | 16.827 |
| 13      | 0.074  | 86.608 | 0.866 | 16.426 | 0.075  | 86.424 | 0.864 | 16.023 |
| 14      | 0.075  | 86.133 | 0.861 | 16.204 | 0.076  | 86.111 | 0.860 | 16.907 |
| 15      | 0.074  | 86.724 | 0.867 | 16.205 | 0.075  | 86.412 | 0.864 | 16.427 |
| 16      | 0.074  | 86.796 | 0.867 | 16.516 | 0.075  | 86.136 | 0.861 | 16.832 |
| 17      | 0.074  | 86.589 | 0.865 | 16.585 | 0.074  | 86.756 | 0.868 | 17.353 |
| 18      | 0.073  | 87.119 | 0.871 | 15.706 | 0.074  | 86.778 | 0.868 | 15.532 |
| 19      | 0.074  | 86.716 | 0.867 | 17.047 | 0.073  | 86.949 | 0.869 | 16.809 |
| 20      | 0.074  | 86.492 | 0.865 | 16.317 | 0.073  | 86.949 | 0.869 | 16.809 |
| 21      | 0.076  | 85.834 | 0.858 | 17.072 | 0.073  | 86.981 | 0.870 | 16.355 |
| 22      | 0.074  | 86.680 | 0.867 | 16.713 | 0.074  | 86.478 | 0.865 | 16.020 |
| 23      | 0.075  | 86.480 | 0.864 | 15.911 | 0.075  | 86.242 | 0.862 | 16.362 |
| 24      | 0.073  | 86.999 | 0.870 | 16.576 | 0.073  | 86.993 | 0.869 | 16.514 |
| 25      | 0.073  | 87.068 | 0.870 | 15.617 | 0.073  | 87.067 | 0.871 | 17.009 |
| 26      | 0.071  | 87.577 | 0.876 | 16.361 | 0.073  | 87.119 | 0.871 | 17.088 |
| 27      | 0.073  | 86.956 | 0.869 | 15.467 | 0.074  | 86.764 | 0.865 | 16.072 |
| 28      | 0.073  | 86.983 | 0.870 | 16.659 | 0.074  | 86.711 | 0.867 | 16.339 |
| 29      | 0.071  | 87.825 | 0.878 | 15.832 | 0.074  | 86.468 | 0.865 | 16.621 |
| 30      | 0.074  | 86.751 | 0.867 | 16.916 | 0.075  | 86.330 | 0.863 | 16.449 |
| 31      | 0.068  | 88.734 | 0.887 | 16.171 | 0.072  | 87.368 | 0.873 | 15.808 |
| 32      | 0.071  | 87.777 | 0.876 | 16.889 | 0.076  | 85.953 | 0.859 | 15.702 |
| 33      | 0.071  | 87.756 | 0.877 | 15.273 | 0.073  | 87.088 | 0.869 | 15.789 |
| 34      | 0.064  | 89.033 | 0.899 | 15.950 | 0.068  | 89.544 | 0.899 | 13.591 |
| 35      | 0.062  | 90.687 | 0.907 | 14.955 | 0.075  | 86.372 | 0.863 | 15.382 |

Although the number of neurons are in the range of 5-35, but from Fig. 1 it is noted that after neuron 15 and from plotted line representing the $R^2$ from the training set is losing similarity with $R^2$ from validation set, indicating over fitting and the model will not be able to generalize the pattern of the data that used as training set during validation (Jeon, 2007).

Thus, the number of neurons was limited to the range between 5-15 neurons and the optimum neuron was selected as shown in Table 1 based on minimum RMSE, maximum VAF, maximum $R^2$ and minimum MAPE.

Logsig-Purelin transfer function indicated 15 neurons is the optimum, while Tansig-Purelin suggested 12 neurons. It is usually preferable to use of simpler models, with fewer number of parameters than more complicated ones with more parameters, whenever feasible (Hamed et al., 2004; Holubar et al., 2002). Thus, tangent sigmoid transfer function (tansig) at hidden layer and a linear transfer function (purelin) at output layer with 12 neurons is the optimum transfer function.

Figure 2 showed the measured experimental data and the predicted using ANN for eighty entries of data that were used for training. Figure 3 showed the measured experimental data and the predicted using...
ANN for eighty entries of data that were used for validation. The best selected model shows significant prediction of actual experiment; hence, it was then used for simulation.

**Simulation results:** The best model with Tansig-Purelin transfer function and 12 neurons was used to simulate random data to find out the optimum efficiency. Figure 4 shows all the hundred and sixty data set that was used for both the training and validation, used here for simulation.

Random data entries ranged from 500 to 10000 was used to simulate the reactor performance. Figure 5 shows the simulated influent and effluent
concentrations in addition to removal efficiency. Highest removal efficiency observed was 82% at influent 1400 mg COD/L and effluent 245 mg COD/L.

CONCLUSION

Raw data from petroleum refinery wastewater treatment with different loads using two UASB reactors were successfully used for modeling. Modeling resulted in a model that used tangent sigmoid transfer function (tansig) at hidden layer and a linear transfer function (purelin) at output layer with 12 neurons as the optimum transfer function. Simulation using the optimum model with random data entries ranged between 500 to 9000 resulted in a pattern that simulates the reactor performance for data that were never really experimentally tested in the lab. Lab experiment was showing highest removal of 82% which confirmed by using the best selected model that developed using mathematical model.

ACKNOWLEDGMENT

This study was supported by the Universiti Teknologi PETRONAS graduate assistantship scheme.

REFERENCES

APHA, 1980. Standard Methods for the Examination of Water and Wastewater. American Public Health Association Inc., New York.
Behling, E., A. Diaz, G. Colina, M. Herrera, E. Gutierrez, E. Chacin, N. Fernandez and C.F. Forster, 1997. Domestic wastewater treatment using a UASB reactor. Bioresource Technol., 61: 239-245.
Elmolla, E.S., M. Chaudhuri and M.M. Eltoukhy, 2010. The use of Artificial Neural Network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. J. Hazard. Mater., 179: 127-134.
Gasim, H.A., S.R.M. Kutty and M.H. Isa, 2012a. Treatment of petroleum refinery wastewater using multi-stage biological reactor. Proceeding of the International Conference on Civil, Offshore and Environmental Engineering (ICCOEE). Kuala Lumpur.
Gasim, H.A., S.R.M. Kutty, M.H. Isa and M.P.M. Isa, 2012b. Treatment of petroleum refinery wastewater by using UASB reactors. Int. J. Chem. Biol. Eng., 6: 174-177.
Gasim, H.A., S.R.M. Kutty and M.H. Isa, 2012c. Anaerobic treatment of petroleum refinery wastewater. Proceeding of International Conference on Waste Management and Environmental Engineering (ICWMEC 2012), Kuala Lumpur.

Gontarski, C.A., P.R. Rodrigues, M. Mori, L.F. Prenem, 2000. Simulation of an industrial wastewater treatment plant using artificial neural networks. Comput. Chem. Eng., 24: 1719-1723.

Hamed, M.M., M.G. Khalafallah and E.A. Hassanian, 2004. Prediction of wastewater treatment plant performance using artificial neural networks. Environ. Modell. Softw., 19: 919-928.

Holubar, P., L. Zani, M. Hager, W. Froschl, Z. Radak and R. Braun, 2002. Advanced controlling of anaerobic digestion by means of hierarchical neural networks. Water Res., 36: 2582-2588.

Jeon, J.K., 2007. Fuzzy and Neural Network Models for Analyses of Piles. Ph.D. Thesis, North Carolina State University, Raleigh, North Carolina.

Khan, M.S., J. Knapp, A. Clemett, M. Chadwick, M. Mahmood and M.I. Sharif, 2006. Managing and Monitoring Effluent Treatment Plants. In: Research for Development, Department for International Development (DFID), Dhaka.

Metcalf, L. and H. Eddy, 2003. Wastewater Engineering, Treatment, Disposal and Reuse. McGraw Hill, New York.

Tay, J.H. and X. Zhang, 1998. Neural Fuzzy Modeling of Anaerobic Biological Wastewater Treatment Systems. Development and Application of Computer Techniques to Environmental Studies, WIT Press, Computational Mechanics Publications, UK, pp: 25-34.

Vohra, M.S., S.A. Malik and M.S. Al-Suwaiyan, 2006. Treatment of synthetic pretreated petroleum refinery wastewater using TiO2-assisted photocatalysis. Proceeding of the 4th Saudi Technical Conference and Exhibition. Riyadh, Saudi Arabia.

Ward, A.J., P.J. Hobbs, P.J. Holliman and D.L. Jones, 2008. Optimization of the anaerobic digestion of agricultural resources. Bioresource Technol., 99: 7928-7940.

Wilcox, S.J., D.L. Hawkes, F.R. Hawkes and A.J. Guwy, 1995. A neural network based on bicarbonate monitoring, to control anaerobic digestion. Water Res., 29: 1465-1470.

Zinatizadeh, A.A.L., A.R. Mohamed, M.D. Mashitah, A.Z. Abdullah and M.H. Isa, 2007. Optimization of pre-treated palm oil mill effluent digestion in an up-flow anaerobic sludge fixed film bioreactor: A comparative study. Biochem. Eng. J., 35: 226-237.