**MODELLING APPROACHES COMPARISON APPLIED TO WASTEWATER TREATMENT PLANT PROJECT**

**Aliyu Muhammad Kamilu**

School of Innovative Technology,

Department of Project Management Technology, Federal University of Technology, Minna.

# *ABSTRACT*

*Wastewater treatment system involves several intricate parallel biological processes, which are quite uncertain and difficult to predict. However, for efficient operation of the system, a reliable and straightforward model capable of accurately describing the behaviour of the system is strongly needed. Most of the existing developed models were applied to industrial wastewater treatment plant. This paper investigates the effectiveness in predicting the organic nutrient removal of four (4) different modelling approaches applied to the domestic step feed activated sludge wastewater treatment plant in Malaysia. Computational efficiency and reliability tempted the selection of the techniques such as autoregressive with exogenous input model (ARX), nonlinear autoregressive with exogenous input model (NARX), artificial neural network (ANN) and adaptive neuro fuzzy inference system (ANFIS). Simulation studies revealed that ANFIS model demonstrated better prediction capability compared to the other models in all the considered variables having root mean square error (RMSE) 0.0133 and mean absolute percentage deviation (MAPD) 3.88% for biochemical oxygen demand (BOD), RMSE 0.0478 and MAPD 5.16% for chemical oxygen demand (COD), RMSE 0.0086 and MAPD 3.42% for suspended solids (SS) and RMSE 0.0310 and MAPD 6.40% for ammonium nitrogen NH4-N during validation. This shows that the ANFIS model may serve as a valuable tool for predicting the effluent quality in the wastewater treatment plant.*

***Keywords:*** *model, prediction, fuzzy inference system, network, principal component, Malaysia*

# INTRODUCTION

Recently, public health and environmental quality enticed scientific research communities and practitioners to focus their attention toward wastewater treatment industries. The prime component in the system is the activated sludge process. Due to efficiency and costeffectiveness, activated sludge process becomes the most widely used technology for removal of organic nutrients from wastewater. The process main concerns are reduction of the pollutants' concentrations to the level nature can handle and rejection of disturbances for smooth operation at minimal cost, to achieve these conveniently; appropriate model cannot be avoided. Suitable model is absolutely crucial not only for control, design, prediction, but also for optimal and trouble-free operation of the system. The existing mathematical models developed by the international association of the water quality (IAWQ) task group (Henze *et al.*, 2000), such as activated sludge model no.1 (ASM1) for organic carbon and nitrogen removal, it's updated version ASM2, which includes phosphorus removal, the ASM2d, an extension of ASM2 and ASM3 the new version of ASM1, have greatly contributed immensely in describing the complex biological activities of the microorganisms responsible for degrading the pollutants in the wastewater and interactions occurring in the system.

Arguably, these models are structurally complex to use and difficult to solve analytically (Tay and X. Zhang, 2000). Therefore, there is an urgent need for a reliable and easy model.

Surveying the literature several simplified models based on different plant's configurations either through simulations or practically were implemented particularly based on the reference model version (ASM1). Linearize models (Smets *et al.*, 2003; Benhalla et al. 2010) demonstrated effectiveness in tracking the responses of ASM1 and maintaining the states biological interpretations. Analytical model (Nagy *et al.* 2010) performed remarkably in predicting the process and conserved the nonlinear nature of the reference model.

Artificial intelligent method (fuzzy logic, neural network and hybrid system) emerged as an alternative and exhibited high level of accuracy (Ragot *et al.,* 2001; Huang et al., 2012; Tay and

X. Zhang, 2000; Perendeci et al. 2008; Civelekoglu et al. 2007; Raduly and Gernaey, 2007; Civelekoglu et al. 2009; Pai *et al.*, 2011) in predicting the process.

However, most of these suggestions were applied to either a completely mixed tank or contact stabilization and industrial wastewater. To date, no report on application to a domestic step feed wastewater treatment plant.

Therefore, it is the objective of this paper to investigate the feasibility and effectiveness in predicting the organic nutrient removal of these modelling techniques applied to the step-feed activated sludge wastewater treatment plant. The methods were proven to be powerful in many real-world applications. Nonlinear autoregressive with exogenous input (NARX) offers greater flexibility in dealing with noisy data. The advantages of NARX are due to its ability to approximate nonlinear dynamics (Nelles, 2000), fast convergence and less computational cost. ANN emerged as an adequate alternative in handling of a complex noisy data. ANN consists of connected neurons that cooperate to execute the desired task. The success of ANN is as the result of fast learning ability and adaptation. The integration of ANN and fuzzy system yielded a robust hybrid system referred as neuro-fuzzy system. One of most commonly used neuro-fuzzy system is ANFIS. The main principle of ANFIS is the hybrid learning algorithm. Imprecision and uncertainties in a complex dynamic data are easily handled by ANFIS. Usually raw data from wastewater treatment plant contained redundant and confusing information. Therefore, it is essential to pre-treat the data using PCA technique to transform the data into trainable and convenient form in order to achieve the desired accurate model. The proposed modelling techniques were evaluated in terms prediction error measures RMSE and MAPD. The paper adopted the judgement of the prediction accuracy defined by (Lewis, 1982) using mean average percentage error (MAPE) also referred as MAPD. Prediction with lower MAPE below 10% is considered to be highly accurate. In this manner, the model that produces effective prediction regarding organic nutrient removal can be found.

# MATERIAL AND METHOD

## Process Description

|  |
| --- |
| 2  3      5    9    10  11  12  1    4    6    7    8    13      Figure 1. The schematic of the IWK treatment plant: (1) Influent wastewater (2) distribution chamber (3) grit chamber (4) primary clarifier (5) aeration tank (6) secondary clarifier (7) effluent discharge (8) returned sludge (9) gravity thickener (10) mechanical thickener (11)Digester tank (12) sludge dewatering (13) sludge cake disposal. |

The plant has daily average capacity of 87,000m3/d, utilizes advance step-feed activated sludge process for BOD, COD, SS and nitrogen removal. The influent wastewater enters the distribution chamber before flowing to the grit unit where the grits contained in wastewater are removed to avoid pump wear and pipe deterioration. Then, the wastewater goes to the primary clarifier in which the wastewater is retained to allow the settle-able organics and floatable solids to settle at the bottom of the clarifier by gravity sedimentation. The solids are withdrawn by the chain flight collector and transferred to the gravity thickener, and the scum is removed by the scum skimmers. The effluent from the primary clarifier flows to the biological reactor unit consisting of two anoxic tanks and two aerobic tanks as shown in Fig. 1. The anoxic tanks have the mixer for agitation while air is provided to the aerobic tanks for the aeration. The solids are separated from the treated water by sedimentation in the secondary clarifier. A portion of the sludge is returned to the biological reactor unit to keep the microorganisms' concentration. The remaining sludge is removed and transferred to the sludge treatment facility.

## Anfis

|  |
| --- |
| *a ti* ( + =1) *a ti* ( )−η*a* ∗∂*E* (1)  *p* ∂*ai*  where *t* is the learning epoch, η*a* is learning rate for *ai* and *p* is the number of input patterns.    Figure 2. The ANFIS structure |

In ANFIS, developing model relied upon the input-output data of a system under consideration. ANFIS structure as shown in Fig. 2 comprises of five layers. Each layer contains several nodes, which perform certain function based on the incoming signal and parameters associated with the nodes (Jang, 1993). The circular nodes are fixed whereas the square nodes contained parameters to be updated. Layer 1- referred as the input layer generates the membership grade which could be bell-shaped or Gaussian or trapezoid or triangular and this layer contained the premise parameters. Layer 2- the product layer, computes the firing strength of the rule through multiplication of the incoming signal from layer 1. Layer 3- normalizes the rules' firing strengths. Layer 4- calculates the rule outputs according to the consequent parameters. Layer 5- calculates the overall output as the sum of contribution from each rule. The mapping of input-output data to realize the desired model is done through updating the parameters (premise and consequent) via the hybrid learning algorithm until the stopping criterion is met. The hybrid learning algorithm uses the least square method to optimize the consequent parameters, and gradient descent to modify the premise parameters. The update rule for premise parameters is expressed as:

The ANFIS model implementation

The full scale data (influent/effluent) of the IWK plant were used to train the model. Input variable selection for model training is essential because it results in realizing a concise and accurate model. The data is converted into a trainable form by normalization using the equation The 180 days data were split into 135 days data for training and 45 days data for validation (applies to all the models). As suggested by Mansi, the choice of training and validation can be done arbitrarily or statistically depending upon the design employed, in this paper the random procedure was used.

*x i*( )−*v*min

*X i*( ) = *v*max −*v*min

(2)

where *x i*( ) is the sample value, *v*min and*v*max refers to the minimum and maximum value in the samples.

The normalized data were analysed using principal component analysis (PCA) in order to remove the redundant and conflicting information in the input vectors.

Based on PCA, four principal components (PCs) were extracted. These four principal components are enough to explain the over 95% variation within the data set. The components that contributed less 5% to the total variation within the data are removed. The percentage of the total variability explained by each principal component is shown in Fig. 3 and table 1. The four principal components were used in building the models. Using the available fuzzy toolbox of the Matlab 7.1 software, the function ‘‘genfis1’’ was applied to generate a first-order Sugeno fuzzy inference system (FIS) using grid partition on the data. The grid partition splits the data space based on the number of membership functions.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Each of the input variables was fuzzified with 4 the final output. As the FIS structure is now Gaussian membership functions which resulted made available, ANFIS utilizes the hybrid into 256 rules and each rule generates one rule learning algorithm to tune (optimize) the output. The aggregate of the rule outputs yielded parameters of the FIS.    Table 1. The percentage of the variance by each principal component   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | PCs | 1 | 2 | 3 | 4 | 5 | 6 | | Variance explained percent (%) | 51.6622 | 22.0591 | 15.6135 | 7.4586 | 2.0264 | 1.1802 |     1  2  3  4  0  10  20  30  40  50  60  70  80  90  Principal Component  Variance Explained (%)  0  %  10  %  20  %  %  30  40  %  50  %  60  %  70  %  80  %  90  %    Figure 3. The percentage of variance explained by the principal components |

Artificial neural network

Artificial neural network (ANN) is designed as a mathematical model of network inspired by biological neurons, which are massively interconnected to process and transmit information. The processing capability is kept in the interconnection strengths (weights) obtained during learning from training methods. ANN is made of input layer, hidden layer and output layer as shown in Fig. 3. The nodes in the layers receive the incoming signals, process them and yield the results (outputs). The direction of the flow of signal is governed by the connections between the nodes. The signal flow could be only in one direction or having feedback to the previous layer. When the flow is unidirectional, the NN is termed as feed-forward. Feedback NN has feedback connections, which help them to learn the temporal behaviour of the training data set (Engelbrecht, 2007) . Feed-forward neural network (FFNN) is the most widely used for input-output mapping for a linear or nonlinear function. The neural network (NN) training involves updating the weights (parameters) using an input-output data of the system to be modelled through learning algorithm such that the NN output is an accurate approximation of the system's output. The training can be a supervised or unsupervised depending on the structure of the network. In supervised training, the inputs and the desired outputs are introduced to the NN. The responses (outputs) of the NN to the given inputs are measured. The parameters are updated by learning principle (algorithm) to minimize the (error) between the actual and the desired output.

The learning principle is deduced by implementing a certain optimization method to a given error measure (Jang et al., 1997) . However, for unsupervised training only inputs are provided, the parameters are adjusted so that the inputs yield the outputs. In both cases, sufficient training data set which comprises of high and low values of the system under study is crucial.

**The ANN model implementation** For simplicity, this paper implemented feedforward neural network (FFNN). The network is formed by three layers with twenty-five neurons. The default TRAINLM training function, TANSIG transfer function in the hidden layer and PURELIN in the output layer were chosen. The parameters (weights) of the FFNN are tuned through learning rule (back propagation) to minimize the means square error between the FFNN output and the target. Fig. 4 illustrates the structure of the network with BOD as the output. The network is trained to realize the desired model. The use of checking data in all the models is to avoid the models from over fitting.

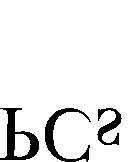
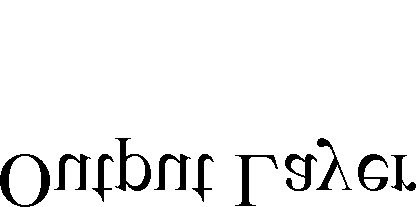
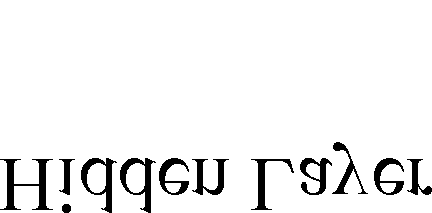
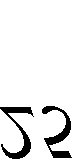
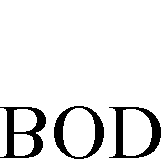
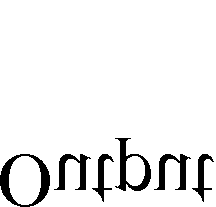
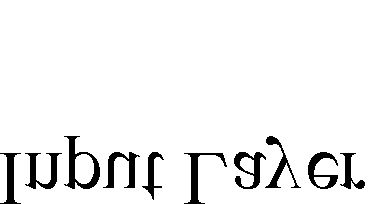


Figure 4. The structure of the feed-forward neural network with output COD

Nonlinear autoregressive with exogenous input NARX portrayed great potentials in representing an input-output relationship of a non-linear system. NARX has received attention due to its ability to capture well the dynamics of nonlinear functions, fast convergence and availability of methods for approximation (Aguirre, 2002). The main task in NARX model is obtaining the order of the polynomial and selecting the important terms from large candidates that exist for the model representation. The expression for NARX model is given as:

and nonlinear combinations of delayed input and output signals, and ϑ is the parameter vector containing the coefficient of polynomial expansion. These regressors are determined from the current and past input and output values, and then mapped by nonlinear estimator using a combination of linear and nonlinear functions to the model output.

NARX implementation

Several nonlinear estimators are provided for NARX model in the system identification toolbox. However, for effectiveness, parallel

*y t*( ) = *f* (*y t*( −1),...., *y t*( −*ny* ),*u t*( −1),....,*u t*( −

(8)

where *nu* and *ny* are the maximum lags of past inputs and outputs entering the model, ξ(*t*) depicts the possible noise and uncertainties, *f* is nonlinear function. In NARX model, *f* (•) is assumed to be a polynomial function, therefore the model (equation 8) can be written as a linear regression

# *y t*( ) =ψ ϑξ*T* (*t*) + (*t*) (9)

where the components of vector ψ(*t*) are referred as regressors, which contain the linear NARX neural network is selected for this *nu* ))+ξ( )*t* application. The NARX is implemented as neural network with output fed back to the input of the network. Apart from specifying the input delay vector as [1 2] and output delay vector as [1 2], the implementation is the same as that of feed-forward neural network. The Fig. 5 shows the structure of the network with COD as the output variable. As shown in the structure of the network, the feed- backing of the estimated value output improves the prediction accuracy of the model.

|  |
| --- |
| Figure 5. The NARX structure with output COD |

## RESULTS AND DISCUSSION

When dealing with several models the on-line comparison is very difficult or often impossible, since, simulation offers greater flexibility to compare various models under different situation easily. The performances of the models were comparatively evaluated through simulation

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Variable | Measure |  | Training | |  |  | Validation | |  |
|  |  | ANFIS | ANN | ARX | NARX | ANFIS | ANN | ARX | NARX |
| BOD | RMSE | 0.0021 | 0.0082 | 0.089 | 0.0061 | 0.0133 | 0.0340 | 0.090 | 0.0285 |
|  | MAPD(%) | 1.59 | 7.82 | 96.22 | 6.26 | 3.88 | 14.11 | 97.20 | 11.91 |
| COD | RMSE | 0.0012 | 0.0198 | 0.1495 | 0.0134 | 0.0478 | 0.0285 | 0.1561 | 0.0157 |
|  | MAPD(%) | 0.42 | 8.78 | 97.3 | 6.21 | 5.16 | 10.52 | 98.6 | 7.15 |
| SS | RMSE | 0.0008 | 0.0043 | 0.0760 | 0.0045 | 0.0086 | 0.0069 | 0.0785 | 0.0082 |
|  | MAPD(%) | 0.98 | 6.68 | 97.81 | 6.55 | 3.42 | 8.70 | 98.30 | 8.56 |
| NH4-N | RMSE | 0.0032 | 0.0066 | 0.1389 | 0.0091 | 0.0310 | 0.0444 | 0.1392 | 0.0268 |
|  | MAPD(%) | 1.70 | 4.99 | 95.60 | 5.64 | 6.40 | 12.46 | 96.17 | 10.37 |

Table 2. Prediction performance

using the variables BOD, COD, SS and NH4-N. The models were trained for each variable; the prediction performances of the models were determined by evaluating the measured and the estimated results through RMSE and MAPD as in Table 2.

Minimal RMSE and MAPD signify how well the predicted values capture the measured values. Fig. 6a shows the validation pattern for BOD variable, the ANFIS, ANN and NARX model demonstrated good agreement with the measured BOD, although the prediction of NARX is quite good but the prediction ANFIS model is better, as it estimated the measured values of the BOD accurately having the MAPD of 3.88%. For the COD variable as illustrated in Fig. 6b, the prediction of the ANFIS, ANN and NARX model is remarkable as both the predicted values of the models follow the measured COD and the performance measures are at their lowest values (MAPD less than 10%).

0

50

100

150

200

-0.2

-0.1

0

0.1

0.2

0.3

0.4

0.5

0.6

Day of operation

Concentration (mgl

-1

)

Measured BOD

ANFIS model

FFNN model

NARX model

Figure 6a. The models validation pattern for the effluent BOD

0

50

100

150

200

-0.1

0

0.1

0.2

0.3

0.4

0.5

0.6

0.7

Day of operation

Concentration (mgl

-1

)

Maesured COD

ANFIS model

FFNN model

NARX model

Figure 6b. The models validation pattern for the effluent COD

0

50

100

150

200

-0.1

0

0.1

0.2

0.3

0.4

0.5

0.6

0.7

0.8

0.9

Day of operation

Concentration (mgl

-1

)

Measured SS

ANFIS model

FFNN model

NARX model

Figure 6c. The models validation pattern for the effluent SS

0

50

100

150

200

-0.4

-0.2

0

0.2

0.4

0.6

0.8

1

Day of operation

Concentration (mgl

-1

)

Measured NH4-N

ANFIS model

FFNN model

NARX model

Figure 6d. The models validation pattern for the effluent NH4-N

The models were able to a certain high degree of accuracy to estimate the measured SS as shown in Fig. 6c; the MAPD is below 10% in both the three models. Similarly, for the measured NH4-N, the models exhibited a reasonable prediction as shown in Fig.6d.

## CONCLUSION

The paper has presented comparison of modelling approaches applied to the Bunu wastewater treatment plant. The results revealed that during the training the performance of ANN and NARX models are comparable to the ANFIS model. However, for the validation the ANFIS outperformed the other models having the lowest error measures in all the considered variables.

This demonstrates that the ANFIS model can serve as valuable tool in predicting the effluent quality in the plant.

## ACKNOWLEDGEMENT

The author wishes to thank the staff of IWK Bunus STP wastewater treatment plants

Malaysia for their support.

## REFERENCES

Aguirre, L.A. (2002). Nonlinearities in NARX polynomial models: representation and estimation. IEE Proc.: Control Theory Appl. 149, 343–348.

Benhalla, A., Mohamed H., & Moussa, C. (2010). Linearization of the full activated sludge model No 1 for interaction analysis. Bioprocess Biosyst. Eng. 33, 759–771.

Civelekoglu, G., Yigit, N.O., Diamadopoulos, E. & Kitis, M. (2009). Modelling of COD removal in a biological wastewater treatment plant using adaptive neuro-fuzzy inference system and artificial neural network. Water Sci. Technol.  60, 1475–1487. .

Civelekoglu, G., Perendeci, A., Yigit, N. O. & Kitis, M. (2007). Modeling carbon and nitrogen removal in an industrial wastewater treatment plant using an adaptive network-based fuzzy inference system. Clean: Soil, Air, Water 35, 617– 625.

Engelbrecht, A.P. (2007). Computational

Intelligence: An Introduction. Wiley,

Henze, Mogens. 2000. *Activated sludge models ASM1, ASM2, ASM2d and ASM3*. IWA Scient. London, UK: IWA Publishing.

Huang, M., Yongwen, M., Jinquan, W., Huiping, Z. & Yan, W. (2012). Modeling a papermaking wastewater treatment process by means of an adaptive network-based fuzzy inference system and principal component analysis. Ind. Eng. Chem.

Res*.* 51, 6166–6174. .

Jang, J.S.R. (1993). ANFIS: Adaptive-networkbased fuzzy inference system. IEEE

Trans. Syst. Man. Cybern. 23, 665–685. Jang, J.S.R, Sun, C.T., Mizutani, E. (1997). Neuro-fuzzy and Soft computing: A

Computational Approach to Learning

and Machine Intelligence. (New

Jersey:Prentice Hall).

Lewis, C.D. (1982). Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting. (Canada:Butterworth).

Nagy, A.M., Gilles M. & Marx, B. (2010). Systematic multimodeling methodology applied to an activated sludge reactor model. Ind Eng. Chem. Res.49, 2790– 2799.

Nelles, O. (2000). Nonlinear System Identification: From Classical Approaches to Neural networks and Fuzzy Models. Springer, Germany.

Pai, T.Y., Yang, P.Y., Wang, S.C., Lo, M.H., Chiang, C.F., Kuo, J.L., Chu, H.H., Su, H.C., Yu, L.F., Hu, H.C. & Chang, Y.H. (2011). Predicting effluent from the wastewater treatment plant of industrial park based on fuzzy network and influent quality. *Appl. Math. Model.* 35, 3674–3684. .

Perendeci, A., Arslan, S., Celebi, S. & Tanyolac, A. (2008). Prediction of effluent quality of an anaerobic treatment plant under unsteady state through ANFIS modeling with on-line input variables. Chem. Eng. J. 145, 78–85.

Raduly, B, & Gernaey, K.V. (2007). Artificial neural networks for rapid WWTP performance evaluation: Methodology and case study. Environ. Modell. Softw. 22, 1208–1216.

Ragot, J., Grapin, G., Chatellier, P. & Colin, F. (2001). Modelling of a water treatment plant: A multi-model representation.

*Environmetrics* 12, 599–611.

Smets, I.Y., Haegebaert, J.V., Ronald, C. & Van Impe, J.F. (2003). Linearization of the activated sludge model ASM1 for fast and reliable predictions. *Water Res.* 37, 1831–51.

Tay, J.H. & Zhang, X. (2000). A fast predicting neural fuzzy model for high-rate anaerobic wastewater treatment

systems. *Water Res.* 34, 2849–2860.