

Development of Artificial Neural Network Models for Biogas Production from Co-Digestion of Leachate and Pineapple Peel

Souwalak Jaroenpoj, Qiming Jimmy Yu^{*} and James Ness

Griffith School of Engineering, Griffith University, Nathan Campus, Brisbane, Queensland 4111, Australia

Abstract: The processes of anaerobic digestion and co-digestion are complicated and the development of computational models that are capable of simulation and prediction of anaerobic digester performances can assist in the operation of the anaerobic digestion processes and the optimization for methane production. The artificial neural network approach is considered to be an appropriate and uncomplicated modelling approach for anaerobic digestion applications. This study developed neural network models to predict the outcomes of anaerobic co-digestion of leachate with pineapple peel using experimental data. The multilayered feed forward neural network model proposed was capable of predicting the outcomes of biogas production from the anaerobic co-digestion processes with a mean squared error for validation of 2.67×10^{-2} and a R value for validation of 0.9942. The approach was found to be effective, flexible and versatile in coping with the non-linear relationships using available information.

Keywords: Artificial neural network models, biogas, anaerobic co-digestion, leachate, pineapple peel.

1. INTRODUCTION

Anaerobic digestion is a well-established bioprocessing technology to produce highly energetic biogas, which contains methane and carbon dioxide, to stabilise municipal sewage sludge and to treat organic wastes and wastewaters. As an important technology, the development of computational models for anaerobic digestion applications can assist in the operation and process control of the anaerobic digestion processes as well as maximising methane production.

Mathematical modelling of anaerobic digestion can be very complex, while more complexity is not necessarily better. Due to the complexities involved, the digestion process is often modelled as a black-box. Many studies have attempted to comprehend the behaviour of the process, but their solutions are still not applicable to all cases. Artificial neural network models can be used to overcome some of the problems of comprehending the anaerobic digestion processes [1-4].

An artificial neural network (ANN) is a computational simulation of the functionalities of the biological neural system that can be configured for particular applications. The key basics of artificial neural networks are based on an information processing paradigm and the learning processes of a biological neural system. Artificial neural networks offer alternative methods of analysing data and recognising patterns within the data in comparison with traditional mechanistic approaches. Learning method applications

of artificial neural networks are carried out by samples, while those of traditional computing approaches are typically achieved by rules. The approach of artificial neural network is particularly effective for modelling processes where the underlying relationships between input and output data are unknown. Artificial neural networks can also process problems of complex and nonlinear data, even though those data are noisy or imprecise. After training with known samples, artificial neural networks can be applied for predicting the outcome of a new set of independent input data as the networks have the ability to generalize [5-9]. However, artificial neural networks also have certain limitations. There is always a small error related to all neural network outputs because the neural network is aimed at finding an approximation of a solution. Errors of neural networks usually vary depending on their architecture. In addition, as the models need to be trained and validated with experimental data sets, ANN models are valid only for the particular systems for which the models were developed [6, 8, 10-11].

The ANN approach has been successfully used for anaerobic digestion processes for many application cases. According to Parthiban *et al.* [12], the ANN approach was applied for modelling the sago wastewater treatment parameters using an anaerobic tapered fluidised bed reactor. Six experimental parameters in the treatment process were considered for modelling. The input parameters were composed of influent flow rate, pH, COD, and hydraulic retention time, and the output parameters included effluent COD and CH₄ gas yield. The ANN simulations were carried out with MATLAB 7.1 using the back propagation algorithm, which was proved to have great adaptability

^{*}Address correspondence to this author at the Griffith School of Engineering, Griffith University, Nathan Campus, Brisbane, Queensland 4111, Australia; Tel: 61 7 37355289; Fax: 61 7 37357459; E-mail: jimmy.yu@griffith.edu.au

to various configurations and operation conditions. The prediction results from the neural network model developed were found to be quite close to the experimental results. The neural network model was also validated by replicative testing. The regression analysis fitting of all test data with the neural network model was 0.99924.

Strik *et al.* [13] developed ANN models to predict trace compounds (hydrogen sulphide and ammonia) in biomass from anaerobic digestion by the use of the Neural Network Toolbox of MATLAB 6.5. The architecture of the neural network models was constructed by applying ten different amounts of hidden neurons (one to ten). The standard back propagation with a gradient descent was used for training the multilayered neural networks. The models developed could predict the trace compounds under dynamic conditions with determination coefficients (R^2) of 0.91 for hydrogen sulphide and 0.83 for ammonia. It was concluded that the neural network models were appropriate as predictive control tools for this application.

A multilayered back propagation artificial neural network model with two hidden layers was constructed to simulate the anaerobic digestion operation as well as to predict the methane production [14]. The operational parameters used in the development of the model were collected and acquired from the Russaifah biogas plant which was fed with fresh organic waste from fruit and vegetable markets, restaurants, slaughterhouses and dairy processing plants. Four input parameters consisted of temperature, pH, total solids and total volatile solids, and the output parameter was CH_4 content. The neural network model was able to recognise the key patterns in methane production from the process and it fitted well with the measurement values. The performance of the neural network model to predict the methane production had a correlation coefficient of 0.87 [14].

Feed forward back propagation neural networks were developed to model the methane production in anaerobic continuous stirred digesters using the Neural Network Toolbox of MATLAB 5.2 [15]. The experimental parameters including pH, COD, volatile fatty acid concentration, volatile suspended solids, redox potential, gas composition and methane production rate were used to train a hierarchical system of the neural networks. The models could simulate the anaerobic digestion of surplus sludge and effectively predict the gas production rate. Thus, the hierarchical

system was introduced into a decision support system for advanced controlling of the anaerobic digestion process. A lab-scale anaerobic continuously stirred digester controlled by the system was capable of maintaining an approximate methane concentration of 60% at a high gas production rate in the range of 5–5.6 $\text{m}^3 \text{m}^{-3} \text{d}^{-1}$.

Ozkaya *et al.* [16] presented artificial neural network models for predicting the methane fraction in landfill gas from anaerobic digestion from field-scale landfill bioreactors at the Odayeri Sanitary Landfill, Turkey. The anaerobic digestion process was conducted with and without leachate recirculation. Two back propagation neural networks were built for the operation with leachate recirculation and for the operation without leachate recirculation, using the tan-sigmoid transfer function for the hidden layer and the linear transfer function for the output layer. The input parameters for the neural network model were selected based on the temperature of waste and leachate components. The eight input parameters included pH, COD, sulphate, chloride, waste temperature, alkalinity, conductivity and refuse age. The CH_4 fraction of landfill gas was used as the only output parameter. The optimal neural network structures of the CH_4 fraction prediction models for the operation with leachate recirculation and for the operation without leachate recirculation contained 12 and 15 hidden neurons in the hidden layers, respectively. The evaluation of the performance of the neural network models found that the network output appeared to well track the corresponding target with R value of 0.951 and 0.957 as well as mean squared errors of 0.00263 and 0.00250 for the predicted CH_4 fraction of the operation with leachate recirculation and the operation without leachate recirculation, respectively. In addition, the study established the optimal structure of the neural network and provided recommendations for further developments of the neural network for predicting the methane production from landfills on an hourly basis and optimising the leachate recirculation operation.

The purpose of this paper was to develop ANN models to predict the outcomes of anaerobic codigestion of leachate with pineapple peel using experimental data.

2. METHODS AND DATA

The neural network model for anaerobic codigestion process of leachate with pineapple peel was developed by using Neural Network Toolbox 7.0 of

MATLAB 7.11.0 (R2010b). The Neural Network Toolbox, which is a built-in tool in MATLAB, provides functions and applications for modelling complex nonlinear problems that are not simply modelled with a closed-form equation, such as data fitting, pattern recognition, time-series prediction, clustering and dynamic system modelling [17]. The neural network modelling procedures are summarised in the following flow chart (Figure 1).

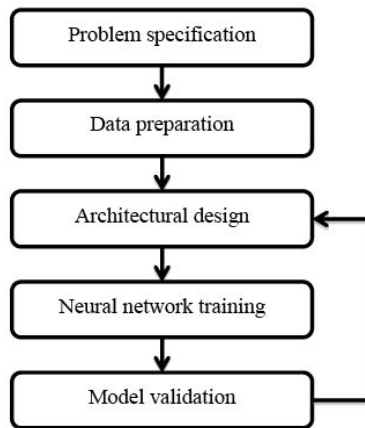


Figure 1: The neural network modelling procedures.

2.1. Problem Specification

It focuses on identifying the problem to be solved by the model and determining the degree of accuracy required. In this study, the model was determined to predict the outcomes of anaerobic co-digestion of leachate with pineapple peel using full scale plant experimental data. The degree of accuracy of the models was evaluated by using the mean squared error (MSE) performance.

2.2. Data Preparation

Data preparation is a step for gathering and choosing a proper set of data in order to ensure proper correlations. This is because the data used have an effect on performance of the neural network model developed in terms of its ability and reliability. The data used for developing the ANN models were obtained from experiments at a full scale plant for biogas production from anaerobic co-digestion of leachate with pineapple peel. The data required for the neural network models included input data and output (or target) data. Not all experimental data were used for the models as too many variables could cause the model to be very sensitive to small changes or noises in data. Thus, the parameters selected as input data for the neural network models included the mass amount of pineapple peel, pH of the inlet, COD of the inlet,

volatile fatty acids (VFA) of the inlet and volatile solids (VS) of the inlet. The selected target data consisted of VS of the outlet, the volume of biogas and the methane fraction of biogas. The input and target data were obtained from experimental results and then prepared in a Microsoft Excel Spreadsheet. The data set of 130 samples was separated into three subsets by random selection, as the training set, the validation set and the testing set.

2.3. Architectural Design

This step involves the construction of the neural network structure in terms of components of the neural network and its operations. It needs to identify the number of layers and the number of neurons in each layer. Trial and error method is commonly used for the determination. In this study, the multilayered feed forward architecture was used for designing the artificial neural network models. The networks were designed to contain three layers which were one input layer, one hidden layer and one output layer. Based on the results of the data preparation, the input layer of the neural network models contained five neurons while the output layer consisted of three neurons. As to the number of hidden neurons, this was determined by using trial and error experimentation, which is a common method for finding the optimal number of hidden neurons. The architecture of the neural network models was decided by applying ten different numbers of hidden neurons in the hidden layer (from one to ten neurons) for three different trials of data separation (% of training set: % of validation set: % of testing set) including 60%:20%:20%, 70%:15%:15% and 80%:10%:10%.

2.4. Neural Network Training

This is a step to adjust values of the connection weights and biases in order to generate the outputs with the given inputs. Training is a very important step since it determines the generalisation of the models. In this study, the Levenberg-Marquardt back propagation training algorithm (trainlm) was used to train the neural networks.

2.5. Model Validation

Performance and validation of the neural network must be achieved to evaluate its ability to solve the problem required. In this study, mean squared error (MSE) and regression R value were calculated to evaluate and validate performance of the neural network models.

3. RESULTS

3.1. Design of Artificial Neural Network Model

The architecture of the ANN model which generated the best result is illustrated in Figure 2.

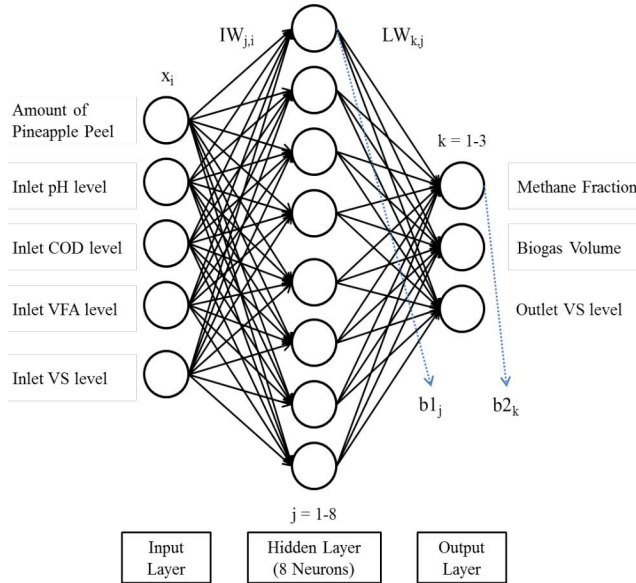


Figure 2: The neural network architecture.

As illustrated in Figure 2, eight hidden neurons with the data separation of 70%: 15%: 15% were found to yield the smallest mean squared error (MSE) for validation. It can be seen that each neuron in a layer is connected with all neurons in the following layer with weights and biases. The transfer functions applied for the neural network models included the hyperbolic tangent sigmoid function (tansig) and the linear transfer function (purelin). Tansig was used for the hidden layer while purelin was used for the output layer. The final form of the equation for the feed forward neural network model proposed can be then expressed as follows.

$$output_k = \sum_{j=1}^{j=8} LW_{k,j} \left[\frac{2}{1 + \exp\left(-2 \cdot \left(\sum_{i=1}^{i=5} (IW_{j,i} \cdot x_i) + b1_j\right)\right)} - 1 \right] + b2_k$$

where x_i is the input of the neuron; $IW_{j,i}$ is the connection weight between an input neuron and a hidden neuron; $LW_{k,j}$ is the connection weight between a hidden neuron and an output neuron; $b1_j$ is bias value of a hidden neuron; $b2_k$ is bias value of an output neuron.

3.2. Performance and Validation of Artificial Neural Network Models for Biogas Production from Co-Digestion

Performance and validation of the ANN models developed were evaluated by using mean squared errors (MSE) and regression R value. Results of the best neural network model with eight numbers of hidden neurons and the data separation of 70%: 15%: 15% are summarised in Table 1.

The chosen neural network generated the smallest MSE for the validation set, being 2.67×10^{-2} . The MSE for the training set was found to be 3.95×10^{-2} while the MSE for the testing set was 1.07×10^{-1} . The R value for validation of the chosen neural network was high, being 0.9942. Regression plots of the chosen neural network generated from the Neural Network Toolbox are shown in Figure 3.

As shown in the regression plots in Figure 3, the proposed neural network has a close correlation between the outputs and the targets. The R values of the training set, the validation set, the testing set, and the entire data set were found to be high, being 0.9944, 0.9942, 0.9800 and 0.9924, respectively.

4. DISCUSSION

The design of neural networks and the data separation were clearly achieved through the trial and error experimentation. According to the results of the different trials of the neural network models developed, the neural network with the smallest mean squared error (MSE) for validation was chosen. The MSE is an estimator of the average squared difference between the outputs and the targets. In other words, it is a sum of the variance and bias. A lower value of MSE thus

Table 1: Results of Neural Network Model which Generated the Smallest Mean Squared Error (MSE) for Validation

% of Separation	Number of Sample	Type of Sample	MSE	R
70%	90	Training	3.95×10^{-2}	0.9944
15%	20	Validation	2.67×10^{-2}	0.9942
15%	20	Testing	1.07×10^{-1}	0.98
100%	130	All	-	0.9924

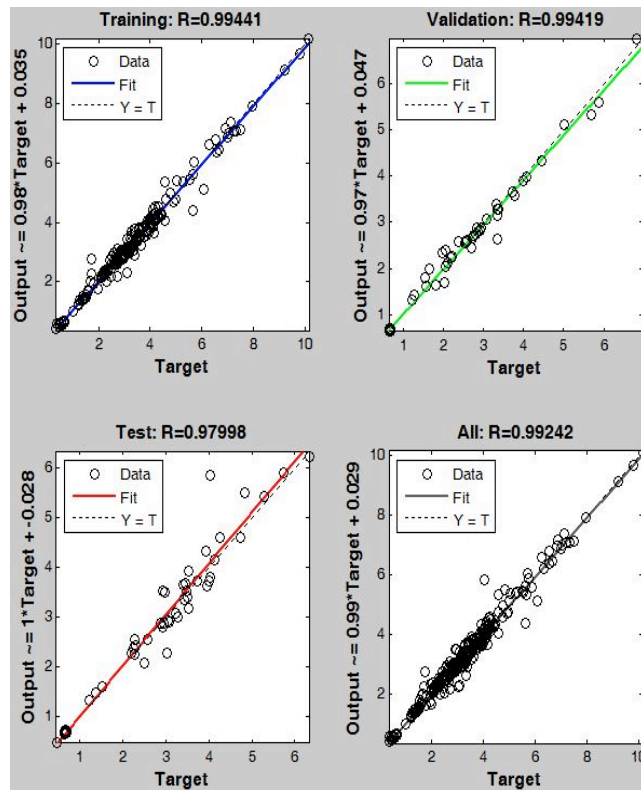


Figure 3: Regression plots of the proposed neural network model.

suggests a better result. The R value is an indication of the correlation between the outputs and the targets. Therefore, a higher value of R indicates a closer relationship and a zero R represents a random relationship. During the training stage, the connection weights of the neural networks were adjusted to minimise the MSE on the training set. Then, the neural networks were validated through the MSE of the validation set. As shown in the regression plots, R values of close to 1 were observed, which suggested that the prediction of the neural network model was linearly correlated with the experimental data. Therefore, it can be concluded that the proposed neural network model was capable of predicting the outcomes of biogas production from the anaerobic co-digestion process in terms of the VS level of the outlet, the volume of biogas and the methane fraction of biogas. However, this neural network model is only valid for the particular experimental conditions in which the data set was trained.

5. CONCLUSIONS

The present work described an artificial neural network approach for the computational predictions of various performance parameters that were influenced by the anaerobic co-digestion processes. The artificial

neural network models developed could be used to predict the outcomes of biogas production from the anaerobic co-digestion processes in terms of VS levels in the outlet, volumes of biogas production and methane fractions of the biogas under particular experimental conditions. Simulation and validation results demonstrated that artificial neural network based modelling was a practical and consistent approach for predicting the complex relationships in anaerobic co-digestion systems. The approach was found to be effective, flexible and versatile in coping with the non-linear relationships using available information.

REFERENCES

- [1] Hanrahan G. Artificial neural networks in biological and environmental analysis: CRC Press 2011. <http://dx.doi.org/10.1201/b10515>
- [2] Harmand J, Pons MN, Dagot C. 2012. Available from: http://apps.ensic.inpl-nancy.fr/COSTWWTP/Work_Group/Wg1/Magdeburg/Harma_nd_pres.pdf.
- [3] Pons M-N, Van Impe J. Computer Applications in Biotechnology 2004: Elsevier; 2005. <http://dx.doi.org/10.1117/3.633187>
- [4] Yu L, Wensel PC, Ma J, Chen S. Mathematical Modeling in Anaerobic Digestion (AD). J Bioremed Biodeg S 2013; 4: 2.
- [5] Kriesel D. A brief introduction to neural networks. Retrieved August 2007; 15: 2011.
- [6] Priddy KL, Keller PE. Artificial neural networks: an introduction: SPIE Press 2005.
- [7] Rabunal JR, Dorado J. Artificial neural networks in real-life applications: IGI Global 2006. <http://dx.doi.org/10.4018/978-1-59140-902-1>
- [8] Yegnanarayana B. Artificial neural networks: PHI Learning Pvt Ltd 2009.
- [9] Rajasekaran S, Pai GV. Neural Networks, Fuzzy Logic and Genetic Algorithm: Synthesis and Applications (With Cd): PHI Learning Pvt Ltd 2003.
- [10] Braspenning PJ, Thuijsman F, Weijters AJMM. Artificial neural networks: an introduction to ANN theory and practice Springer 1995.
- [11] Sivanandam S, Deepa S. Introduction to neural networks using Matlab 6.0: Tata McGraw-Hill Education 2006.
- [12] Parthiban R, Parthiban L. Back propagation Neural network modeling approach in the anaerobic digestion of wastewater treatment. International Journal of Environmental Sciences 2012; 2(4): 1944-51.
- [13] Strik DP, Domnanovich AM, Zani L, Braun R, Holubar P. Prediction of trace compounds in biogas from anaerobic digestion using the MATLAB Neural Network Toolbox. Environmental Modelling Software. 2005; 20(6): 803-10. <http://dx.doi.org/10.1016/j.envsoft.2004.09.006>
- [14] Abu Qdais H, Bani Hani K, Shatnawi N. Modeling and optimization of biogas production from a waste digester using artificial neural network and genetic algorithm. Resources. Conservation and Recycling. 2010; 54(6):359-63. <http://dx.doi.org/10.1016/j.resconrec.2009.08.012>
- [15] Holubar P, Zani L, Hager M, Frschi W, Radak Z, Braun R. Advanced controlling of anaerobic digestion by means of hierarchical neural networks. Water Research 2002; 36(10): 2582-8. [http://dx.doi.org/10.1016/S0043-1354\(01\)00487-0](http://dx.doi.org/10.1016/S0043-1354(01)00487-0)

- [16] Ozkaya B, Demir A, Bilgili MS. Neural network prediction model for the methane fraction in biogas from field-scale landfill bioreactors. *Environmental Modelling and Software* 2007; 22(6): 815-22.
<http://dx.doi.org/10.1016/j.envsoft.2006.03.004>
- [17] MathWorks. Neural Network Toolbox 2013. Available from: <http://www.mathworks.com/help/nnet/index.html>.

Received on 12-11-2014

Accepted on 22-11-2014

Published on 10-01-2015

DOI: <http://dx.doi.org/10.15377/2410-3624.2014.01.02.2>

© 2014 Jaroenpoj *et al.*; Avanti Publishers.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.