Artificial Neural Network Modelling for Biogas Production in Biodigesters

Artur S. C. Rego[a], Sibele Augusta F. Leite[b], Brenno S. Leite[b], Alexandre V. Grillo[c], Brunno F. Santos[a*]

[a]Department of Chemical and Materials Engineering (DEQM), Pontifical Catholic University of Rio de Janeiro (PUC-Rio). Rua Marquês de São Vicente, 225 – Gávea, Rio de Janeiro – RJ, 22430-060, Brazil.
[b]Institute of Science and Technology, Universidade Federal de Viçosa (UFV – *Campus* Florestal), Rodovia LMG 818, km 6, CEP: 35690-000, Florestal-MG, Brazil.
[c]Federal Institute of Education, Science and Technology (IFRJ), Rua Lúcio Tavares, 1045 – Centro, Nilópolis – RJ, 26530-060, Brazil.
bsantos@puc-rio.br

The use of the biodigestion is considered promising for the energetic valorization of agriculture biomass such as swine farm sewage and lignocelulosic residues. The understanding of biodigesters operation and the control of their main operational variables are of great importance to improve the performance of anaerobic digestion process in order to increase biogas production. In this context, mathematical modelling can be used as a tool to increase process efficiency. This work presents the development of Artificial Neural Network (ANN) and Adaptive Neuro Fuzzy Inference System (ANFIS) to predict volume of biogas. The variables from process were temperature (ºC), pH, FOS/TAC ratio and type of biodigesters. A database was constructed with the information of the experiments, dividing them into groups of training (67 %) and test (33 %). The models were obtained using MATLAB R2018b toolbox. In the developed neural models, the data obtained from process were used as neurons in the input layer and the volume of biogas was used as the only neuron in the output layer. The performance of the neural models was evaluated by determination coefficient (R^2) and error index (RMSE). The model developed from ANN and ANFIS modelling were satisfactory, showing the R^2 value giving the system's complexity. In addition, the RMSE values of both models were close to each other, showing agreement of the methods used.

1. Introduction

Anaerobic digestion is a well-established technology for biogas production. The process is carried out in the absence of oxygen in order to stabilize organic matter. Different types of biomass can be used in the process. For example, Surra et al. (2018) reported the use of municipal solid waste and maize cob waste as feedstock for biogas generation. The use of this technology with these feedstocks is environmentally attractive since it treats effluents and produces biogas that are a potential source of heat and energy (Leite et al., 2018).

Despite of its simplicity, anaerobic digestion is highly dependent of its environment variables (parameters) for instance pH, temperature, organic load, acidity/alkalinity and ammonia concentration (Cuadros et al., 2011; Mata-Alvarez et al., 2014; Yadvika et al., 2004). Hence, control of the main operational parameters is a priority when it comes to bioprocess since small changes in the operation can result in different biogas yields and composition.

The need to control and to optimize the anaerobic digestion process has led to the development of mathematical models that can be used to promote a more efficient process (Kythreotou et al., 2014). To develop these models, alkalinity deserves a special attention. This parameter is related to the buffering capacity of anaerobic digestion systems and it is considered an indicator of the process imbalance of the. Because of its importance, different alkalinity measurement methods were developed to monitor the biodigestion process (Wang et al., 2018).

One alternative way to measure the alkalinity is using FOS/TAC ratio (also known as IA/PA). The first parameter (FOS) represents the accumulation of volatile fatty acids and the second parameter (TAC) is a measure of the buffer capacity in the digester due to bicarbonates alkalinity. Stable operating conditions have a FOS / TAC ratio smaller than 0.30. The increase in the ratio indicates inhibition of methanization, due to accumulation of organic acids, which are intermediates in biodigestion process (Drosg, 2013).

There is a great amount of mathematical models for anaerobic digestion available in the literature. They can be divided in the following categories: the ones based on kinetic parameters such as microorganisms growth or substrate consumption; the statistical models; computational fluid dynamics (CFD); chemical composition of the feedstock and the artificial neural networks (Kythreotou et al., 2014; Xie et al., 2016). The last one has great potential to be used in the biodigestion process modelling since it requires a minimum knowledge of the reaction mechanisms and experimental measures of the variables accessed in order to optimize the biogas yield (Xie et al., 2016).

Some works using neural networks have been carried out in order to predict biogas yield using substrates characteristics and composition (Kana et al., 2012; Beltramo et al., 2016; Verdaguer et al., 2016). Ghatak et al. (2018) investigated a neural network model to predict the behavior of biogas production curve at various temperatures. Neural network technique has also been used to predict methane production, to avoid shock-loads and to optimize anaerobic digestion, using operational data from the process (Holubar et al. 2002; Qdais et al, 2010). Temperature, pH, chemical oxygen demand (COD), volatile fatty acid (VFA), alkalinity, volatile solids (VS) and biogas/methane flow are the mostly used process variables to feed the models (Holubar et al. 2002; Qdais et al, 2010; Guwy et al., 1997; Yetilmezsoy et al., 2013). However, as far as we are concerned, neural networks models mostly used bicarbonate alkalinity as operational data instead of FOS/TAC ratio (Guwy et al., 1997; Yetilmezsoy et al., 2013; Ozkaya et al., 2007), which may not be the best option to indicate process stability.

In the present study, a model was proposed concerning the biogas volume produced according to different operational conditions (reactor type, temperature, pH and FOS/TAC). In order to solve the optimization problem, two types of neural networks were tested: Artificial Neural Network (ANN) and Adaptive Network-based Fuzzy Inference System (ANFIS), a combination of artificial neural networks and fuzzy logic. To ensure the quality of the results, determination coefficient (R^2) and root mean square error (RMSE) were used for both methods.

2. Material and methods

2.1 Sampling

The swine manure (SM) was collected in the finishing stage at the swine breeding unit (lineage Agroceres) of University of Viçosa (UFV - Campus Florestal, Brazil). Samples of rice husk (RH) were also collected at UFV - Campus Florestal, during a period of four months. They were pre-dried at 65 °C in an oven-dry for two days to eliminate extrinsic moisture and to obtain samples in the same operating conditions and able of being handled. Thereafter, they were cut in a knife mill (Marconi-MA280) to obtain fibers between 0.5-1.0 mm length. After that, they were stored for conducting physicochemical analysis and biodigestion experiments.

2.2 Experimental

Three biodigestion sets were operated in order to gain data to feed the neural networks model. All of them were operated at the same conditions. The biodigesters had different design: 1st) jacketed inox reactor, without recirculation of the effluent, 5 L nominal capacity; 2nd) reactor made of PVC, with internal heating coil, without recirculation of the effluent, 7 L nominal capacity and 3rd) reactor made of PVC, with internal heating coil, with recirculation of effluent to provide mixing, 7 L nominal capacity.

They were fed with a mixture of swine sewage and rice husk, in order to provide organic load between 1.0 and 1.5 g (VS) L^{-1} d^{-1}. Rice husk was used to provide more carbon source and it was added in a proportion of 2 wt. %.

2.3 Process monitoring

Biogas production was monitored and measured at a regular interval through water displacement method. The biogas volume was converted to the volume in Normal Temperature and Pressure Conditions (NTP). In order to set the organic load, rice husk and swine sewage were submitted to VS analysis (American Public Health Association, APHA, 1998; Rendeiro, 2008).

To obtain FOS/TAC ratio, 20 mL of effluent samples (periodic withdraw from de biodigester after treatment) were centrifuged and the supernatant was treated with sulfuric acid 0.05 mol L^{-1}. Titration was first carried out until pH 5.0 (bicarbonate alkalinity) and then until pH 4.4 (alkalinity caused by VFA). The FOS/TAC alkalinity ratio was calculated as proposed by Drosg (2013).

The pH measurements were performed by direct reading using a TEKNAL/ T1000 brand benchmark potentiometer.

2.4 Artificial neural network (ANN) architecture

The model was built using four operation conditions (reactor type, temperature, pH and FOS/TAC) as inputs and biogas volume as output. In ANN, the number of hidden layers, the number of neurons in each hidden layer, the training algorithms (*trainbr, trainlm*) and the activation functions (*logsig, tansig* – Eq(1) and Eq(2), respectively) were varied in order to reach the best topology in the optimization. Both training algorithms used are a variation of Backpropagation. In this algorithm, firstly the information goes forward from the input layer to the output layer and then goes backward in order to update the network parameters minimizing the error. Both training algorithms tested use the Levemberg-Marquadt optimization method, but the trainbr also uses Bayesian regularization. To evaluate the model's performance, RMSE, showed in Eq(3), and R^2, showed in Eq(4), values were used in Matlab 2018b.

$$logsig(x) = 1/{1 + e^{-x}} \tag{1}$$

$$tansig(x) = 2/{1 + e^{-2x}} - 1 \tag{2}$$

In which x is the value that comes from the input layer.

$$R^2 = \sum_{i=1}^{n}(\hat{x}_i - \bar{x})^2 \Big/ {\sum_{i=1}^{n}(\hat{x}_i - \bar{x})^2 + \sum_{i=1}^{n}(x_i - \hat{x}_i)^2} \tag{3}$$

$$RMSE = \sqrt{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2 \Big/ n} \tag{4}$$

In which x_i is the experimental value, $\hat{x}_i$ is the estimated value, $\bar{x}$ is the mean value and n is the data size in the input layer.

2.5 Adaptive neuro fuzzy inference system (ANFIS) architecture

Sugeno fuzzy inference method was used in the system modelling in which the fuzzy rules are related to the input variables and the defuzzification method is given by a Membership Function (MF). In this case, the neuro fuzzy system based in Subtractive Clustering (SC), with linear function in the output, using hybrid training algorithm was tested. SC is a method to determine the amount of clusters and clusters canters needed to solve the problem. MF used in the SC is the Gaussian curve function (*gaussmf* – Eq(5)). The same database used in the neural network model was given to the neuro-fuzzy model (four inputs and one output). Four parameters (Range of influence, Squash factor, Accept ratio and Reject ratio) inside SC were varied in order to reach the best configuration to fit in the dataset. To evaluate the model's performance, RMSE and R^2 values were used in Matlab 2018b.

$$gaussmf(x, \sigma, c) = e^{-\frac{(x-c)^2}{2\sigma^2}} \tag{5}$$

In which c is the center of the Gaussian function, σ is the width of the curve and x is the experimental value.

3. Results and discussion

3.1 ANN modelling of biodigestion

The ANN modelling concerned prediction of the biogas volume produced. The results of experimental biogas volume produced in the biodigesters were used as database for the training (67 %) and testing (33 %) of the models.

Table 1 displays some ANN modelling results for the biogas volume produced, with different activation functions, training algorithms, number of hidden layers and number of neurons in each hidden layer. The R² and RMSE values were analyzed to select the best topology.

Table 1: Artificial neural network topologies used in biogas volume prediction

Topology	Hidden layer 1		Hidden layer 2		Training algorithm	R²	RMSE
	Neurons	Activation function	Neurons	Activation function			
1	10	*logisg*	-	-	*trainbr*	0.60701	0.143178
2	8	*logsig*	-	-	*trainbr*	0.61114	0.1249
3	16	*tansig*	-	-	*trainlm*	0.5663	0.205426
4	20	*logsig*	-	-	*trainbr*	0.69391	0.087579
5	24	*tansig*	-	-	*trainbr*	0.63891	0.124499
6	10	*logsig*	10	*tansig*	*trainbr*	0.77704	0.088487
7	7	*logsig*	10	*tansig*	*trainlm*	0.62051	0.118322
8	12	*logsig*	10	*logsig*	*trainbr*	0.73549	0.108628
9	12	*tansig*	10	*logsig*	*trainlm*	0.55618	0.107703
10	12	*logsig*	12	*tansig*	*trainbr*	0.72905	0.096747

The results indicate that the process has a non-linear tendency given that the performance indicators of the artificial neural networks were not very adequate, with R^2 values diverging from 1.0 and relatively high RMSE values. This could occur due to the unpredictable nature of living microorganisms' metabolism (Campos et al., 2018). However, the best proposed topology was capable of predicting biogas volume with a determination coefficient of 0.77704 using two hidden layers with 10 neurons in each of them, and with *logsig* in the first hidden layer and *tansig* in the second hidden layer as activation functions. The training algorithm was *trainbr*. Furthermore, RMSE value for this topology is one of lowest, indicating that this is the best proposed topology to predict biogas volume.

Figure 1 shows the R^2 value for the best topology alongside a comparison of experimental and modelled data. Despite the calculated values not having the same value as the experimental data, it is noticeable that both the values follow the same pattern, showing neural network prediction capability.
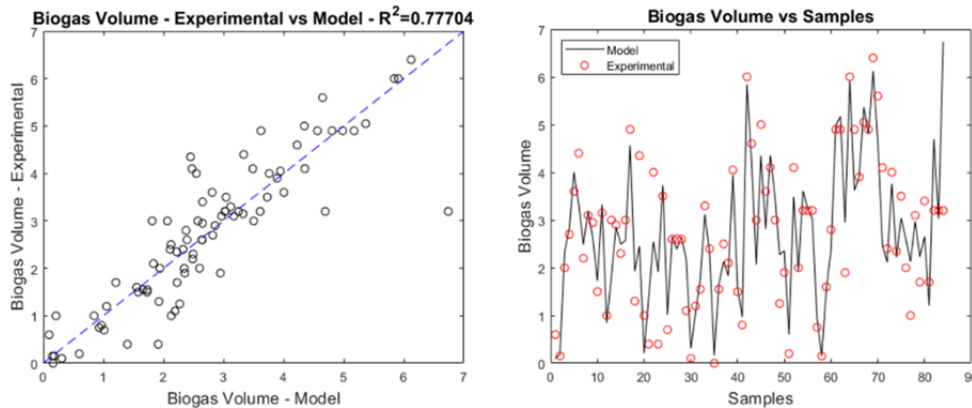

Figure 1: Result of the regression plot of biogas volume prediction (left) and biogas volume for each sample in the test data (right) for ANN.

3.2 ANFIS modelling of biodigestion

As well as ANN, ANFIS also concerned the prediction of biogas volume produced, following the same methodology used in ANN modelling, using the same database for training (67 %) and testing (33%). All the tested configurations of ANFIS models are depicted in Table 2, testing different values of range of influence, squash factor, accepted ratio and rejected ratio. The variation of these parameters causes the set of rules to change. The first configuration shows the standard values proposed by Matlab 2018b, resulting in 16 rules. A high number of rules can cause an overfitting problem, which can avoid the model to be a good representation of the process.

Table 2: ANFIS configurations used in biogas volume prediction

Configuration	Range of influence	Squash factor	Accepted ratio	Rejected ratio	Number of rules	R²	RMSE
1	0.5	1,25	0,5	0,15	16	0.75262	0.12071
2	0.7	1	0.7	0.15	8	0.79704	0.10827
3	0.6	1	0.5	0.15	17	0.81209	0.10423
4	1	1	0.5	0.15	4	0.52096	0.16565
5	0.5	1	0.5	0.3	13	0.76171	0.11704
6	0.7	1.25	0.5	0.15	7	0.77338	0.11469
7	0.5	1.25	0.5	0.3	8	0.67352	0.13703

It is important to observe that a high amount of rules do not mean that the model will have a better performance over another, for instance topology nº 2 had a better performance when compared with topology nº 1, but had half the amount of rules.

The results of ANFIS modelling also shows a non-linear tendency of the problem, in agreement with the results found in ANN modelling, with R² values distant from 1.0 and RMSE relatively high. In general, the neuro fuzzy modelling showed slightly better results when compared with the neural network, with higher R² values and RMSE values in the same range. The best ANFIS configuration was capable of predicting the biogas volume with a determination coefficient of 0.81209, showing that despite of process complexity, the model had a good prediction capability.

Figure 2 displays the comparison of modelled and experimental data for the best ANFIS configuration. Even though calculated and experimental values did not show total agreement, the model was capable to follow the pattern showed by experimental data, displaying the model's prediction capability.
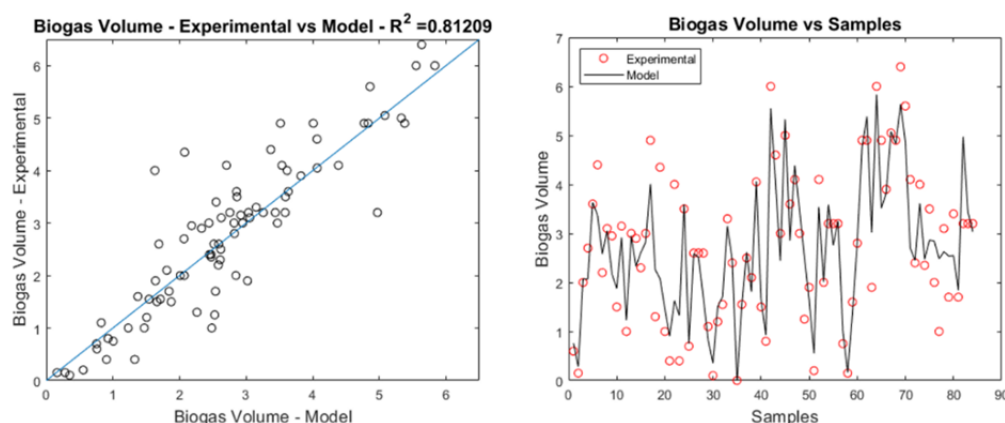

Figure 2: Result of the regression plot of biogas volume prediction (left) and biogas volume for each sample in the test data (right) for ANFIS.

Due to the non-linear tendency of the problem, it would be necessary to increase the database of the inputs in order to enhance of both ANN and ANFIS performances. That would cause the network to have a better idea of possible scenarios regarding biodigestion, hence increasing the networks capability to train in different situations and having a better performance in the test phase.

4. Conclusions

Swine manure and rice husk can be a renewable feedstock to be used in the biodigestion process in order to generate biogas. The mathematical complexity of the process can be expressed in terms of neural networks modelling, allowing the process to be better understood, hence enhancing process control and monitoring. This study showed that both ANN and ANFIS modelling have good prediction capability of biogas volume produced in a given condition. ANN had R² value of 0.77704 at its best topology meanwhile ANFIS had R² value of 0.81209 at its best configuration, showing better performance than ANN. In order to have a better knowledge of all the variables influences, it would be desirable to acquire more data to input in the networks.

References

American Public Health Association (APHA), American Water Works Association (AWWA), Water Environment Federation (WEF), 1998, Standard Methods for the Examination of Water and Wastewater, Washington, DC.

Beltramo, T., Ranzan, C., Hinrichs, J., & Hitzmann, B., 2016, Artificial neural network prediction of the biogas flow rate optimised with an ant colony algorithm. Biosystems Engineering, 143, 68-78.

Campos A.L., Nogueira J., Coelho F.A., Fileti A.M.F., Santos B., 2018, Genetic algorithm optimization of the parameters involved in biosurfactant production from beet peel as substrate, Chemical Engineering Transactions, 65, 469-474 DOI: 10.3303/CET1865079

Cuadros, F., López-Rodríguez, F., Ruiz-Celma, A., Rubiales, F., & González-González, A., 2011, Recycling, reuse and energetic valuation of meat industry wastes in Extremadura (Spain). Resources, Conservation and Recycling, 55(4), 393-399.

Drosg, B., 2013, Process monitoring in biogas plants. In IEA Bioenergy Task (Vol. 37).

Ghatak, M.D., & Ghatak, A., 2018, Artificial neural network model to predict behavior of biogas production curve from mixed lignocellulosic co-substrates. Fuel, 232, 178-189.

Guwy, A.J., Hawkes, F.R., Wilcox, S.J., Hawkes, D.L., 1997, Neural network and on-off control of bicarbonate alkalinity in a fluidised-bed anaerobic digester. Water Research, 31(8), 2019-2025.

Holubar, P., Zani, L., Hager, M., Fröschl, W., Radak, Z., Braun, R., 2002, Advanced controlling of anaerobic digestion by means of hierarchical neural networks. Water Research, 36(10), 2582-2588.

Kana, E.G., Oloke, J.K., Lateef, A., Adesiyan, M.O., 2012, Modeling and optimization of biogas production on saw dust and other co-substrates using artificial neural network and genetic algorithm. Renewable Energy, 46, 276-281.

Kythreotou N., Florides G., Tassou S.A., 2014, A review of simple to scientific models for anaerobic digestion, Renew. Energy, 71, 701–714.

Leite S., Leite B., Figueiredo M., Dell'Isola A.T., Dangelo J.V., 2018, Biogas production on a small swine farm: study of prediction using different models, Chemical Engineering Transactions, 65, 85-90 DOI: 10.3303/CET1865015

Mata-Alvarez, J., Dosta, J., Romero-Güiza, M.S., Fonoll, X., Peces, M., Astals, S., 2014, A critical review on anaerobic co-digestion achievements between 2010 and 2013. Renewable and Sustainable Energy Reviews, 36, 412-427.

Ozkaya, B., Demir, A., Bilgili, M.S., 2007, Neural network prediction model for the methane fraction in biogas from field-scale landfill bioreactors. Environmental Modelling & Software, 22(6), 815-822.

Qdais, H. A., Hani, K.B., Shatnawi, N., 2010, Modeling and optimization of biogas production from a waste digester using artificial neural network and genetic algorithm. Resources, Conservation and Recycling, 54(6), 359-363.

Rendeiro, G., 2008, Combustão e gasificação de biomassa sólida. Ministério de Minas e Energia (MME).

Surra E., Bernardo M., Lapa N., Esteves I., Fonseca I., Mota J.P., 2018, Enhanced biogas production through anaerobic co-digestion of ofmsw with maize cob waste pre-treated with hydrogen peroxide., Chemical Engineering Transactions, 65, 121-126 DOI: 10.3303/CET1865021

Verdaguer, M., Molinos-Senante, M., Poch, M., 2016, Optimal management of substrates in anaerobic co-digestion: an ant colony algorithm approach. Waste Management, 50, 49-54.

Wang, X., Bai, X., Li, Z., Zhou, X., Cheng, S., Sun, J., Liu, T., 2018, Evaluation of artificial neural network models for online monitoring of alkalinity in anaerobic co-digestion system. Biochemical Engineering Journal, 140, 85-92.

Xie, S., Hai, F. I., Zhan, X., Guo, W., Ngo, H.H., Price, W. E., Nghiem, L.D., 2016, Anaerobic co-digestion: A critical review of mathematical modelling for performance optimization. Bioresource Technology, 222, 498-512.

Yadvika S., Sreekrishnan T.R., Kohli S., Rana, V., 2004, Enhancement of biogas production from solid substrates using different techniques–a review. Bioresource Technology 95, 1–10.

Yetilmezsoy, K., Turkdogan, F. I., Temizel, I., & Gunay, A., 2013, Development of ann-based models to predict biogas and methane productions in anaerobic treatment of molasses wastewater. International journal of green energy, 10(9), 885-907.