Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

ISSN: 0972-6268	Vol. 14	No. 4	pp. 811-816	2015

Original Research Paper

Research of Interpolation and Prediction by Elman NN on Anaerobic Digestion Processes Parameter

Yong Liang*(**), Ling Qiu*(**)†, Junting Pan*(**) and Wen Lu***
*College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, China
**The West Experimental Station of Rural Renewable Energy Exploitation and Utilization of Ministry of Agriculture, Yangling, China
***Guilin University of Electronic & Technology, Guilin, China
†Corresponding author: Ling Qiu

Nat. Env. & Poll. Tech.
Website: www.neptjournal.com
Received: 29-12-2014
Accepted: 03-03-2015

Key Words:
Silage cornstalk
Anaerobic digestion
Interpolation
Elman neural network
Simulation

ABSTRACT

Parametric data from anaerobic digestion processes are normally collected once every couple of days and not daily. As a result, only a small amount of data could be collected and this is not sufficient for the neural network analysis. In this research, interpolations were used during the modelling process to increase the sample data used for Elman neural network (Elman NN) modelling. Laboratory digestion of silage cornstalk was conducted for 54 days, and a portion of the biogas data was used for training the Elman neural network (Elman NN) model, while the remaining biogas data were used to verify the prediction capability of the model. Compared to the Elman NN model without interpolations, using an interpolation coefficient of 0.2 increased the number of experimental data from 54 to 266 and the correlation coefficient of prediction data and sampling data from 0.7966 (for no-interpolation) to 0.9962 (for cubic spline interpolation) and 0.9942 (for piecewise linear interpolation). In addition, the mean square error decreased from 0.1190 (for no interpolation) to 0.001 (for cubic spline interpolation) and 0.001 (for piecewise linear interpolation), while the average relative error decreased from 63.04% (for no interpolation) to 3.93% (for cubic spline interpolation) and 4.01% (for piecewise linear interpolation). The Elman NN simulation results thus showed that the interpolation algorithm can greatly improve the prediction accuracy of biogas production from an anaerobic digestion process.

INTRODUCTION

Anaerobic digestion is a process where organic matter is converted into methane and carbon dioxide through the metabolic activity of various microbes in an anaerobic environment under favourable conditions (Li et al. 2011). More attention is paid towards anaerobic digesters since the process produces clean energy and high quality organic fertilizer that, in turn, could protect the environment. With the ability to produce stable quantities of gas efficiently, anaerobic digestion could play an important role in the production of renewable energy in the future (Appels et al. 2011). Anaerobic digestion is an ecosystem of interdependence and mutual restraint among various microorganisms. Both, non-methanogenic and methanogenic bacteria rely on nutrients, maintained homoeostasis, and interactions such as mutualism, competition, and product inhibition in three-stage theory (Yan 2007). Therefore, anaerobic digestion is a complex system of multivariate coupling, multi-substrate, and multi-operating conditions, and includes a number of biological, physical and chemical processes that interact with each other. In order to accurately describe the anaerobic digestion process, the establishment of an appropriate mathematical model is needed.

There are two kinds of anaerobic digestion model: mechanism and black-box. Mechanism model consists of equations derived by deduction, while black-box model is inferred from a set of experimental data (Strik et al. 2005). The typical mechanical model is the ADM1 model developed by IWA (Batstone et al. 2002). The model takes into account various processes including decomposition, hydrolysis, acidogenic, acetogenic, methane production, etc. (Bollon et al. 2011). Since the digestion system is complex and contains numerous processes and principles, some of which are still undetermined, ADM1 model has been modified greatly to incorporate specific applications. The black-box model has unique advantages when the anaerobic digestion process could not be understood well or parameters are difficult to obtain (Striket et al. 2004). Artificial neural network is a typical black box model that could adapt to changes of system by learning without the prior knowledge of the structure and relationships between variables (Merkel et al. 1999). Because this model does not consider the chemical, physical, and microbial processes of anaerobic digestion (in most cases), this method would be attractive when the target is a prediction of a few specific output variables (Lauwers et al. 2013). However, neural networks need a large set of data for machine learning to achieve better statistical verification and high reliability (Jin et al. 2011). In addition, while anaerobic digestion period is generally 30-90 days or even longer, most experimental studies only obtained the digestion process parameters once every couple of days (Souza et al. 2013, Yuanfang et al. 2012, Kafle et al. 2013, Krishania et al. 2013). Hence, the amount of data did not meet the conditions for the application of neural network model. To resolve this contradiction, data interpolation could be used. Common interpolation methods include linear interpolation, polynomial interpolation, and spline interpolation. Ben et al. (2012), for instance, used interpolation method to reconstruct the input sample continuity for a fed-batch fermentation producing lactic acid. Bhowmik et al. (2011) compared three interpolation methods (spline, inverse distance weighting, kriging) to create continuous surfaces that describe temperature trends. In this paper, Elman neural network model (Elman NN) was used to predict and compare daily gas production, using the data obtained from piece-wise linear interpolation and cubic spline interpolation.

MATERIALS AND METHODS

Experimental device: The small-scale laboratory anaerobic fermentation device consisted of six parts: bottle for anaerobic digestion (2000 mL plastic bottle), gas collection bottle (1000 mL conical flask), sample collection tube, three-way glass stopcock, glass tube, beaker for collecting displaced water, and other auxiliary equipment (Fig. 1). These components were connected through rubber tubes and rubber stoppers.

Experimental materials: The substrate used for the experiment was corn straw that had a silage time of two weeks. Inoculum (pH: 7.2) was a mixture of digestate obtained from a well-run biogas digester and pig manure (pH: 7.0) with weight ratio of 1:1. The inoculum was incubated and acclimated for 15 days to obtain high quality and abundant species. The physical and chemical properties of silage corn straw and inoculum can be found in Table 1.

Experimental design and methods: The fermentation bottle was filled with 122.23 g of silage corn straw, which was cut into short pieces, and 450 g of inoculum. Urea was added to adjust the carbon-nitrogen ratio to 25:1. Finally, distilled

Table 1: Physical and chemical characteristics of silage corn straw and inoculum.

Materials	TS/%	VS/%	TOC/%	TN/%
Silage corn straw	68.25	62.45	46.50	0.75
Inoculum	10.00	5.60	3.19	0.2

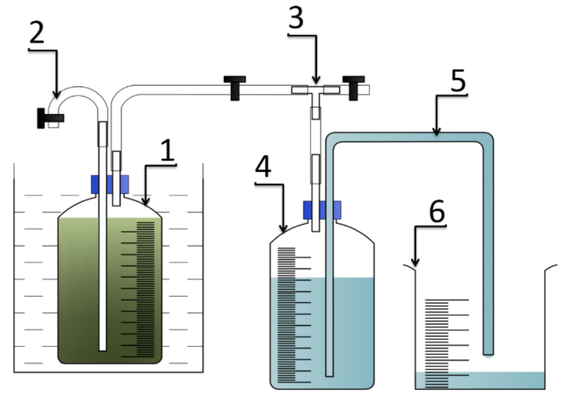

Fig 1: Schematic diagram of the experiment.
1. Bottle for anaerobic digestion 2. Sample collection tube
3. Three-way glass stopcock 4. Gas collecting bottle 5. Glass tube
6. Beaker for collecting water

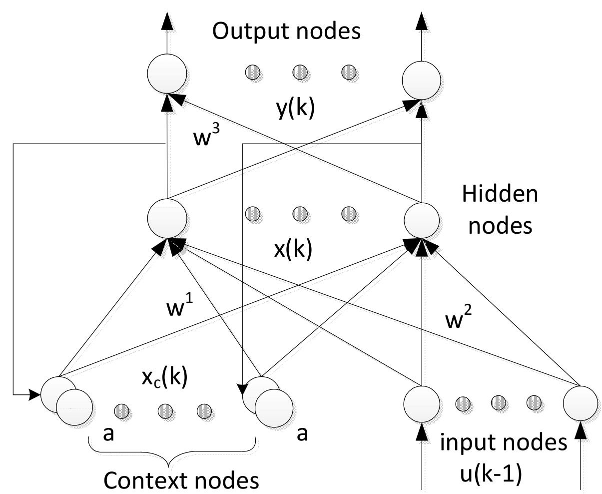

Fig. 2: Structure of Elman NN.

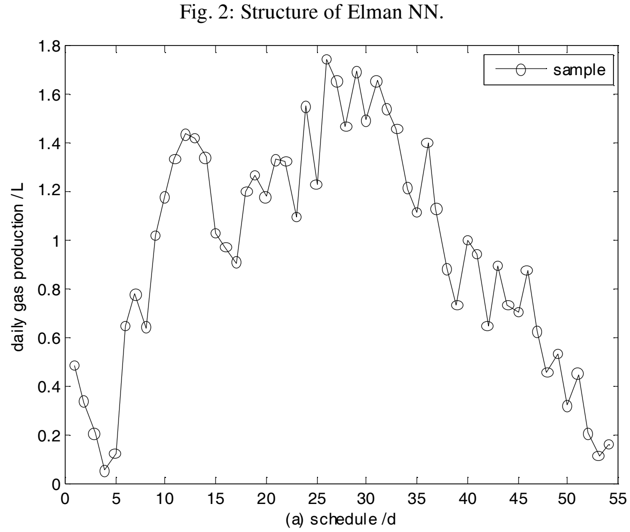

Fig. 3: Daily gas production of anaerobic digester.

water was added to bring the concentration of total solids (TS) to 8%. The experiment was conducted in triplicate, and the average of their biogas production was presented as the final result. The digestion process lasted for 54 days and biogas production was measured every day.

Determination index and method: Total solids (TS) was determined by heating the fermentation material at 105°C for six hours, and volatile solids (VS) was determined by heating the materials from TS analysis in a muffle furnace at 600°C for one hour. Biogas yield was determined by water-displacement method, in which the water displaced by produced biogas was measured using a graduated cylinder. Total organic carbon (TOC) was measured using a TOC analyser, while total nitrogen (TN) was measured using Kjeldahl nitrogen instrument.

Structure of Elman NN: Elman NN is a dynamic feedback network. In addition to the input layer, hidden layer and output layer, it has a special context layer. The special context layer is considered a delay operator and is used to store the previous moment output value of the hidden layer, creating a feedback network. The feedback network is sensitive to historical data, which improves the dynamic information processing capacity of Elman NN, allowing it to carry out dynamic modelling. Elman NN also can approximate any nonlinear function with arbitrary precision and need not consider the specific form of external noise acting on the system. The structure of Elman NN is shown in Fig. 2.

Elman NN mathematical model: The following mathematical model can be used to describe the Elman NN (Shi et al. 2004):

$$x(k) = f(w^{I1}x_c(k) + w^{I2}u(k-1)) \quad ...(1)$$

$$x_c(k) = \alpha x_c(k-1) + x(k-1) \quad ...(2)$$

$$y(k) = g(w^{I3}x(k)) \quad ...(3)$$

Where, W^{I1} = the matrix of the connecting context layer with the hidden layer; W^{I2} = the matrix of connecting input layer with the hidden layer; W^{I3} = the matrix of connecting output layer with the hidden layer; $x_c\ (k)$ = output of the context layer; $x\ (k)$ = output of the hidden layer; $u\ (k\text{-}1)$ = output of the input layer; $y\ (k)$ = output of the output layer; α = the feedback gain factor with a range of $0 \leq \alpha < 1$; $f(x)$ = sigmoid function similar to equation (4).

$$f(k) = \frac{1}{1+e^{-x}} \quad ...(4)$$

Where e = the natural logarithm; x = the independent variable.

RESULTS

Experimental results: The duration of anaerobic digestion experiment lasted for 54 days, and a total of 54 biogas production data were collected. The daily gas production data are shown in Fig. 3.

Daily biogas values predicted by Elman NN: The 54 daily gas production data as a whole were divided into two parts: the first part contained 38 data (70%) and was used for training, while the second part contained 16 data (30%) and was used for prediction purposes. Elman NN was trained 1000 times and the resulting training error was 10^{-4}. No interpolation was used in this prediction run. Daily gas production and relative errors are shown in Fig. 4.

The simulation results can be seen in Fig. 4. The correlation coefficient between predicted data with measured data was 0.7966, while the mean square error was 0.119. The average relative error was 63.04%.

Interpolation coefficient at 0.5: The simulation results showed that the prediction ability of Elman NN was poor, and this was mainly due to the limited amount of training data. Elman NN could not configure a suitable network structure to reflect the daily gas production behaviour from the limited sample data used to train the network. To solve this problem, two interpolation algorithms of piecewise linear interpolation and cubic spline interpolation were used to increase the number of the sample data in this experiment. An interpolation coefficient of 0.5 was first used. This meant that data would be inserted between two adjacent measured data by interpolation algorithms. In this way, a total of 107 data were obtained, and these data were divided into two parts: the first 75 data (70% of the data) were used to train the neural network, while the remaining 32 data (30% of the data) were used to predict biogas production (the proportion of data used for training and prediction was the same with that used in the Elman NN model with no interpolation previously described). Elman NN was trained 1000 times and had an error of 10^{-4}. The predicted daily gas production and the relative errors are shown in Figs. 5 and 6.

When the cubic spline interpolation was used to generate data for the Elman NN, a correlation coefficient of 0.9790, a mean square error of 0.005, and an average relative error of 13.31% was observed (Fig. 5). Similar results were seen when the piecewise linear interpolation method was used, with correlation coefficient of 0.9483, a mean square error of 0.010, and an average relative error of 16.01% observed. Therefore, the use of interpolation algorithms improved the indicators significantly, with the correlation coefficient increasing from 0.7966 to 0.9790 and 0.9483, the mean square error decreasing from 0.119 to 0.005 and 0.010, and the average relative error decreasing from 63.04% to 13.31% and 16.01%.

Interpolation coefficient at 0.2: The relative errors, when the two types of interpolations were used, were still large.

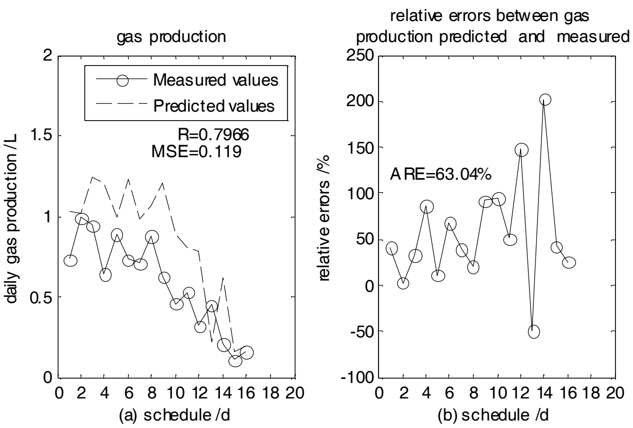

Fig. 4: Predicted biogas values and errors when no interpolation was used.

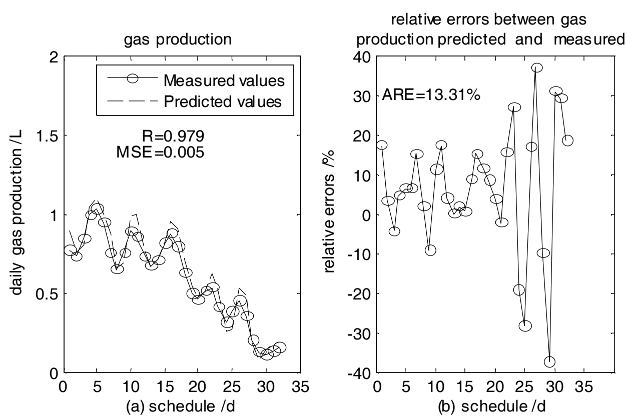

Fig. 5: Predicted biogas values and errors when cubic spline interpolation with an IC of 0.5 was used

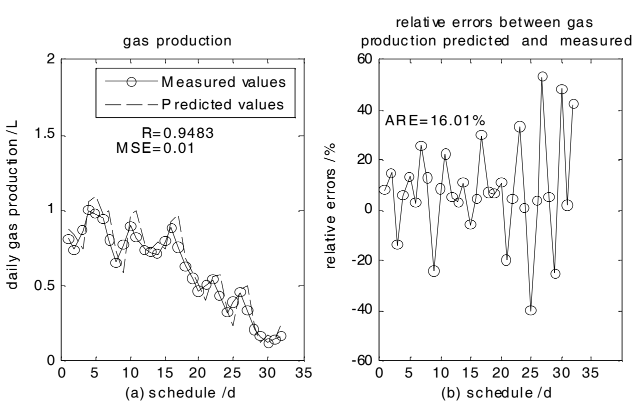

Fig. 6: Predicted biogas values and errors when piecewise linear interpolation with an IC of 0.5 was used.

However, considering that interpolations could improve the indicators significantly, it is expected that the same methods could reduce the average relative error to less than 5%. An error value of 5% is generally accepted in the engineering field. Therefore, the two interpolation algorithms of piecewise linear interpolation and cubic spline interpolation were used to increase the numbers of the sample data by decreasing the interpolation coefficient to 0.2. This meant that the interpolation algorithms inserted four data between two adjacent measured data. In this way, a total of 266 data were obtained. The data were divided into two parts: the first 186 data (70% of data) were used to train the Elman NN, while the remaining 80 data (30% of data) were used for predicting the biogas production values. The training was conducted 1000 times and the error was observed to be 10-4. The predicted daily gas production values and the relative errors are shown in Figs. 7 and 8.

The correlation coefficient of the Elman NN prediction results and measured results for the cubic spline and piecewise linear interpolations were 0.9962 and 0.9942, respectively. Both of the mean square errors were 0.001, while the average relative error for the cubic spline and piecewise linear interpolations were 3.93% and 4.01%, respectively.

After the interpolation coefficient was decreased from 0.5 to 0.2, the correlation coefficient, when cubic spline algorithm was used, increased from 0.9790 to 0.9962, while the mean square error was reduced from 0.005 to 0.001. The average relative error was reduced from 13.31% to 3.93%. Similar results were observed when the piecewise linear algorithm was used, with correlation coefficient increasing from 0.9483 to 0.9942, the mean square error decreasing from 0.010 to 0.001, and the average relative error decreasing from 16.01% to 4.01%. Minor improvements were seen for the first two indicators (correlation coefficient and average relative error), but the average relative error improved significantly to values less than the desired 5%.

The cubic spline algorithm was observed to be better than the piecewise linear algorithm in this experiment, but both algorithms can meet the requirement of 5% average relative error.

DISCUSSION AND CONCLUSIONS

Anaerobic digestion is a nonlinear dynamic system, which makes the development of a general prediction model for the system difficult. Anaerobic digestion generally has long fermentation periods, and the collection of parametric data during the process through automatic means is difficult. Therefore, there are limited amount of data that can be collected during an anaerobic digestion process. The neural

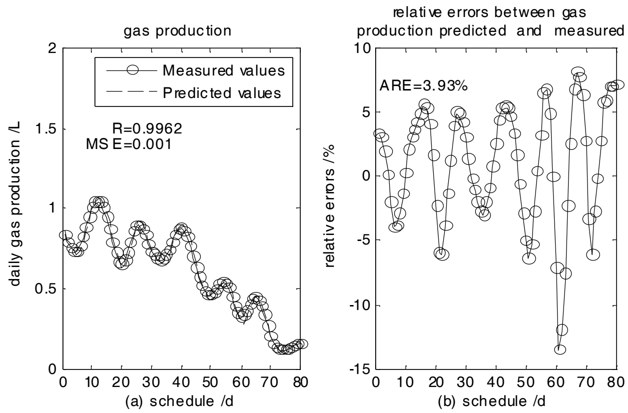

Fig. 7: Predicted biogas values and errors when cubic spline interpolation with an IC of 0.2 was used.

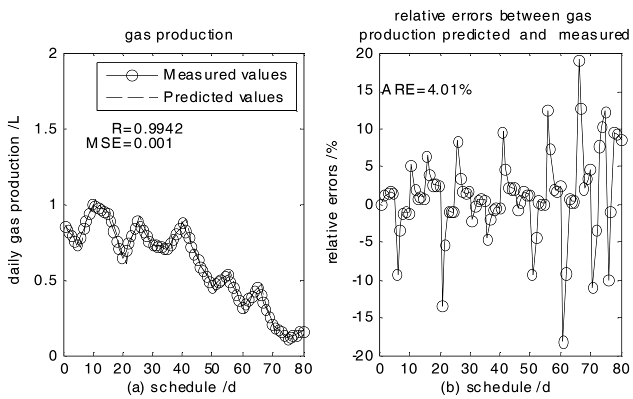

Fig. 8: Predicted biogas values and errors when piecewise linear interpolation with an IC of 0.2 was used.

network model was shown to be able to predict biogas production well but only when a large amount of data were supplied for training. Due to this contradiction, it is difficult to apply the neural network model for the anaerobic digestion process. Fortunately, interpolation algorithm can reconcile the contradiction to some extent. As it can be seen from this experiment, piecewise linear interpolation did not change the nonlinear characteristics of the data, and cubic spline interpolation was more consistent with the nonlinear characteristics of the data, so the results from the cubic spline algorithm were better than those observed when piecewise linear algorithm was used. When the interpolation coefficient was decreased from 0.5 to 0.2, the simulation results showed an improvement in the correlation coefficient, mean square error, and the relative error, which decreased to less than 5%. The conclusions of this paper are:

1. An increasing amount of data could improve the Elman NN prediction method. Since the Elman NN prediction model is a simple method with good versatility and high precision accuracy, it is suitable for modelling anaerobic digestion systems.
2. Compared to the model without interpolation, the use of an interpolation coefficient of 0.5 increased the experimental data from 54 to 107 and the correlation coefficient of predicted data and measured data from 0.7966 to 0.9790 (for cubic spline interpolation) and 0.9483 (for piecewise linear interpolation). In addition, the mean square error decreased from 0.1190 to 0.005 (for cubic spline interpolation) and 0.010 (for piecewise linear interpolation), while the average relative error decreased from 63.04% to 16.01% (for cubic spline interpolation) and 13.31% (for piecewise linear interpolation).
3. Compared to the model without interpolation, the use of an interpolation coefficient of 0.2 increased the experimental data from 54 to 266 and the correlation coefficient of predicted data and measured data from 0.7966 to 0.9962 (for cubic spline interpolation) and 0.9942 (for piecewise linear interpolation). In addition, the mean square error decreased from 0.1190 to 0.001 (for cubic spline interpolation) and 0.001 (for piecewise linear interpolation), while the average relative error decreased from 63.04% to 3.93% (for cubic spline interpolation) and 4.01% (for piecewise linear interpolation).

ACKNOWLEDGMENTS

This work was financially supported by the Rural Energy Project of the Ministry of Agriculture of China (2014-30), the Science and Technology Project of the China Renewable Energy Engineering Institute (KY-J2013-122), and National Natural Science Foundation of China (Nos. 51576167).

REFERENCES

Appels, L., Lauwers, J., Degrève, J., Helsen, L., Lievens, B., Willems, K., Impe Jan Van and Dewil, R. 2011. Anaerobic digestion in global bioenergy production: potential and research challenges. Renew Sust. Energ. Rev., 15(9): 4295-301.

Batstone, D. J., Keller, J., Angelidaki, I., Kalyuzhny, S. V., Pavlostathis, S. G., Rozzi, A., Sanders, W. T. M., Siegrist, H. and Vavilin, V. A. 2002. Anaerobic Digestion Model No.1. International Water Association, London.

Ben, Youssef C. and Zepeda, A. 2012. Iterative learning estimation of a parameterized input trajectory to control fedbatch fermentation process: a case study. Revista Mexicana de Ingenieria Quimica, 2: 351-362.

Bhowmik, A.K. and Cabral, P. 2011. Statistical evaluation of spatial interpolation methods for small-sampled region: a case study of temperature change phenomenon in Bangladesh. Lecture Notes in Computer Science, 6782: 44-59.

Bollon, J., Le-Hyaric, R., Benbelkacem, H. and Buffière, P. 2011. Development of akinetic model for anaerobic dry digestion processes: focus on acetate degradation and moisture content. Biochemical Engineering Journal, 56(3): 212-218.

Jin, G., Guofu, L. and Guangqing L. 2011. A method of state diagnosis for rolling bearing using support vector machine and BP neural network. Lecture Notes in Electrical Engineering, 2: 127-134

Kafle, G.K. and Kim, S.H. 2013. Anaerobic treatment of apple waste with swine manure for biogas production: batch and continuous operation. Applied Energy, 103: 61-72.

Lauwers, J., Appels, L., Thompson, I. P., Degrève, J., Van Impe, J. F. and Dewil, R. 2013. Progress in Energy and Combustion Science, 39: 383-402.

Krishania, M., Vijay, V.K. and Chandra, R. 2013. Methane fermentation and kinetics of wheat straw pretreated substrates co-digested with cattle manure in batch assay. Energy, 57: 359-367.

Li, Y.B., Stephen, Y.P. and Zhu, J. 2011. Solid-state anaerobic digestion for methane production fromorganic waste. Renewable and Sustainable Energy Reviews, 15(1): 821-826.

Merkel, W. and Krauth K. 1999. Mass transfer of carbon dioxide in anaerobic reactors under dynamic substrate loading conditions. Water Research, 33(9): 2011-2020.

Shi, X. H., Liang, Y. C., Lee, H. P., Lin, W. Z., Xu, X. and Lim, S. P. 2004. Improved Elman networks and applications for controlling ultrasonic motors. Applied Artificial Intelligence, 18(7): 603-629.

Souza, T. S., Ferreira, L. C., Sapkaite, I., Pérez-Elvira, S. I., and Fdz-Polanco, F. 2013. Thermal pretreatment and hydraulic retention time in continuous digesters fed with sewage sludge: assessment using the ADM1. Bioresource Technology, 148: 317-324.

Strik, D. P., Domnanovich, A. M., Zani, L., Braun, R. and Holubar, P. 2005. Prediction of trace compounds in biogas from anaerobic digestion using the MATLAB neural network toolbox. Environmental Modelling & Software, 20: 803-810.

Strik, D. P. B. T. B., Domnanovich, A. M., Zani, L., Braun, R. and Holubar, P. 2004. A benchmark study on modelling of sulfate reduction using artificial neural networks and anaerobic digestion model No.1. In: Lens, P., Stuetz, R. (Eds.), Water Environment Management Series, Young Researchers, 2004 IWA Publishing, London, 333-340.

Yan, Z. 2007. Study of the partitional variety-box model of anaerobic digestion reactor based on system dynamic. Beijing Municipal Research Institute of Environmental Protection, 13-14.

YuanFang, D., Ling, Q., QuanPing, S. and LiangXiao, J. 2012. Influence of anaerobic co-digestion of mushroom cultivation wastes and pig manure on the biogas production and its quality. Journal of Agro-Environment Science, 31(3): 613-619.