Watermatex 2015
9th IWA Symposium on Systems Analysis and Integrated Assessment
14-17 June 2015, Gold Coast, Australia

Predicting biogas production rate with Black/Grey-Box-Models based on artificially produced datasets

Hien, S.* and Hansen, J.*
*Urban Water Management and Hydraulic Engineering, Université du Luxembourg, 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg

Keywords: Biogas Prediction; Artificial Datasets; Black-Box-Modelling

Abstract

For the integration of wastewater treatment plants (WWTP) and biogas plants (BP) in the concept of virtual power plants a reliable biogas production forecast is necessary. This research will show, that it is possible to develop suitable and precise Black/Grey–Box–Models to predict biogas production rate on WWTPs/BPs from artificially produced datasets.

In this framework software based on the reference model Anaerobic Digestion Model No.1 (ADM1), was used to produce a large calibration dataset for Black/Grey-Box-Modelling approaches with artificial neural networks (ANN), Fuzzy-Logic, case based reasoning (CBR) as well as adaptive neuro fuzzy inference system (ANFIS). In a final step, the model should be evaluated considering their prediction accuracy and variables used for the prediction.

Results obtained up to now are very promising and imply a successful modelling approach with further experiments.

Introduction

There are profound changes in the energy market at the moment. The importance of renewable energy sources will increase in the near future because of exhaustible energy sources, climate change as well as the dangers of nuclear energy.

A lot of renewable energy sources are decentralized and show weather related fluctuation in energy production. This problem leads to the need of additional balancing power. One suitable solution for this problem is the virtual power plant (VPP), combining decentralized renewable energy producers by the use of a central control station with the aim to provide energy and system services to the market in the same or better quality than a conventional large power plant.

One class of renewable energy sources with a very high potential for being integrated in VPPs are biogas plants (BPs) and wastewater treatment plants (WWTPs), where biogas is generated by anaerobic digestion of biodegradable materials. They have the advantage of a highly flexible power generation since the biogas can either be converted to electrical power or stored.

However, the integration of BPs and WWTPs in VPPs also requires detailed knowledge about the available amount of electrical power in a short and medium term perspective which mainly depends on the current and future production of biogas. Therefore, a reliable prediction of the biogas generation in BPs and WWTPs would be essential for an efficient integration in VPPs.

There are already a large number of biogas process models available (Overview e.g. in [3]). Most of these models are so called White–Box–Models. These models are usually very complex and require an enormous amount of input variables, making them hardly applicable on real plants. But when it comes to multidimensional input and output cases, solutions based on soft computing methods like fuzzy logic or neural networks become handy. These are so called Black–Box–Models and Grey–Box–Models, respectively, and are very attractive in cases where no prior knowledge about the structure and relationship between used parameters exits [6].

In literature, evidence can be found (Overview see [7]), that Black/Grey–Box–Models can be used to predict the biogas production. Usually these models are only rated according to their accuracy and are not related to other factors. In this framework an indexing of Black/Grey–Box–Models with different variables with reference to their influencing factors seems to be necessary to provide a viable guideline for operators and use in step with actual practice is necessary.

Methods and Materials

To avoid the naturally occurring extrapolation problems with usually small experimentally produced datasets in Black-Box-Modelling the ifak SIMBA software will be used to create a large calibration dataset. SIMBA is based on the ADM1, nowadays representing the most comprehensive description of the anaerobic digestion process and the standard model in this field. In SIMBA a WWTP and a BP, respectively, is built up and used to produce datasets containing relevant biogas variables artificially.

With these datasets several Black/Grey–Box–Models with different input variables are developed using Black/Grey-Box-Modelling approaches. Coding and simulations are done in the MatLab Toolboxes including the following techniques:

- Artificial neural networks (ANN): is by definition "a system of simple processing elements, called neurons, which are connected to a network by a set of weights" [6]. A model neuron can be seen as a threshold unit, receiving input from other units and/or external sources.
- Fuzzy–Logic: is working similar to human being's feeling and inference process. Implementing fuzzy logic technique to a real application requires the following three steps: Fuzzification, Fuzzy Inference Process and Defuzzification [1].
- Case based reasoning (CBR): is powerful in learning dynamic system behaviour over a period of time [5] and comprises four process steps, Retrieve, Reuse, Revise and Retain.
- Adaptive neuro fuzzy inference system (ANFIS): is a neuro fuzzy technique combining the fuzzy inference system and the neural network. The Fuzzy–Logic can deal with linguistic expressions whereas the neural network gives the ability to self–learn and self–improve [4].

In the final step an index of the different suitable variables will be gathered to evaluate the different models based on the variables used for prediction. An advanced attempt could be calculated on the base of different factors, e.g.:

1. Duration until the variable measurement is available
2. Costs of the variable measurement
3. Complexity of the variable measurement

Results and Discussion

The upcoming results show selected Black-Box-Modelling approaches with a rather small training dataset (adapted from [2], monofermentation of grass silage under mesophilic conditions in a lab–scale 10 L reactor, 102 data–points).

Figure 1 shows the result for a feed–forward neural network with two layers and one neuron. The Figure illustrates the measurements (black dotted line) and the ANN-Model prediction (red line) of the biogas production rate in ml per day on a normalized scale against time in days. The used independent variables were temperature, pressure, organic loading rate, H2S, acetic acid, propionic acid, butyric acid, valeric acid and total inorganic carbonate. Here the data interval up to the first vertical blue line indicates the training set, the second vertical blue indicates line the validation set and the last data set represents the test data set for the ANN.

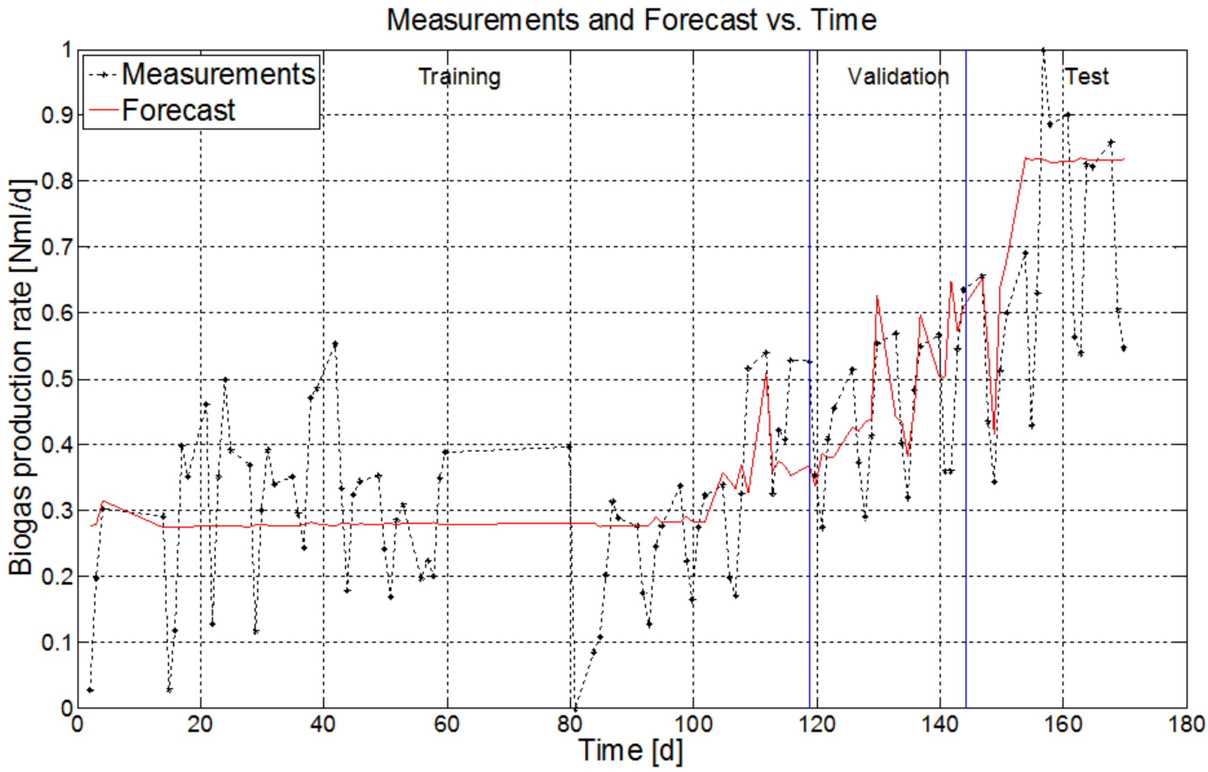

Figure 1: Example of an ANN MatLab simulation with 2 layers and 1 neuron. (Abscissae = Time [d], Ordinate = Biogas production rate [NL/d]); Red line = Prediction; Black line = Measurements; First blue line indicates the training set, Second blue line the validation set, last set is the test set)

As can be seen, there is hardly any adaption to the biogas production rate dynamic. This was an expected outcome, as the prediction power of a one-neuron ANN is limited.

In contrast, Figure 2 shows an ANN with 5 neurons based on the same training dataset and variables as Figure 1. In this case the ANN reproduces the measurement data well. As can be seen from comparing Figure 1 and 2, the prediction accuracy increases with increasing neurons. But at one point the generalization ability of the ANN will be lost resulting in an over–fitting. With a larger dataset including the maximum and minimum values of each variable, this problem should be fixed.

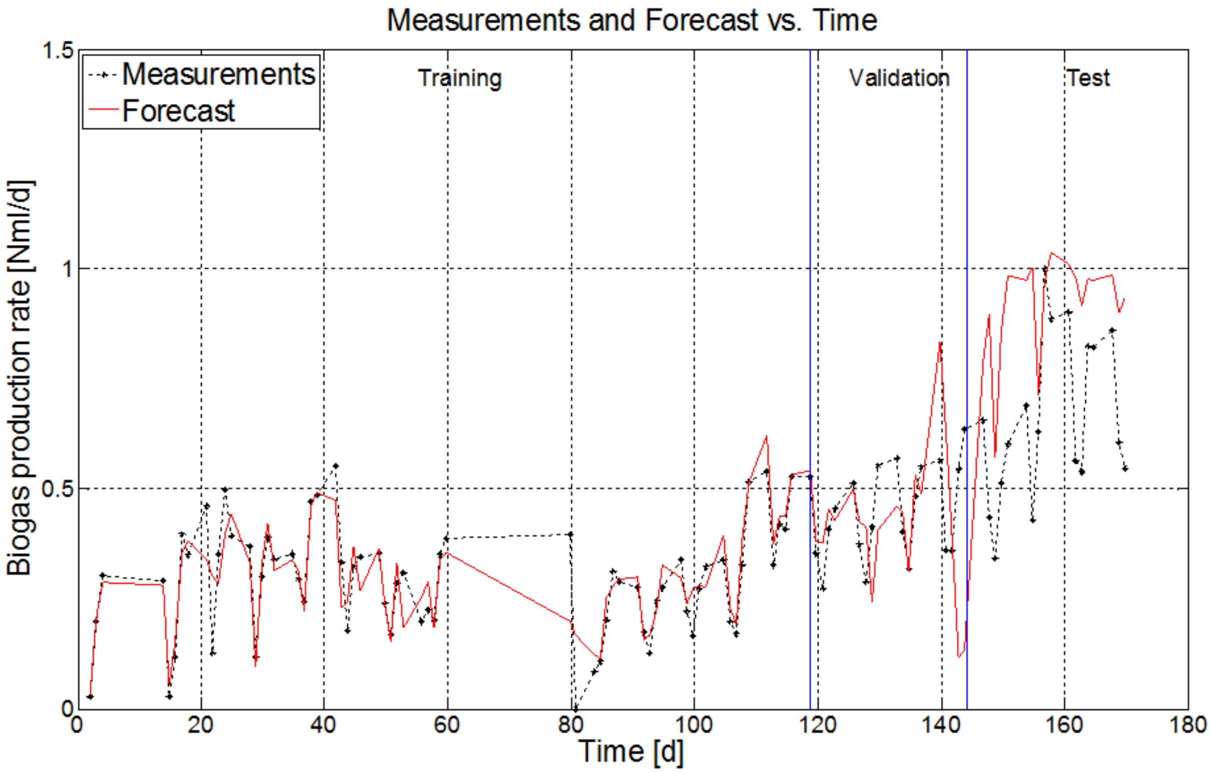

Figure 2: Example of an ANN MatLab simulation with 2 layers and 5 neurons. (Abscissae = Time [d], Ordinate = Biogas production rate [NL/d]); Red line = Prediction; Black line = Measurements; First blue line indicates the training set, Second blue line the validation set, last set is the test set)

Conclusion

Overall, the results seem to be very promising and it could be shown, that the biogas production rate can be predicted with different variables using Black/Grey-Box-Models. Further experiments will show, that it is possible to substitute experimentally produced datasets with artificially generated ones and still create accurate Black/Grey-Box-Models to predict the biogas production rate.

References

1 Bai, Y., et al. 2010. Advanced Fuzzy Logic Technologies in Industrial Applications. Springer London.

2 Benito, P., et al. 2014. Digestion of grass silage in a semi-continuously fed reactor at increasing loading rates. Venice 2014 Symposium. Venice.

3 Gerber, M. et al. 2008. An Analysis of Available Mathematical Models for Anaerobic Digestion of Organic Substances for Production of Biogas. IGRC, 2008. Paris.

4 Hamdan, H. 2013. An Exploration of the Adaptive Neuro–Fuzzy Inference System (ANFIS) in Modelling Survival. Dissertation. University of Nottingham, School of Computer Science.

5 Im, K. H et al. 2007. Case–based reasoning and neural network based expert system for personalization. Expert Syst Appl 32(1), 77–85.

6 Strik, D. P. B. T. B., et al. 2005. Prediction of trace compounds in biogas from anaerobic digestion using the MATLAB Neural Network Toolbox. Environ. Model. Softw. 20(6), 803–810.

7 Thorin, E., et al. 2012. Modeling of the Biogas Production Process – A Review. ICAE. Suzhou.