Available at www.sciencedirect.com

INFORMATION PROCESSING IN AGRICULTURE 6 (2019) 349–356

journal homepage: www.elsevier.com/locate/inpa

Prediction of the biogas production using GA and ACO input features selection method for ANN model

Tanja Beltramo [a,*], *Michael Klocke* [b], *Bernd Hitzmann* [a]

[a] *Institute of Food Science and Biotechnology, University of Hohenheim, Stuttgart 70599, Germany*
[b] *Leibniz Institute for Agricultural Engineering and Bioeconomy, Potsdam 14469, Germany*

ARTICLE INFO

Article history:
Received 6 June 2018
Received in revised form
19 December 2018
Accepted 7 January 2019
Available online 11 January 2019

Keywords:
Ant colony optimization
Artificial neural networks
Biogas
Genetic algorithm

ABSTRACT

This paper presents a fast and reliable approach to analyze the biogas production process with respect to the biogas production rate. The experimental data used for the developed models included 15 process variables measured at an agricultural biogas plant in Germany. In this context, the concentration of volatile fatty acids, total solids, volatile solids acid detergent fibre, acid detergent lignin, neutral detergent fibre, ammonium nitrogen, hydraulic retention time, and organic loading rate were used. Artificial neural networks (ANN) were established to predict the biogas production rate. An ant colony optimization and genetic algorithms were implemented to perform the variable selection. They identified the significant process variables, reduced the model dimension and improved the prediction capacity of the ANN models. The best prediction of the biogas production rate was obtained with an error of prediction of 6.24% and a coefficient of determination of $R^2 = 0.9$.

1. Introduction

Biogas belongs to the renewable energy sources, which underwent a worldwide expanding technological development over the last decades. The basis of the biogas production is anaerobic digestion (AD) and subsequent biomethanation of biomass and organic wastes, e.g. from husbandry or industrial production processes. AD is a complex process involving complex microbial consortia with numerous metabolic processes and kinetic reactions [1,2]. While the biogas production is a complex and long lasting biological process, the use of the conventional analytical methods, e.g. according to German VDI norm 4630 [3], is time- and equipment-consuming and, hence, expensive. Numerous process variables must be taken into consideration and be controlled to evaluate the process.

Mathematical modelling represents a quick and cheap alternative to the conventional analytics, among others especially chemometric methods and metaheuristics. These powerful tools are helpful to identify the process structure and to analyze the correlations between the process components without any prior knowledge about the process correlations [4]. One of the most popular mathematic methods applied for the optimization of biological systems are the artificial neural networks (ANN) [5,6]. In agricultural sciences it was used for visual identification of orange varieties [7], to detect plant diseases [8], to improve milk service platform [9] and to estimate the biophysical variables [10]. This approach was successfully implemented in field of AD systems in order to predict the process intermediates, to optimize the bioreactor performance and to improve the process conditions. As an example, Strik et al. used ANN to predict the trace compounds in biogas from anaerobic digestion [11]. Here the ANN models were successfully used to predict hydrogen

* Corresponding author.
E-mail address: tanja.beltramo@gmail.com (T. Beltramo).
Peer review under responsibility of China Agricultural University.
https://doi.org/10.1016/j.inpa.2019.01.002

sulfide and ammonia concentrations in biogas under dynamical conditions. The developed approach can be used to control and to reduce the toxic trace compounds in fuel cells. In terms of the prediction of the bioreactor performance, Sahinkaya et al. implemented the neural networks to analyze the thermophilic sulfidogenic fluidized-bed reactor [12], which help to control the operational conditions improving the process cycles. Also based on ANN, Ozkaya et al. predicted the methane fraction in biogas [13]. Here the anaerobic digestion process was evaluated using the leachate characteristics measured in two different time periods used to predict the biogas output. Kanat et al. evaluated the biogas production rate [14] using the ANN models based on the data measured by steady-state and abnormal process conditions to control the thermophilic bioreactors. Bernari et al. used an automated medium scale prototype for anaerobic co-digestion of olive mill wastewater [15]. The described implementations of ANN used for the modelling of biogas digestion process based on different process variables used for control and evaluation either of the separate process parameters or of the entire process, proved to be a reliable technique to optimize and to control AD processes. In comparison to other existing models, such as the anaerobic digestion model No.1 (ADM1) [16], neural networks have a simple structure. The ADM1 model includes 19 process rate equations, six acid-base reactions, three gas transfer reactions, a number inhibition balances and over 30 water phase equations for soluble and particulate matter. Moreover, ADM1 requires a number of kinetic parameters and rates need to be estimated for each process. Thus the complexity of the microbiological degradation processes limits the application to a certain course of the process. In comparison to ADM1, ANN models are data-driven and do not require any prior knowledge about the kinetics and the microbiological digestion processes [17]. ANN models are represented by the multilayered perceptron, that enables to evaluate the correlation between the independent process input variables and the dependent output variables without any prior knowledge about their interrelations.

Metaheuristic techniques, such as genetic algorithms (GA) [18] and ant colony optimization (ACO) [19] have been applied to solve complex problems within biological systems. For example, ACO was implemented in the medical sciences for solving problems in protein folding. Here ACO was used to predict the protein conformation based on the amino acids sequences [20]. The used ACO represents a successful tool to solve the problems in bioinformatics, where other state-of-the-art methods failed. In field of bioinformatics, GA was used to detect recombination problems [21] and to analyze an enzyme kinetic process [22]. In food science ACO and GA were implemented for the flour characterization based on the NIR spectral filter data [23] and to predict pH and lactate in porcine meat [24]. In agricultural sciences Silalahi et al. used GA based on infrared spectral data for the identification of ripeness of oil palm [25], while Mehdizadeh et al. for simulation of greenhouse processes [26]. The interdisciplinary use of metaheuristics proved it to be reliable optimization tools and opens the perspective to be also applied for the optimization of AD systems.

The main objective of this study was to develop a reliable tool able to predict the bioreactor performance with respect to the biogas production rate. It should improve the biogas production process and could select the significant process features in order to simplify the analytical evaluation procedure. The developed methodology represents a fast and adaptable method to be used for different AD processes under real conditions based on the real anaerobic digestion process data. In detail, it should be able to identify the correlations within the process and predict accurately the process evolution. For this approach, ANN was used to predict the biogas production rate. As variable selection tools, GA and ACO were applied, which were compared according to their feasibility and usefulness in field of AD optimization. In that sense, the optimization step is unique due to the algorithm combination method and the application of ACO for the real data of an anaerobic digestion process.

2. Materials and methods

2.1. *Origin of experimental data*

The experimental data used in this study are off-line measurements collected at a biogas plant in Germany over a time-period of ten months with a frequency of one sampling per week. The used data originated from the joint project BIOGAS-ENZYME supported by the German Federal Ministry of Food and Agriculture (BMEL), grant no. 22027707 [27]. The measurements were done in July 2011 and from February till October 2012. In this biogas plant four feedstocks, i.e. maize and grass silages together with pig and cattle manure, were co-digested at mesophilic temperature (42 °C). From digestate samples, the concentrations of acetic acid, propionic acid, n- and iso-butyric acid, n- and iso-valeric acid and ammonium (NH_4^+-N) were determined as well as the contents of total solids (TS), volatile solids (VS), neutral detergent fibre (NDF), acid detergent fibre (ADF), the organic loading rate, alkalinity ratio and acid detergent lignin (ADL). Hydraulic retention time (HRT) was additionally calculated. The measured variables are represented in Table 1. TS and VS serve generally as

Table 1 – The measured process variables (measured with a frequency one measurement per week).

Variable name (abbreviation)	Unit
Acetic acid (AcA)	$g\,l^{-1}$
Acid detergent fibre (ADF)	$g\,kg^{-1}\,VS^{-1}$
Acid detergent lignin (ADL)	$g\,kg^{-1}\,VS^{-1}$
Alkalinity ration (FOS/TAC)	–
Ammonium (NH_4^+-N)	$g\,l^{-1}$
n-Butyric acid (nBA)	$g\,l^{-1}$
Iso-butyric acid (iBA)	$g\,l^{-1}$
Total solids (TS)	% Fresh matter (FM)
Hydraulic retention time (HRT)	day
Neutral detergent fibre (NDF)	$g\,kg^{-1}\,VS^{-1}$
Volatile solids (VS)	% TS
Organic loading rate (OLR)	$kg\,VS\,m^{-3}\,d^{-1}$
Propionic acid (PA)	$g\,l^{-1}$
n-Valeric acid (nVA)	$g\,l^{-1}$
Iso-valeric acid (iVA)	$g\,l^{-1}$

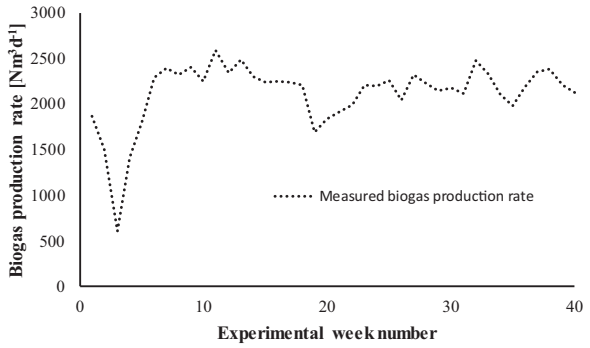

Fig. 1 – Evolution of the biogas production rate during sampling period.

reference values for other substrate compounds, such as proteins, nitrogen and trace elements. ADF, ADL and NDF describe the content of nutrients. HRT is the average period that a given quantity of biomass material remains in the digester. Organic loading rate indicates the quantity of dry solids loaded per m^3 of digester volume and unit of time. Alkalinity ratio indicates the quantity of volatile organic acids in relation to the buffer capacity of carbonate. Biogas production rate measured during the sampling period is presented in Fig. 1. The succession of organic acids, TS, VS, NDF, ADL, ADF, alkalinity ratio, NH_4^+-N, HRT, and OLR is shown in Supplementary, Figs. S1–S4.

2.2. Computational platform

For the development of the neural network models the Matlab® Neural Network Toolbox Version 7.10 2010a (The MathWorks Inc., Natick, USA) provided with a user's guide [28] was used. A Matlab script generated by the Neural Network Toolbox loaded the data, trained, validated and tested the models. The data were normalized before application in the neural network using a script written in Matlab. The implemented optimization tools, here the ant colony optimization and the genetic algorithm were written in Matlab. All Matlab scripts are available free of costs from the corresponding author. The treatment of the experimental data and the model calculations were carried out on a processor AMD Phenom ™ II X2 B57 with 3.2 GHz.

2.3. Pre-processing technique

Data pre-processing is an essential step of any data analysis [29]. In this work a normalization method, the standard normal variate (SNV) was used. Eqs. (1) and (2) represent the mathematical computation of SNV-transformation.

$$\sigma = \sqrt{\sum_{k=1}^{n} \left(x_k - \bar{x}\right)^2 / (n-1)} \quad (1)$$

$$x_{k,norm} = \left(x_k - \bar{x}\right)/\sigma \quad (2)$$

Here x_k is a measured value, while $\bar{x}$ is the mean value of x_k. The computed σ represents the standard deviation of x_k.

2.4. Artificial neural networks

For the prediction of the biogas production rate, a two layered feedforward neural network from the Matlab® ANN Toolbox was applied. The ANN network was trained using the Levenberg-Marquardt algorithm, embedded in Matlab ANN Toolbox. ANN included an input layer, a hidden layer with a sigmoid activation function and an output layer with a linear output function. A random initialization was used to provide more accurate prediction results. The sigmoid transfer function at the hidden layer enables to treat the data with non-linear features. It takes the input (any values between plus and minus infinity) and squashes the output into the range 0–1 [30]. The sigmoid activation function is presented in Eq. (3).

$$f(x) = 1/(1 + \exp^{-x}) \quad (3)$$

The data of the measured variables were used as input neurons, while the biogas production rate was calculated by the output neuron. To identify the optimal model structure the number of hidden neurons was varied from 3 to 20. The data set was split into three data sets for training (70%, 28 samples), validation (15%, 6 samples) and test (15%, 6 samples). The network is trained to find the optimal weights, which minimize the cost functional. Here the RMSE training was used as a cost functional and was minimized. The validation data served to overcome overfitting and is used to stop training, which was determined by the failure of the error decrease for six consecutive iterations. Using the test data the prediction of the biogas production rate was done. The assessment of the calculated models was done using the root mean square error (RMSE) and the coefficient of determination (R^2), which are presented in Eqs. (4) and (5). RMSE was used to evaluate the accuracy of the models, while R^2 to evaluate their robustness. Here $\widehat{y_k}$ refers to the predicted value, whereas y_k represents the measured value and $\bar{y}$ is the mean value of y_k. n is the number of samples.

$$RMSE = \sqrt{\left(\sum_{k=1}^{n} \left(\widehat{y_k} - y_k\right)^2\right)/n} \quad (4)$$

$$R^2 = 1 - \left(\left(\sum_{k=1}^{n} \left(\widehat{y_k} - y_k\right)^2\right) / \left(\sum_{k=1}^{n} \left(y_k - \bar{y}\right)^2\right)\right) \quad (5)$$

RMSE and R^2 were computed for training, validation and test models.

2.5. Metaheuristic methods

For the variable selection purpose two metaheuristic methods were implemented, the genetic algorithm and an ant colony optimization. GA represents a mathematical interpretation of the Darwin's evolution theory, which is based on the natural rule of survival and the Mendel's principles of inheritance [31]. Due to the selection of the best individuals and further mutation and cross-over transformation of their genome only the best qualities will be transferred to the next generation [32]. In this work, the implemented GA was used to identify the significant process variables [33]. The principle of the

GA variable selection is based on the lowest cross-validated RMSE of the calculated models. Thus, the significant process variables are those used in the partially least square regression models (PLSR) with the lowest cross-validated RMSE. For the applied GA approach the following algorithm parameters were set: PLS as regression model type, cross over points of 2, mutation rate of 0.005, 15 individuals with a population size of 256 and 100 runs. PLSR is a multivariate statistical tool, which identifies the latent variables (factors) regarding the most of variance in data and linear correlation between the process variables. The cross-over operator is related to the natural reproduction cross-over and will be set randomly. At each step GA selects randomly 15 individuals out of a set population during all runs.

The principle of ACO was inspired by natural behavior of ants seeking for food [23]. They leave on the way to food and back to the nest a pheromone trail, which serves as a natural guide for other ants. Accordingly the "shortest ways" have the highest pheromone concentration and will be frequently used by ants. Hence, the pheromone concentration can serve as an independent criterion for the variable selection [19]. In the approach used here, the measured process variables represent the possible ways to the "food source", namely the biogas production rate. The virtual ants use different combinations of measured process variables to find out the best possible solution. The pheromone concentration will be calculated for each process variable separately and depends on the quality of correlation with the predicted variable. Generally, pheromone concentration is in range between 0 and 1. The measured process variables, which correlated best with the biogas production rate and are mostly used by the virtual ants, have a higher pheromone concentration in comparison to those correlating less with the predicted variable. The measured variables which had the highest pheromone concentration are the ones most important for the prediction of the biogas production rate. For ACO the following algorithm parameters were applied: 100 ants, PLS as regression method with a principal component number von 1–15, initial pheromone concentration of 10^{-6} and pheromone evaporation rate of 0.5 per iteration, 50 iterations. Here the pheromone trail is an independent evaluation parameter, which will be adapted by ants to determine the solutions to the problem. It will be updated for each variable separately regarding the evaporation rate during the ACO implementation. The implemented ACO algorithm proceeds in three stages, the 0th stage is the initiation, where the model parameters will be set. In the 1st stage the objective function will be calculated and the best global output variables will be defined, following by a number of cycles. In the last stage the results will be presented (Table 2).

3. Results & discussion

3.1. *Ann prediction of the biogas production rate*

The input layer of the implemented ANN included 15 neurons, which were represented by the measured process variables; the number of hidden neurons was varied from 3 to 20 in order to get the optimal model structure. The biogas production rate resulted as an output neuron. In general, the results of 5–20 hidden neurons did not show significant differences. The best prediction results with a small-dimensioned model structure are shown in Table 3. The less successful prediction of the biogas production rate is not shown.

As can be seen in Table 3, the prediction of biogas production rate using 15 input neurons was successful. The best results were achieved using 5 and 10 hidden neurons, although the results obtained with more than 5 neurons are not significantly different and were not shown. The models with 10 hidden neurons were more robust and accurate in comparison to models with 5 hidden neurons. Here the training RMSE was 4.81% and the training R^2 was 0.90. The test RMSE was 9.66%, while the test R^2 reached 0.80. The evolution of the predicted biogas production rate done with ANN using 15 input neurons and 10 hidden neurons is presented in Fig. 2.

3.2. *Optimized ANN prediction of the biogas production rate*

The metaheuristic tools implemented in the model were used to identify the significant process variables and to improve model performance. The calculated pheromone concentration served as an independent evaluation criterion to select the significant process variables. It depends on the correlation with the predicted biogas production rate. The calculated value equivalent to the pheromone concentration in the ant

Table 2 – Stages of the ACO algorithm.

Stage	Routines
1. Initiation	Initial pheromone concentration Model type Pheromone evaporation rate Number of ants Number of iterations Number of principal components
2. Calculation and iterations	Objective function will be calculated The best global output variables will be defined Iterations will be performed The best ants will be selected and compared with the best stored ones Update of the pheromone trail Selection of the best variable combinations
3. Results	Display of the results

Table 3 – ANN prediction of the biogas production rate using 15 input neurons.

Number of input neurons	15	15
Number of hidden neurons	5	10
Number of iterations	12	8
RMSE training [%]	5.89	4.81
RMSE validation [%]	13.66	5.15
RMSE test [%]	**13.08**	**9.66**
R^2 training [–]	0.90	0.88
R^2 validation [–]	0.46	0.70
R^2 test [–]	**0.76**	**0.80**

The results of the prediction were highlighted in bold. Other results belong to validation and training models.

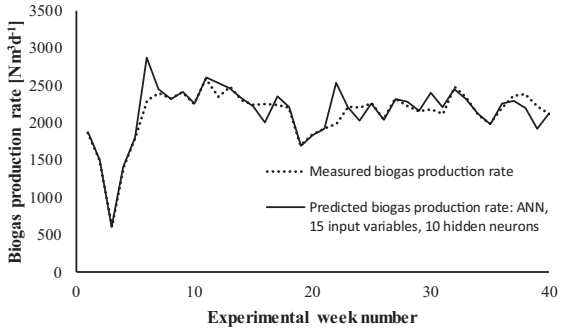

Fig. 2 – Prediction of the biogas production rate done with ANN using 15 input variables and 10 hidden neurons.

model of all variables is shown in Supplementary (Table S1). The measured process variables with a virtual pheromone concentration value higher than 0.5 were identified as the most significant ones. The ACO model identified seven significant process variables, namely TS, HRT, NDF, ADL, alkalinity ratio, n-butyric acid, and iso-valeric acid.

The GA algorithm selected nine process variables, namely HRT, TS, NDF, ADF, ADL, propionic acid, n-butyric acid, iso-butyric acid, and n-valeric acid. Five process variables, namely HRT, TS, NDF, ADL, and n-butyric acid were identified as process relevant by both optimization algorithms.

The selected variables were used for the further prediction of the biogas production rate. Therefore, nine significant process variables selected by GA, seven significant process variables selected by ACO and five significant process variables selected by both optimization algorithms were used. The prediction results are presented in Table 4.

Regarding in Table 4 presented results a good prediction performance could be achieved using a small-dimensioned model structure. The ACO- and GA-optimized ANN models showed generally similar results. For the GA-optimized ANN models more hidden neurons were required. A more accurate prediction was obtained using the ACO selected process variables and 10 hidden neurons. Here, the training RMSE was 3.53%, while the training R^2 was 0.98. The test RMSE was 10.37%, and the test R^2 was 0.83. The models with the significant process variables selected by both algorithms showed the best result. Here, the best model performance was achieved using 5 input neurons and 5 hidden neurons. The training models had RMSE of 3.49% and an R^2 of 0.98. By the test models RMSE was reduced to 6.24%, the R^2 reached 0.9. The regression performance of the GA-ACO-optimized ANN models is shown in Fig. 3.

The dynamic evolution of the predicted biogas production rate done with the optimized ANN models are shown in Figs. 4–6.

The achieved results proved the implemented approach to be a feasible methodology to analyze the biogas production process. The implemented ANN models predicted successfully the biogas production rate. In contrast to the approach used in this study, previously published studies used more complex ANN to achieve a good prediction performance in context of biomethanation processes. For example, Ozkaya et al. used eight input neurons and 15 hidden neurons to predict the methane fraction of biogas [13]. In another study published by Sahinkaya et al., six input neurons and 20 hidden neurons were required to predict sulphate, acetate and sulphide concentrations in a thermophilic sulfidogenic fluidized-bed reactor [12]. In our study we applied five input neurons and five hidden neurons to predict the biogas production rate. Thus, an effective simplification of the model dimension was done due to the implemented optimization algorithms. In addition, the used metaheuristics enabled an accurate variable selection by defining the significant measured process variables. The approach using an ACO algorithm is unique in the field of assessment of AD systems. Several publications used a genetic algorithm to optimize ANN models. Abu Qdais et al. used GA to optimize the methane output of the bioreactor and to define the best

Table 4 – Prediction results done using the optimized ANN models.

	GA optimized		ACO optimized		GA-ACO optimized
Number of input neurons	9	9	7	7	5
Number of hidden neurons	10	15	5	10	5
Number of iterations	14	12	16	13	18
RMSE training [%]	6.98	13.90	11.21	3.53	3.49
RMSE validation [%]	11.68	13.89	6.97	7.07	3.30
RMSE test [%]	**13.70**	**11.86**	**12.37**	**10.37**	**6.24**
R^2 training [–]	0.95	0.85	0.94	0.98	0.98
R^2 validation [–]	0.62	0.74	0.56	0.56	0.96
R^2 test [–]	**0.80**	**0.89**	**0.75**	**0.83**	**0.90**

The results of the prediction were highlighted in bold. Other results belong to validation and training models.

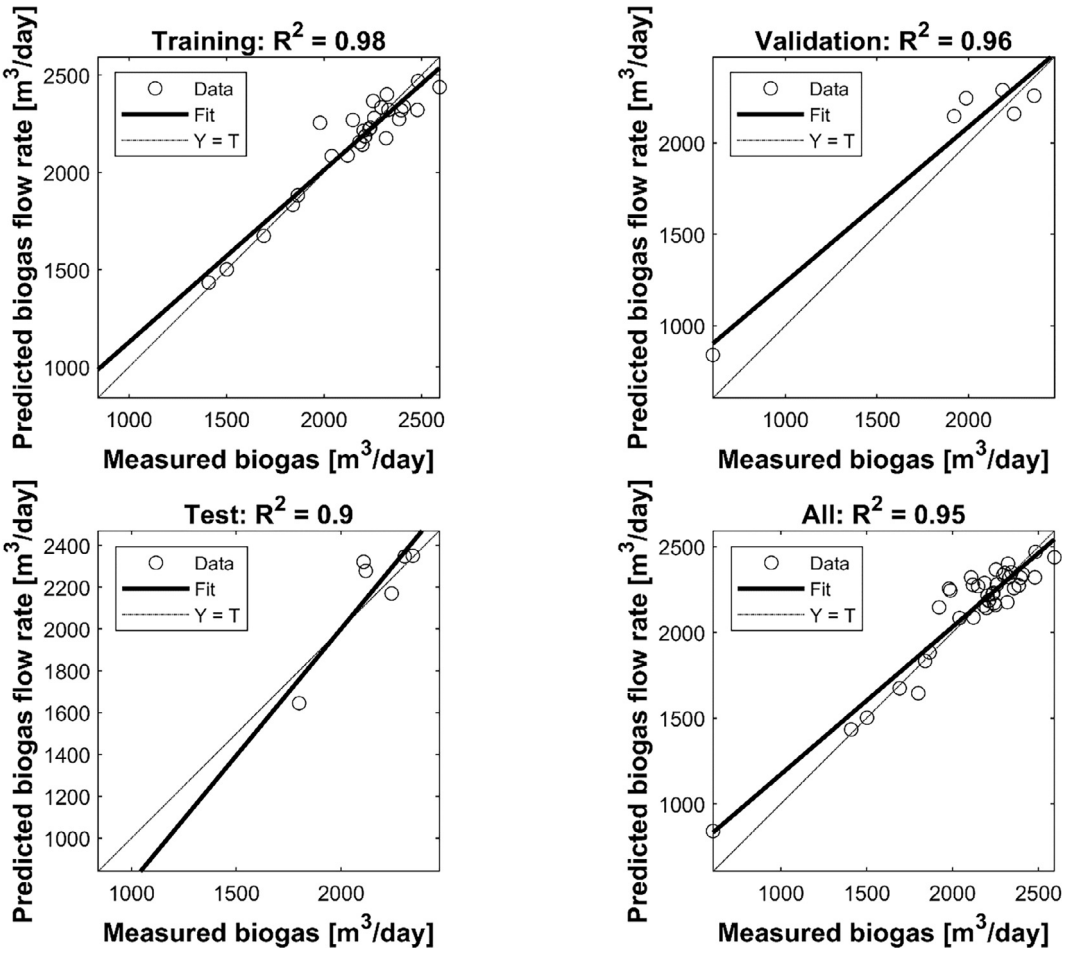

Fig. 3 – Regression performance of the GA-ACO-optimized ANN models: (1): training; (2): validation; (3): test; and (4): all data.

operational conditions while Gueguim Kana et al. (2012) applied GA to improve the biogas production on saw dust and other co-substrates [34,35]. The use of two optimization algorithms, namely ACO and GA, was aimed to assess their feasibility in field of variable selection by AD systems. Thus five significant measured process variables could be identified, which enabled an accurate prediction of the biogas production rate. Thus, the developed approach represents a fast and robust method to analyze the process evolution. In comparison to other commonly accepted models, such as ADM1, the methodology developed in this study requires only a small number of process variables to perform a successful evaluation of the process. In contrast to it, ADM1 needs a comparatively huge number of process variables and kinetic parameters as well as rates additionally determined for each kind of substrate, that makes its application complex.

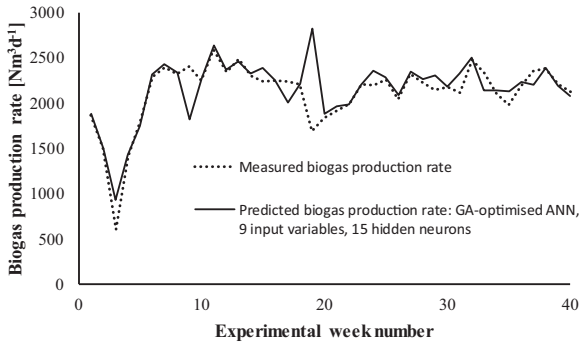

Fig. 4 – Prediction of the biogas production rate calculated using the GA-optimized ANN with nine input neurons and 15 hidden neurons.

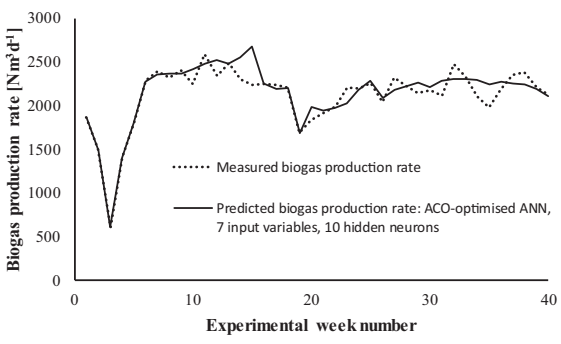

Fig. 5 – Prediction of the biogas production rate calculated using the ACO-optimized ANN with seven input variables and 10 hidden neurons.

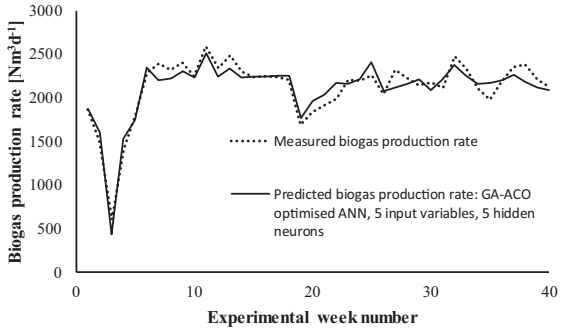

Fig. 6 – Prediction of the biogas production rate calculated using ACO-GA optimized ANN with five input variables and 5 hidden neurons.

4. Conclusion

The approach developed in this study represents a fast and reliable method to evaluate the biogas production process. Herewith the evaluation and prediction of typical process variables namely the biogas production rate could be performed. For the prediction of the biogas production rate, ANN models were implemented. The used ANN models had a simple structure [36] performing a robust and accurate prediction of the biogas production rate. The optimization tools were used to evaluate the process variables selecting the significant ones. The used variable selection tools made it possible to reduce the model dimension and to improve its performance. The best results were gained using ACO-GA optimized ANN models. Here the prediction error was reduced to 6.24% and R^2 increased to 0.90. The developed approach can be further used to develop an on-line control, which will help to improve the process conditions and to prevent possible process failures. The variable evaluation tools can support the operating engineer with information about the main process correlations. The developed approach demonstrates that ANN in combination with GA and ACO optimization tools showed reliable results in evaluation of the biogas production rate. Here the neural logic could predict the process development, while the optimization tools could improve the prediction capacity by selection of the significant process variables. Moreover, for the modelling a small number of data sets was required. This strategy can be used as a control operator to evaluate the process development based on the measured data. It can be used as an alternative approach to replace the computationally intensive and time-consuming ADM1 as well as to speed up the simulation procedure of biological processes [37]. Another way to use the developed approach is to evaluate the composition of the substrates. The intelligent model can rapidly estimate the best process conditions, accurately analyzing the process variables regarding the complex non-linear process behavior. It will help to improve the process development, to gain the highest biogas output, saving time, costs and to avoid the time-consuming and expensive analytics. The developed approach can be used for different AD processes regarding chemical and technical influence coefficients. Nevertheless, it is recommended in future studies to focus on the evaluation of the effects of the individual input variables. For that a more elaborated data is required.

Conflict of interest

None.

Acknowledgement

This work was part of the joint projects BIOGAS-ENZYME and BIOGAS-BIOCOENOSIS supported by the German Federal Ministry of Food and Agriculture (BMEL), grant nos. 22027707, 22010711 and 22028911 [27].

Our sincere gratitude to Robert Kausmann and Susanne Theuerl for the support by this research.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.inpa.2019.01.002.

REFERENCES

[1] Klang J, Theuerl S, Szewzyk U, Huth M, Tölle R, Klocke M. Dynamic variation of the microbial community structure during the long-time mono-fermentation of maize and sugar beet silage. Microb Biotechnol 2015;8:764–75.
[2] Theuerl S, Kohrs F, Benndorf D, Maus I, Wibberg D, Schlüter A, et al. Community shifts in a well-operating agricultural biogas plant: how process variations are handled by the microbiome. Appl Microbiol Biotechnol 2015;99:7791–803.
[3] Friedmann H, Görtz G, Helm M, Kaltschmitt M, Knopf U, et al. Fermentation of organic materials – characterisation of the substrate, sampling, collection of material data fermentation tests (Vdi 4630). In: Vdi-Handbuch Energietechnik, Verein Deutscher Ingenieure. Düsseldorf: Beuth; 2006. p. 1–48.
[4] Kessler W. Multivariate Datenanalyse: Für Die Pharma-, Bio- Und Prozessanalytik; Ein Lehrbuch. Weinheim: Wiley-Vch; 2007. in German.
[5] Hitzmann B, Ritzka A, Ulber R, Schöngarth K, Broxtermann O. Neural networks as a modeling tool for the evaluation and analysis of FIA signals. J Biotechnol 1998;65:15–22.
[6] Hitzmann B, Kullick T. Evaluation of ph field effect transistor measurement signals by neural networks. Anal Chim Acta 1994;294:243–9.
[7] Sabzi S, Abbaspour-Gilandeh Y, Garcia-Mateos GA. New approach for visual identification of orange varieties using neural networks and metaheuristic algorithms. Inform Process Agric 2018;5:162–72.
[8] Golhani K, Balasundram SK, Vadamalai G, Pradhan B. A review of neural networks in plant disease detection using hyperspectral data. Inform Process Agric 2018;5:354–71.
[9] Ma W, Fan J, Li Q, Tang YA. Raw milk service platform using BP neural network and fuzzy inference. Inform Process Agric 2018;5:308–19.
[10] Sehgal VK, Charkraborty D, Sahoo RN. Inversion of radiative transfer model for retrieval of wheat biophysical parameters from broadband reflectance measurements. Inform Process Agric 2018;3:107–18.
[11] 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. Environ Modell Softw 2005;20:803–10.
[12] Sahinkaya E, Ozkaya B, Kaksonen AH, Puhakka JA. Neural network prediction of thermophilic (65 °C) sulfidogenic fluidized-bed reactor performance for the treatment of metal-containing wastewater. Biotechnol Bioeng 2007;97 (4):780–7.
[13] Ozkaya B, Demir A, Bilgili M. Neural network prediction model for the methane fraction in biogas from field-scale landfill bioreactors. Environ Modell Softw 2007;22:815–22.
[14] Kanat G, Saral A. Estimation of biogas production rate in a thermophilic UASB reactor using artificial neural networks. Environ Model Assess 2009;14:607–14.
[15] Bernardi B, Zema DA, Tamburino V, Zimbalatti G. An automated medium scale prototype for anearobic co-digestion of olive mill wastewater. Inform Process Agric 2017;4:316–20.
[16] Batstone DJ, Keller J, Angelidaki I, Kalyuzhnyi SV, Pavlostathis SG, Rozzi A, et al. Anaerobic digestion model No. 1: [(Adm1)]. London: Iwa Publ; 2002.
[17] Beltramo T, Ranzan C, Hinrichs J, Hitzmann B. Artificial neural network prediction of the biogas flow rate optimised with an ant colony algorithm. Biosyst Eng 2016;143:68–78.
[18] Goldberg DE, Holland JH. Genetic algorithms and machine learning. Mach Learn 1988;3:95–9.
[19] Dorigo M, Blum C. Ant colony optimization theory: a survey. Theoret Comput Sci 2005;344:243–78.
[20] Shmygelska A, Hoos HH. An ant colony optimisation algorithm for the 2d and 3d hydrophobic polar protein folding problem. BMC Bioinf 2005;6:30.
[21] Kosakovsky Pond Sl, Posada D, Gravenor MB, Woelk CH, Frost SDW. GARD: a genetic algorithm for recombination detection. Bioinformatics 2006;22:3096–8.
[22] Lindner PFO, Hitzmann B. Experimental design for optimal parameter estimation of an enzyme kinetic process based on the analysis of the fisher information matrix. J Theor Biol 2006;238:111–23.
[23] Ranzan C, Strohm A, Ranzan L, Trierweiler LF, Hitzmann B, Trierweiler JO. Wheat flour characterization using NIR and spectral filter based on ant colony optimization. Chemometr Intell Lab 2014;132:133–40.
[24] Nache M, Scheier R, Schmidt H, Hitzmann B. Non-invasive lactate- and pH-monitoring in porcine meat using Raman spectroscopy and chemometrics. Chemometr Intell Lab 2015;142:197–205.
[25] Silalahi DD, Reano CE, Lansigan FP, Panopio RG, Bantayan NC. Using genetic algorithm neural network on near infrared spectral data for ripeness grading of oil palm (Elaeis Guineensis Jacq.) fresh fruit. Inform Process Agric 2016;4:252–61.
[26] Taki M, Mehdizadeh SA, Rahnama M, Rahmati-Joneidabad MR. Applied machine learning in greenhouse simulation; new application and analysis. Inform Process Agric 2018;2:253–68.
[27] Klocke M, editor. Prozessmikrobiologie In Landwirtschaftlichen Biogasanlagen: Ermittlung Der Mikrobiellen Diversität Sowie Von Hauptsächlichen Verfahrenstechnischen Einflussfaktoren Auf Die Mikroflora. Potsdam-Bornim: Leibniz-Institut Für Agrartechnik Potsdam-Bornim E.V.; 2014 [in German].
[28] Demuth H, Beale M. Neural network toolbox user's guide. Natick: The Mathworks; 2000.
[29] Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemometr Intell Lab 2001;58:109–30.
[30] Hagan MT, Demuth HB, Jesus OD. An introduction to the use of neural networks in control systems. Int J Robust Nonlinear Control 2002;12:959–85.
[31] Mccall J. Genetic algorithms for modelling and optimisation. J Comput Appl Math 2005;184:205–22.
[32] Leardi R. Application of a genetic algorithm to feature selection under full validation conditions and to outlier detection. J. Chemometrics 1994;8:65–79.
[33] Tsai C-F, Eberle W, Chu C-Y. Genetic algorithms in feature and instance selection. Knowl-Based Syst 2013;39:240–7.
[34] Gueguim Kana EB, Oloke JK, Lateef A, Adesiyan MO. Modeling and optimization of biogas production on saw dust and other co-substrates using artificial neural network and genetic algorithm. Renew Energ 2012;46:276–81.
[35] 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. Resour Conserv Recy 2010;54:359–63.
[36] Li C, Li Z, Wu J, Zhu L, Yue J. A hybrid model for dissolved oxygen prediction in aquaculture based on multi-scale features. Inform Process Agric 2018;5:11–20.
[37] Samadianfard S, Ghorbani MA, Mohammadi B. Forecasting soil temperature at multiple-depth with a hybrid artificial neural network model coupled-hybrid firefly optimizer algorithm. Inform Process Agric 2018;5:465–76.