Estimation of biogas and methane yields in an UASB treating potato starch processing wastewater with backpropagation artificial neural network

Philip Antwi, Jianzheng Li, Portia Opoku Boadi, Jia Meng, En Shi, Kaiwen Deng, Francis Kwesi Bondinuba

PII: S0960-8524(16)31715-1
DOI: http://dx.doi.org/10.1016/j.biortech.2016.12.045
Reference: BITE 17421

To appear in: *Bioresource Technology*

Received Date: 1 November 2016
Revised Date: 8 December 2016
Accepted Date: 11 December 2016

Please cite this article as: Antwi, P., Li, J., Opoku Boadi, P., Meng, J., Shi, E., Deng, K., Kwesi Bondinuba, F., Estimation of biogas and methane yields in an UASB treating potato starch processing wastewater with backpropagation artificial neural network, *Bioresource Technology* (2016), doi: http://dx.doi.org/10.1016/j.biortech.2016.12.045

Estimation of biogas and methane yields in an UASB treating potato starch processing wastewater with backpropagation artificial neural network

Philip Antwi[a], Jianzheng Li[a]*, Portia Opoku Boadi[b], Jia Meng[a], En Shi[a], Kaiwen Deng[a], Francis Kwesi Bondinuba[c].

a. State Key Laboratory of Urban Water Resource and Environment, School of Municipal and Environmental Engineering, Harbin Institute of Technology, 73 Huanghe Road, Harbin 150090, P.R. China
b. School of Management, Harbin Institute of Technology, 92 West Dazhi Street , Nan Gang District, Harbin 150001 , P.R. China
c. School of Energy, Geoscience, Infrastructure and Society, Institute for Social Policy, Housing, Environment and Real Estate, Heriot-Watt University, UK.

* Corresponding author (J. Li). Tel.: +86 451 86283761; Fax: +86 451 86283761; E-mail:ljz667@163.com

Authors:

Philip Antwi: Tel.: +86 451 86283761; E-mail: kobbyjean@yahoo.co.uk

Jianzheng Li: Tel.: +86 451 86283761; E-mail: ljz667@163.com

Portia Opoku Boadi: Tel.: +86 451 86414010; E-mail: portiaopokuboadi@yahoo.com

Jia Meng: +1 336 334 5388; E-mail: mengjia2726688@126.com

En Shi: +86 451 86283761; E-mail: shien851216@126.com

Kaiwen Deng: +86 451 86283761; E-mail: dengkw_hit@126.com

Francis Kwesi Bondinuba: +44 131 451 4434; E-mail: fkb30@hw.ac.uk

ABSTRACT

Three-layered feedforward backpropagation (BP) artificial neural networks (ANN) and multiple nonlinear regression (MnLR) models were developed to estimate biogas and methane yield in an upflow anaerobic sludge blanket (UASB) reactor treating potato starch processing wastewater (PSPW). Anaerobic process parameters were optimized to identify their importance on methanation. pH, total chemical oxygen demand, ammonium, alkalinity, total Kjeldahl nitrogen, total phosphorus, volatile fatty acids and hydraulic retention time selected based on principal component analysis were used as input variables, whiles biogas and methane yield were employed as target variables. Quasi-Newton method and conjugate gradient backpropagation algorithms were best among eleven training algorithms. Coefficient of determination (R^2) of the BP-ANN reached 98.72% and 97.93% whiles MnLR model attained 93.9% and 91.08% for biogas and methane yield, respectively. Compared with the MnLR model, BP-ANN model demonstrated significant performance, suggesting possible control of the anaerobic digestion process with the BP-ANN model.

Keywords: potato starch processing wastewater; upflow anaerobic sludge blanket; methane yield; optimized; artificial neural networks

1 Introduction

Energy recovery through biological processes is an environmentally sensitive means to generate energy and reduce greenhouse gases that has the potential to impacts negatively on the environment (Angenent et al., 2004; Şentürk et al., 2010; Akkaya et al., 2015). Anaerobic wastewater treatment can yield methane, hydrogen or other scarce biochemicals that can effectively be used as energy. Potato starch processing generates tons of wastewater which contains organic by-products such as starch, proteins, amino acids sugars, and potassium (Dabestani et al., 2017). These organic by-products that are biodegradable contributes to the high records of chemical oxygen demand (COD), 5-day biochemical oxygen demand (BOD_5) and suspended solids (SS) in the potato starch processing wastewater (PSPW) (Dabestani et al., 2017). Regarding the biodegradability characteristics, valuable energy resources such as methane or biogas could be harnessed from the wastewater (PSPW) through anaerobic digestion (AD) (Arhoun et al., 2013).

AD has not only been employed to treat sewage and industrial wastewater but also generate biogas (Şentürk et al., 2010; Zheng et al., 2012; Arhoun et al., 2013). So far, various types of processes have been proposed and reported in the treatment of potato wastewater, among which AD has proven to be very effective and improved the final effluent quality (Wang, 2013). As known, process modeling can be employed as a tool for predicting and describing performance of biological processes (Hu et al., 2002). Artificial neural networks (ANN) could be developed into process models and used successfully due to its capacity to capture the non-linear relationships that might exist among variables (multi-input/output) in a complex system (Kanat and Saral, 2009; Delnavaz et al., 2010; Khataee and Kasiri; 2011

Sun et al., 2012; Ghosh et al., 2013; Yetilmezsoy et al., 2013; Gong and Ordieres-Meré, 2016; Nair et al., 2016). Nasr and coworkers (Nasr et al., 2013) were successful in predicting hydrogen production profile with an ANN model. Khataee and coworkers also investigated the biological treatment of a dye solution by macro algae *Chara* sp., where 97% of the variations in the output variable were well explained by the input variables within the ANN framework (Khataee et al., 2010). Mechanistic modeling has also been implemented successfully, although means to acquire kinetic parameters is often laborious and difficult (Nasr et al., 2013; Brooks et al., 2016). Comparatively, the ANN methodology and framework can investigate and model AD processes without dependence on kinetic parameters acquired from the anaerobic process or system. However, few researches could be found employing ANN modeling to estimate biogas and methane yield in an upflow anaerobic sludge blanket (UASB) reactor treating PSPW.

Herein, the aim was; to develop a rapid and efficient methodology able to estimate biogas and methane production processes given initial substrate compositions and operational parameters; to identify and optimize essential process variables capable of making reliable predictions; and to develop a process that could possibly reduce cost and time of analysis. pH, COD, ammonium (NH_4^+), alkalinity (ALK), total Kjeldahl nitrogen (TKN), total phosphorus (TP), volatile fatty acids (VFAs) and hydraulic retention time (HRT) obtained from the anaerobic process were selected based on principal component analysis and used as input variables to develop three-layered ANN models (8:N_H:1) and multiple non-linear regression models. The anaerobic process parameters were optimized to identify their effects on methanation from the UASB. The efficiency of the developed ANN-based

models was compared with the multiple nonlinear regression models to make reliable simulations and predictions about biogas and methane yields within the UASB.

2 Materials and Methods

2.1 Experimental setup and operation

Experiments were conducted in a 120 cm high UASB constructed with a Plexiglas column (Fig. 1). The reactor had a total and effective working volume of 8.8 L and 7 L, respectively. Five sampling ports at approximately 25 cm interval were allocated along the vertical height of the cylinder under the gas-liquid-solid separator. The reactor was operated at 35±1°C which was maintained with a controller. Excess activated sludge collected from a local anaerobic-anoxic-oxic process treating municipal sewage was used to inoculate the UASB. At the started up of the reactor, the mixed liquor suspended solid (MLSS) and mixed liquor volatile suspended solid (MLVSS) was 11.5 and 5.6 g/L, respectively. PSPW was collected from a local starch producing industry and kept under 4^{o}C.

The concentration of the wastewater in terms of COD, NH_4^+, TP, TKN, ALK and VFAs averaged 49179, 302, 190, 1023, 4945 and 534 mg/L, respectively. The raw wastewater was diluted to a favorable quality and fed to the UASB by a peristaltic pump (BT100-2J, Langer Instruments, UK). The average feed concentration in terms of COD, NH_4^+, pH, ALK, TKN, TP and VFAs was 4029, 110, 7, 2152, 511, 45, 103 mg/L, respectively . Within the startup period, operation of the UASB was divided into two stages in terms of HRT. The first 49 days was the first stage with a HRT of 48 h. HRT was subsequently

curtailed to 24 h in the following 63 days as the second stage. The evolved biogas were collected by the gas-solid-liquid separator and was measured daily by a wet gas meter (Model LML-1, Changchun Filter Co., Ltd., China).

2.2 Analytical methods

All chemical analysis were conducted in accordance with Standard Methods for the Examination of Water and Wastewater, APHA (APHA, 2007). Influent and effluent COD, ALK (in terms of $CaCO_3$), TKN, NH_4^+ and TP were analyzed daily. pH was determined using a DELTA 320 (Mettler Toledo, USA). VFAs in liquid samples were measured by a gas chromatograph (SP6890, Shandong Lunan Instrument Factory, China) equipped with a 30 m capillary column (Stabilwax-DA, i.d.0.32 mm, 11054, Restek) and a flame ionization detector (FID) (Liu et al., 2015). The operational temperatures of the injection port, oven and detector were 210°C, 180°C, and 210°C, respectively. Nitrogen gas was used as the carrier gas, with a 0.75 MPa column head pressure. The split ratio was 1:50. Liquid sample of 1 mL was centrifuged at 13000 rpm for 3 min. A 0.5 mL of the supernatant after centrifuge was pipetted and acidified with 25% H_3PO_4, and then 1 μL of the final solution injected. For biogas fraction, 0.5 mL biogas was sampled from the headspace of the UASB to determine methane (CH_4) and carbon dioxide (CO_2) fractions by another gas chromatograph (SP-6800A, Shandong Lunan Instrument Factory, China). The gas chromatograph was equipped with a thermal conductivity detector (TCD) and a 2 m stainless column packed with Porapak Q (60/80 mesh) (Liu et al., 2015). Temperatures of the injector, column and the TCD were 80°C, 50°C and 80°C, respectively.

2.3 Optimization and selection of input and output variables

Optimization of anaerobic parameters was carried out by setting methane recovery target at 65-75% to the biogas production.

Parameters above the targeted limits (65-75%) were easily identified with scatter plots. Based on the variable selection process output, thus principal component analysis (PCA), pH, COD, ammonium (NH_4^+), alkalinity (ALK), total Kjeldahl nitrogen (TKN), total phosphorus (TP), volatile fatty acids (VFAs) and biogas yield were optimized against the targeted methane proportion. Furthermore, the experimental data set was divided into input (I_P) and target (T_P) variables and loaded into the MATLAB workspace (Matrix Laboratory R2014a, version 8.3 by MathWorks, Inc., USA) to appropriately identify and select the most effective variables. The input and target data given in matrices [I_P] and [T_P] were normalized using *prestd* algorithm code. Prior to the training of the network, clear definitions were given to the mean input data, mean target data, standard deviations of input data, standard deviations of target data, transformed input vectors and principal component transformation matrix as *meanIp*, *meanTp*, *stdIp, stdTp*, *Iptrans* and *transMat*, respectively.

Principal component analyses (PCA) were carried out to reduce the number of variables. The PCA transformed a number of correlated variables into a smaller number of uncorrelated variables which could sufficiently explain the data structure. The principal components that contributed less than 0.1% to the variation in the data set were eliminated (Yetilmezsoy and Sapci-Zengin, 2009). As reported, predictive effects of anaerobic digestion processes highly depends on the variable selection process (Yetilmezsoy et al.,

2013). Based on the PCA, eight process-related parameters were selected as shown in Table 1 and considered as I_P in the ANN model development (Faul et al., 2009) while, biogas and methane yield were selected as T_p.

2.4 Description of the artificial neural network

The MATLAB backpropagation (BP) algorithm was used to develop the ANN model. The ANN model of input vector (8×112) and target vector (2×112) consisted of neurons ordered in 3 layers, thus input layer, hidden layer and output layer as illustrated in Fig.2. The input neurons represented the independent process variables. The output neurons were the dependent predicted variables.

The hidden layer transformed the input information (Beltramo et al., 2016). As data set was trained, the input pattern given to the input layers of the network would compute the output in the output layer (Liu et al., 2016). The BP learning rule defined a method to adjust the weights of the networks. The network then gave outputs that could match the desired output pattern given any input pattern in the training set (Cheng et al., 2016). The outputs of the hidden neurons acted as inputs to the output neuron and then underwent another transformation. The output of the BP-ANN with a hidden layer and one output neural network was estimated with Eq.1.

$$\hat{Y} = f_o\left[\sum_{j=1}^{HN} WO_j \times f_h\left(\sum_{i=1}^{m} WH_{ij} \times X_{it} + b_j\right) + b_o\right] \quad (1)$$

where, WH_{ij} is the weight of the link between the i^{th} input and the j^{th} hidden neuron, m is the number of input neurons, WO_j is the weight of the link between the j^{th} hidden neuron and the output neuron, f_h is the hidden neuron activation function, f_o is the output neuron

activation function, b_j is the bias of the j^{th} hidden neurons, b_o is the bias of the output neuron, X_{it} is the input variable, and *HN* is the number of hidden neurons.

Tangent sigmoid transfer function (*tansig*) (Eq.2) and linear transfer function (*purelin*) (Eq.3) were employed at the hidden and output layer, respectively.

$$f(x) = \frac{2}{(1+exp^{-2x})} - 1 \tag{2}$$

$$f(x) = x \tag{3}$$

where x is the vector of inputs.

The original data set was divided randomly into three ANN subsets (train, validation and test). Out of 112 data set points obtained to develop the 3-layered ANN model, 17 data points representing 15% of the original data set were respectively selected for the validation and testing subsets, whiles 78 data set points representing 70% were allocated for the training set. The test set was used for the prediction. The BP-ANN models were subsequently validated with the index of agreement (*IA*) and the fractional variance (*FV*) as shown in Eq.4 and Eq.5.

$$IA = 1 - \frac{\sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}(|P_i - O_m| + |O_i - O_m|)^2} \tag{4}$$

$$FV = \frac{2(\delta_o - \delta_p)}{(\delta_o + \delta_p)} \tag{5}$$

where *O, P,* δ and *m* indicates experimental data, predicted values, standard deviation and arithmetic mean of the observed data points, respectively.

2.5 ANN Training algorithm selection and optimization process

A benchmark comparison was conducted to facilitate the selection of the optimum neural network in the ANN modeling process (Almasri and Kaluarachchi, 2005). The mean square error (MSE) was used to justify the learning effects of the BP-ANN. The hidden layer was firstly assigned with two neurons as an initial assumption. As neuron numbers were increased stepwisely, the corresponding MSEs obtained were used for the comparison. The training continued until the MSEs were below some tolerance level. 10 neurons were finally set as default number of neurons at the hidden layer for each training algorithms. Networks selection was primarily centered on the highest performed training algorithm. The relationship between the MSE values and the number of neurons in the hidden layer is given in Eq.6.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(T_i - A_i)^2 \quad (6)$$

where N is the number of data point, T_i is the network predicted value at the i^{th} data, A_i is the experimental value at the i^{th} data and i is an index of the data. Since there is the tendency of underfitting or overfitting per the number neurons, the early stopping method was employed in this study. The training set was used as the first subset to compute the gradient and update the network weights and biases. The validation set was the second subset and the error obtained in this set was constantly monitored during the training process.

2.6 Multiple non-linear regression analysis

A multiple nonlinear regression models (MnLRM) by residual analysis was also developed within the MINITAB (version 17), SIGMAPLOT (version 12.5) and XLSTATS statistical computing environment. First, Pearson's correlation analysis was performed using the correlation function in MINITAB. Significances of variables were corrected based on *p*-values less than 0.05 (Ramette, 2007). The general form of the MnLRM used in this study is as shown in Eq.7. The output variable *y*, written as a function of *k,* had input variables x_1, x_2…. x_k and a random error term $\hat{\varepsilon}$ that was added to make the model probabilistic rather than deterministic. The coefficients β_0, β_1... β_k which were usually unknown were subsequently estimated (Mac Nally, 2000; Huang and Chen, 2001; Yetilmezsoy and Sakar, 2008; Singh et al., 2010; Turkdogan-Aydınol and Yetilmezsoy, 2010).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \hat{\varepsilon} \quad (7)$$

where x_1, x_2, and x_k represented terms for quantitative predictors. Assumptions, including linearity, independence among errors, non-multicollinearity, homoscedasticity, non-autocorrelation and normal distribution of errors, were considered (Wold et al., 2001).

2.7 MnLRM selection

The optimum MnLRM were selected based on the following statistical performance criterion: coefficient of multiple determination (R^2) (Eq.8), adjusted coefficient of multiple determination ($Adj\text{-}R^2$) (Abdul-Wahab et al., 2005) (Eq.9), residual average (*RA*) (Eq.10), sum of squared residuals (*SSR*) (Eq.11), standard error of the estimate (*SEE*) (Xu et al.,

2015) (Eq.12), *VIF* (Eq.13), Durbin-Watson statistics (*d*) (Eq.14) and *p*-value (Yetilmezsoy et al., 2013) (Eq.15).

$$R^2=\frac{\sum_{i=1}^{n}\left(Y_p-\bar{Y}\right)^2}{\sum_{i=1}^{n}(Y_o-\bar{Y})^2} \tag{8}$$

$$R_{adj}^2=\left[\frac{\left(1-R^2\right)(n-1)}{n-k-1}\right] \tag{9}$$

$$RA=\sum_{i=1}^{n}\left(Y_o-Y_p\right) \tag{10}$$

$$SSR=\sum_{i=1}^{n}\left(Y_o-Y_p\right)^2 \tag{11}$$

$$SEE=\sqrt{\frac{\sum_{i=1}^{n}\left(Y_o-Y_p\right)^2}{n-m}} \tag{12}$$

$$VIF=\frac{1}{1-R^2} \tag{13}$$

$$d=\frac{\sum_{i=1}^{n}(e_i-e_{i-1})^2}{\sum_{i=1}^{n}e_i^2} \tag{14}$$

$$p=2\times P(TS>|ts||H_o \text{ is true})=2\times\left(1-cdf(|ts|)\right) \tag{15}$$

where, Y_o, Y_p and $\bar{Y}$ denotes experimental data, predicted values and arithmetic mean of the observed data; n and m is the number of data points and parameters in the regression model, respectively; k is the number of independent regressors excluding the constant term; $e_i = y_i - \hat{y}_l$, and y_i and $\hat{y}_l$ were, respectively, the observed and predicted values of the response variable for individual i; *TS* is random variable associated with the assumed distribution; *ts* is the test statistics calculated from sample, and *cdf* is the cumulative density function of the assumed distribution.

3 Results and Discussion

3.1 Performance of the UASB and optimized anaerobic parameters

The UASB was operated for a period of 112 days at different HRTs (48 h and 24 h). The operation commenced with a 48 h HRT along with an organic loading rate (OLR) of 1.5 kgCOD/m^3·d. Subsequently, HRT was stepwisely shortened to 24 h with an increased OLR of 4.23 kgCOD/m^3·d. The OLRs employed had no inhibition effect on the UASB operation. Accordingly, Fang and coworkers had earlier reported that, an UASB could maintain a stable process once operation is within the normal OLR boundaries which ranges between 1.5 and 16.0 kgCOD/m^3·d (Fang et al., 2011; Jing et al., 2013). Fig. 3 presented the performance of the reactor during the 112 days of operation. As HRT of 48 h and influent COD of about 3799 mg/L was employed, an average COD removal of 87% was observed in the UASB (Fig.3a). However, when HRT was curtailed to 24 h along with increased OLR (4.23 kgCOD/m^3·d), COD removal achieved was about 92%. This could be ascribed to the high adaptability of the responsible microbes to the environment (Lu et al., 2015).

Observably, biogas yield increased along with an increasing OLR suggesting positive correlation among biogas yield and OLR. At both HRTs, biogas produced ranged from 3.4 to 17.4 L/d. Methane (CH_4) fractions were maintained in the range of 56.2% to 84.5% at both HRTs (Fig.3b). It was found that the average influent and effluent ALK at HRT 48 h were 6010 and 10948 mg/L, while that of 24 h HRT were 3592 and 8638 mg/L for HRT 24 h, respectively (Fig.3c). The ALK could be prime factor as the UASB reactor exhibited a better buffering capacity regardless of the influent pH. pH of about 8.0 was maintained in

the effluent although influent pH (5.2 - 8.0) fluctuated remarkably (Fig.3c). The feasible pH and ALK enhanced the acetogenesis and methanogenesis in the reactor, resulting in few VFAs (<150 mg/L) in the effluent shown in Fig.3d (De Sousa et al., 2008).

The illustration in Fig.3e depicts the average influent and effluent TKN. At HRT of 48 h and 24 h, the influent TKN were respectively, 466 and 518 mg/L, while that in the effluent at 48 h and 24 h were 307 and 507 mg/L, respectively. The degradation of organic nitrogen compounds elevated the NH_4^+ concentration in the effluent as similarly reported by Park and coworkers (Park et al., 2010). Accordingly, effluent NH_4^+ observed averaged 241mg/L whiles that in the influent feed (PSPW) averaged 109 mg/L (Fig.3f). The UASB showed little TP removal with almost same concentration (45 mg/L) in both influent and effluent at both HRTs.

The anaerobic parameters optimized by setting CH_4 content in biogas at 60–80% is shown in Fig S1. The optimum range of the corresponding anaerobic parameters were determined and the results are as follows: pH between 6.43 and 7.74, COD of 3485-4964 mg/L, ammonium (NH_4^+) of 80.9-137.2 mg/L, alkalinity (ALK) of 3010-6889 mg/L, total Kjeldahl nitrogen (TKN) of 399-604 mg/L, total phosphorus (TP) of 34-54 mg/L, volatile fatty acids (VFAs) of 259-809 mg/L and biogas yield of 3.40-16.8 L/d (Fig S1).

3.2 Optimization of the neural network structure

Speed of a network algorithm primarily depends on the characteristics of the data set, the complexity of the problem and the number of neurons specified in the network (Yetilmezsoy and Sapci-Zengin, 2009).

Selection of a suitable training algorithm is vital to defining the optimal architecture of the ANN model. As a result, several training algorithms and their variations have been proposed in the literature (Giwa et al., 2016; Nasr et al., 2013). The benchmark comparison conducted among the 11 different algorithms in the present research revealed that, (Table 2), the Broyden–Fletcher–Goldfarb–Shanno (BFGS) Quasi-Newton (*trainbfg*) and Polak-Ribiere conjugate gradient backpropagation (CGP) algorithms (*traincgp*) for biogas and methane, respectively, manifested as best algorithms for making predictions (Fig.4). Compared with the other 9 algorithms, smaller mean squared errors (MSE) of 0.567 and 0.617 (Table 2) were obtained in Broyden–Fletcher–Goldfarb–Shanno (BFGS) and Polak–Ribiere conjugate gradient backpropagation (CGP), respectively, for the estimation of biogas (Fig.4a) and methane (Fig.4b) yield. The BFGS and the CGP algorithms showed optimum performance in the ANN architectures for biogas and methane predictions. The worst performed algorithms in terms of MSE (Table 2) were the batch gradient descent (*traingd*) and the scaled conjugate gradient backpropagation (*trainsgc*) in biogas and methane predictions, respectively. Evidence of the loss on optimality in the 9 algorithms may be ascribed to the combinatorial nature and non-linear conditions that existed in the data set (Yetilmezsoy et al., 2013).

As illustrated in Fig.4, the number of neurons in the hidden layer (N_H) for methane neural network was optimized with 3 neurons at a minimum MSE of 0.541 (Fig.4c), while that of biogas occurred with 4 neurons with a minimum MSE value of 0.282 (Fig.4d). Cheng et al had discussed that, number of hidden neurons which is an important feature is not selected based on any formulae but rather on the relative mean squared error from the different

nodes within the range (Cheng et al., 2016). Herein, MSE increased tremendously as number of neurons exceeded the global minimum. Observably, when 10 neurons were assigned within the BFGS algorithm, MSE increased to 1.21 as compared to the 4 neurons. Similarly with the methane data set, MSE reached a peak of 1.884 within the CGP algorithm when neurons were increased to 16.

Large size of nodes in hidden layer may lead to over-fitting (Gong and Ordieres-Mere, 2016). In this study, the early stopping employed to evaluate underfitting and overfitting indicated that, the training set error and the validation error decreased at the initial training phase. However, the validation set error increased as the network began to overfit the data. When the validation error increased with a specified number of iterations, the training was pulsed, and the weights and biases at the minimum of the validation error were returned. No significant overfitting, underfitting and negative ANN estimations were observed in the output data sets although a linear transfer function (*purelin*) was used in the output layer. This phenomenal observation could be attributed to the characteristics of the input vectors used.

The illustration in Fig.5 and Fig.6 depicted the correlations and corresponding visual agreement between the experimental data and the BP-ANN output. The proposed BP-ANN model demonstrated very satisfactory performance in predicting biogas and methane yield. Coefficient of determination in all ANN subset including training, validation and testing data sets for biogas predictions reached 98% (Fig.5a, Fig.5c, and Fig.5e). Validation and testing data sets in the methane prediction were greater than 97% (Fig.6c, Fig.6e), but that in the training set reached a little above 95% (Fig.6a). This phenomenal performance could

be attributed to the fact that, BP-ANN model had the ability to capture the complex behavior that existed among the variables obtained from the anaerobic digestion process (Giwa et al., 2016).

3.3 Optimization of the multiple regression models

Biogas and methane yield from the UASB were also estimated with the multiple nonlinear regression models. The model coefficients, constants term, input variables and the result on variable analysis including standard error and the *p*-values are given in Table 3. Statistical analysis of the regression input variables revealed that some variables were statistically significant which confirmed their importance in the model development over the others. Considering the *p*-values obtained (Table 3) for all input variables in the MnLR models, it was evident that COD, VFAs, NH_4^+ and HRT were more statistically significant compared to that of pH, TKN, ALK and TP. This observation demonstrated that COD, VFAs, NH_4^+ and HRT had greater importance in estimating biogas and methane yield compared to the other input variables. Obviously, pH, TKN, ALK and TP with *p*-values of 0.75, 0.74, 0.8 and 0.9, respectively, were greater than the 0.05 threshold value specified. Similar observation was made in the case of methane prediction by the MnLR models. pH, NH_4^+, TKN, ALK and TP were not significant except COD, VFA and HRT (Table 3).

The results from the regression model are presented in Table 4. The best fit models for both biogas and methane predictions were the linear model. Wider deviations were noticed between experimental data and the predicted values obtained from the exponential model. However, relatively smaller deviations were demonstrated by the linear model during

predictions. The linear models were defined as a function of HRT and seven anaerobic process parameters including COD, VFAs, NH_4^+, TKN, ALK, PH and TP. To appreciate the performance of the MnLR models, residual analysis were conducted and the results presented in Table 4. The multiple coefficient of determination (R^2) obtained in predicting biogas and methane were 93.30% and 91.08%, respectively. Obviously, the deviations were seen in the visual agreements between MnLR output and the experimental data as illustrated in Fig S2.

In terms of standard error of the estimates (*SEE*), the linear model in both biogas and methane predictions recorded the lowest values of 0.82 and 1.10, respectively (Table 4), suggesting a more precise evaluation of the variation in the estimated mean for the set of predictor values. The values obtained in estimating the squared sum residual (*SSR*) were the lowest compared to the exponential models. The values 69.7 and 125.2 for biogas and methane, respectively (Table 4), represented least variation or deviation of predictions from the mean.

3.4 Comparison of BP-ANN and MnLR Models

The prediction accuracy of the BP-ANN and MnLR models were evaluated with R^2, index of agreement (*IA*) and the fractional variance (*FV*) (Table 5). At biogas predictions, the obtained coefficient of determination (R^2) in BP-ANN (98.72%) was relatively higher compared to that in MnLR model (95.31%). Similar higher performance was also observed with BP-ANN (97.93%) during methane prediction. However, that noticed in the MnLR model (92.62%) was relatively low suggesting the inability of the MnLR to make reliable

predictions in complex biological systems. Further comparing the models efficiency with the R^2, BP-ANN model demonstrated higher predicting efficiency over the MnLR model. It was noticed that, only 1.28 % of the total variation existing in the biogas data sets were not explained by the BP-ANN model as opposed to 4.69% of the MnLR model. With respect to estimating methane with the BP-ANN model, similar lower percentage (2.07%) of the total variation did not fit (unexplained) the experimental methane data set. However, as much as 7.38% was unexplained by MnLR model indicating its low efficiency in predicting methane. The observation could be ascribed to the advantage of the ANNs capability in explaining complex interactions between inputs and output parameters (Yetilmezsoy and Sapci-Zengin, 2009). The index of agreement (*IA*) obtained with the BP-ANN model at biogas and methane prediction was 0.9941 and 0.9806, respectively. Comparing the *IA*s obtained in MnLR model (biogas; 0.9725, methane; 0.9611) to that of the BP-ANN model, it was obvious *IA*s in BP-ANN were higher, suggesting that BP-ANN model could make reliable prediction.

The fractional variance (*FV*) will be 1 if the explanatory variables (x) tell nothing about variable (Y), thus, the predicted values of Y do not co-vary with Y. On the contrary, *FV* is 0 if the explanatory variables (x) are able to make perfect predictions of variable Y (Yetilmezsoy et al., 2008). Relatively lower *FV*s were observable in BP-ANN model. At biogas predictions, *FVs* obtained in BP-ANN and MnLR model were 0.0075 and 0.0821, respectively, whiles that obtained at methane prediction was 0.00284 and 0.073, respectively. The results confirmed that, BP-ANN model could not estimate only 0.75% of the biogas data set and 0.28% of the methane data set. However, the MnLR model

manifested about 8.2% and 7.3% for biogas and methane prediction, respectively. The overall performance of the models in terms of R^2, *IA* and *FV* suggested that, the BP-ANN model had a stronger predictive power compared to the MnLR models.

4 Conclusions

Biogas and methane generated from mesophilic UASB was evaluated and modeled. Optimized methanation (target at 65-75%) could be achieve at pH between 6.43 and 7.74, COD of 3485-4964 mg/L, ammonium (NH_4^+) of 80.9-137.2 mg/L, alkalinity (ALK) of 3010-6889 mg/L, volatile fatty acids (VFAs) of 259-809 mg/L and biogas yield of 3.40-16.8 L/d. Quasi-Newton method and conjugate gradient backpropagation were best algorithms among eleven training algorithms. R^2, *IA* and *FV* obtained in BP-ANN model indicated that, only 1.2% and 2.07% of the biogas and methane data set respectively, were unexplained suggesting BP-ANN's potential to simulate nonlinear relationship in wastewater treatment systems.

Acknowledgements

The authors are very grateful to the following supporting institutions for their financial support: The Major Science and Technology Program for Water Pollution Control and Management (Grant No. 2013ZX07201007), and The State Key Laboratory of Urban Water Resource and Environment, Harbin Institute of Technology (Grant No. 2016DX06).

References

Abdul-Wahab, S. A., Bakheit, C. S., Al-Alawi, S. M. 2005. Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. Environmental Modelling & Software, 20(10), 1263-1271.

Akkaya, E., Demir, A., Varank, G., 2015. Estimation of Biogas Generation from a Uasb Reactor via Multiple Regression Model. International Journal of Green Energy, 12, 185-189.

Almasri, M. N., Kaluarachchi, J. J., 2005. Modular neural networks to predict the nitrate distribution in ground water using the on-ground nitrogen loading and recharge data. Environmental Modelling & Software, 20, 851-871.

Angenent, L. T., Karim, K., AL-Dahhan, M. H., Wrenn, B. A., Domíguez-Espinosa, R., 2004. Production of bioenergy and biochemicals from industrial and agricultural wastewater. TRENDS in Biotechnology, 22, 477-485.

APHA, A., 2007. WEF (2005) Standard methods for the examination of water and wastewater. American Public Health Association, American Water Works Association, and Water Environment Federation.

Arhoun, B., Bakkali, A., EL Mail, R., Rodriguez-Maroto, J., Garcia-Herruzo, F., 2013. Biogas production from pear residues using sludge from a wastewater treatment plant digester. Influence of the feed delivery procedure. Bioresource technology, 127, 242-247.

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.

Brooks, W., Corsi, S., Fienen, M., Carvin, R., 2016. Predicting recreational water quality advisories: A comparison of statistical methods. Environmental Modelling & Software, 76, 81-94.

Cheng, J., Wang, X., Si, T., Zhou, F., Zhou, J., Cen, K., 2016. Ignition temperature and activation energy of power coal blends predicted with back-propagation neural network models. Fuel. 173, 230-238

De Sousa, J. T., Santos, K. D., Henrique, I. N., Brasil, D. P., Santos, E. C., 2008. Anaerobic digestion and the denitrification in UASB reactor. Journal of Urban and Environmental Engineering, 2, 63-67.

Dabestani, S., Arcot, J., Chen, V., 2017. Protein recovery from potato processing water: Pre-treatment and membrane fouling minimization. Journal of Food Engineering, 195, 85-96.

Delnavaz, M., Ayati, B., Ganjidoust, H., 2010. Prediction of moving bed biofilm reactor (MBBR) performance for the treatment of aniline using artificial neural networks (ANN). Journal of hazardous materials, 179, 769-775.

Fang, C., Boe, K., Angelidaki, I. 2011. Biogas production from potato-juice, a by-product from potato-starch processing, in upflow anaerobic sludge blanket (UASB) and expanded granular sludge bed (EGSB) reactors. Bioresource technology, 102(10), 5734-5741.

Faul, F., Erdfelder, E., Buchner, A., Lang, A.-G., 2009. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. Behavior research methods, 41, 1149-1160.

Ghosh, A., Sinha, K., Saha, P. D., 2013. Central composite design optimization and artificial neural network modeling of copper removal by chemically modified orange peel. Desalination and Water Treatment, 51, 7791-7799.

Giwa, A., Daer, S., Ahmed, I., Marpu, P., Hasan, S. 2016. Experimental investigation and artificial neural networks ANNs modeling of electrically-enhanced membrane bioreactor for wastewater treatment. Journal of Water Process Engineering, 11, 88-97.

Gong, B., Ordieres-Meré, J., 2016. Prediction of daily maximum ozone threshold exceedances by preprocessing and ensemble artificial intelligence techniques: Case study of Hong Kong. Environmental Modelling & Software, 84, 290-303.

Hu, W., Thayanithy, K., Forster, C., 2002. A kinetic study of the anaerobic digestion of ice-cream wastewater. Process Biochemistry, 37, 965-971.

Huang, L., chen, J. C., 2001. A multiple regression model to predict in-process surface roughness in turning operation via accelerometer. Journal of Industrial Technology, 17, 1-8.

Jing, Z., Hu, Y., Niu, Q., Liu, Y., Li, Y.-Y., Wang, X.C. 2013. UASB performance and electron competition between methane-producing archaea and sulfate-reducing bacteria in treating sulfate-rich wastewater containing ethanol and acetate. Bioresource technology, 137, 349-357.

Kanat, G., Saral, A., 2009. Estimation of biogas production rate in a thermophilic UASB reactor using artificial neural networks. Environmental Modeling & Assessment, 14, 607-614.

Khataee, A., Dehghan, G., Ebadi, A., Zarei, M., Pourhassan, M., 2010. Biological treatment of a dye solution by Macroalgae Chara sp.: Effect of operational parameters, intermediates identification and artificial neural network modeling. Bioresource technology, 101, 2252-2258.

Khataee, A. R., Kasiri, M. B., 2011. Modeling of biological water and wastewater treatment processes using artificial neural networks. CLEAN–Soil, Air, Water, 39, 742-749.

Liu, C., Li, J., Zhang, Y., Philip, A., Shi, E., Chi, X., Meng, J., 2015. Influence of glucose fermentation on CO2 assimilation to acetate in homoacetogen Blautia coccoides GA-1. Journal of industrial microbiology & biotechnology, 42, 1217-1224.

Liu, S., Xu, L., Li, D., 2016. Multi-scale prediction of water temperature using empirical mode decomposition with back-propagation neural networks. Computers & Electrical Engineering, 49, 1-8.

Lu, X., Zhen, G., Estrada, A.L., Chen, M., Ni, J., Hojo, T., Kubota, K., Li, Y.-Y. 2015. Operation performance and granule characterization of upflow anaerobic sludge blanket (UASB) reactor treating wastewater with starch as the sole carbon source. Bioresource technology, 180, 264-273.

Mac Nally, R., 2000. Regression and model-building in conservation biology, biogeography and ecology: the distinction between–and reconciliation of–'predictive'and 'explanatory'models. Biodiversity & Conservation, 9, 655-671.

Nasr, N., Hafez, H., El Naggar, M. H., Nakhla, G., 2013. Application of artificial neural networks for modeling of biohydrogen production. international journal of hydrogen energy, 38, 3189-3195.

Nair, V. V., Dhar, H., Kumar, S., Thalla, A. K., Mukherjee, S. & Wong, J. W. 2016. Artificial neural network based modeling to evaluate methane yield from biogas in a laboratory-scale anaerobic bioreactor. Bioresource technology, 217, 90-99.

Park, J., Jin, H.-F., Lim, B.-R., Park, K.-Y., Lee, K. 2010. Ammonia removal from anaerobic digestion effluent of livestock waste using green alga Scenedesmus sp. Bioresource technology, 101(22), 8649-8657.

Ramette, A., 2007. Multivariate analyses in microbial ecology. FEMS Microbiology Ecology, 62, 142-160.

Şentürk, E., Ince, M., Engin, G. O., 2010. Kinetic evaluation and performance of a mesophilic anaerobic contact reactor treating medium-strength food-processing wastewater. Bioresource technology, 101, 3970-3977.

Singh, K. P., Basant, N., Malik, A., Jain, G., 2010. Modeling the performance of "up-flow anaerobic sludge blanket" reactor based wastewater treatment plant using linear and nonlinear approaches—a case study. Analytica chimica acta, 658, 1-11.

Sun, X., Turng, L. S., Dougherty, E., Gorton, P., 2012. Artificial neural network–based supercritical fluid dosage control for microcellular injection molding. Advances in Polymer Technology, 31, 7-19.

Turkdogan-Aydınol, F. I., Yetilmezsoy, K., 2010. A fuzzy-logic-based model to predict biogas and methane production rates in a pilot-scale mesophilic UASB reactor treating molasses wastewater. Journal of hazardous materials, 182, 460-471.

Wang, H. F., 2013.Treatment of Sweet Potato Starch Wastewater with UASB. Advanced Materials Research, Trans Tech Publ, 234-237.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and intelligent laboratory systems, 58, 109-130.

Xu, Y., Ma, C., Liu, Q., Xi, B., Qian, G., Zhang, D., Huo, S. 2015. Method to predict key factors affecting lake eutrophication–A new approach based on Support Vector Regression model. International Biodeterioration & Biodegradation, 102, 308-315.

Yetilmezsoy, K., Sakar, S., 2008. Development of empirical models for performance evaluation of UASB reactors treating poultry manure wastewater under different operational conditions. Journal of Hazardous materials, 153, 532-543.

Yetilmezsoy, K., Sapci-Zengin, Z., 2009. Stochastic modeling applications for the prediction of COD removal efficiency of UASB reactors treating diluted real cotton textile wastewater. Stochastic environmental research and risk assessment, 23, 13-26.

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, 885-907.

Zheng, M., Wang, K., Zuo, J., Yan, Z., Fang, H., Yu, J., 2012. Flow pattern analysis of a full-scale expanded granular sludge bed-type reactor under different organic loading rates. Bioresource technology, 107, 33-40.

Figure captions

Fig. 1: Schematic diagram of the upflow anaerobic sludge bed reactor

Fig.2: Schematic flowchart of the proposed feedforward BP-ANN methodology

Fig.3: (a) COD and COD removal; (b) biogas and methane; (c) ALK and pH; (d) influent and effluent VFAs in total; (e) influent and effluent TKN; (f) influent and effluent ammonium

Fig 4: Comparison between training algorithms based on MSE, (a) biogas yield, (b) Methane yield; Optimum number of neurons required at hidden layer based on MSE; (**c**) Polak–Ribiere conjugate gradient backpropagation algorithm for methane yield (**d**) BFGS Quasi–Newton backpropagation algorithm for biogas yield.

Fig.5: Correlations (a,c,e) and visual agreements between biogas experimental data and ANN predictions(b,d,f)

Fig.6: Correlations (a,c,e) and visual agreements between methane experimental data and ANN predictions(b,d,f)

Tables

Table 1. Descriptive statistics of input and output variables

Variable		Term	Mean
Input parameters	COD (mg/L)	x_1	4028
	NH_4^+ (mg/L)	x_2	110
	pH	x_3	7.1
	ALK (mg/L, $CaCO_3$)	x_4	4944
	TKN (mg/L)	x_5	510
	TP (mg/L)	x_6	45
	Total VFAs (mg/L)	x_7	534
	HRT (h)	x_8	--
Target parameters	Biogas yield (L/d)	*Y1*	10.9
	Methane yield (L/d)	*Y2*	7.4

Table 2: Comparison of backpropagation training algorithms

Training Algorithm	Function	Abrv	Target sets used in the ANN study					
			Biogas data set			Methane data set		
			R^2	*IN*	*MSE*	R^2	*IN*	*MSE*
Broyden–Fletcher–Goldfarb–Shanno Quasi-Newton	*trainbfg*	BFG	97.7	109	0.567	97.05	119	1.405
Powell–Beale conjugate gradient backpropagation	*traincgb*	CGB	98.4	115	1.126	93.92	110	1.056
Fletcher–Reeves conjugate gradient backpropagation	*traincgf*	CGF	98.7	162	0.914	96.6	174	0.895
Polak–Ribiere conjugate gradient backpropagation	*traincgp*	CGP	98.6	128	0.643	97.04	229	0.617
Batch gradient descent	*traingd*	GD	14.6	100	3.21	96.57	1000	1.157
Batch gradient descent with momentum	*traingdm*	GDM	98.0	1000	1.414	94.67	237	1.620
Variable learning rate backpropagation	*traingdx*	GDX	97.6	172	0.929	94.84	202	2.250
Levenberg Marquardt backpropagation	*trainlm*	LM	96.7	76	2.140	95.16	105	1.940
One step secant backpropagation	*trainoss*	OSS	98.8	193	0.645	95.26	112	0.922
Resilient backpropagation	*trainrp*	RP	99.0	164	0.992	95.76	109	0.841
Scaled conjugate gradient backpropagation	*trainsgc*	SGC	98.7	139	0.781	95.47	128	2.826

R^2 - coefficient of determination; *IN* - number of iterations; *MSE* - mean squared errors; Abrv- abbreviation

Table 3: Descriptive statistics of MnLR model variables

Model coefficient and constant term	Input parameters	Standard error	[a]*p*-value
Biogas			
$\beta_0 = 18.28$	constant term	1.67	0.000
$\beta_1 = 0.00115$	$x_1 =$ COD	0.000157	0.000
$\beta_2 = 0.0113$	$x_2 = NH_4^+$	0.00583	0.055
$\beta_3 = -0.050$	$x_3 =$ pH	0.162	0.757
$\beta_4 = 0.000061$	$x_4 =$ ALK	0.000188	0.747
$\beta_5 = -0.00028$	$x_5 =$ TKN	0.0011	0.801
$\beta_6 = -0.001$	$x_6 =$ TP	0.0137	0.941
$\beta_7 = 0.00197$	$x_7 =$ VFA	0.000668	0.004
$\beta_8 = -0.4085$	$x_8 =$ HRT	0.0206	0.000
Methane			
$\beta_0 = 9.60$	constant term	2.24	0.000
$\beta_1 = 0.000892$	$x_1 =$ COD	0.00021	0.000
$\beta_2 = 0.00086$	$x_2 = NH_4^+$	0.0078	0.913
$\beta_3 = 0.160$	$x_3 =$ pH	0.217	0.463
$\beta_4 = 0.000261$	$x_4 =$ ALK	0.000252	0.304
$\beta_5 = 0.00138$	$x_5 =$ TKN	0.00148	0.353
$\beta_6 = 0.0175$	$x_6 =$ TP	0.0183	0.341
$\beta_7 = 0.00186$	$x_7 =$ VFA	0.000895	0.040
$\beta_8 = -0.304$	$x_8 =$ HRT	0.0276	0.000

[a]*p-values* < 0.05 were considered significant.

Table 4: Performance statistics of the Multiple Nonlinear Regression models

Rank	Structure of Equations	NNI	R^2 (%)	Adj-R^2 (%)	SEE	SSR	DWS	RA
Biogas yield (Output 1)								
1	$\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_5X_5+\beta_6X_6+\beta\ X_7+\beta\ X_8$		93.30	92.03	0.82	69.7	2.05	8.45×10^{-15}
2	$Exp(\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_5X_5+\beta_6X_6+\beta\ X_7+\beta\ X_8)$	89	88.40	88.06	2.17	491.9		0.26
Methane yield (Output 2)								
1	$\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_5X_5+\beta_6X_6+\beta\ X_7+\beta\ X_8$		91.08	90.39	1.10	125.2	1.83	1.03×10^{-15}
2	Exp $(\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_5X_5+\beta_6X_6+\beta\ X_7+\beta\ X_8)$	92	86.60	85.70	1.58	260.6		0.12

SEE standard error of the estimate; SSR sum of squared residuals; R^2 coefficient of multiple determination; Adj-R^2 adjusted coefficient of multiple determination; NNI number of nonlinear iterations; DWS Durbin–Watson statistics; SR sum of residuals; RA residuals average

Table 5: Summary of ANN and MnLR models performance

Performance criterion indicators[a]	Testing data set			
	Biogas yield		Methane yield	
	BP-ANN	MnLR	BP-ANN	MnLR
Multiple coefficient of determination (R^2):	98.72 %	95.31%	97.93%	92.62%
Index of agreement (*IA*):	0.9941	0.9723	0.9806	0.9611
Fractional Variance (*FV*):	0.0075	0.0821	0.00284	0.073

Figures

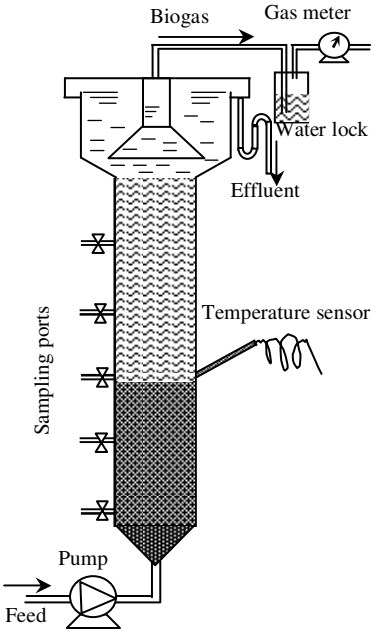

Fig. 1 Schematic diagram of the upflow anaerobic sludge bed reactor

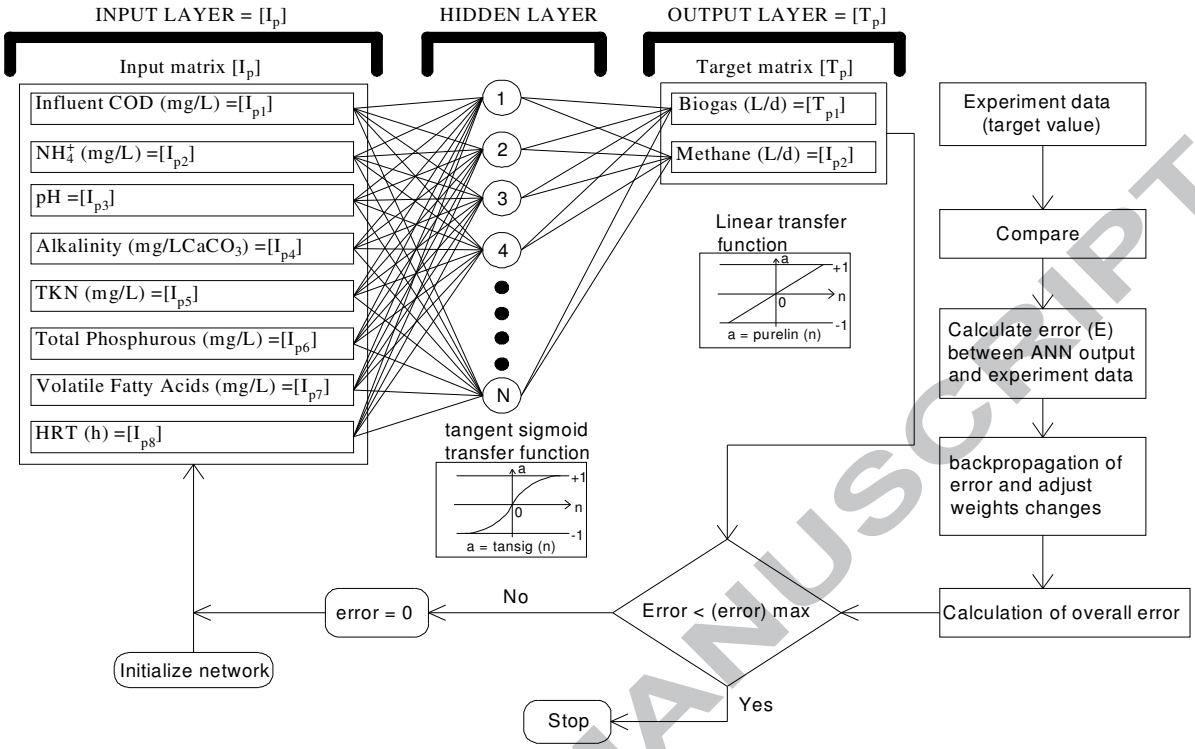

Fig.2: Schematic flowchart of the proposed feedforward BP-ANN methodology

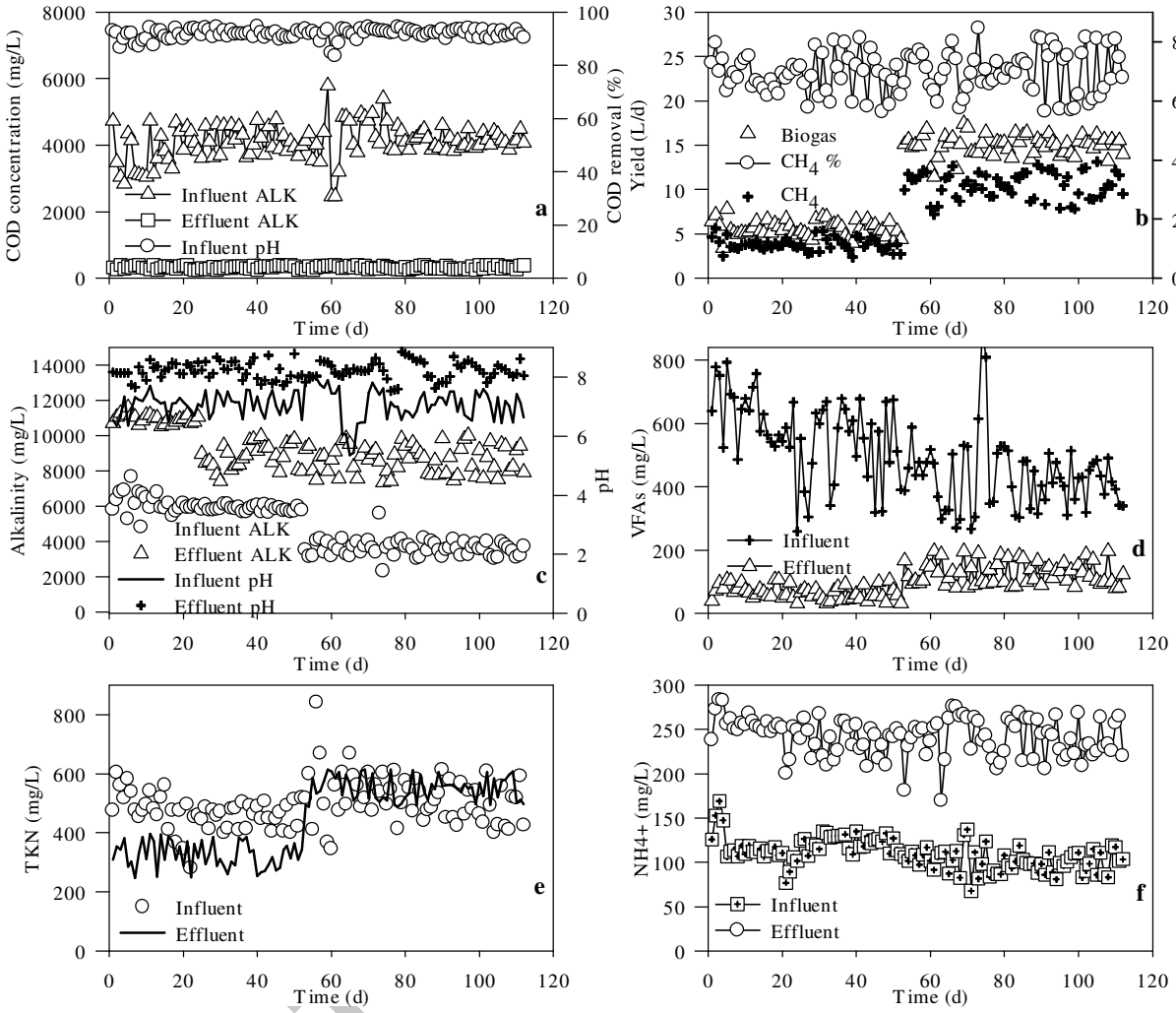

Fig.3: (a) COD and COD removal; (b) biogas and methane; (c) ALK and pH; (d) influent and effluent VFAs in total; (e) influent and effluent TKN; (f) influent and effluent ammonium

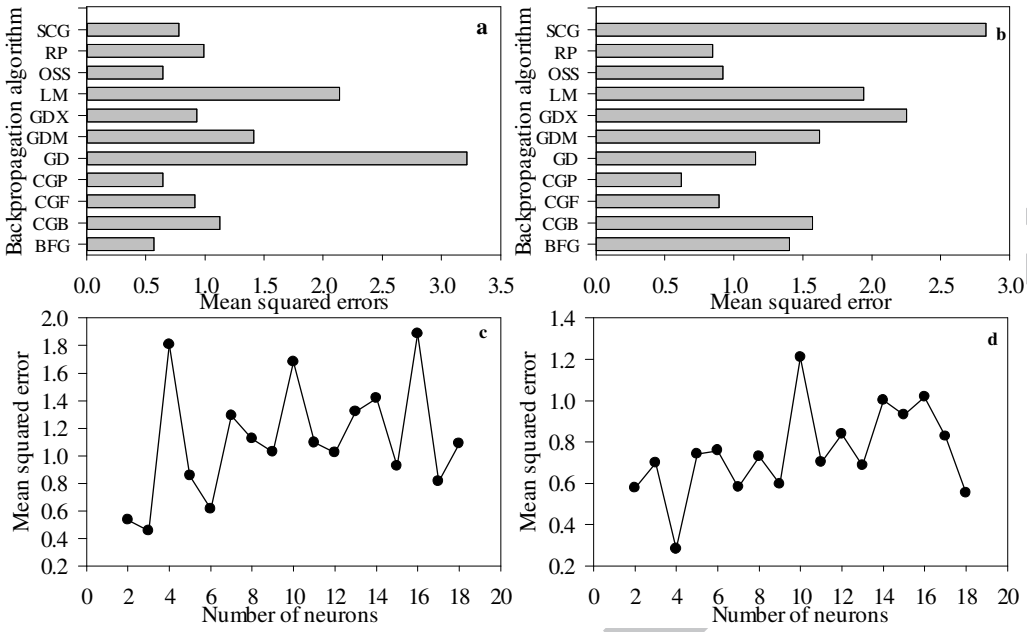

Fig 4. Comparison between training algorithms based on MSE, (a) biogas yield, (b) Methane yield; Optimum number of neurons required at hidden layer based on MSE; (**c**) Polak–Ribiere conjugate gradient backpropagation algorithm for methane yield (**d**) BFGS Quasi–Newton backpropagation algorithm for biogas yield.

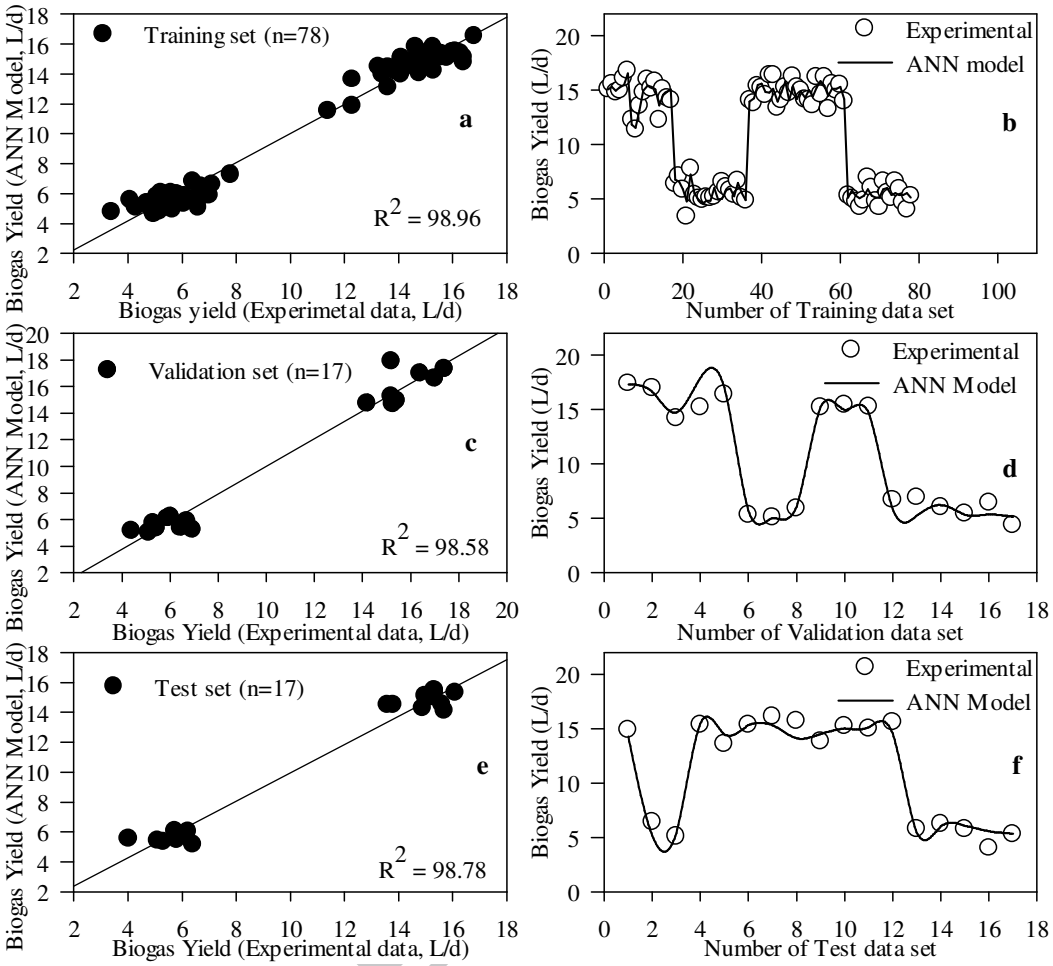

Fig.5: Correlations (a,c,e) and visual agreements between biogas experimental data and ANN predictions(b,d,f)

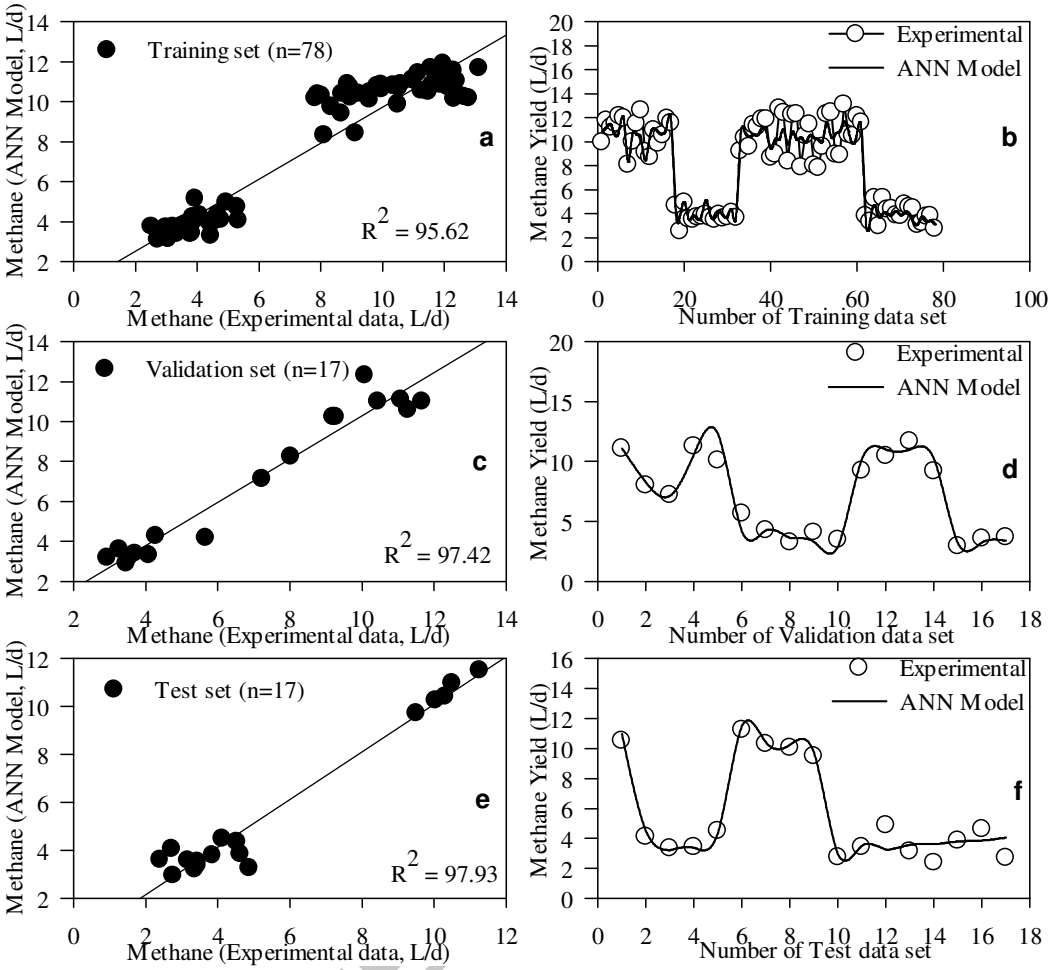

Fig.6: Correlations (a,c,e) and visual agreements between methane experimental data and ANN predictions(b,d,f)

Highlights:

- Estimation of CH_4 and biogas yield from a UASB with BP-ANN and MnLR.
- Evaluation and selection of optimum algorithm from eleven training algorithms.
- Optimization of anaerobic parameters to identify their effects on methanation.
- BP-ANN models predictions were more reliable compared to MnLR.