Feedforward neural network model estimating pollutant removal process within mesophilic upflow anaerobic sludge blanket bioreactor treating industrial starch processing wastewater

Philip Antwi, Jianzheng Li, Jia Meng, Kaiwen Deng, Frank Koblah Quashie, Jiuling Li, Portia Opoku Boadi

PII: S0960-8524(18)30264-5
DOI: https://doi.org/10.1016/j.biortech.2018.02.071
Reference: BITE 19580

To appear in: *Bioresource Technology*

Received Date: 28 November 2017
Revised Date: 10 February 2018
Accepted Date: 14 February 2018

Please cite this article as: Antwi, P., Li, J., Meng, J., Deng, K., Koblah Quashie, F., Li, J., Opoku Boadi, P., Feedforward neural network model estimating pollutant removal process within mesophilic upflow anaerobic sludge blanket bioreactor treating industrial starch processing wastewater, *Bioresource Technology* (2018), doi: https://doi.org/10.1016/j.biortech.2018.02.071

Feedforward neural network model estimating pollutant removal process within mesophilic upflow anaerobic sludge blanket bioreactor treating industrial starch processing wastewater

Philip Antwi[a, b, c] , Jianzheng Li[c,*], Jia Meng[c], Kaiwen Deng[c], Frank Koblah Quashie[c], Jiuling Li[d], Portia Opoku Boadi[e]

a. Jiangxi Key Laboratory of Mining & Metallurgy Environmental Pollution Control, School of Resources and Environmental Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, P.R. China
b. Department for Management of Science and Technology Development, Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City, Vietnam
c. State Key Laboratory of Urban Water Resource and Environment, School of Environmental, Harbin Institute of Technology, 73 Huanghe Road, Harbin 150090, P.R. China
d. Advanced Water Management Centre, Gehrmann Building, Research Road, The University of Queensland, St Lucia, Brisbane, QLD 4072, Australia
e. School of Management, Harbin Institute of Technology, 92 West Dazhi Street , Nan Gang District, Harbin 150001 , P.R. China

* Corresponding authors. E-mail: ljz6677@163.com (J. Li); philip.antwi@tdt.edu.vn (P. Antwi).

Authors:

Philip Antwi: Tel.: +86 451 86283761; E-mail: philip.antwi@tdt.edu.vn

Jianzheng Li: Tel.: +86 451 86283761; E-mail: ljz6677@163.com

Jia Meng: Tel.: +86 451 86283761; E-mail: mengjia2726688@126.com

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

Frank Koblah Quashie: Tel.: +86 451 86282008; E-mail: ololofrank@yahoo.com

Jiuling Li: Tel.: + +61 733469042; Email: jiuling.li@awmc.uq.edu.au

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

Abstract

In this study, three-layered feedforward-backpropagation artificial neural network (BPANN) model was developed and employed to evaluate COD removal in upflow anaerobic sludge blanket (UASB) reactor treating industrial starch processing wastewater. At the end of UASB operation, microbial community characterization revealed satisfactory composition of microbes whereas morphology depicted rod-shaped archaea. pH, COD, NH_4^+, VFA, OLR and biogas yield were selected by principal component analysis and used as input variables. Whilst tangent sigmoid function (*tansig*) and linear function (*purelin*) were assigned as activation functions at the hidden-layer and output-layer, respectively, optimum BPANN architecture was achieved with Levenberg-Marquardt algorithm (*trainlm*) after eleven training algorithms had been tested. Based on performance indicators such as mean squared errors, fractional variance, index of agreement and coefficient of determination (R^2), the BPANN model demonstrated significant performance with R^2 reaching 87%. The study revealed that, control and optimization of an anaerobic digestion process with BPANN model was feasible.

Keywords: Industrial starch processing wastewater; Upflow anaerobic sludge blanket; Feedforward backpropagation artificial neural network; Microbial community characterization; Anaerobic digestion

1 Introduction

Agricultural and industrial and activities in recent years have been the main source of pollutants (organic and inorganic) which find its way into waterbodies that subsequently lead to water pollution (Schweitzer & Noblet, 2018; Wu et al., 2017). For instance, potatoes cultivation, production and processing have increased exponentially in recent years. However, processing potato into starch and related products mostly yields huge volumes of wastewater (Antwi et al., 2017a; Przetaczek-Rożnowska, 2017) which is characterized by high level of organic pollutants (Table 1). Such wastewater emanating from potatoes starch processing may contribute to water pollution particularly when discharged untreated. That notwithstanding, anaerobic digestion (AD) process, a potential feasible process proposed for solving waste problems has successfully been employed over the years to treat wastewater that emerges from sectors including domestic, industrial and agriculture fields (Barua & Dhar, 2017). Besides treatment, AD has also been successful in: (1) bioenergy generation; (2) production of stabilized materials used as organic composting; and (3) destruction of pathogens (Lloret et al., 2013; Lizama et al., 2017). Upflow anaerobic sludge blanket (UASB), a high-rate AD reactor is one the most competitive and preferred AD technology that has the tendency to treat industrial effluents.

The efficacy of AD process however depends mainly on complex biological activities of functional population. Biological activities consequently turn out to be challenging when enhancement of the AD digestion is required through process optimization and control. However, mathematical modelling (process simulations and predictions) has been suggested as a means to optimize and control the performance of an AD process (Revilla et al., 2016) could learn the complex and non-linear relationships existing in an AD process. Developed

math models such as neural networks were successful when employed to capture the non-linear relationships existing in AD process (Podder & Majumder, 2016; Antwi et al., 2017c). ANN have been trained to perform complex functions in various fields of application including pattern recognition, identification, classification, speech, vision, and control systems. Other models such as anaerobic digester model 1 (ADM1 of IWA) which was initially developed purposely for continuous stirred tank reactor (CSTR) has been modified severally by many other researchers besides the main developers. So far, the ADM1 have been a successful model when employed in expanded granular sludge blanket (EGSB) or UASB reactor. On the contrary, ADM1 requires kinetic parameters to achieve optimality. Thus, the model requires about 26 or more dynamic state variables and many other parameters as well as all related processes under practical conditions. Acquiring these kinetic parameters is very challenging as extensive, laborious and relatively expensive experiments are needed to be conducted (Lee et al., 2016; Xie et al., 2016). The ADM1 have been simplified in recent years by reducing state variables and parameters (López & Borzacconi, 2011) where the identification of parameter is more straightforward. But the simplified model: (1) could not depict the complexity of anaerobic process; and (2) cannot be applied to other reactors or conditions.

Compared with ADM1, artificial neural network (ANN) modelling could simulate and predict complex relationship between independent and dependent variables associated with AD process with high efficiency without requiring detailed mechanisms of anaerobic process. Thus, provided model building parameters such as number of layers, type training function, number of neurons assigned in the hidden layer, initial adaptive value, minimum gradient, maximum fail, maximum number of epochs, initial weights/biases and training

goal are optimized to achieve a robust ANN architecture (Rosales-Colunga et al., 2010; Yusof et al., 2014). Hu Yi-Fan and coworkers developed a ANN model to predict the performance of an expanded granular sludge bed (EGSB) reactor and their study indicated that the proposed ANN model exhibited superior predictive accuracy for the forecast of chemical oxygen demand (COD) removal performance by EGSB system (Yi-Fan et al., 2017). In another work, a new configuration of an electrically-enhanced membrane bioreactor was introduced to treat medium strength wastewater to reduce wastewater contaminant concentrations. Artificial neural networks (ANNs) based ensemble model was used to model the experimental findings of COD, PO_4^{3-}-P and NH_4^+-N removal given the initial mixed liquor compositions. Comparison between the model results and experimental data set gave high correlation coefficients for COD (r = 0.9942), PO_4^{3-}P (r = 0.9998) and NH_4^+-N (r = 0.9955) (Giwa et al., 2016). Although in relevant literature, significant amount of experimental and numerical analysis has been conducted on pollutant removal in wastewater treatment, an ANN-based prediction model presented in this study to evaluate COD removal efficiency within an UASB treating industrial starch processing wastewater has not been developed so far.

This study seeks to describe the application of artificial neural networks for modeling wastewater treatment processes within an UASB. In this study, the feasibility of an ANN model to estimate, predict and simulate experimental results obtained from a mesophilic UASB treating industrial starch wastewater was investigated. Herein, a novel three layered feedforward backpropagation artificial neural network (BPANN) model (6:N_H:1) was developed with anaerobic process parameters for the estimation of chemical oxygen demand (COD) removal in a mesophilic UASB treating potato starch processing wastewater

(PSPW). Six anaerobic process parameters such as influent COD, Ammonium (NH_4^+-N), influent pH, organic loading rate (OLR), effluent volatile fatty acid (VFA) and biogas yield were selected based on principal component analysis and used as input variables for the model development. Furthermore, the network architecture parameters including number of neurons, initial adaptive value and initial value of weights/biases were initially optimized using response surface methodology in order to achieve optimum performance of the proposed model. Microbial communities and sludge morphology within activated sludge obtained after startup period and end of UASB operation were also probed to comprehend the performance of the UASB in COD removal.

2 Materials and methods

2.1 Wastewater characteristics and experimental setup

PSPW was collected from a local starch factory and stored under 4^{o}C in a deep freeze refrigeration unit (XINGX BD/BC-142CH, Guangdong Xingxing Refrigeration Equipment Co., Ltd., China). Characteristics associated with the raw wastewater after characterization is presented in Table 1. The UASB reactor with a gas-liquid-solid separator was constructed with Plexiglas material (Fig.1A). Height, working volume and total volume of the reactor was 120 cm, 7 L and 8.8 L, respectively, with 5 sampling ports spaced at 25 cm interval. Activated sludge (mixed liquor suspended solid of 11.5 g/L and mixed liquor volatile suspended solid of 5.6 g/L) was collected from a local treatment plant and used to inoculate the UASB (Antwi et al., 2017d). Peristaltic pump (BT10032J, Langer Instruments, United Kingdom) was used to feed wastewater to the UASB. Mesophilic condition (35±1°C) of the reactor was maintained with a thermistor and a controller (Fig.1A). pH of the raw feed was adjusted to about 6±1 with sodium bicarbonate ($NaHCO_3$) prior to feeding.

HRT of 36 h followed by a 24 h was implemented at the startup phase. After the startup period had obtained stability (83 days), the main operation of the UASB was initiated under 9 different stages (Table 1). Each phase had a unique combination of HRT (stage 1, 4 and 7: 72 h; stage 2, 5 and 8: 48 h and; stage 3, 6 and 9: 36 h) and organic loading rates (OLR) ranging from 2.7 – 3.75, 5.24 – 7.16 and 6.7 – 13.2 kgCOD/m^3.d under phases I, II and III (Table 1). At each stage, a steady state in the performance was obtained to warrant another set of HRT and OLR to be introduced. Biogas yield from the UASB was collected by the gas-solid-liquid separator and measured daily using wet gas meter (LML31, Changchun Filter Co., Ltd., China).

2.2 Analytical methods

All physical and chemical analysis presented in Table 1 were conducted in accordance with Standard Methods for the Examination of Water and Wastewater (APHA, 2012). NH_4^+-N was determined by Nessler method and the final solution was measured with a spectrophotometer (UV-1800 UV-VIS, Shimadzu Corporation, Japan) at a wavelength of 420nm. TP was determined by ascorbic acid method as prescribed in the manual of the Standard Methods for the Examination of Water and Wastewater. pH was determined with a pH meter (DELTA 320, Mettler Toledo, USA). Carbohydrate was measured by the phenol-sulfuric acid method using glucose standard. Protein was analyzed by Lowry's method using bovine serum albumin as standard (Antwi et al., 2017a). Volatile fatty acids (VFAs) were measured by a chromatograph (SP6890, Shandong Lunan Instrument Factory, China) equipped with an RTX-Stabilwax glass column (30 m×0.32 mm×1 μm) and a flame ionization detector. Detailed protocol is described in our previous report (Antwi et al., 2017a; Liu et al., 2015; Shi et al., 2016). Biogas fractions were determined as done

previously by another gas chromatograph (SP-6800A, Shandong Lunan Instrument Factory, China) equipped with a thermal conductivity detector (TCD) and a 2 m stainless column packed with Porapak Q (60/80 mesh) (Antwi et al., 2017b). Temperatures of the gas chromatograph's injector, column and the TCD were 80°C, 50°C and 80°C, respectively.

2.3 Selection of input and output variables for BPANN model

Anaerobic parameters to be used as input variables were first filtered by examining their contribution on COD removal (output variable). COD removal was set at a target of 80-95% against the input variables. Input variables above the targeted limits (80-95%) were easily identified with scatter plots. Variables which fitted better with the targeted COD removal were considered for further scrutiny by principal component analyses (PCA) to further reduce number of variables. PCA was conducted within the MATLAB workspace (Matrix Laboratory R2014a, version 8.3 by MathWorks, Inc., USA) and principal components that contributed less than 0.1% to the variation in the data set were eliminated (Yetilmezsoy & Sapci-Zengin, 2009). PCA revealed pH, COD, ammonium (NH_4^+), OLR, volatile fatty acids (VFAs) and biogas yield as variables that had optimum effect on the targeted COD removal (Table 2). The input and target data set given in matrices [I_P] and [T_P] (Table 2) were normalized using *prestd* algorithm code. 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 were given clear definitions as *meanIp*, *meanTp*, *stdIp*, *stdTp*, *Iptrans* and *transMat*, respectively, before commencement of the training (Antwi et al., 2017b).

2.4 Description of the artificial neural network architecture

Feed forward backpropagation (BP) algorithm was employed in the development of the ANN model on the MATLAB platform. The ANN model with an input vector (6×218) and target vector (1×218) had its architecture comprised of neurons ordered in 3 layers (input, hidden and output) as illustrated in Fig.1B. Whilst the input and output neurons represented the independent variables and dependent variables, respectively, the hidden layer was tasked to transformed the input information (Beltramo et al., 2016). The BP learning rule defined a method to adjust the weights of the networks (Cheng et al., 2016). The structure of the network was designed to render hidden neurons outputs to be used as an input to the output neuron after hidden layers output had undergone transformation in the process. Output of the BPANN was estimated with Eq.1. Tangent sigmoid transfer function (*tansig*) (Eq.2) and linear transfer function (*purelin*) (Eq.3) were introduced at the hidden and output layer, respectively.

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

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

$$f(x) = x \quad (3)$$

Where: $COD\ rem$, output of BPANN (COD removal efficiency); WH_{ij}, weight of the link between the i^{th} input and the j^{th} hidden neuron; m, number of input neurons; WO_j, weight of the link between the j^{th} hidden neuron and the output neuron; f_h, hidden neuron activation function; f_o, output neuron activation function; b_j, bias of the j^{th} hidden neurons; b_o, bias of the output neuron; X_{it}, input variable; HN, number of hidden neurons; and x, vector of inputs.

Three functions viz., *dividerand*, *divideblock* and *divideint* established on the MATLAB platform were tested to unraveled at the most efficient method for dividing dataset. *Dividerand* was employed to randomly divide the entire data set into three subsets viz., training, validation and testing data set in this study based on its low MSE observed. Out of 218 data set points, 15% each, comprising 33 data points were selected to represent the validation and testing subsets, respectively. On the other hand, 70% (152 data points) was assigned for the training subset. The efficacy of the BPANN learning was validated with mean square error (MSE) (Eq.4). Besides MSE, index of agreement (IA) and fractional variance (FV) as given in Eq.5 and Eq.6 was also used for further validation.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(T_i - A_i)^2 \tag{4}$$

$$IA = 1 - \frac{\sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}(|P_i - O_m| + |O_i - O_m|)^2} \tag{5}$$

$$FV = \frac{2(\delta_o - \delta_p)}{(\delta_o + \delta_p)} \tag{6}$$

Where: N, number of data point; T_i, network predicted value at the i^{th} data; A_i, experimental value at the i^{th} data and i is an index of the data; O, P, δ and m indicates experimental data, predicted values, standard deviation and arithmetic mean of the observed data points, respectively.

2.5 Optimization of the BPANN Training algorithm

A benchmark comparison was carried out to facilitate the selection of the optimum neural networks in the ANN modeling process. 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.

2.6 Optimization of BPANN model topology for optimum performance

After the optimum training algorithm was identified, other model parameters including initial value of weights and biases, initial adaptive value and number of neurons in the hidden layer were also optimized by response surface methodology (RSM) to help achieve a relatively optimal performance of the BPANN model, thus [X1] number of neurons in the hidden layer, [X2] initial adaptive value and [X3] initial weights and biases were investigated for optimum values. First, Box Behnken Design (BBD) methodology proposed a set of 15 experimental runs to be conducted (Table 3). MSE obtained from the experimentation was used as response [Y] for the optimization process. Further evaluation of the significance of the model building parameters [X1], [X2] and [X3] was examined with multiple nonlinear regression models (MnLRM) by residual analysis with MINITAB (version 17). The general form of the MnLRM used is given in Eq.7.

$$COD\ rem\ (\%) = b_0 + \sum_{i=1}^{3} b_i X_i + \sum_{i=1}^{3} b_{ii} X_i^2 + \sum_{i=1}^{2} \quad \sum_{j=i+1}^{3} b_{ij} X_i X_j + \varepsilon \qquad (7)$$

where x_1, x_2, and x_k represented terms for quantitative predictors, b_0, b_i, b_{ii}, b_{ij} represents constant, linear, quadratic and interaction coefficients, respectively, and ε is random error.

Statistical assumptions such as linearity, independence among errors, non-multicollinearity, homoscedasticity, non-autocorrelation and normal distribution of errors

were considered during regression analyses (Wold et al., 2001). The efficiency of MnLR model was validated with coefficient of determination (R^2) (Eq.8), adjusted coefficient of determination (Adj-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), variance inflation factor (VIF) (Eq.13), Durbin-Watson statistics (d)(Antwi et al., 2017b) (Eq.14) and p-value (Yetilmezsoy et al., 2013).

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

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

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

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

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

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

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

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

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; $e_i = y_i - \hat{y_i}$, and y_i and $\hat{y_i}$ 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.

2.7 Microbial community analysis by high-throughput sequencing

2.7.1 Sampling, DNA extraction, PCR amplification and analysis of sequences

To establish the performance of the reactor, sludge samples were taken from the UASB at the following periods: (1) end of the startup phase (83 days) of the UASB; and (2) end of main operation (218 days) of the UASB and their microbial communities and morphology were characterized and compared. Sampled activated sludge was stored at -20ºC until DNA extraction was due. Extraction and purification of the total DNA of the samples were carried out with bacteria DNA Isolation Kit (Power Soil DNA Isolation Kit-MOBIO Laboratories, Inc., Carlsbad, CA) in accordance with the manufacturer's manual. Agarose gel electrophoresis was conducted to check the integrity of extracted DNA (Antwi et al., 2017a). Qubit 2.0 DNA kit (Qubit ssDNA Assay Kit, Life Technologies) was used to quantify Genomic DNA for the PCR reaction. Details of PCR amplification protocol is reported in our previous research (Antwi et al., 2017b). Amplicons were finally sequenced on an Illumina Miseq sequencing platform (Sangon Biotech Shanghai Co. Ltd, China). After sequencing, statistical analysis was conducted on the obtained sequences as reported in our previous work (Antwi et al., 2017b).

2.8 Characterization of sludge morphology

Morphology of sampled sludge was examined with scanning electron microscopy (SEM). Samples were taken in 10 mL aliquots, washed three times with ultra-pure water to remove impurities and then fixed overnight at 4ºC with 2.5% vol/vol glutaraldehyde (pH of

6.8). Rinsing was performed again after which dehydration was conducted with ethanol solutions at 25, 50, 70, 80, 90 and 100% (Wu et al., 2010; Ding et al., 2015) and dried till full dehydration. Morphology of sludge was finally observed with SEM (FEI Quanta-200).

3 Results and discussion

3.1 Pollutant removal within the UASB

The employed UASB for treating the industrial starch processing wastewater was operated under different organic loading rates (Table 1) and the performance in terms of COD removal rate was investigated. Organic loading rates (OLR) ranging between 1.5–4.23 kgCOD/m^3·d was introduced at the start-up period of the UASB operation. Highest COD removal observed within the startup period was about 96%. The observed pH determined in the effluent reached an average of 8.08 suggesting reactor stability. No major washout or inhibition phenomenon was noticed. Sludge granules were observable on the 63rd day after startup indicating the effectiveness of the proposed HRT and upflow velocity. The stability could also be ascribed to the effective acclimatization of the sludge irrespective of culturing conditions (Luo et al., 2016).

Besides the startup period, the UASB was further operated in accordance with the treatment scheme presented in Table 1. As illustrated in Fig.2, COD removal encountered transient decline during the early periods of each stage (16th, 35th, 57th, 71st, 100th, 132nd, 159th, 188th day). Thus, anytime OLR was increased, COD removal declined to some threshold but subsequently regained higher efficiencies after time had elapsed for a while. The observed reduction in COD removal could be ascribed to possible shock loading effect on the functional population. On the other hand, COD removal was stable when employed OLR was low particularly during the startup phase. For the avoidance of such shock loading

effect and effective acclimatization of the activated anaerobic sludge, a higher HRT of 72 h accompanied with an OLR of 2.7 kg COD/m^3·d was first introduced in stage (1). At the commencement of stages (2) and (3), thus, when OLR was elevated from 2.7 kg COD/m^3·d in stage (1) to 3.73 kgCOD/m^3·d and 5.02 kgCOD/m^3·d, respectively, COD removal declined to about 86% and 74%, respectively (Fig.2). Phase (I) as presented in Table 1 lasted for about 56 days and COD removal achieved could reach about 97%. Similarly in stages (5), (6), (8) and (9), an abrupt decline in COD removal was also observable when OLR was further elevated to a range of 5.27-7.16 kgCOD/m^3·d and 9.88-13.27 kgCOD/m^3·d (Table 1). Notably, COD removal declined to about 42.1, 58.1, 64.5 and 70.1%, in in stages 5, 6, 8 and 9, respectively. However, the UASB recovered from the suspected shock loading effects and subsequently elevated the COD removal to about 93.69, 95.6, 95.1 and 92.0%, respectively.

The highest COD removal rate was recorded in stage 4 (97.7%) when OLR of 3.65 kg COD/m^3·d was introduced. On the contrary, stage 9 with HRT of 36 h and OLR elevated to 13.27 kg COD/m^3·d recorded the lowest COD removal of 92.0% (Fig.2). The relatively low performance at higher OLR could also be attributed to the high starch or polysaccharide content in the PSPW. Lu et al., had reported that, starch becomes sticky upon contact with heated water and that may bind to the surfaces of anaerobic granules and lower the mass transfer rate (Lu et al., 2015). Therefore, higher OLRs was suspected to have high proportion of starch which bound sludge particle together to lower mass transfer.

3.2 Characterization of sludge and microbial communities at the end of startup and main operation

3.2.1 Sludge morphology as revealed by scanning electron microscopy

The morphology of the anaerobic sludge sampled at the end of startup and main operation of the UASB was characterized and compared using scanning electron microscopy (SEM). Granular sludge was sampled at the middle belt of the UASB after startup (83 days) and main operation (218 days) and their respective micrographs are presented in the supplementary information. The average size of sludge granules obtained after startup ranged 1-3 mm whereas that observed after the main operation of the UASB ranged between 2-5 mm. Granules observed in the sludge obtained after startup were predominantly irregular in shape with some few having an elliptical shape. On the other hand, sludge at the end of the main operation was predominantly elliptical in shape with quite a few having irregular shape. However, granules in both samples revealed cavities on their surfaces indicating an escape route for biogas produced.

Among the sludge samples, randomly intertwined cell morpho-types were observable, except their shape and structure that varied from the startup period through to the end of the UASB operation. Again, the morphology as revealed by the micrographs indicated the presence of *Methanosaeta*-like cells, *Methanosarcina*-like cells, rods and cocci colonies at both periods as observed and indicated in other reports (Lu et al., 2016). Acetoclastic methanogen such as *Methanosaeta* are often identified by its bamboo-shaped filament, fluorescence-emitting, rod shaped shell (0.6-0.8 × 2.0-3.5μm) with flat ends, and an ultrastructure with outer and inner cell walls (Subramanyam & Mishra, 2013).

Notably, the predominance of congeries of rod-shaped archaebacterial and cocci-shaped archaea were more visible in the sludge obtained at the end of the UASB operation as opposed to that of the startup period. Similar to this study, rod-shaped archaebacterial and

cocci-shaped archaea belonging to *Methanothrix* and *Methanosarcina* were reported by Nizami and Murphy, respectively (Nizami & Murphy, 2011).

3.2.2 Microbial communities as revealed by high throughput sequencing

Microbial community analysis was conducted to elucidate their functionalities to the anaerobic digestion process within the UASB. Three genera including *Methanosaeta*, *Methanosarcina* and *Methanobacterium* belonging to phylum Euryarchaeota was the most noticeable genera within the archaeal kingdom (Fig 3). It has been established that, *Methanosaeta*, *Methanosarcina* and *Methanobacterium* with acetoclastic- hydrogenotrophic functionalities is suggested to have played a major role in metabolizing VFAs produced during acetogenesis phase in the UASB treating PSPW (Kundu et al., 2012). No accumulation of VFA was observed suggesting the ratio of abundance of acetoclastic-hydrogenotrophic and methanogens were in good agreement to produce and consume VFAs to balance pH and acidity in the reactor.

At the end of the startup period (Fig 3), genus *Methanosaeta* (10.06%) was dominant followed by *unclassified_Anaerolineaceae* (8.76%), *Incertae_Sedis* (6.41%), *unclassified_Porphyromonadaceae* (6.12%), *Longilinea* (5.42%) and *Syntrophomonas* (4.98%). Compared with the startup sludge, growth was observed in microbial communities in sludge sampled at the end of UASB operation. Genus *Methanosaeta* still dominated the population with 18.19% followed by *unclassified_Anaerolineaceae* (15.31%), *Longilinea* (9.26%), *Acinetobacter* (6.88%), *Incertae_Sedis* (6.72%), *Syntrophomonas* (5.46%), *unclassified_Veillonellaceae* (5.3%) and *unclassified_Porphyromonadaceae* (5.09%) (Fig 3). It has been reported that, genera affiliated to phylum Firmicutes such as *Syntrophomonas*, *unclassified_Veillonellaceae*, *unclassified_Christensenellaceae*,

unclassified_Porphyromonadaceae and *vadinBC27_wastewater-sludge_group* are some of the main bacteria with protein and/or amino acid degradation functionalities (Forchhammer, 2007). Therefore, these genera observed in this study may have contributed to the successful degradation of the protein available in the PSPW.

Again, the presence of *Longilinea*, *unclassified_Anaerolineaceae*, *Leptolinea*, *Anaerolinea* and *Ornatilinea* affiliated to Chloroflexi could be attributed to cellulose and glucose degradation in the PSPW since it has hydrolytic fermentative functionalities (Hagen et al., 2014). The synergistic coexistence of Chloroflexi and Firmicutes is suggested to have mainly contributed to the hydrolysis and acidogenesis phase of the AD process. *Acinetobacter* and *Incertae_Sedis* belonging to Proteobacteria is also believed to have performed heterotrophic functionalities (Fig 3). The overall observation suggested that, sludge sampled at the end of the UASB operation was composed with microbes that contributed effectively to hydrolysis, acidogenesis, acetogenesis and methanogenesis as opposed to that sampled at the end of the startup.

3.3 Optimization of the BPANN architecture

3.3.1 Selection of backpropagation algorithm by benchmark comparison

The effectiveness, speed and accuracy of a BPANN architecture depend on factors such as selection training algorithm, characterization of data set, number of neurons specified in the network, initial adaptive value and initial values of weights and biases (Hu et al., 2017). A benchmark comparison was conducted based on the mean squared errors (MSE) to select optimum training algorithm (Table 4). Eleven training algorithms were tested for efficacy and Levenberg-Marquardt (*trainlm*) training algorithm manifested as best so far as COD removal prediction and simulation are concern (Table 4). Compared with the other 10

algorithms, smaller MSE of 0.131 was obtained with the Levenberg-Marquardt (*trainlm*) algorithm whereas the worst performed algorithm was the batch gradient descent (*traingd*). The poor performance associated with the other 10 algorithms may be ascribed to their inability to explain variation that exists in the data set (Antwi et al., 2017c). Comparison made between predicted and experimental output values show that ANN is a successful technique to predict COD removal from an UASB.

3.3.2 Optimizing BPANN model building parameters

The number of neurons to be assigned in the hidden layer required optimization based on MSE recorded when different nodes were assigned (Cheng et al., 2016). As a result, neurons in the hidden layer (X1), initial adaptive value (X2) and initial weights and biases (X3) were optimized by RSM. Multiple nonlinear regressions (MnLR) modelling was conducted with MSE (Table 3) as response (Y) to evaluate the precision and significance of the model and variables, respectively. The MnLR equation (Eq.16) in uncoded units was fitted as a second-order response.

$$Y = 0.03533 - 0.001989X1 - 0.001995X2 - 0.01711X3 + 0.000082X1^2 + 0.000300X2^2 + 0.00821X3^2 + 0.000005X1X2 + 0.000074X1X3 - 0.000304X2X3 \quad (16)$$

The operations (negative and positive) associated to the coefficient indicated the impact of each factor on the response variable (Y = MSE). ANOVA revealed that, P-value was < 0.05 and R^2 was 0.972 indicating high significance of the model. The percentage contributions (PC) for individual term estimated based on sum of squares revealed variable $X1^2$ showing the highest level of significance with a contribution of 53.21% (Table 5).

Optimized values estimated by response optimizer indicated that, number of neurons, initial adaptive value and initial value of weights/biases were 11.4 approximated at 12, 3.75

approximated at 4 and 1.05, respectively. The experimental value of MSE was about 0.0115, whilst that obtained as predicted value under the optimized conditions was 0.011. With number of neurons set at 3, 14 and 25, initial adaptive value set at 0.5, 4 and 7.5, and initial values of weight/biases set at 0.2, 1 and 1.8, 3D response surface plots were also employed to highlight how input parameters interacted and influenced MSE (Fig.4A, B and C). The number of neurons in hidden layer demonstrated quadratic effects on the response. Notably, when the number of neurons assigned at the hidden layer was < 11 or > 12, MSE increased significantly to about 0.02. The interactions observed among the selected independent variables were also found to be insignificant herein.

3.4 Performance of the BPANN model during COD removal simulation

To avoid possible overfitting as a result of large size of nodes in a hidden layer (Srivastava et al., 2014), early stopping methodology was employed in this study to evaluate underfitting and overfitting and thereby resolving them accordingly. It was observed that, the estimated errors associated with the training set and validation set decreased at the initial training phase. However, error of the validation set increased as the network began to overfit the data. In this regard, whilst the validation error increased with a specified number of iterations, the training was pulsed to initiate the weights and biases at the minimum of the validation error to be returned. As a result, overfitting or underfitting tendencies were resolved. Again, BPANN predictions had negative operations irrespective of the fact that a linear transfer function (*purelin*) was engaged at the output layer.

As illustrated in Fig.5, the corresponding visual agreement and correlations between the experimental data and the BPANN output was presented. The proposed BP-ANN model demonstrated very satisfactory performance during COD removal predictions and

simulation. The performance was demonstrated with the correlation plots (Fig.5B, Fig.5D, and Fig.5F) where coefficient of determination (R^2) achieved in the training, validation and testing subset reached 0.86, 0.93 and 0.83, respectively, indicating the efficacy of the BPANN capable of explaining over 93% of the variation existed in the entire COD removal data set. This phenomenal performance could be attributed to the ability of the BPANN model to capture complex behavior or trend existed among the variables obtained from the anaerobic digestion process (Giwa et al., 2016). Again, the performance efficacy of the BPANN could be ascribed to the advantage of the ANNs capability in explaining complex interactions between inputs and output parameters.

The overall performance of the BPANN was further demonstrated with another visual agreements and correlations as illustrated in Fig.5G and Fig.5H, respectively. Notably, the overall data set (experimental) could agree well with the predicted data set (Fig.5G) with an R^2 of 0.87 as revealed with the correlation plot in Fig.5H. Notably, only 13% of the total variation existing in COD removal data sets was not explained by the BPANN model.

Besides the coefficient of determination (R^2), the prediction accuracy of the BPANN models was further validated with index of agreement (IA) and the fractional variance (FV) (Table 6). The index of agreement (IA) obtained with the BPANN model was 0.9117, suggesting BPANN model could make reliable prediction. As reported earlier, fractional variance (FV) will yield 1 when explanatory variables (x) reveal nothing about the dependent variable (Y). On the other hand, FV will be zero (0) when explanatory variables (x) are able to make perfect predictions of variable Y (Antwi et al., 2017c). In this study therefore, relatively lower FVs were associated with the BPANN model. At COD removal predictions, the evaluated FV yielded 0.00113 confirming satisfactory efficiency of the

BPANN model. 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 during COD removal predictions and simulations.

4 Conclusion

Three-layered-BPANN model was developed to estimate COD removal in an UASB treating PSPW. Microbial community and sludge characterization revealed hydrolytic bacteria and methanogens in good ratio to achieve optimum COD removal during UASB operation. Levenberg-Marquardt algorithm emerged the best tested algorithm. Optimum number of neurons, initial adaptive value and initial value of weights/biases were 12, 4.0 and 1.0, respectively, in accordance to RSM optimization. Only 13% (R^2=0.87) of the variation in the COD removal data set could not be explained by BPANN model. Result from modelling and optimization showed high forecast accuracy by the BPANN model.

E-supplementary data for this work can be found in e-version of this paper online.

Acknowledgements

The authors are grateful to the following institutions for funding and supporting this research: Major Science and Technology Program for Water Pollution Control and Management, China (Grant No.: 2013ZX07201007); Harbin Institute of Technology Environment and Ecology Innovation Special Funds, China (Grant No.: HSCJ201614); and Ton Duc Thang University (TDTU-DEMASTED), Vietnam.

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.

Antwi, P., Li, J., Boadi, P.O., Meng, J., Koblah Quashie, F., Wang, X., Ren, N., Buelna, G. 2017a. Efficiency of an upflow anaerobic sludge blanket reactor treating potato starch processing wastewater and related process kinetics, functional microbial community and sludge morphology. Bioresource Technology, 239, 105-116.

Antwi, P., Li, J., Boadi, P.O., Meng, J., Shi, E., Chi, X., Deng, K., Ayivi, F. 2017b. Dosing effect of zero valent iron (ZVI) on biomethanation and microbial community distribution as revealed by 16S rRNA high-throughput sequencing. International Biodeterioration & Biodegradation, 123, 191-199.

Antwi, P., Li, J., Boadi, P.O., Meng, J., Shi, E., Deng, K., Bondinuba, F.K. 2017c. Estimation of biogas and methane yields in an UASB treating potato starch processing wastewater with backpropagation artificial neural network. Bioresource Technology, 228, 106-115.

Antwi, P., Li, J., Opoku Boadi, P., Meng, J., Shi, E., Xue, C., Zhang, Y., Ayivi, F. 2017d. Functional bacterial and archaeal diversity revealed by 16S rRNA gene pyrosequencing during potato starch processing wastewater treatment in an UASB. Bioresource Technology, 235, 348-357.

Apha, A. 2012. WEF 2012. Standard methods for the examination of water and wastewater, 22.

Barua, S., Dhar, B.R. 2017. Advances towards understanding and engineering direct interspecies electron transfer in anaerobic digestion. Bioresource Technology, 244, 698-707.

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.

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.

Ding, A., Pronk, W., Qu, F., Ma, J., Li, G., Li, K., Liang, H. 2015. Effect of calcium addition on sludge properties and membrane fouling potential of the membrane-coupled expanded granular sludge bed process. Journal of Membrane Science, 489, 55-63.

Forchhammer, K. 2007. 1. Abstract 2. Introduction 3. The role of glutamine in bacterial metabolism 3.1. General features of glutamine metabolism 3.2. Glutamine synthesis, the primary reaction in ammonia assimilation 4. The P II signal transduction protein, a common module in bacterial nitrogen control 4.1. P II signalling proteins. Frontiers in Bioscience, 12, 358-370.

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.

Hagen, L.H., Vivekanand, V., Linjordet, R., Pope, P.B., Eijsink, V.G., Horn, S.J. 2014. Microbial community structure and dynamics during co-digestion of whey permeate and cow manure in continuous stirred tank reactor systems. Bioresource Technology, 171, 350-359.

Herlemann, D.P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., Andersson, A.F. 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. The ISME journal, 5(10), 1571-1579.

Hu, Y.-f., Yang, C.-z., Dan, J.-f., Pu, W.-h., Yang, J.-k. 2017. Modeling of expanded granular sludge bed reactor using artificial neural network. Journal of Environmental Chemical Engineering.

Hugerth, L.W., Wefer, H.A., Lundin, S., Jakobsson, H.E., Lindberg, M., Rodin, S., Engstrand, L., Andersson, A.F. 2014. DegePrime, a program for degenerate primer design for broad-taxonomic-range PCR in microbial ecology studies. Applied and environmental microbiology, 80(16), 5116-5123.

Kundu, K., Sharma, S., Sreekrishnan, T. 2012. Effect of operating temperatures on the microbial community profiles in a high cell density hybrid anaerobic bioreactor. Bioresource technology, 118, 502-511.

Lee, K.Y., Chung, N., Hwang, S. 2016. Application of an artificial neural network (ANN) model for predicting mosquito abundances in urban areas. Ecological Informatics, 36, 172-180.

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(9), 1217-1224.

Lizama, A.C., Figueiras, C.C., Herrera, R.R., Pedreguera, A.Z., Ruiz Espinoza, J.E. 2017. Effects of ultrasonic pretreatment on the solubilization and kinetic study of biogas production from anaerobic digestion of waste activated sludge. International Biodeterioration & Biodegradation, 123, 1-9.

Lloret, E., Pastor, L., Pradas, P., Pascual, J.A. 2013. Semi full-scale thermophilic anaerobic digestion (TAnD) for advanced treatment of sewage sludge: Stabilization process and pathogen reduction. Chemical Engineering Journal, 232, 42-50.

López, I., Borzacconi, L. 2011. Modelling of an EGSB treating sugarcane vinasse using first-order variable kinetics. Water Science and Technology, 64(10), 2080-2088.

Lu, X., Zhen, G., Chen, M., Kubota, K., Li, Y.-Y. 2015. Biocatalysis conversion of methanol to methane in an upflow anaerobic sludge blanket (UASB) reactor: long-term performance and inherent deficiencies. Bioresource technology, 198, 691-700.

Lu, X., Zhen, G., Ni, J., Hojo, T., Kubota, K., Li, Y.-Y. 2016. Effect of influent COD/SO 4 2− ratios on biodegradation behaviors of starch wastewater in an upflow anaerobic sludge blanket (UASB) reactor. Bioresource technology, 214, 175-183.

Luo, G., Li, J., Li, Y., Wang, Z., Li, W.-T., Li, A.-M. 2016. Performance, kinetics behaviors and microbial community of internal circulation anaerobic reactor treating wastewater with high organic loading rate: Role of external hydraulic circulation. Bioresource technology, 222, 470-477.

Nizami, A.-S., Murphy, J.D. 2011. Optimizing the operation of a two-phase anaerobic digestion system digesting grass silage. Environmental science & technology, 45(17), 7561-7569.

Podder, M.S., Majumder, C.B. 2016. The use of artificial neural network for modelling of phycoremediation of toxic elements As(III) and As(V) from wastewater using Botryococcus braunii. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 155, 130-145.

Przetaczek-Rożnowska, I. 2017. Physicochemical properties of starches isolated from pumpkin compared with potato and corn starches. International Journal of Biological Macromolecules, 101, 536-542.

Revilla, M., Galán, B., Viguri, J.R. 2016. An integrated mathematical model for chemical oxygen demand (COD) removal in moving bed biofilm reactors (MBBR) including predation and hydrolysis. Water Research, 98, 84-97.

Rosales-Colunga, L.M., García, R.G., Rodríguez, A.D.L. 2010. Estimation of hydrogen production in genetically modified E. coli fermentations using an artificial neural network. international journal of hydrogen energy, 35(24), 13186-13192.

Schweitzer, L., Noblet, J. 2018. Chapter 3.6 - Water Contamination and Pollution. in: Green Chemistry, (Eds.) B. Török, T. Dransfield, Elsevier, pp. 261-290.

Shi, E., Li, J., Leu, S.-Y., Antwi, P. 2016. Modeling the dynamic volatile fatty acids profiles with pH and hydraulic retention time in an anaerobic baffled reactor during the startup period. Bioresource Technology, 222, 49-58.

Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. 2014. Dropout: a simple way to prevent neural networks from overfitting. Journal of machine learning research, 15(1), 1929-1958.

Subramanyam, R., Mishra, I.M. 2013. Characteristics of methanogenic granules grown on glucose in an upflow anaerobic sludge blanket reactor. Biosystems engineering, 114(2), 113-123.

Wold, S., Sjöström, M., Eriksson, L. 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and intelligent laboratory systems, 58(2), 109-130.

Wu, C.-Y., Peng, Y.-Z., Wang, S.-Y., Ma, Y. 2010. Enhanced biological phosphorus removal by granular sludge: from macro-to micro-scale. Water research, 44(3), 807-814.

Wu, Z., Guo, X., Lv, C., Wang, H., Di, D. 2017. Study on the quantification method of water pollution ecological compensation standard based on emergy theory. Ecological Indicators.

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

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., 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(1), 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(9), 885-907.

Yi-Fan, H., Chang-Zhu, Y., Jin-Feng, D., Wen-Hong, P., Jia-Kuang, Y. 2017. Modeling of expanded granular sludge bed reactor using artificial neural network. Journal of environmental chemical engineering, 5(3), 2142-2150

Yusof, T.R.T., Man, H.C., Rahman, N.A.A., Hafid, H.S. 2014. Optimization of Methane Gas Production From Co-Digestion of Food Waste and Poultry Manure Using Artificial Neural Network and Response Surface Methodology. Journal of Agricultural Science, 6(7), 27.

Figure captions

Fig.1 Schematic diagram of an UASB (A); flowchart of the feedforward BPANN architecture (B)

Fig.2 Performance of UASB at various HRTs: COD removal with respect to organic loading rate

Fig.3 Comparison of functional population at genus level between end of startup phase and end of UASB operation

Fig.4 3D curvature plots from RSM showing interaction effects of: initial adaptive value and number of neurons on MSE (A); initial values of weights and biases and number of neurons on MSE (B); initial values of weights and biases and initial adaptive value on MSE (C)

Fig.5 Performance of the BPANN model predicting COD removals: Visual agreements between experimental and predicted data sets (train set [A], validation set [C], testing set [E] and overall data set [G]); correlations between experimental and predicted data sets (train set [B], validation set [D], testing set [F] and overall data set [H])

Tables

Table 1
Characteristics of raw wastewater and UASB operation after startup

Wastewater characteristics		Operation of the UASB after startup			
Parameters	Mean	Stage (HRT)	Influent COD (g/L)	Average OLR ($kgCOD/m^3.d$)	Period (days)
TCOD (mg/L)	26385	1-(72h)	7.5-8.0 (Phase I)	2.70	1-15
NH_4^+-N (mg/L)	349	2-(48h)		3.73	16-34
pH (mg/L)	5	3-(36h)		5.02	35-56
ALK (mg $CaCO_3$/L)	2945	4-(72h)	10.5-12 (Phase II)	3.65	57-70
TP (mg/L)	96	5-(48h)		5.27	71-99
Protein (mg/L)	6322	6-(36h)		7.16	100-131
Carbohydrate (mg/L)	7834	7-(72h)	19-20 (Phase III)	6.70	132-158
Total solids (g/L)	24.63	8-(48h)		9.88	159-187
Volatile solids (g/L)	19.6	9-(36h)		13.27	188-218

h, hours

Table 2
Input and output variables and related descriptive statistics

Statistical parameter	Input variables						Output variable
	[Ip1]	[Ip2]	[Ip3]	[Ip4]	[Ip5]	[Ip6]	[Y]
	Inf. COD (mg/L)	NH4+-N (mg/L)	Inf. pH	OLR (Kg COD/m^3/d)	Eff.VFA (mg/L)	Biogas yield (L/d)	COD removal (%)
Mean	13579.52	136.2511	6.60	6.994177	311.5052	18.02061	84.06073
Median	11198.96	134.1396	6.61	6.581633	314.694	18.60063	87.46304
Minimum	5099.502	90.03433	6.05	2.411714	29.28911	2.6	42.1783
Maximum	21806.85	181.4866	7.1	13.8322	626.247	32.7561	97.72828
Counts	218	218	218	218	218	218	218

Ip, input variable; Inf, influent; Eff, effluent

Table 3
Experimental runs from Box Behnken Design (BBD)

Runs	Variables						Mean squared errors	
	Number of neurons (X1)		Initial adaptive value (X2)		Initial value of w and b (X3)			
	Uncoded	Coded	Uncoded	Coded	Uncoded	Coded	Experimental	Predicted
1	14	0	7.5	1	0.2	-1	0.0245	0.0226
2	25	1	4	0	1.8	1	0.0321	0.0311
3	25	1	4	0	0.2	-1	0.0311	0.0312
4	25	1	7.5	1	1	0	0.0286	0.0303
5	14	0	7.5	1	1.8	1	0.0203	0.0195
6	14	0	0.5	-1	1.8	1	0.0182	0.0200
7	3	-1	7.5	1	1	0	0.0203	0.0211
8	3	-1	4	0	0.2	-1	0.0227	0.0236
9	14	0	4	0	1	0	0.0114	0.0115
10	25	1	0.5	-1	1	0	0.0297	0.0288
11	14	0	4	0	1	0	0.0121	0.0115
12	14	0	4	0	1	0	0.0112	0.0115
13	14	0	0.5	-1	0.2	-1	0.0191	0.0197
14	3	-1	0.5	-1	1	0	0.0221	0.0203
15	3	-1	4	0	1.8	1	0.0211	0.0209

w, weights; b, bias, RSM, response surface methodology

Table 4
Benchmark comparison of backpropagation training algorithms

Backpropagation algorithm	Training function	Target sets (COD removal %) used in the ANN study		
		R^2	IN	MSE
BFGS quasi-Newton	*trainbfg*	78.2	138	12.42
Powell–Beale conjugate gradient	*traincgb*	49.24	146	1.047
Fletcher–Reeves conjugate gradient	*traincgf*	12.50	201	56.92
Polak–Ribi'ere conjugate gradient	*traincgp*	66.91	98	5.231
Batch gradient descent	*traingd*	19.89	172	67.98
Batch gradient descent with momentum	*traingdm*	7.12	181	15.96
Variable learning rate	*traingdx*	36.09	128	3.015
Levenberg-Marquardt	*trainlm*	85.76	143	0.131
One step secant	*trainoss*	81.65	211	3.451
Resilient	*trainrp*	39.87	152	6.013
Scaled conjugate gradient	*trainscg*	56.81	102	1.782

IN - number of iterations;

Table 5
Performance statistics of multiple nonlinear regression model of model building parameters

Term	Effect	Coef	SE Coef	T-Value	SS	PC (%)	P-Value
X1	0.008825	0.004412	0.000674	6.55	0.000156	22.807	0.001
X2	0.001175	0.000587	0.000674	0.87	0.000003	0.4385	0.423
X3	-0.001400	-0.000700	0.000674	-1.04	0.000004	0.5847	0.347
X1*X1	0.019858	0.009929	0.000992	10.01	0.000364	53.216	0.000
X2*X2	0.007358	0.003679	0.000992	3.71	0.000050	7.309	0.014
X3*X3	0.010508	0.005254	0.000992	5.29	0.000102	14.912	0.003
X1*X2	0.000350	0.000175	0.000953	0.18	0.000000	0.001	0.862
X1*X3	0.001300	0.000650	0.000953	0.68	0.000002	0.292	0.526
X2*X3	-0.001700	-0.000850	0.000953	-0.89	0.000003	0.438	0.413
Constant		0.01157	0.00110	10.51			0.000

*, multiplication function; p-values<0.05 were considered significant; Coef, coefficient; X1, neurons in the hidden layer; X2, initial adaptive value; and X3, initial weights and biases

Table 6
Performance summary of the BP-ANN model

Performance indicators	Testing data set
	COD removal
Coefficient of determination(R^2)	0.87
Index of agreement (IA):	0.9117
Fractional Variance (FV):	0.00113

Figures

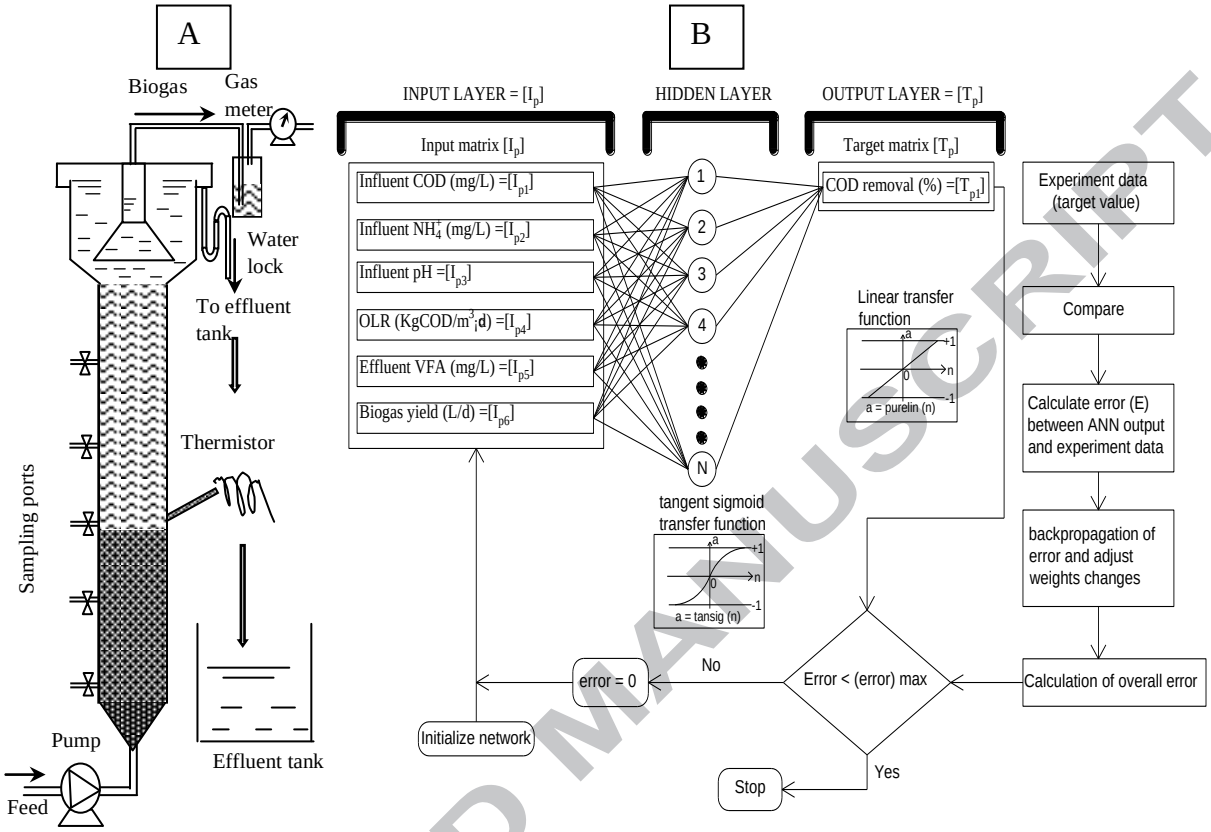


Fig.1 Schematic diagram of an UASB (A); flowchart of the feedforward BPANN architecture (B)

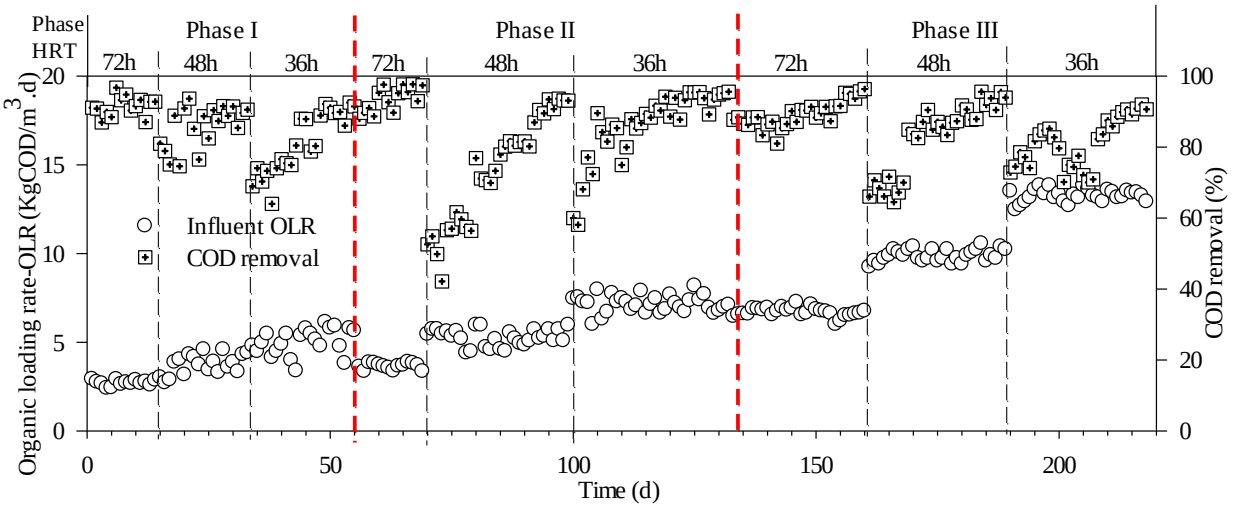


Fig.2 Performance of UASB at various HRTs: COD removal with respect to organic loading rate

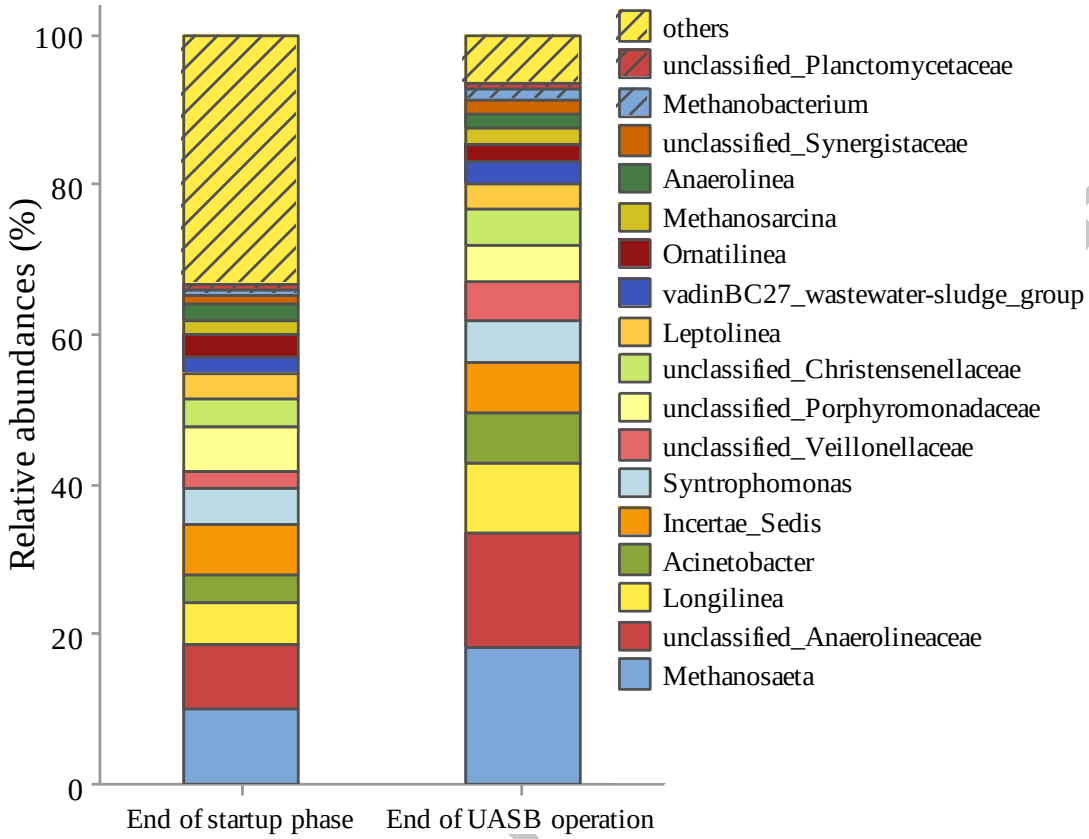


Fig.3 Comparison of functional population at genus level between end of startup phase and end of UASB operation

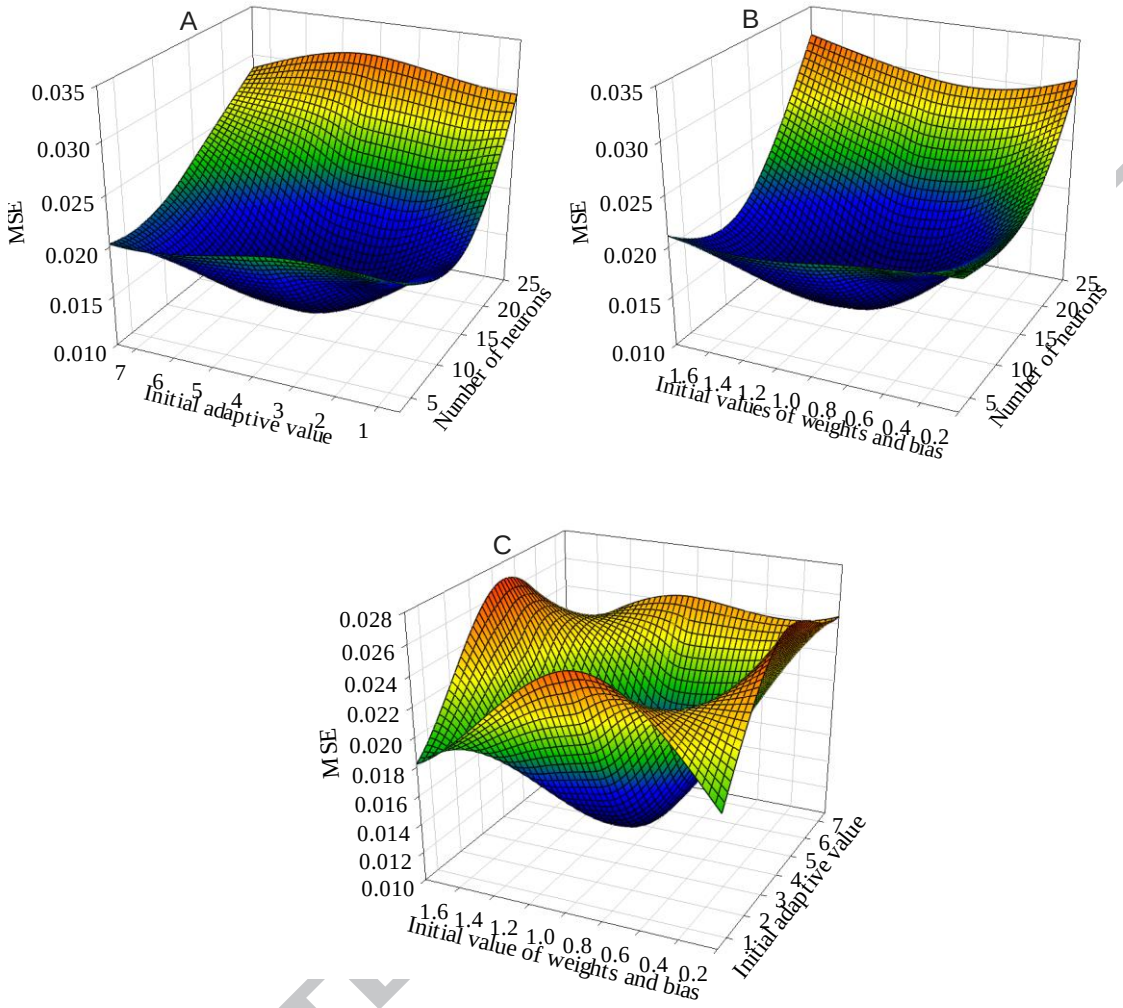


Fig.4 3D curvature plots from RSM showing interaction effects of: initial adaptive value and number of neurons on MSE (A); initial values of weights and biases and number of neurons on MSE (B); initial values of weights and biases and initial adaptive value on MSE (C)

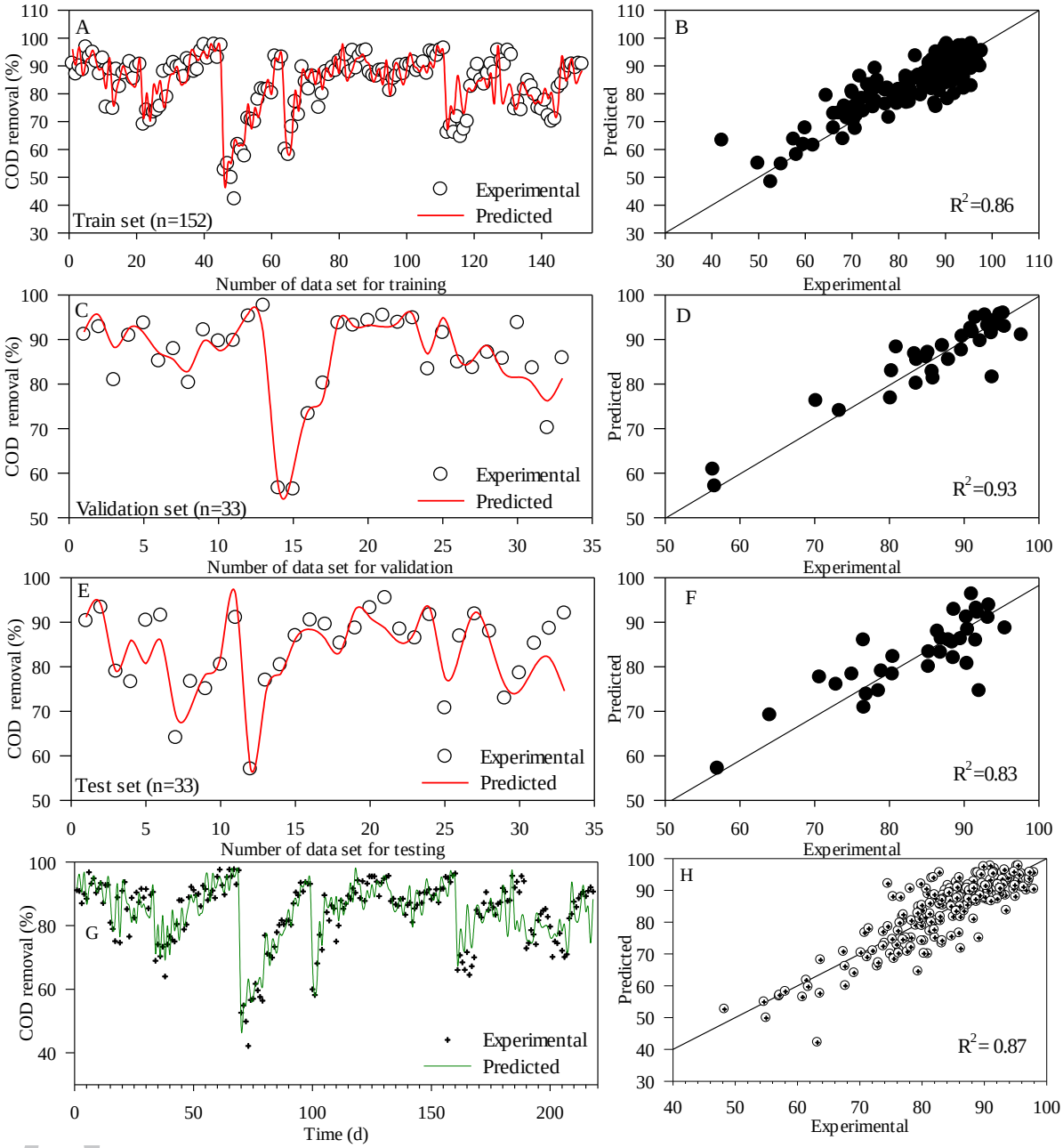
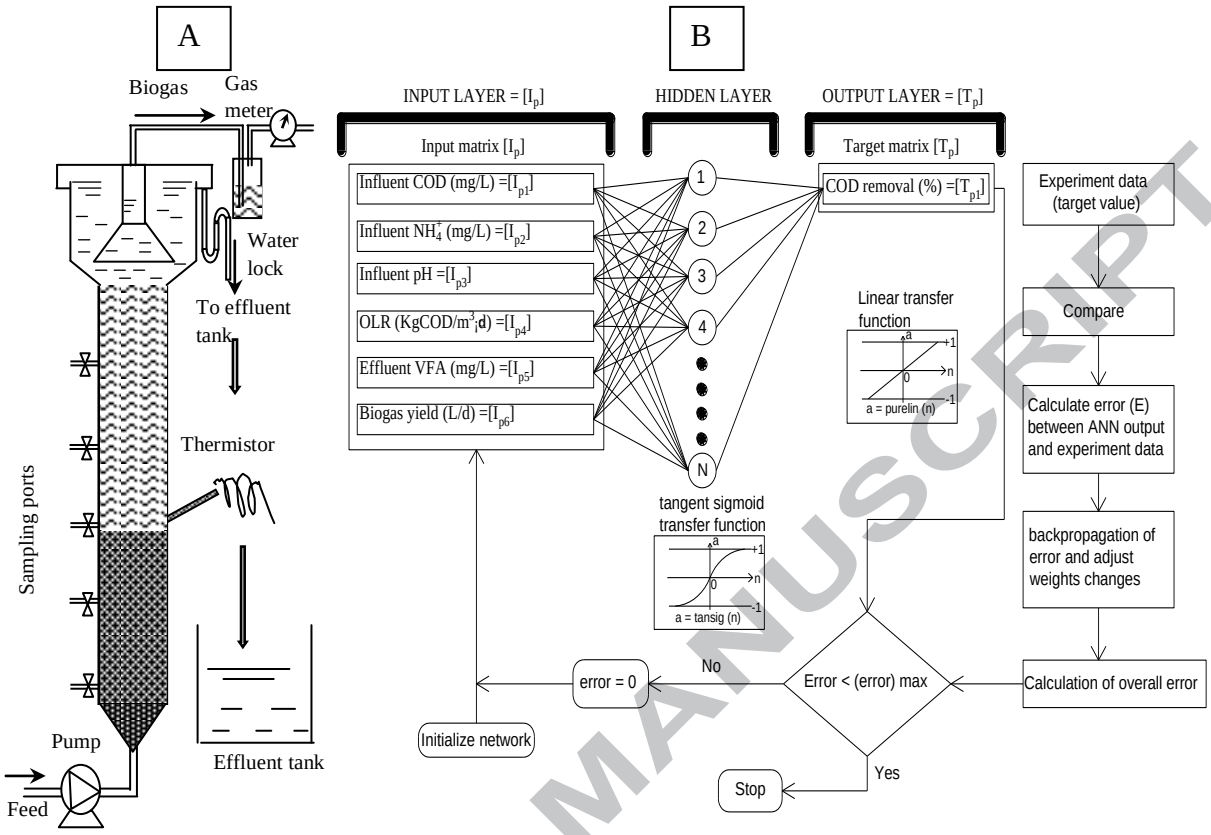


Fig.5 Performance of the BPANN model predicting COD removals: Visual agreements between experimental and predicted data sets (train set [A], validation set [C], testing set [E] and overall data set [G]); correlations between experimental and predicted data sets (train set [B], validation set [D], testing set [F] and overall data set [H])

Highlights:

- COD removal process was predicted with feedforward artificial neural networks.
- Training algorithm and model building parameters were optimized before employed.
- Model performance suggested feasibility to control and optimize AD process with BPANN.
- Microbial communities coexisted without evidence of inhibition on the AD process.



Schematic diagram of: (A) an UASB operated under mesophillic condition; (B) flowchart of the feedforward BPANN architecture