Article

Artificial Neural Network (ANN) Modelling for Biogas Production in Pre-Commercialized Integrated Anaerobic-Aerobic Bioreactors (IAAB)

Wei-Yao Chen [1], **Yi Jing Chan** [1,*], **Jun Wei Lim** [2], **Chin Seng Liew** [2], **Mardawani Mohamad** [3], **Chii-Dong Ho** [4], **Anwar Usman** [5], **Grzegorz Lisak** [6,7], **Hirofumi Hara** [8] and **Wen-Nee Tan** [9]

1 Department of Chemical and Environmental Engineering, University of Nottingham Malaysia, 43500 Kuala Lumpur, Selangor, Malaysia; ebycw1@nottingham.edu.my
2 Department of Fundamental and Applied Sciences, HICoE-Centre for Biofuel and Biochemical Research, Institute of Self-Sustainable Building, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak Darul Ridzuan, Malaysia; junwei.lim@utp.edu.my (J.W.L.); liew_20001952@utp.edu.my (C.S.L.)
3 Faculty of Bioengineering and Technology, Jeli Campus, Universiti Malaysia Kelantan, Jeli Campus, 17600 Jeli, Kelantan, Malaysia; mardawani.m@umk.edu.my
4 Department of Chemical and Materials Engineering, Tamkang University, Tamsui, New Taipei 251, Taiwan; cdho@mail.tku.edu.tw
5 Department of Chemistry, Faculty of Science, Universiti Brunei Darussalam, Jalan Tungku Link, Gadong BE1410, Brunei; anwar.usman@ubd.edu.bn
6 Residues and Resource Reclamation Centre, Nanyang Environment and Water Research Institute, Nanyang Technological University, 1 Cleantech Loop, CleanTech One, Singapore 637141, Singapore; g.lisak@ntu.edu.sg
7 School of Civil and Environmental Engineering, College of Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore
8 Department of Biotechnology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1, Tokyo 113-8657, Japan; ahhara@g.ecc.u-tokyo.ac.jp
9 Chemistry Section, School of Distance Education, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia; tanwn@usm.my
* Correspondence: yi-jing.chan@nottingham.edu.my

Abstract: The use of integrated anaerobic-aerobic bioreactor (IAAB) to treat the Palm Oil Mill Effluent (POME) showed promising results, which successfully overcome the limitation of a large space that is needed in the conventional method. The understanding of synergism between anaerobic digestion and aerobic process is required to achieve maximum biogas production and COD removal. Hence, this work presents the use of artificial neural network (ANN) to predict the COD removal (%), purity of methane (%), and methane yield ($LCH_4/gCOD_{removed}$) of anaerobic digestion and COD removal (%), biochemical oxygen demand (BOD) removal (%), and total suspended solid (TSS) removal (%) of aerobic process in a pre-commercialized IAAB located at Negeri Sembilan, Malaysia. MATLAB R2019b was used to develop the two ANN models. Bayesian regularization backpropagation (BR) showed the best performance among the 12 training algorithms. The trained ANN models showed high accuracy ($R^2 > 0.997$) and demonstrated good alignment with the industrial data obtained from the pre-commercialized IAAB over a 6-month period. The developed ANN model is subsequently used to create the optimal operating conditions which maximize the output parameters. The COD removal (%) was improved by 33.9% (from 68.7% to 92%), while the methane yield was improved by 13.4% (from 0.23 $LCH_4/gCOD_{removed}$ to 0.26 $LCH_4/gCOD_{removed}$). Sensitivity analysis shows that COD inlet is the most influential input parameters that affect the methane yield, anaerobic COD, BOD and TSS removals, while for aerobic process, COD removal is most affected by mixed liquor suspended solids (MLSS). The trained ANN model can be utilized as a decision support system (DSS) for operators to predict the behavior of the IAAB system and solve the problems of instability and inconsistent biogas production in the anaerobic digestion process. This is of utmost importance for the successful commercialization of this IAAB technology. Additional input parameters such as the mixing time, reaction time, nutrients (ammonium nitrogen and total phosphorus) and concentration of microorganisms could be considered for the improvement of the ANN model.

1. Introduction

The palm oil industry in Malaysia is substantial, contributing nearly 3.6% to Malaysia's gross domestic product (GDP) in 2020 [1]. Moreover, the global market share of Malaysia in palm oil production and export is 25.8% and 34.3% respectively [2]. Hence, it is important that the waste produced by the palm oil industry, namely the palm oil mill effluent (POME), is regulated, and treated safely before discharging it into the environment. POME contains several polluting characteristics, including phosphorus, sulphate, volatile solid, and total organic carbon [3]. It is estimated that 2.5 to 3.5 tons of palm oil mill effluent is produced from every 1 ton of crude palm oil produced locally [4]). Therefore, a proper and appropriate operation of wastewater treatment plant (WWTP) must be utilized [5].

There are many methods of treating POME, the most common method being anaerobic digestion and facultative digestion [6]. There are some recent studies on the treatment of POME that suggest a more innovative way of treating POME. A study done by Chong et al. [7] highlights the benefit of using both aerobic and anaerobic processes in a single bioreactor, namely integrated anaerobic-aerobic bioreactor (IAAB) which are able to perform much better than conventional treatment of palm oil mill effluent methods. The simulation of integrated anaerobic and aerobic bioreactor shows that removal efficiency of chemical oxygen demand (COD) and biochemical oxygen demand (BOD) of up to 99% can be achieved while the cost of net expenditure can be reduced by 5.8%. However, there are two main concerns that have to be addressed in order to convince the palm oil millers to adopt this newly invented IAAB technology to generate biogas and producing high quality treated effluent simultaneously. This includes the stability of the anaerobic digestion process and consistent biogas generation. POME treatment is usually prone to have stability issue due to the varying characteristics of POME which depend on the crop season and the loading rate of the milling process. Besides, lack of skilled personnel for monitoring and control of anaerobic digester could also lead to slow feedback time to address any potential instability issue [8]. Therefore, the development of a prediction model of anaerobic digestion of POME for IAAB is crucial to address the aforementioned issues.

Despite the development of several mechanistic model such as International Water Association (IWA) Anaerobic Digestion Model No.1 (ADM1) [9], the need for a large number of parameters and estimations [10] makes it difficult to quantify the organic composition of the effluent feed stream, which is a piece of essential information needed to utilize the model. In recent years, there has been an increased use of machine learning such as artificial neural networks (ANNs) to model and optimize the process as an alternative to mechanistic models [11]. The ANN modelling can be used as a monitoring framework for WTTP operation that can increase cost optimization and identification of the quality or stability of wastewater effluent. The practical benefit of using ANN in WWTP is its ability to identify the complex non-linear relationship between input and output parameters that are useful for prediction and estimation of WWTP effluent properties without the need for prior knowledge on theoretical and physical laws that govern the biological process. However, the mentioned benefits of using an ANN is not applicable if the simulation were done solely with engineering software like HYSYS or SuperPro. ANN modelling can be used in WWTP to predict the performance of WTTP with a high degree of accuracy, provided that the quality of historical data is fairly good [12]. Numerous studies that use ANN to optimize the anaerobic digestion have been reported with high R^2 of up to 0.998. For instance, Güçlü, Yılmaz and Ozkan-Yucel [13] achieve a $R^2 = 0.89$ using the ANN model. The parameters used to train the model were pH value, gas flow rate, VFAs, temperature, organic matter, dry matter, and alkalinity. Similarly, Sathish and Vivekanandan [14] also used ANN to optimize the anaerobic digestion process using 4 input variables, which are temperature, agitation time, pH level, and concentration of substrate. They are able

to obtain a satisfying result with an accuracy of $R^2 = 0.998$ for the prediction of biogas production. Therefore, ANN is preferred in this study due to their faster and simpler computation modelling, particularly for large scale operations like IAAB.

However, the potential use of ANN to accurately model the methane production and chemical oxygen demand (COD) removal of IAAB that treats POME has not been reported. This is necessary as POME has a varying characteristics (COD of 44,300 mg/L–112,000 mg/L) where close monitoring and efficient feedback are required to ensure stable operation [8]. Most importantly, process optimization can be conducted at a lower cost and shorter period of time with the aid of the prediction model prior to the real application of IAAB in industry scale. Therefore, this work aims to model the POME treatment process using ANN to further optimize the production of biogas and COD removal efficiencies. MATLAB software is used to develop an artificial neural network that is based on the principle of Feedforward Neural Network (FFNN). A set of training data (96 data sets) obtained from the pre-commercialized IAAB over a six-month period will be used to train the feedforward network using 12 different back propagation training algorithms. It is hypothesized that this ANN model can serve as a platform to estimate the behavior of IAAB in treating POME and provide decisive information to the plant engineers in applying suitable control approaches and preventing digester upset.

2. Methodology

The ANN will be trained using 12 different training algorithms in MATLAB R2019b software. The best training algorithm will be determined by considering several performance criteria. Two ANN models will be developed for anaerobic digestion and aerobic process respectively. The training data set for the ANN model will be obtained from the pre-commercialized IAAB located at Negeri Sembilan, Malaysia. The operation of the IAAB can be found elsewhere [7]. The ANN model with the highest prediction accuracy and acceptable correlation coefficient (R), mean squared error (MSE), mean absolute error (MAE) and mean absolute percent error (MAPE) will be used. Sensitivity analysis is conducted to evaluate the relative importance of input parameters on the output parameters. The first ANN model developed is used to determine the optimum operating condition for the removal rate of COD, purity of methane and methane yield for the anaerobic digestion. The second ANN model developed will be used to determine the optimum operating condition for the COD removal, biochemical oxygen demand (BOD) removal and total suspended solids (TSS) removal for the aerobic process. The industry data (experimental data) will be compared with the predicted data obtained from the ANN model.

2.1. Artificial Neural Network Model (ANN)

Artificial neural networks (ANNs), or neural networks, are a complex computational model technique that can be used under machine learning to predict the output of a process. The artificial neural network structure is similar to the biological neural networks in our human brains in the way that neurons signal is being transferred from one node to another [15]. Numerous studies have been done on the application of ANN for use in improving the anaerobic reaction. For instance, the feed-forward backpropagation ANN is used to help in investigating the influence of substrates such as food waste, vegetable waste and organic loading rate on the methane generation to achieve methane purity of around 60 to 70% [16]. Moreover, an ANN model with R value of 0.997 was developed by Mougari et al. [17] to predict the production of methane and biogas from anaerobic digestion process where its substrates are organic wastes. A typical artificial neural network architecture is made up of a single input layer, a hidden layer or multiple hidden layers, and a single output layer. Each layer in the neural network architecture contains several nodes or neurons in biological terms that connect to the next layer. The nodes are given a weight that can be adjusted during the training process. The weights are increased or decreased to strengthen or weaken the strength of the interconnected neuron's signal.

In a single layer architecture, the connection between the nodes can be shown in Equation (1) [18].

$$Y_i = f\left(\sum_{j=1}^{N} W_{ij}X_j + b_i \right) \tag{1}$$

where Y is the output value, and X is the input layer and W and b represents the weights and bias respectively. The subscript i and j denote the previous layer and current layer respectively. The function, f is the activation function or transfer function of the single layer architecture. There are multiple types of input and output activation functions that can be used to model a single neuron, the three most commonly used are log-sigmoid transfer function, tan-sigmoid transfer function and linear transfer function. In the case of modelling a neural network that is used for pattern recognition problems, sigmoid output neurons are the most suitable, whereas linear function output neurons are more suitable to be used for function fitting problems [19]. A non-linear input function is needed for the artificial neural network to learn non-linear relationships between the input variable and output variable. Between a non-linear function like tan-sigmoid transfer function (tansig) and log-sigmoid function (logsig), tan-sigmoid transfer function is preferred due to the [−1, 1] range of tansig as compared to [0, 1] for log-sigmoid function. Hence, tan-sigmoid transfer function is able to provide a stronger gradient. To build an ANN model, three phases are needed which is training, testing and validation. Hence, the dataset is typically divided into three categories where 70% to 80% are used for training, which the remaining data is used for testing and validation purposes. In a training process of a feedforward neural network, two main processes are involved: (a) feed-forward process and (b) back-propagation process. In the feed-forward process, the input is propagated into the neural network specifically the hidden layer and then to the output layer where all the weights are randomly assigned, bias and activation function is used to provide a predictive value [20]. Next, the back-propagation phase is initiated where the weights are adjusted to minimize the error between the predicted value and the experimental value of the network inputted from the training data. This is done by a specific set of backpropagation training algorithms.

There is a normalization of data before proceeding to train the network so that each input data contributes equally and is scaled to a standard range which helps in the efficiency of the learning process [21]. The input and output data can be normalized using the Equation (2).

$$y = y_{min} + \frac{(x - x_{min})(y_{max} - y_{min})}{(x_{max} - x_{min})} \tag{2}$$

where the x_{min} and x_{max} represent the value of the minimum and maximum data in the dataset, while y_{min} and y_{max} represent the range for normalization. Finally, y is the normalized value of x. The range of normalization is between −1 to 1.

The hidden layers comprise several hidden neurons that are adjustable. The learning performance of the artificial neural network usually performs better if the number of hidden neurons is increased. However, a high number of hidden neurons could result in worse performance as the weights are given too much freedom to adjust and overfitting problems may arise. On the other hand, if the number of hidden neuron is too low, it will limit the performance of the artificial neural network.

The anaerobic digestion remains a kind of a black box as it is difficult to control the process due to its complex mechanism and the number of process parameters is high and variable [22]. Therefore, mechanistic modelling of the anaerobic digestion process is too simplistic and unreliable because there is limited information regarding gas-liquid mass transfer coefficients that is relevant to the biogas formation process [23]. The complexity of anaerobic digestion mechanism makes black-box modelling like artificial neural network a more attractive option as compared to mechanistic modelling. The advantage of using artificial neural network is that it can provide accurate prediction without needing any prior information or concept about the relationship between process parameters and the output [24]. Moreover, the learning abilities of artificial neural network modelling allow

them to be adaptive to the complex non-linear behavior of the many complex processes in the wastewater treatment plant [25].

There are different types of ANNs models that can be classified according to their purpose and relevant features. For example, the Hopfield network is a type of recurrent ANN model that is best used for image recognition and detection, enhancement of X-ray images, etc. [26]. Kohonen network is a two layer (input and output layer) neural network that forms data cluster, which is also referred to as a self-organizing map. It is typically used for compression of higher dimensional data to a lower dimensional data while maintaining the content [27]. Both Kohonen and Hopfield network uses unsupervised learning algorithm. An unsupervised learning algorithm does not need to be labelled or classified as input data as the learning algorithm will identify the datasets based on patterns. On the other hand, supervised learning algorithms labelled and structured input data that correspond to the output data. The most common neural network that uses supervised learning algorithm is the feedforward backpropagation network (BP), also referred to as the Multi-layer Perceptron (MLP). The BP network is made up of 3 layers, mainly the input layer (independent variable), hidden layer, and output layer (dependent variable). The number of hidden layers can be varied, and its function is to assist in capturing the nonlinearity relationship between the input and output data through a supervised learning algorithm. A BP network is versatile and flexible and can be used for classification, data modelling, pattern recognition and forecasting [28]. Hence, feedforward BP network is used as the parameters (input data) in a wastewater treatment plant are structured and labelled that has an effect on the output data. However, it should be acknowledged that the accuracy of the ANN model is dependent on the quality of the training data. Hence, increasing the number of good quality training data will improve the accuracy of the ANN model.

2.1.1. Activation Function

In a neural network, the activation function determines how the weighted sum of the input neuron signal is transferred into an output node. There are generally only two types of activation functions that will be used to model the neural network. The activation function can be divided into two categories, non-linear and linear function. The linear function (purelin) is calculated based on the multiplication of the input value and a constant shown below in Equation (3).

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

The linear activation function is used at the output layer to find a linear approximation to a non-linear function. A non-linear function is required to introduce non-linearity to the network [29]. As such, a non-linear function will be used in the hidden layer. There are more variety of non-linear activation functions that can be used such as the radial basis function (Radbas), tangent sigmoid activation (tansig) function, logistic sigmoid (logsig) activation function, rectified linear unit (ReLu) activation function and symmetric saturating linear (Hardlims) activation function. Non-linearity is essential in the network to allow the model to generate complex mapping between the input and output layer. Hence, the neural network is able to predict a target variable that varies non-linearly with its input variables. A literature study done by [30] that compares different combination of activation function in the hidden layer and output layer conclude that the prediction accuracy for tansig non-linear activation function in the hidden layer and purelin linear activation function in the output layer produces the highest accuracy of predictive value of wastewater effluent. Tansig activation function used in ANN modelling also showed the highest accuracy in predicting the susceptibility of shallow landslides with an accuracy of 90% [20]. The tan-sigmoid function shown in Equation (4) has an output range of −1 to 1.

A hyperbolic tangent function is preferred as it can give a stronger gradient and positive and negative output values.

$$\text{tansig}(x) = \frac{2}{(1 + e^{-2x})} - 1 \tag{4}$$

The linear activation function was chosen for the output layer as it is suitable for a continuous valued target such as the chemical oxygen demand (COD) and biogas production.

Hence, the final form of the equation in the feedforward neural network model proposed can be written as shown in Equation (5).

$$\text{output}_k = \sum_{j=1}^{j=8} \left[LW_{k,j} \left(\frac{2}{1 + \exp\left(-2\left(\sum_{i=1}^{i=5}(IW_{j,1}x_i) + b1_j\right)\right)} - 1 \right) \right] + b2_k \tag{5}$$

2.1.2. Training Algorithms

One of the more important values in an artificial neural network model is the weight and bias values. A training algorithm is used to determine the best value of weights and bias that is able to predict the output value. The most commonly used training algorithm is the Levenberg-Marquardt (LM) method due to lesser computational time and better performance [31,32].

The LM training algorithm is a type Gauss-Newton method, which is an optimization technique that utilized steepest descent for complex non-linear patterns. Generally, training algorithms that use Quasi-Newton method requires less computational time. However, the drawback of using LM second order optimization technique is it requires storing and loading the inverse Hessian approximation matrix. This means that extra storage is required in the computer's memory in used. The combination of inverse of Hessian algorithm in Equation (6) and Gauss-Newton algorithm in Equation (7), the LM algorithm can be represented in the below Equation (8).

$$H = J_k^T J_k + \mu I \tag{6}$$

where, μ represent the combination coefficient, and I is the identity matrix

$$w_{k+1} = w_k - \left(J_k^T J_k\right)^{-1} J_k e_k \tag{7}$$

$$w_{k+1} = w_k - \left(J_k^T J_k + \mu I\right)^{-1} J_k e_k \tag{8}$$

The combination of Gaus-Newton and steepest descent algorithm enable LM to train ANN model using two algorithms enable high efficiency and less computation time. It is typically used for small size network training where computation time is the shortest instead of large networks such as image recognition that will result in longer computation time.

The BR training algorithm can help overcome the problem of overfitting during the training of the neural network. Hence, the prediction accuracies can be enhanced. The adjustment of weights and biases is based on the LM optimization [33]. Hence, it will determine the least combination of squared error to generate a network that generalizes well. The BR trained neural network is almost identical to the back-propagation network with the difference being an additional ridge parameter included in the objective function. There are many advantages of using BR as the training algorithm for producing the neural network. For instance, there is a low probability of neural network being over-trained and over-fit, it is also not affected by the size of the network, and require minimal data set as Bayesian neural network generate consistent result with data [34]. Hence, the BR training algorithms were chosen to develop the ANN model for anaerobic digestion and aerobic process.

The measurement of the performance of the ANN model is done by using the mean squared error (MSE), coefficient of correlation (R), mean absolute error (MAE) and mean absolute percentage error (MAPE) shown in Equations (9)–(12).

$$\text{MSE} = \frac{\sum (E_i - P_i)^2}{N} \tag{9}$$

$$\text{MAE} = \frac{1}{N} \left(\sum_{i=1}^{N} |E_i - P_i| \right) \tag{10}$$

$$\text{MAPE} = \frac{1}{N} \left(\sum_{i=1}^{N} \frac{|E_i - P_i|}{|E_i|} \right) \tag{11}$$

$$R = \frac{\sum_{i=1}^{N} \left(P_i - \overline{P_i} \right) \left(E_i - \overline{E_i} \right)}{\sqrt{\sum_{i=1}^{N} \left(E_i - \overline{E_i} \right)^2 \sum_{i=1}^{N} \left(P_i - \overline{P_i} \right)^2}} \tag{12}$$

where,

E_i:	Experimental value
P_i:	Predicted value
$\overline{P_i}$:	Mean of observed data
$\overline{E_i}$:	Mean of experimental data
N:	Number of datasets/model

The value of MSE provides the information regarding the accuracy of the ANN model and hence is a crucial performance factor in deciding if the model can be regarded as usable. It indicates the difference between the predicted values and experimental data. The lower the MSE value, the higher the accuracy of the fit. The value of R is used to evaluate and indicate the closeness of the data fit to the regression line. R value ranges from 0 to 1, a value closer to 1 indicates a smaller difference between the observed data and the fitted values. A R^2 value of more than 0.8 is considered to be satisfactory.

2.2. Selection of Input and Outputparameters

There are several factors that are important and greatly affect the performance of the anaerobic digester. These factors are useful in determining the production of biogas and corresponding efficiency which ultimately will help to further understand and optimize our anaerobic digestion process. In addition to that, with the use of artificial neural network, process parameters can be adjusted to maximize methane production yield while minimizing cost.

The most critical process parameters that can affect the efficiency of anaerobic digestion include the organic loading rate (OLR), pH level, presence of toxic compounds, temperature and hydraulic retention time (HRT) and others [35]. The input parameters that were used to train the artificial neural network for anaerobic digestion were the inlet flowrate (Q_{in}), OLR, COD_{in}, BOD_{in}, TSS_{in}, and pH inlet, which are recorded periodically by the plant operators. These parameters can represent the real situation in the IAAB.

On the other hand, the input parameters used for aerobic process were OLR, COD_{in}, BOD_{in}, TSS_{in}, and MLSS, DO, and F/M ratio. There are a number of parameters that could affect the performance of aerobic reaction. One of the parameters is the operating temperature of the aerobic process as it will have an effect on the aerobic bacteria ability to digest the organic waste. Different species of aerobic bacteria will have different optimum temperatures. For instance, the optimum temperature for psychrophilic microorganisms is between 12 °C to 18 °C, for mesophilic microorganisms it is between 25 °C to 40 °C, and finally for thermophilic microorganisms, the optimum temperature for best performance for the bacteria is 55 °C to 65 °C [36].

The second factor is the ratio of food to microorganisms (F/M). The ratio simply means the amount of substrate available to the microorganism in relation to the number of microorganisms present in the aeration tank. The unit can be expressed as the kg COD per

kg of MLSS per day. A high F/M ratio indicates a higher amount of food as compared to the number of microorganisms that will lead to the increased growth of aerobic bacteria in the bioreactor. A high level of aerobic bacteria activity is not always the best-case scenario as excessive food will degrade the flocculation process, causing the sludge to not settle easily [37]. On the other hand, a low F/M ratio is not preferable as well due to the rapid settling rate of sludge that would lead to an increased waste rate [37]. In order to ensure that bacteria functioned optimally, the operating pH level must be adjusted to increase microbial growth. Typical bacteria will have optimum microbial growth at pH level of between 6.5 and 7.5 [36]. As mentioned, aerobic bacteria require oxygen to grow. Hence, dissolved oxygen (DO) is one of the critical parameters to tune for optimum microbial growth. The optimum concentration of DO is around 2 mg/L and above [7].

For an anaerobic digestion reactor, the most important output parameters are the production of biogas and the removal of COD. This will reflect the profitability and sustainability of the IAAB plant. Hence, the output for the ANN model will be COD removal (%), percentage of methane (CH_4) in the biogas, and methane yield ($LCH_4/gCOD_{removed}$). The aerobic process is typically the secondary treatment to remove the remaining biodegradable organic matter. Hence, the output parameters for the aerobic process would be the percentages of COD, BOD, and TSS removal. These parameters are critical to ensure the final treated effluent is able to meet the discharge limit.

2.3. Optimization of Artificial Neural Network

There are 4 characteristics of the ANN's architecture that are important elements in optimizing the ANN model. First, it is the number of hidden layers present in the ANN architecture. Secondly, it is the number of neurons in each of the input, hidden, and output layers. The third element is the type of activation function in each of the layers respectively. Finally, the training algorithm used to train the ANN model is crucial as well. However, it is also worth noting that the quality and number of training data available could also impact the performance of the ANN model. The typical method to determine the best ANN architecture is by a trial-and-error process where different ANN architecture is compared to one another. The general rule of thumb is that the number of neurons in the hidden layer is increased when the time taken to train is longer and performance parameter such as the mean squared error is large. There is no definitive number of neurons or hidden layers that are needed to produce an ANN model that has high predictive value. The number of neurons present in the input layer and output layer is determined by the number of input parameters and output parameters respectively.

The activation functions are selected based on the various type of data that are present and for which layer. For example, identity function is usually utilized in the input layer, while non-linear activation function such as the hyperbolic tangent sigmoid function is used in the hidden layer. The performance of the ANN is affected by the type of training algorithm as it adjusts the weights and bias to generate the ANN model. The speed at which the training is done is also affected by the training algorithm. In general, the influencing factor of an ANN model is the number of neurons in the hidden layer and the type of training algorithm. The artificial neuron network will train 70% of the input data, 15% as the testing data, and the remaining 15% as the validation data. There are a total number of 96 datasets for anaerobic and aerobic digestion, where each dataset contains 6 input parameters and 3 output parameters.

2.4. Determination of the Number of Hidden Neurons

One of the key network parameters is the number of hidden neurons that are used in the development of artificial neural network (ANN) model. A high number of hidden neurons can cause an overfitting problem, which is that the model overestimated the target value due to the complexity of the input parameters [38]. The determination of the number of hidden neurons is done through trial and error. There have been several equations

that are developed by various authors that supposedly are able to calculate the optimum number of hidden neurons. The findings are summarized in Table 1.

Table 1. Equations to determine the number of hidden neurons required.

Equation	Reference
$N_h = \frac{(4n^2+3)}{n^2-8}$	[39]
$N_h = 2^n - 1$	[40]
$N_h = \sqrt{N_i N_o}$	[41]
$N_h = C_f \left(\frac{N}{d} \log N\right)^{0.5}$	[42]
$N_h = \frac{(N_{in} + \sqrt{N_p})}{L}$	[38]
$N_h = \frac{2^n}{n} + 1$	[43]
$N_h = \frac{K \log \|P_c Z\|}{\log S}$	[44]
$N_h = N - 1$	[45]
$N_h = \frac{(\sqrt{1+8n}-1)}{2}$	[37]

Nevertheless, the equations mentioned may not be accurate in determining the optimum number of hidden neurons required. Another method that is helpful is by calculating the mean squared error (MSE) of the training and validation data at varying numbers of hidden neurons.

As shown in Figure 1, the training and validation set both showed the least MSE value when the number of hidden neurons is 10 or 12. Nonetheless, 12 training algorithms are trained using the trial and error method by changing the number of hidden neurons to further verify the optimum number of hidden neurons.
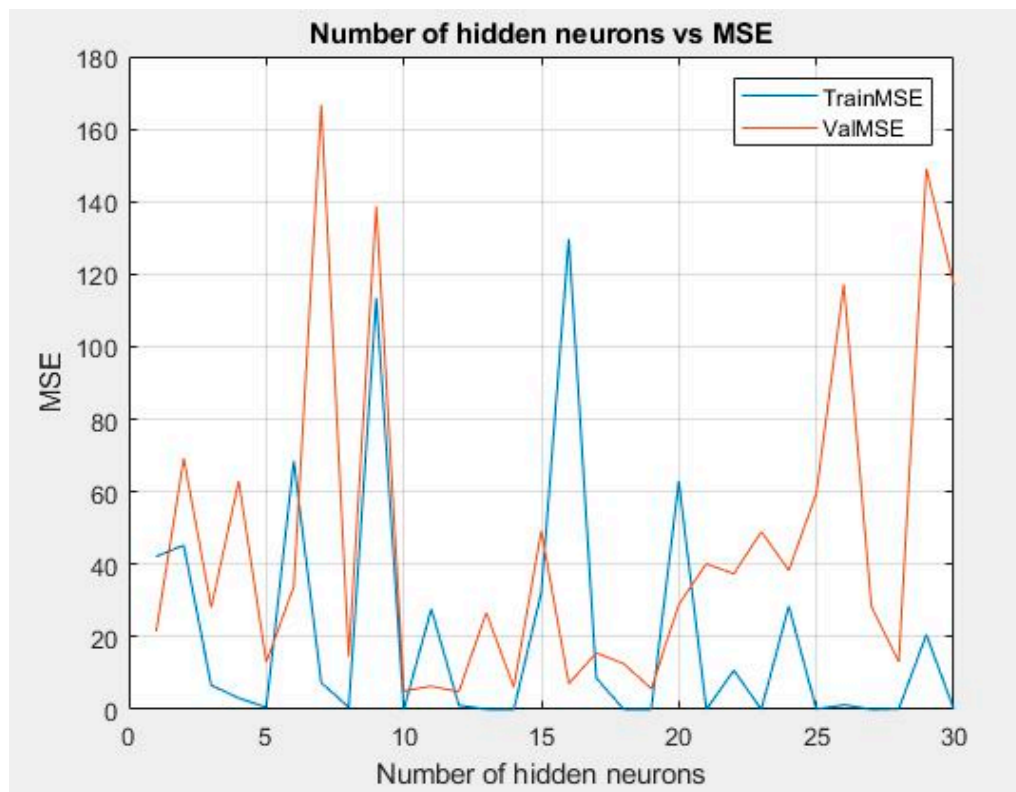
Figure 1. MSE vs. number of hidden neurons.

2.5. Sensitivity Analysis

Two methods are employed in this work to evaluate the relative importance of the input parameters, including connection weight approach and Garson algorithm [46,47].

2.5.1. Connection Weight Approach

The equation used for connection weight approach is shown as below:

$$V = \sum_{j=1}^{N} (W_{ij} \times W_{jo}) \tag{13}$$

where the W represent the value of connection weight, the superscript, I, j, o represent the input, hidden, and output neuron, respectively.

2.5.2. Garson Algorithm Method

The equation used for 1.1.1 Garson algorithm method is shown as below:

$$I_j = \frac{\sum_{m=1}^{m=Nh} \left(\left(\frac{\left| W_{jm}^{ih} \right|}{\sum_{k=1}^{Ni} \left| W_{km}^{ih} \right|} \right) \times \left| W_{mn}^{ho} \right| \right)}{\sum_{k=1}^{k=Ni} \left\{ \sum_{m=1}^{m=Nh} \left(\frac{\left| W_{km}^{ih} \right|}{\sum_{k=1}^{Ni} \left| W_{km}^{ih} \right|} \right) \times \left| W_{mn}^{ho} \right| \right\}} \tag{14}$$

where I_j represent the relative importance of jth input parameter on the output parameter, and the input and hidden neuron is represented by Ni and Nh respectively. The connection weight is represented by W. The subscript o, h, and i represent the output, hidden and input neuron respectively.

3. Results and Discussion

3.1. Performance of ANN Model

The performance of all 12 trained algorithms is analyzed for the anaerobic process and aerobic processes. The results will be discussed in the next section.

3.1.1. Anaerobic Digestion

Tables A7 and A18 (Appendix A) summarized the values of R, MAE, MAPE and MSE of anaerobic digestion.

Table A7 showed the highest R value that can be achieved in a trained ANN model for COD removal response, while Tables A8–A10 showed the lowest MAE, MAPE, and MSE values for COD removal response of different training algorithms utilized at different number of hidden neurons. The training algorithm that achieves the highest R value (0.998) for the COD removal processes is the Bayesian Regularization backpropagation (BR) algorithm. The lowest MAE value of 0.211 was obtained from BR using 12 hidden neurons; the lowest MAPE value of 0.009 from BR using 10 hidden neurons; the lowest MSE value of 0.43 using 12 hidden neurons. In general, LM, BR, RP, and BFG performed much better than 8 other training algorithms. However, LM and BR have consistently ranked higher in terms of producing the best performances ANN model. Moreover, by using BR training algorithm, the MAPE error has never exceeded 10% from one neuron to the twenty neurons. On the other hand, GD has the worst performance among the other training algorithms, where it has the highest MAE value of 95.015, MAPE value of 0.77, and MSE of 9769.

The values of R value, MAE, MAPE, and MSE value for methane purity response are summarized in Tables A11–A14 (Appendix A) respectively. The highest R value achieved after training the neural network is 0.997 by using BR training algorithm with 13 hidden neurons. Interestingly, the highest and lowest MAE value that was obtained is from BR training algorithm at 13 hidden neurons and 1 hidden neuron respectively. For MAPE performance criteria, BR training algorithm clearly performs far better than others as the error for each added neuron were less than 1%. Again, BR and LM trained the best ANN

model compared to others with the lowest MAE, MAPE, and MAE values. The MSE values for BR were less than 1 from 3 hidden neurons to 20 hidden neurons, indicating good performance.

The highest and lowest value of R, MAE, MAPE, and MSE for methane yield response can be found in Tables A15–A18 (Appendix A). The highest R value across all training algorithms is again obtained from BR with a value of 0.996 at 15 hidden neurons, while GD performed the worst at R value of 0.267 using 11 hidden neurons. Table A16 shows the lowest MAE that is from BR, training algorithms, where it achieves a value of less than 0.001 at 3 different cases, while GDX has the highest MAE value at 1.389. In general, all the training algorithms performed well to predict the methane yield, with MSE error ranges from 0.0001 to 0.04000. For the results shown below, 6 input parameter and 3 output parameters are used to train the ANN model. Using the Bayesian regularization (BR) backpropagation training algorithm, the input and output weights and biases are shown in Tables A8 and A9 (Appendix A).

3.1.2. Aerobic Process

Following the consistently better performance of BR and LM than the 10 other training algorithms in predicting the output from anaerobic digestion, only BR and LM training algorithm is used to train the data for the aerobic process. The performance criteria are summarized in Table A19, Table A20, and Table A21 (Appendix A) for aerobic process. The use of 6 hidden neurons gives an overall lower MSE value for the 3 output parameters. MSE values of 1.0288, 0.6047, and 0.5560 for the prediction of COD removal, BOD removal and TSS removal respectively.

3.2. *Optimization of ANN Model Parameters*

To determine the final ANN model parameters, the consideration of MSE value and R value between experimental and predicted values are used [48]. The MSE value obtained from using 12 neurons is 0.48, 0.43 and 0.0000206 for the prediction of COD removal, purity of methane, and methane yield respectively. The R value achieved is 0.998, 0.983 and 0.990 for the prediction of COD removal, purity of methane, and methane yield respectively, which indicates good fitting between the predicted values and experimental data. As shown in Figure 2a, an ANN architecture of 6-12-3 were developed where bias for the hidden layer and output layer is represented by b1 and b2 respectively.
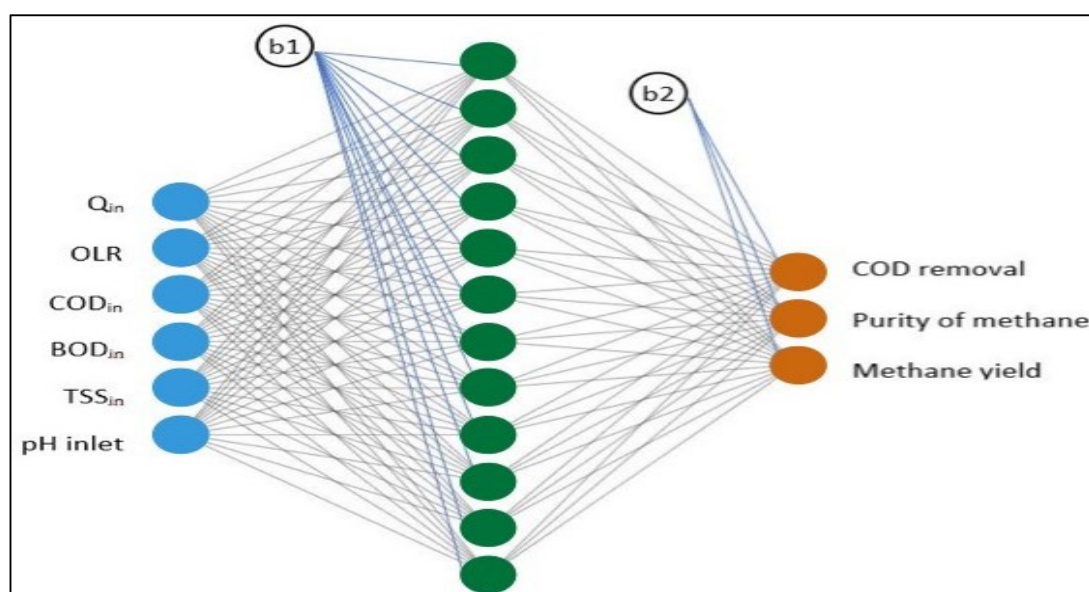
(**a**)

Figure 2. *Cont.*
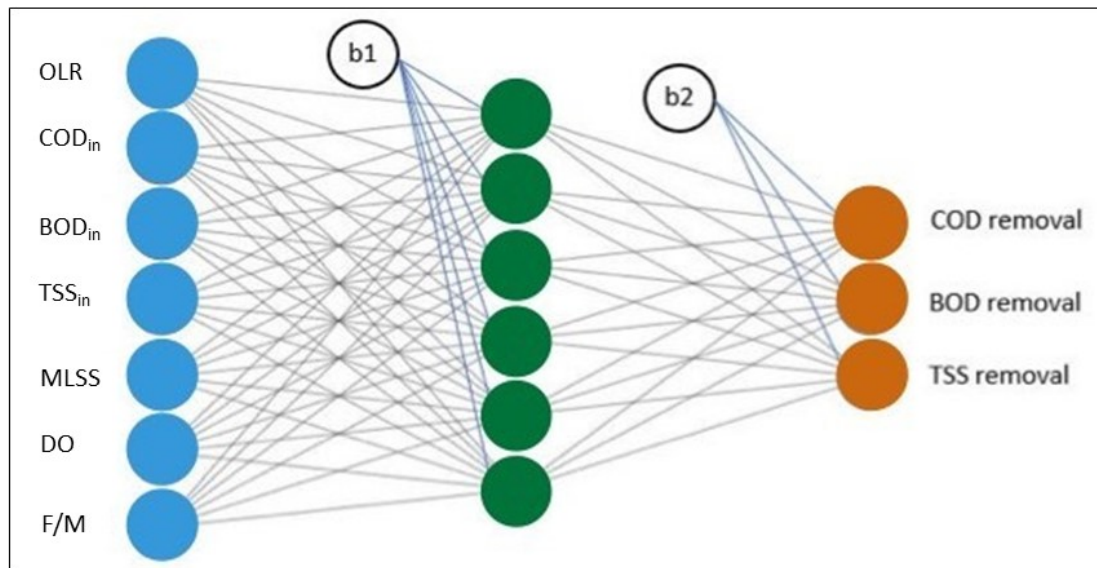
(**b**)

Figure 2. Neural network architecture for (**a**) anaerobic digestion, (**b**) aerobic process.

For aerobic process, the MSE value obtained from using 6 neurons is 1.0288, 0.6047 and 0.5560 for the prediction of COD removal, purity of methane, and methane yield respectively. The R value achieved is 0.997, 0.981 and 0.995 for the prediction of COD removal, purity of methane, and methane yield respectively. The high value of R and low value of MSE using 6 hidden neurons indicates good fitting and minimal deviation between the predicted values and experimental data. As shown in Figure 2b, an ANN architecture of 7-6-3 were developed where bias for the hidden layer and output layer is represented by b1 and b2 respectively.

3.3. Validation of the ANN Model

The illustration in Figures 3–8 depicted the correlations and corresponding visual agreement between the experimental data and the BR-ANN output. The proposed BR-ANN model demonstrated very satisfactory performance in predicting the COD removal, methane yield, and methane purity for anaerobic process and COD, BOD and TSS removals for aerobic process.

The R value for predicting the value of methane purity for each hidden neuron can be found in Table A7. A high R value close to 1 indicates a good accuracy of the COD removal. It can be seen that by using 12 hidden neurons, the highest R value of 0.998 can be achieved compared to other combinations. The R^2 value of 0.997 was obtained for predicting the COD removal using 12 hidden neurons and BR training algorithm (Figure 3a). The result is comparable with the neural network obtained from Dibaba et al. [49] where he obtained a $R^2 = 0.906$ and $R^2 = 0.87$ obtained from Antwi et al. [50] on the removal of COD during up flow anaerobic sludge blanket bioreactor (UASB) operation. Therefore, it can be seen from Figure 3b that the predicted value obtained from the ANN model and experimental value of COD removal has a slight deviation, proving the capability of the ANN model to predict the outcome of anaerobic digestion.

The R value for predicting the value of methane purity for each hidden neuron can be found in Table A11. The R^2 value obtained from the BR trained ANN model with 12 hidden neurons is 0.972 (Figure 4a), which is slightly lower than the $R^2 = 0.985$ achieved by Yu, Jaroenpoj and Griffith [51] using 5 input neurons with 8 hidden neurons. Similar to COD removal, the predicted value obtained from the ANN model and experimental value of methane purity has slight deviation only (Figure 4b).
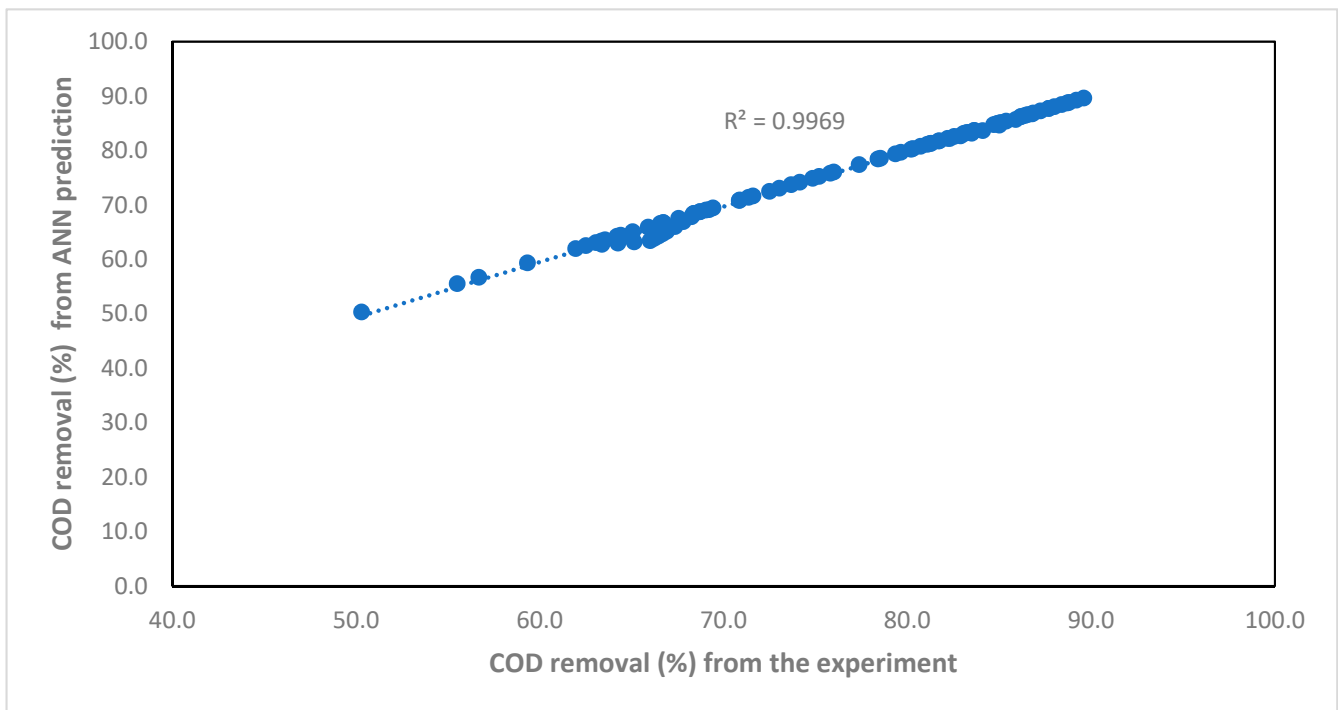
(**a**)
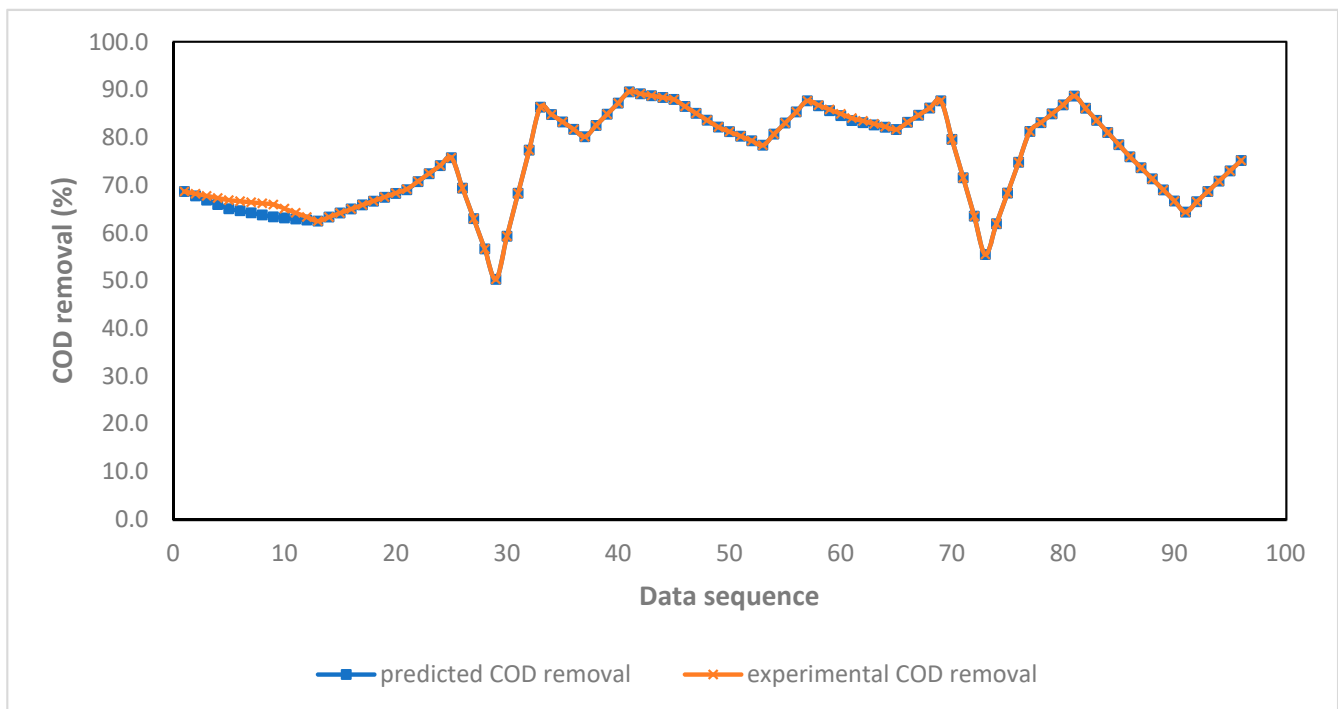
(**b**)

Figure 3. (**a**) Correlation between experimental data and predicted value of COD removal. (**b**) Prediction value and measured data by ANN model for COD removal of anaerobic digestion.
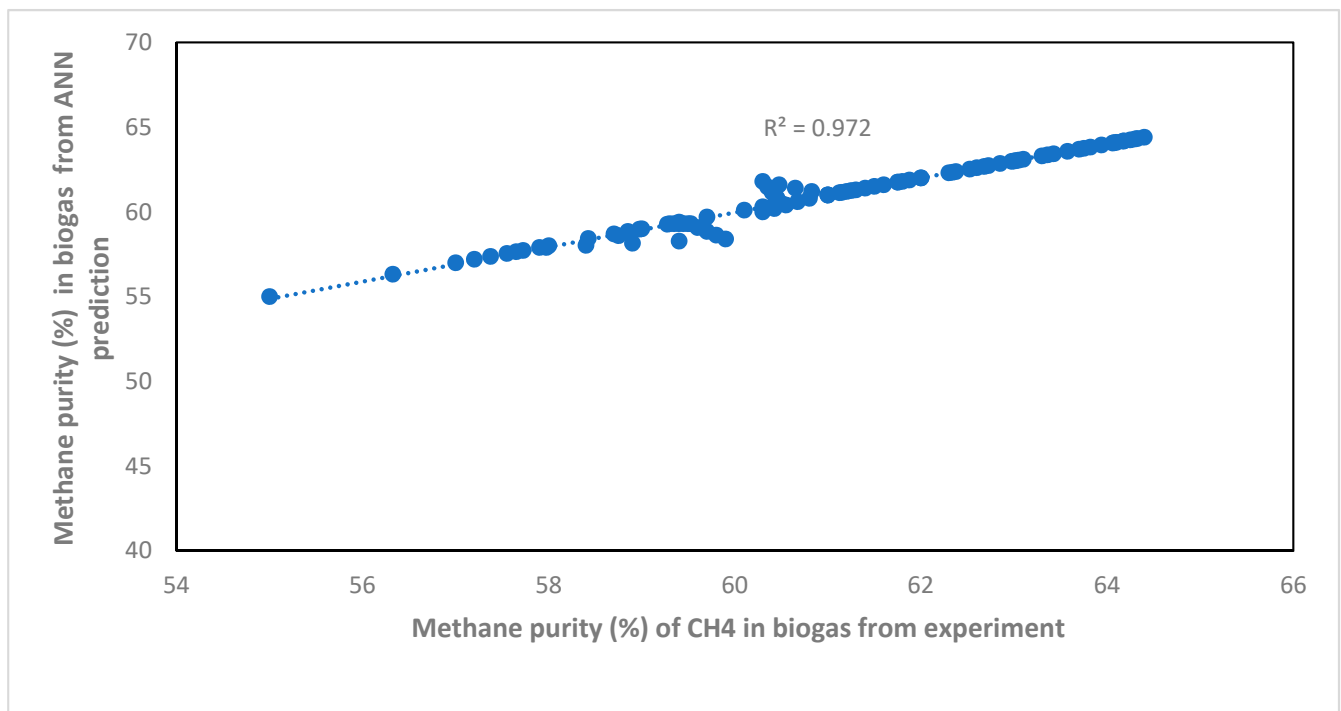
(**a**)
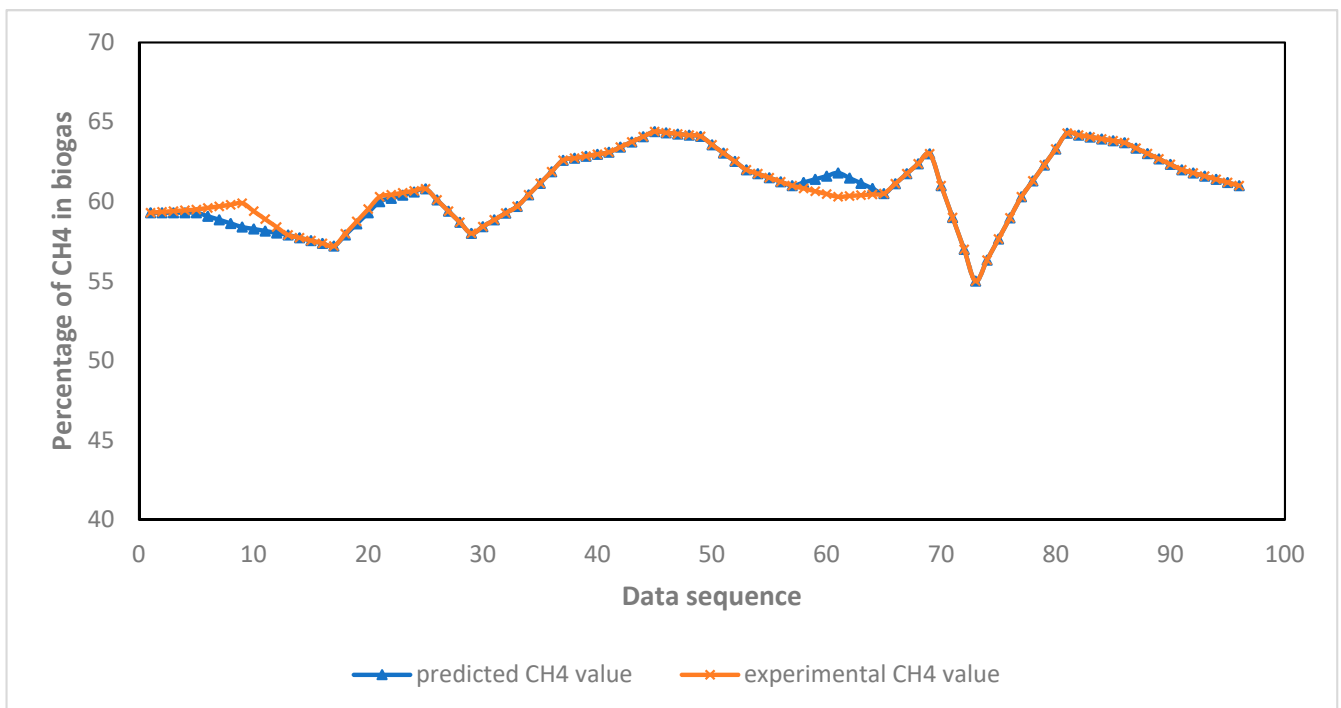
(**b**)

Figure 4. (**a**) Correlation between experimental data and predicted value of percentage of CH4 in biogas. (**b**) Prediction value and measured data by ANN model for percentage of CH4 in biogas.
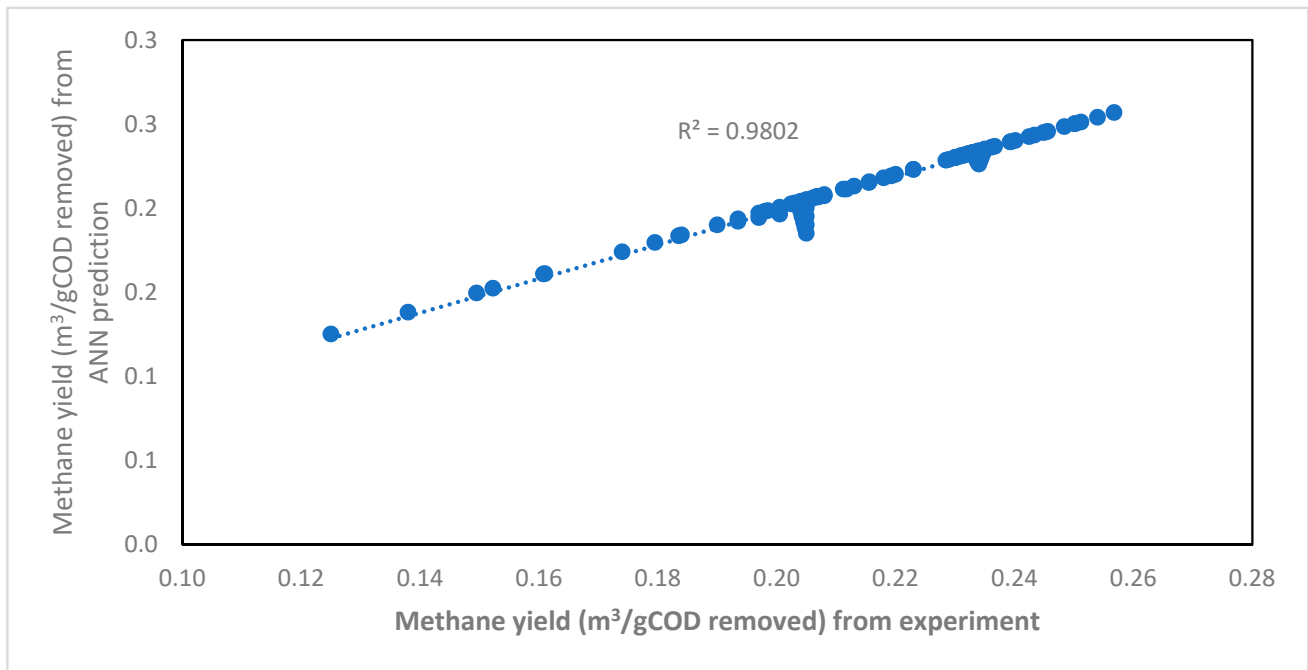
(**a**)
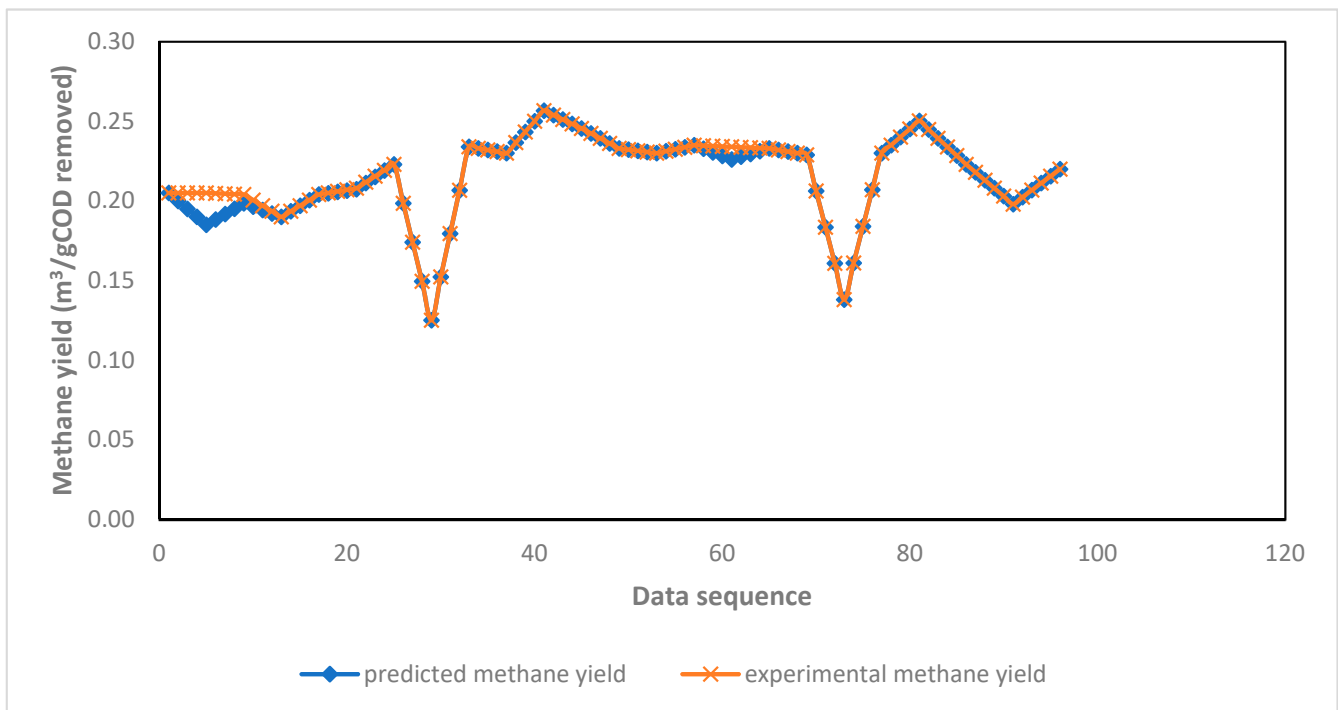
(**b**)

Figure 5. (**a**) Correlation between experimental data and predicted value of methane yield, (**b**) Prediction value and measured data by ANN model for methane yield (anaerobic digestion).
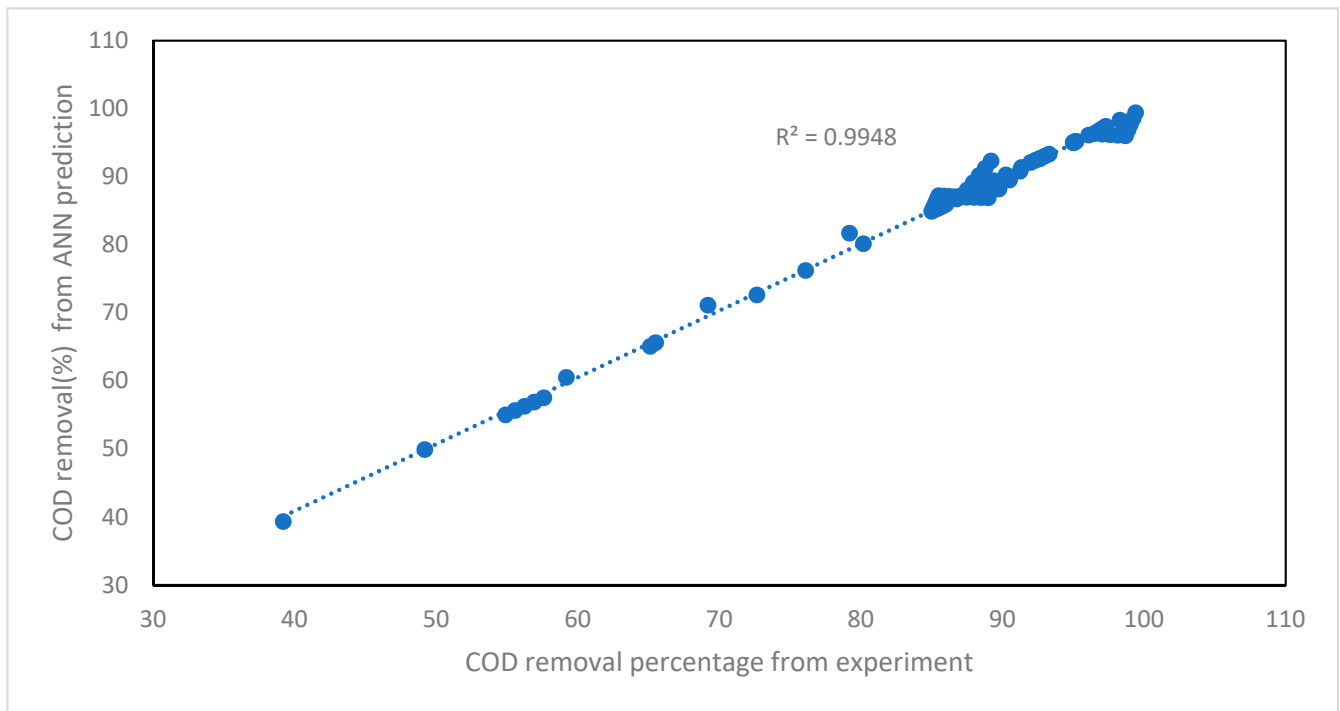
(**a**)
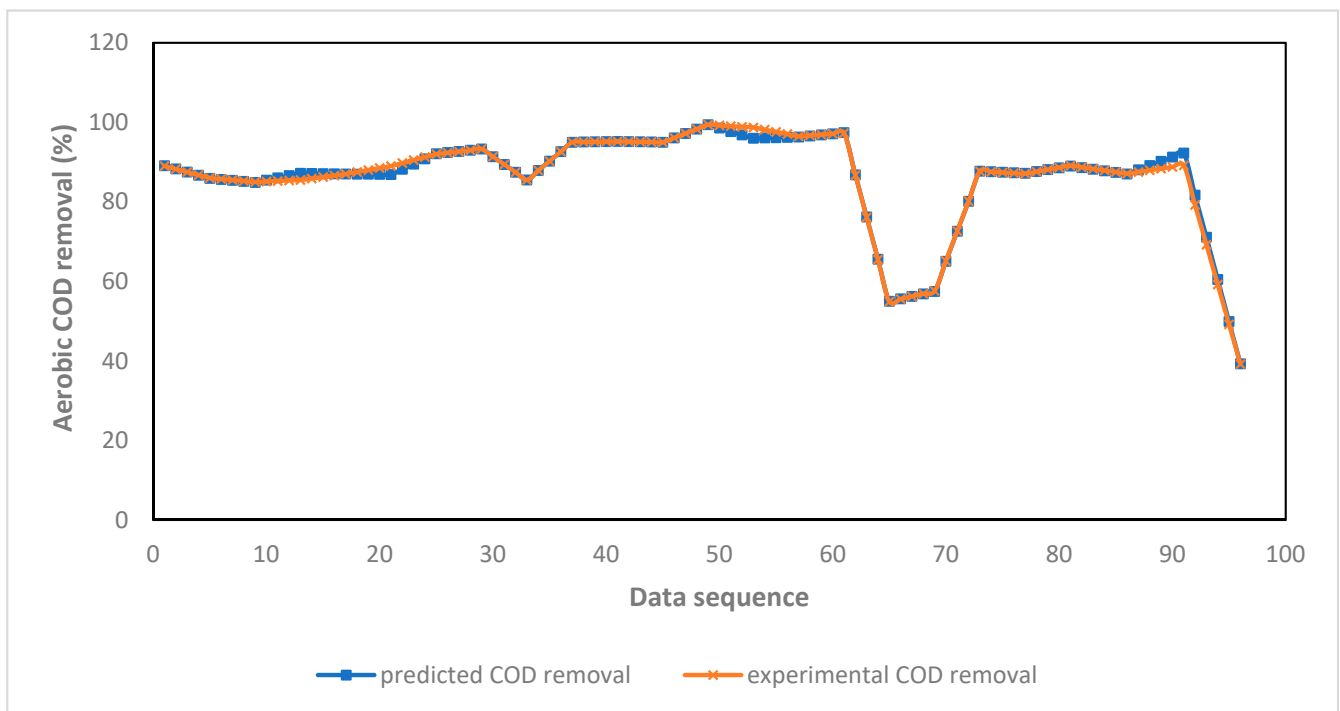
(**b**)

Figure 6. (**a**) Correlation between experimental data and predicted value of COD removal. (**b**) Prediction value and measured data by ANN model for COD removal (aerobic process).
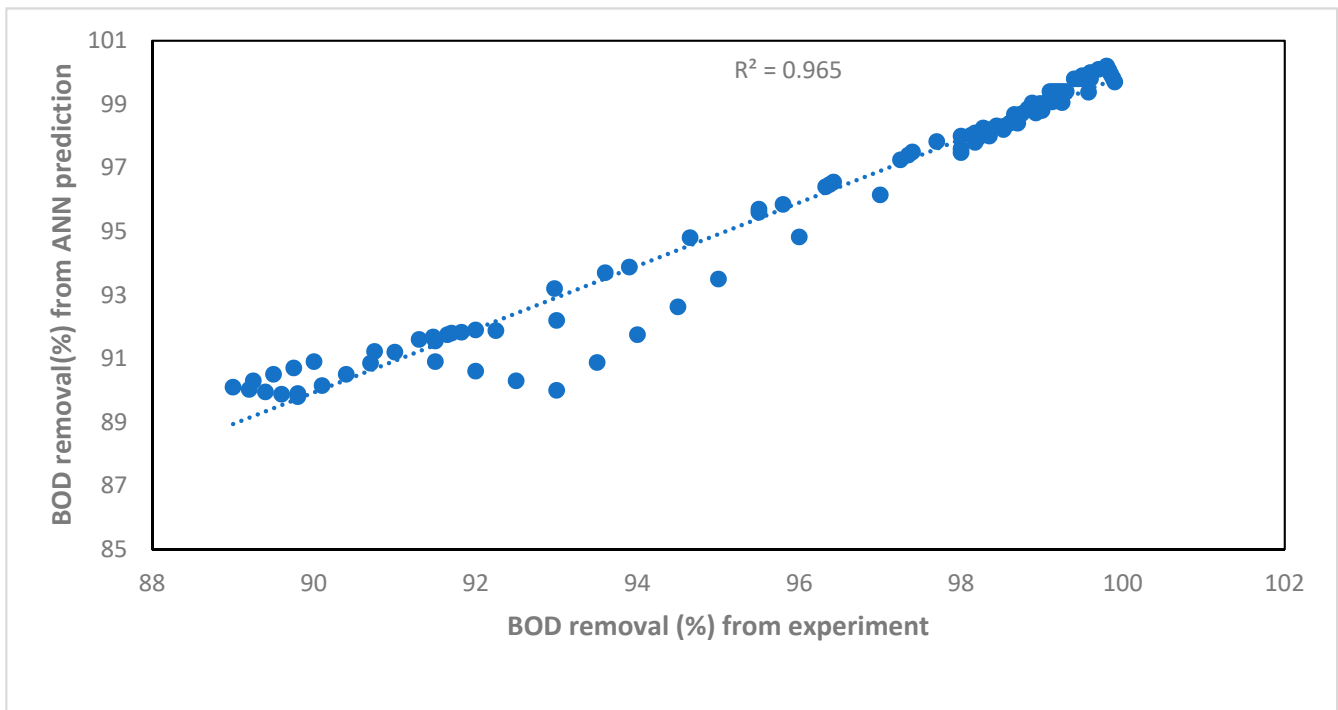
(**a**)
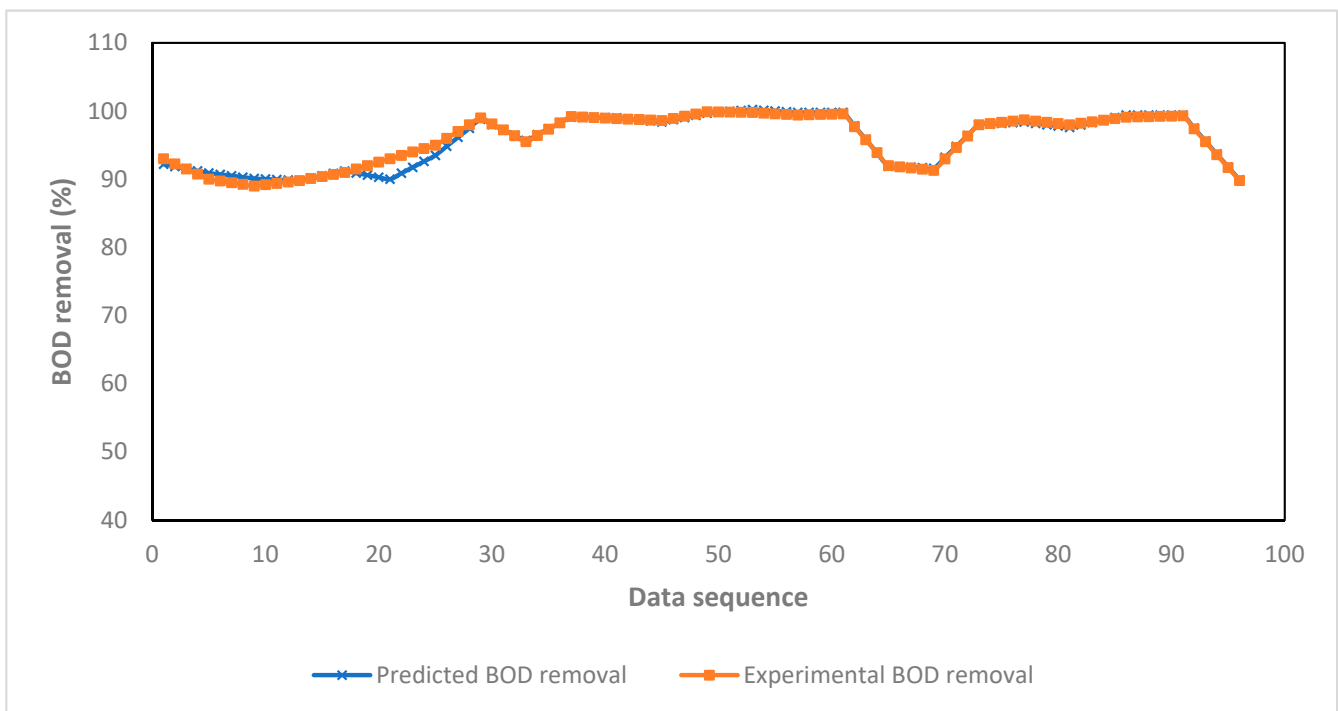
(**b**)

Figure 7. (**a**) Correlation between experimental data and predicted value of BOD removal. (**b**) Prediction value and measured data by ANN model for BOD removal (aerobic process).
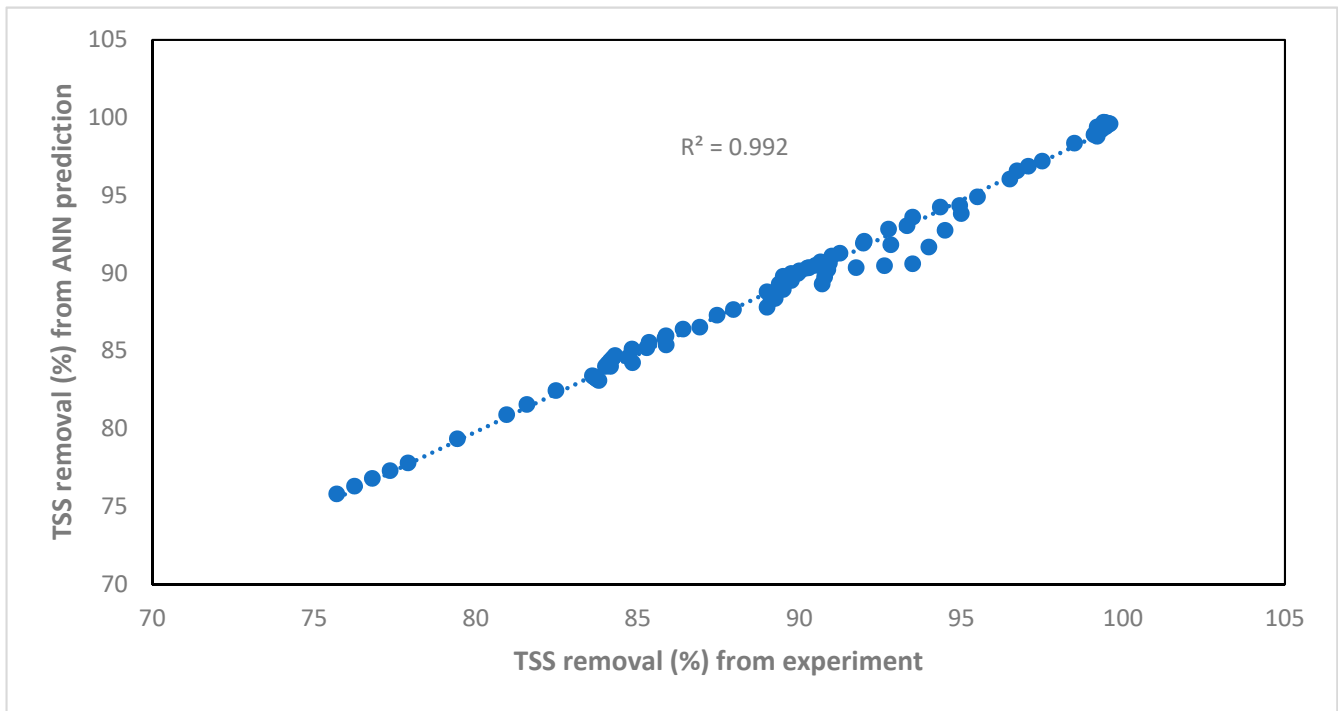
(**a**)
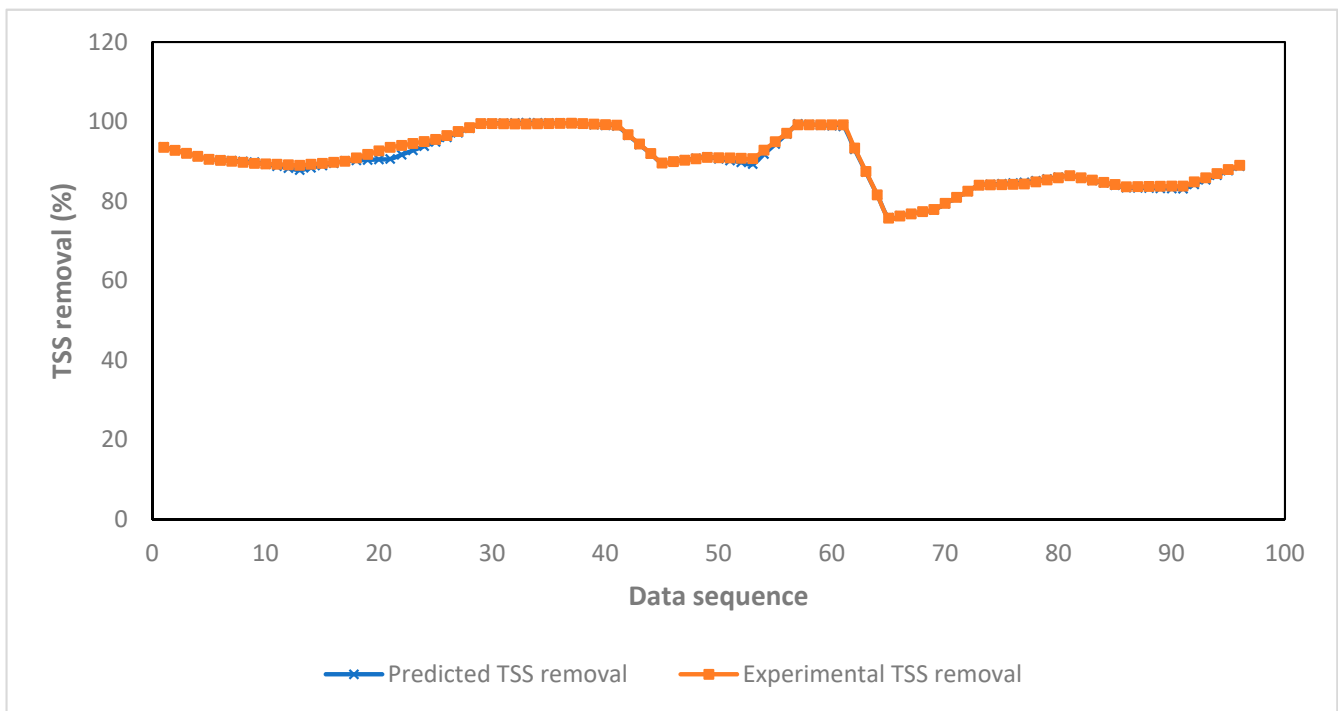
(**b**)

Figure 8. (**a**) Correlation between experimental data and predicted value of TSS removal. (**b**) Prediction value and measured data by ANN model for TSS removal (aerobic process).

The R value for predicting the value of methane purity for each hidden neuron can be found in Table A15 (Appendix A). The BR algorithm produced the best performance ANN model for predicting the methane yield with R value of 0.990, and lowest MAE, MAPE, and MSE value at 0.001, 0.002, 2.06×10^{-5} respectively (Figure 5). The results are better

with the ANN developed by Xu, Wang and Li [52], which achieve a R value of 0.937 using 8 input neurons and 6 hidden neurons to predict methane yield from anaerobic digestion of plant biomass.

For the ANN model to predict the COD removal, BOD removal, and TSS removal of aerobic process, two of the best training algorithm based on the simulation conducted on anaerobic was chosen to train the dataset, which is BR and LM algorithm. The values of the performance criteria such as R, MAE, MAPE, and MSE value can be found in Tables A19–A21.

The overall performance of BR is better than LM in terms of R, MAE, MAPE, and MSE values across the 3 output parameters. BR training algorithm with 6 hidden neurons was chosen as it gives overall lower MAE, MAPE and MSE values at 0.495, 0.005 and 1.0288 respectively for the prediction of COD removal. As shown in Figure 6, the R^2 value of 0.995 indicates a good fitting between the experimental COD value and the predicted COD value.

The high value of determination coefficient, R^2 value (Figure 7a) indicates that the ANN model is able to learn the non-linear relationship between the input and output parameters well. In Figure 7b, minimal deviation can be observed between the experimental BOD removal and predicted BOD removal values.

High R^2 value and minimal deviation between the predicted TSS removal and experimental TSS removal was observed in Figure 8 respectively. While all three R^2 values for COD, BOD and TSS removals are more than 0.9, this shows that ANN model can be used to predict the output parameters accurately.

3.4. Sensitivity Analysis

The connection weight approach and Garson method are used to perform the sensitivity analysis and the results are summarized in Tables 2 and 3 respectively. It can be seen from Table 2 that the inflow of biochemical oxygen demand (BOD_{in}) and organic loading rate (OLR) are the least influential parameters on the COD removal process, while the value of inlet chemical oxygen demand (COD_{in}), inlet flowrate (Q_{in}), and pH inlet, and inlet total suspended solid (TSS) have more influence towards the COD removal process in anaerobic digestion. Besides, COD_{in} is the most influential factor in determining the COD removal and the methane yield according to Table 2. The methane yield in anaerobic digestion is known to increase with increasing COD strength [53]. The input parameters strength of influence on purity of methane is as follows: Q_{in} > COD_{in} > pH inlet > BOD_{in} > TSS_{in} > OLR. As for methane yield, COD_{in} ranked the highest in terms of relative importance. The organic waste is the source of the production of methane [54], where level of COD is high, hence COD_{in} is one of the more important chemical properties that could impact the methane yield.

As for aerobic process, the efficiency of COD removal is most affected by mixed liquor suspended solids (MLSS) value, and least affected by F/M ratio. Basim [55] showed that increasing the quantity of MLSS would lead to higher concentration of COD in the sludge, increasing the COD removal efficiency. For output parameter of BOD removal (%), concentration of DO is the most influential parameter, while F/M ratio is the least influential parameter relatively. The amount of DO required by the aerobic microorganisms is the value of biochemical oxygen demand (BOD) [56]. Therefore, there is a strong relationship between the two parameters. Finally, the MLSS is the most influential parameter on the removal efficiency of TSS, and BOD_{in} is the least influential parameter in aerobic process.

However, it is worth noting that the importance of TSS and pH inlet is the highest in anaerobic digestion when using the Garson method (Table 3). The deviation between the two sensitivity analysis methods can be explained by the poor estimation of Garson algorithm. According to Olden, Joy and Death [57], the connection weight approach is superior to the Garson algorithm, where the mean similarity between the estimated variable and true ranks was 92% as compared to less than 50% for Garson's algorithm.

Table 2. Influence of input factors on anaerobic process using connection weight approach.

Reaction	Output	Input	Connection Weights	Rank
Anaerobic reaction	COD removal	Q_{in}	0.726	2
		OLR	0.491	5
		COD_{in}	0.752	1
		BOD_{in}	0.320	6
		TSS_{in}	−0.558	4
		pH inlet	−0.668	3
	Purity of methane	Q_{in}	1.001	1
		OLR	−0.094	6
		COD_{in}	−0.568	2
		BOD_{in}	0.380	4
		TSS_{in}	0.263	5
		pH inlet	−0.559	3
	Methane yield	Q_{in}	0.466	3
		OLR	−0.332	4
		COD_{in}	0.592	1
		BOD_{in}	0.135	6
		TSS_{in}	−0.209	5
		pH inlet	−0.575	2
Aerobic process	COD removal	OLR	1.892	5
		COD_{in}	−1.146	6
		BOD_{in}	2.454	4
		TSS_{in}	3.256	3
		MLSS (mg/L)	−4.998	1
		DO (mg/L)	4.020	2
		F/M (kgCOD/kg MLVSS. Day)	0.497	7
	BOD removal	OLR	3.197	3
		COD_{in}	−2.482	6
		BOD_{in}	3.877	2
		TSS_{in}	2.656	4
		MLSS (mg/L)	−2.630	5
		DO (mg/L)	4.293	1
		F/M (kgCOD/kg MLVSS. Day)	−0.364	7
	TSS removal	OLR	−0.712	5
		COD_{in}	−1.215	3
		BOD_{in}	0.197	7
		TSS_{in}	−1.193	4
		MLSS (mg/L)	2.400	1
		DO (mg/L)	2.198	2
		F/M (kgCOD/kg MLVSS. Day)	−0.558	6

Table 3. Relative importance of input variables for anaerobic process using Garson method.

Reaction	Output	Input		Importance (%)
		Q_{in}		18.0
		OLR		18.0
	COD removal	COD_{in}		12.8
		BOD_{in}		12.5
		TSS_{in}		20.3
		pH inlet		18.4
			Total:	100
		Q_{in}		15.8
		OLR		19.8
	Purity of methane	COD_{in}		13.1
Anaerobic digestion		BOD_{in}		12.6
		TSS_{in}		16.4
		pH inlet		22.3
			Total:	100
		Q_{in}		17.6
		OLR		17.2
	Methane yield	COD_{in}		13.3
		BOD_{in}		12.6
		TSS_{in}		20.3
		pH inlet		19.0
			Total:	100
		OLR		13.2
		COD_{in}		12.0
		BOD_{in}		13.8
	COD removal	TSS_{in}		9.0
		MLSS		22.0
		DO		20.6
		F/M		9.4
			Total:	100.0
		OLR		14.0
		COD_{in}		12.3
		BOD_{in}		13.6
Aerobic process	BOD removal	TSS_{in}		7.9
		MLSS		24.0
		DO		19.1
		F/M		9.1
			Total:	100
		OLR		15.1
		COD_{in}		13.1
		BOD_{in}		12.4
	TSS removal	TSS_{in}		8.4
		MLSS		24.4
		DO		15.7
		F/M		10.9
			Total:	100.0

3.5. Optimum Operating Conditions

The developed ANN model for anaerobic digestion is used to create the optimal operating conditions which maximize the output parameters. Following the optimal values, the COD removal (%) increases from 68.7% to 92% as shown in Table 4. The maximum methane yield of 0.26 $LCH_4/gCOD_{removed}$ was able to achieve from 0.23 $LCH_4/gCOD_{removed}$, which represent a 13.4% increase in the methane yield as shown in Table 5.

Table 4. Optimal operating condition for maximum COD removal (%) from anaerobic digestion.

Parameters	Optimal Values
Q_{in} (m^3/h)	153.81
OLR (g COD/L day)	7.00
COD_{in} (mg/L)	97,200.06
BOD_{in} (mg/L)	49,090.82
TSS_{in} (mg/L)	37,399.8
pH inlet	4.54

Table 5. Optimal operating condition for maximum methane yield from anaerobic digestion.

Parameters	Optimal Values
Q_{in} (m^3/h)	154.27
OLR (g COD/L day)	7.00
COD_{in} (mg/L)	50,500.08
BOD_{in} (mg/L)	25,250.07
TSS_{in} (mg/L)	29,200.28
pH inlet	4.00

The optimum operating value is shown below in Table 6, where aerobic COD removal (%) increases from 85% to 99% using the developed ANN model. The recommended operating conditions on IAAB based on the literature study are summarized in Table 7. The operating conditions based on ANN model were within the range of recommended operating condition provided by Chan et al. (2020). The purity of methane obtained from Chan et al. [58] were 63%, while ANN based operating condition were able to achieved purity of 66.4%. This indicates that higher purity of methane is possible for IAAB configuration. Finally, the most important parameter that determines the quality of the final discharge waste is the COD level. Hence, it is important to achieve maximum COD removal (%) in the aerobic process in this IAAB plant. The developed ANN model predicted the value of COD removal (%) of 99% based on the operating condition in Table 6. Similar results (85–99.6% COD removal) were achieved from literature [58].

Table 6. Optimum operating condition for maximum COD removal (%) from aerobic process.

Parameters	Optimal Values
OLR (g COD/L day)	5.8
COD_{in} (mg/L)	30,351.00
BOD_{in} (mg/L)	19,103.93
TSS_{in} (mg/L)	17,971.41
MLSS (mg/L)	37,007.56
DO (mg/L)	7.27
F/M (kg COD/kg MLVSS • Day)	0.2776

Table 7. Recommended operating condition for anaerobic digestion and aerobic process [58].

Operating conditions	Anaerobic	Aerobic
OLR (g COD/L day)	0–20	0–9.5
HRT (day)	4.59–27.7	4.1–22.7
MLSS (mg/L)	9000–49,600	9000–40,500
DO (mg/L)	-	≥ 2
pH	6.5–7.4	7.5–8.5

In summary, COD_{in} of raw POME (the most influential parameter in anaerobic process) needs to be closely monitored and maintained in the range of 50,000–97,000 mg/L to achieve high COD removal and methane yield based on the results presented in the sensitivity analysis (Section 3.4). For aerobic process, it is essential to maintain the most influential

parameters, i.e., MLSS and DO at the ranges of 37,000–40,500 mg/L and 2.0–7.3 mg/L respectively to ensure the compliance with the discharge limit. These show that the synergism created between anaerobic and aerobic processes is the key to achieve maximum biogas production and overall COD removal in the IAAB plant.

4. Conclusions

Two ANN models were developed to predict the effluent parameters of anaerobic digestion and aerobic processes in a pre-commercialized IAAB. A total of 6 input parameters that is Q_{in}, OLR, COD_{in}, BOD_{in}, and TSSin and pH inlet were used to predict the 3 output parameters for anaerobic digestion. For aerobic process, 7 input parameters that are OLR, COD_{in}, BOD_{in}, TSS_{in}, MLSS, DO and F/M ratio were used to predict 3 output parameters for aerobic process. The R value obtained for COD removal (%), purity of methane (%) and methane yield were 0.998, 0.983, and 0.990 respectively for anaerobic digestion. As for aerobic process, accurate prediction for COD removal, BOD removal and TSS removal was obtained with high R value of 0.997, 0.981, and 0.995 respectively. The ANN architecture used for anaerobic reaction was 6-12-3 and 7-6-3 for aerobic process with both using BR training algorithm and tan-sig activation function. The developed ANN models successfully predicted the properties of effluent for the IAAB plant accurately with minimal errors. The developed ANN model is used to optimize the output parameters. Under optimum operating conditions, the anaerobic COD removal (%) and methane yield were successfully improved by 33.9% and 13.4% respectively. Sensitivity analysis shows that COD inlets the most influential input parameters that affect the anaerobic COD removal (%) and methane yield. The trained ANN model can be utilized as a decision support system (DSS) for operators to predict the behavior of the IAAB system. Therefore, it will enable user to perform cost analysis using the ANN model to achieve optimum performance of the IAAB system, which can bring one step closer to successful commercialization of this IAAB technology. For further studies, input parameters such as the mixing time, reaction time, nutrients (ratio C:N:P) and concentration of microorganisms can be used to further develop the ANN model.

Author Contributions: Conceptualization, W.-Y.C.; methodology, C.S.L. and J.W.L.; software, W.-Y.C.; validation, M.M.; formal analysis, C.-D.H.; investigation, A.U.; resources, G.L.; data curation, H.H. and W.-N.T.; writing—original draft preparation, W.-Y.C.; writing—review and editing, Y.J.C.; visualization, W.-Y.C.; supervision, Y.J.C.; project administration, Y.J.C.; funding acquisition, J.W.L. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A.

Appendix A.1. Tables

Table A1. Input weights and biases of the ANN model for anaerobic digestion.

Position of Neurons	Weights (Wi)						
	Q_{in}	OLR	COD_{in}	BOD_{in}	TSS_{in}	pH Inlet	Bias (b1)
1	1.024	1.272	0.888	1.098	−0.962	−0.508	0.115
2	−0.186	−0.093	0.101	0.162	0.072	0.234	−0.224
3	−0.627	−0.517	−0.073	0.099	0.851	0.721	0.497
4	−0.253	−0.793	−0.902	−0.717	0.150	1.080	−0.482
5	−0.237	0.126	0.127	−0.082	−1.021	1.599	−0.095
6	0.979	0.100	−0.352	−0.397	−1.760	0.567	0.255
7	−0.196	−0.654	−1.175	−1.024	−0.655	1.396	0.233
8	0.205	0.211	−0.338	−0.379	0.307	−0.493	0.476
9	−1.154	−1.319	−0.189	−0.074	0.527	0.516	−0.981
10	−0.708	−0.976	−0.560	−0.472	−0.145	−0.142	0.711
11	−0.140	0.124	−0.275	−0.279	−0.464	−1.125	−0.515
12	0.811	0.624	−0.171	−0.124	−0.659	0.274	−0.308

Table A2. Output weights and biases of the ANN model for anaerobic digestion.

Outputs	Weights (Wi)												
	W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	W_9	W_{10}	W_{10}	W_{12}	Bias (b2)
COD removal (%)	−1.063	0.429	1.101	−0.703	−0.120	1.323	−0.969	−0.919	−1.369	0.268	0.140	−0.356	−0.582
Methane purity	1.264	0.305	0.734	1.137	−1.341	0.053	0.194	−0.322	−0.805	1.082	−0.448	0.025	−0.642
Methane yield (m³CH₄/gCOD removed)	−0.681	0.022	−0.423	0.088	−0.078	1.028	−1.140	−0.306	−0.788	0.716	−0.929	−0.979	−0.515

Table A3. Best performance result of the ANN model for anaerobic digestion.

% of Separation	Number of Sample	Type of Sample	MSE
70%	16	Training	2.26×10^{-14}
15%	4	Validation	1.246496
15%	4	Testing	N/A
100%	24	All	-

Table A4. Input weights and biases of the ANN model for aerobic process.

Position of Neurons	Weights (Wi)							
	OLR	CODin	BODin	TSSin	MLSS (mg/L)	DO (mg/L)	F/M (kgCOD/kg MLVSS. Day)	Bias (b1)
1	−2.047	0.130	−1.591	−0.541	0.631	0.595	−0.882	−1.786
2	−1.069	0.913	−0.580	0.076	−1.382	−0.366	0.373	−1.803
3	0.149	−0.504	0.333	−1.105	1.741	−2.891	−0.174	−1.809
4	−0.021	0.591	0.022	0.758	−1.117	0.761	−0.692	0.722
5	0.024	1.695	0.313	−0.086	−1.043	0.041	0.665	0.793
6	0.875	−0.339	−1.303	0.010	1.688	0.684	0.585	0.510

Table A5. Output weights and biases of the ANN model for aerobic process.

Outputs	Weight (Wi)						
	W_1	W_2	W_3	W_4	W_5	W_6	Bias (b2)
COD removal	−1.67677	0.391637	−2.03913	−0.07522	−1.51773	−0.8974	−0.35335
BOD removal	−1.40607	−1.66764	−1.98104	−0.33949	−1.19136	−1.31442	−1.69595
TSS removal	1.159768	−2.19573	−0.8436	−1.68153	0.573482	−0.69684	−0.45734

Table A6. Best performance result of the ANN model for aerobic process.

% of Separation	Number of Sample	Type of Sample	MSE
70%	16	Training	0.115
15%	4	Validation	3.804727
15%	4	Testing	N/A
100%	24	All	-

Table A7. R value for COD removal (%) for different training algorithms at different number of hidden neurons.

Hidden Neurons	Training Algorithm											
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	0.823	0.657	0.759	0.767	0.820	0.773	0.732	0.775	0.770	0.523	0.762	0.733
2	0.940	0.922	0.930	0.900	0.925	0.938	0.912	0.935	0.900	0.795	0.962	0.635
3	0.953	0.863	0.930	0.930	0.909	0.901	0.915	0.928	0.938	0.565	0.958	0.658
4	0.908	0.943	0.951	0.896	0.903	0.884	0.931	0.895	0.886	0.534	0.962	0.649
5	0.964	0.891	0.948	0.935	0.935	0.929	0.967	0.921	0.908	0.721	0.988	0.651
6	0.953	0.832	0.943	0.891	0.941	0.939	0.924	0.946	0.897	0.618	0.981	0.773
7	0.973	0.924	0.918	0.826	0.962	0.914	0.960	0.904	0.954	0.393	0.994	0.634
8	0.899	0.950	0.903	0.906	0.912	0.924	0.936	0.921	0.962	0.452	0.990	0.542
9	0.953	0.917	0.852	0.872	0.899	0.918	0.901	0.940	0.927	0.531	0.994	0.625
10	0.921	0.762	0.863	0.944	0.936	0.904	0.946	0.873	0.922	0.574	0.987	0.698
11	0.945	0.848	0.940	0.874	0.862	0.899	0.946	0.902	0.881	0.734	0.994	0.568
12	0.916	0.924	0.924	0.880	0.930	0.863	0.910	0.922	0.922	0.682	**0.998**	0.419
13	0.960	0.819	0.930	0.903	0.880	0.849	0.964	0.882	0.830	0.599	0.989	0.711
14	0.973	0.878	0.888	0.914	0.960	0.928	0.905	0.889	0.900	0.510	0.995	0.871
15	0.964	0.821	0.933	0.931	0.894	0.928	0.883	0.885	0.904	0.498	0.996	0.574
16	0.933	0.824	0.887	0.907	0.906	0.911	0.954	0.841	0.846	0.554	0.989	0.615
17	0.959	0.891	0.956	0.818	0.928	0.795	0.919	0.905	0.935	0.626	0.995	0.543
18	0.955	0.880	0.909	0.928	0.972	0.946	0.945	0.963	0.783	0.603	0.998	**0.344**
19	0.944	0.910	0.923	0.884	0.967	0.892	0.910	0.955	0.943	0.346	0.992	0.495
20	0.968	0.873	0.921	0.840	0.930	0.935	0.928	0.895	0.905	0.499	0.997	0.693

Bolded number are the lowest and highest R value.

Table A8. MAE for COD removal (%) for different training algorithms at different number of hidden neurons.

Hidden Neurons	Training Algorithm											
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	5.208	0.026	4.252	4.403	5.083	4.442	5.338	4.740	4.819	7.765	4.123	4.973
2	3.224	3.386	3.248	3.797	3.716	2.843	3.758	3.333	4.105	29.537	2.098	18.369
3	12.120	3.730	3.140	3.111	3.524	3.853	3.531	3.285	2.890	15.200	1.616	9.097
4	2.487	3.071	2.970	3.263	4.319	4.150	2.985	3.497	4.189	10.127	1.057	19.697
5	2.119	3.668	2.931	3.372	3.260	3.351	2.047	2.611	3.665	37.161	0.656	23.856
6	1.981	4.685	2.948	3.829	2.337	2.797	2.668	2.828	4.045	34.257	0.834	9.163
7	1.345	3.238	3.529	4.521	2.331	3.031	2.349	3.569	1.939	50.897	0.453	11.841
8	2.320	2.243	3.802	3.058	2.952	2.905	2.657	3.143	2.295	40.457	0.501	15.996
9	1.880	3.135	4.485	4.250	4.225	3.290	3.294	3.008	2.555	49.306	0.476	14.285
10	2.004	5.698	4.169	2.488	3.124	3.494	2.146	3.763	3.135	15.346	0.741	6.472
11	2.132	3.821	1.961	2.839	3.695	3.822	1.982	2.877	3.206	34.591	0.434	9.620
12	1.850	2.769	3.516	3.171	2.648	3.366	3.274	2.918	2.670	27.077	**0.211**	27.139
13	1.476	3.947	2.105	3.805	4.438	4.195	1.869	3.096	2.781	32.379	0.509	11.513
14	1.320	3.029	3.929	2.710	2.245	2.352	2.505	3.755	3.744	23.346	0.416	5.579
15	1.690	4.082	2.656	1.918	2.808	3.195	3.229	2.735	2.554	32.215	0.358	11.765
16	2.286	3.770	3.555	2.533	3.784	3.000	1.962	4.028	3.515	17.154	0.644	23.895
17	2.152	3.646	2.431	3.573	2.869	4.038	2.861	2.435	2.818	51.956	0.377	7.989
18	1.140	3.748	3.107	2.687	1.528	2.428	1.902	1.888	4.336	41.445	0.220	30.201
19	2.307	3.223	2.828	3.514	2.118	3.643	3.570	2.399	2.661	**95.015**	0.462	30.494
20	1.454	3.149	2.186	3.469	2.399	2.721	2.535	3.099	2.803	32.193	0.238	31.910

Bolded number are the lowest and highest MAE value.

Table A9. MAPE for COD removal (%) for different training algorithms at different number of hidden neurons.

| Hidden Neurons | Training Algorithm | | | | | | | | | | | |
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	0.07	0.1	0.06	0.06	0.07	0.06	0.08	0.07	0.07	0.11	0.066	0.07
2	0.04	0.05	0.04	0.05	0.05	0.04	0.06	0.05	0.06	0.41	0.034	0.27
3	0.16	0.05	0.04	0.04	0.05	0.05	0.05	0.04	0.04	0.21	0.025	0.13
4	0.04	0.04	0.04	0.05	0.06	0.06	0.04	0.05	0.06	0.15	0.022	0.26
5	0.03	0.05	0.04	0.05	0.04	0.05	0.03	0.03	0.05	0.51	0.014	0.33
6	0.03	0.07	0.04	0.05	0.03	0.04	0.04	0.04	0.06	0.47	0.016	0.13
7	0.02	0.04	0.05	0.06	0.03	0.04	0.03	0.05	0.02	0.68	0.011	0.17
8	0.03	0.03	0.05	0.04	0.04	0.04	0.04	0.04	0.03	0.55	0.017	0.22
9	0.02	0.04	0.06	0.06	0.06	0.04	0.05	0.04	0.03	0.68	0.018	0.19
10	0.03	0.08	0.06	0.03	0.04	0.05	0.03	0.05	0.04	0.21	**0.009**	0.09
11	0.03	0.05	0.03	0.03	0.05	0.05	0.03	0.04	0.05	0.48	0.006	0.14
12	0.02	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.39	0.003	0.37
13	0.02	0.06	0.03	0.05	0.06	0.06	0.02	0.05	0.04	0.45	0.006	0.16
14	0.02	0.05	0.05	0.03	0.03	0.03	0.03	0.05	0.05	0.32	0.006	0.07
15	0.02	0.06	0.04	0.03	0.04	0.04	0.04	0.04	0.03	0.46	0.005	0.15
16	0.03	0.06	0.05	0.03	0.05	0.04	0.02	0.05	0.05	0.23	0.009	0.32
17	0.03	0.05	0.03	0.05	0.04	0.06	0.04	0.04	0.04	0.71	0.005	0.12
18	0.01	0.05	0.04	0.03	0.02	0.03	0.03	0.02	0.06	0.56	0.003	0.39
19	0.03	0.04	0.04	0.05	0.03	0.05	0.05	0.03	0.04	**0.77**	0.006	0.42
20	0.02	0.04	0.03	0.05	0.03	0.04	0.03	0.04	0.03	0.44	0.003	0.42

Bolded number are the lowest MAPE value.

Table A10. MSE for COD removal (%) for different training algorithms at different number of hidden neurons.

| Hidden Neurons | Training Algorithm | | | | | | | | | | | |
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	39.37	72.07	50.14	51.13	41.25	49.97	55.26	47.20	48.22	97.30	52.79	59.20
2	14.22	18.03	16.17	23.44	19.69	15.09	26.29	15.00	22.39	924.88	9.30	426.35
3	13.95	30.34	16.70	16.75	21.25	22.39	22.69	16.65	15.44	321.74	10.81	125.63
4	21.99	16.24	11.56	25.80	23.18	30.01	16.00	24.50	26.58	185.53	9.61	477.13
5	9.69	26.75	12.21	15.85	14.75	17.01	7.80	18.43	21.22	1442.88	2.83	799.09
6	13.95	38.86	13.58	24.97	15.91	15.68	17.47	13.06	24.18	1257.65	5.13	112.85
7	7.90	20.31	20.34	37.70	9.08	20.92	10.47	23.30	10.98	2773.36	1.70	199.71
8	23.24	11.57	22.00	21.45	21.46	17.36	14.91	18.08	9.08	2060.30	2.53	441.06
9	12.52	19.59	32.82	31.10	25.68	19.67	25.95	14.15	16.82	2551.04	1.47	432.19
10	18.66	55.96	31.23	13.72	16.06	21.73	52.87	28.60	21.58	337.34	4.03	67.22
11	13.24	33.81	15.46	31.13	32.03	23.43	12.64	24.86	29.99	1251.08	1.45	156.54
12	21.69	17.52	19.06	27.49	16.14	30.96	20.70	18.28	17.75	1251.11	**0.43**	1106.73
13	11.95	40.31	16.34	23.61	43.13	33.02	8.37	27.08	39.50	1151.90	2.63	215.13
14	7.09	29.06	25.47	20.52	10.34	16.63	23.32	27.41	22.65	841.85	1.17	49.48
15	9.64	39.72	16.80	15.95	26.48	17.25	26.72	27.73	21.89	1571.34	1.05	259.51
16	17.96	38.19	25.85	23.69	24.83	20.15	11.35	38.12	38.28	390.73	2.88	831.74
17	11.02	25.80	10.52	42.45	17.09	47.26	19.29	22.07	15.62	2797.18	1.27	99.43
18	11.08	27.99	22.01	16.75	6.70	12.45	12.59	9.00	81.80	1933.79	0.59	982.16
19	13.09	20.65	19.75	26.84	7.95	26.28	21.74	10.74	13.23	**9769.54**	2.20	1096.88
20	8.43	30.78	18.04	39.58	19.50	15.82	19.11	23.55	23.10	1232.83	0.70	1261.02

Bolded number are the lowest MSE value.

Table A11. R value for CH4 percentage in biogas for different training algorithms at different number of hidden neurons.

Hidden Neurons	Training Algorithm											
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	0.571	0.653	0.565	0.542	0.568	0.529	0.600	0.517	0.517	0.523	0.569	0.531
2	0.810	0.710	0.720	0.794	0.810	0.764	0.809	0.745	0.789	0.795	0.824	0.691
3	0.901	0.852	0.801	0.761	0.836	0.755	0.890	0.779	0.807	0.565	0.904	0.424
4	0.909	0.739	0.802	0.846	0.813	0.835	0.856	0.798	0.814	0.534	0.915	0.634
5	0.864	0.745	0.834	0.816	0.783	0.813	0.852	0.845	0.711	0.721	0.989	0.454
6	0.901	0.826	0.785	0.769	0.865	0.845	0.857	0.799	0.823	0.618	0.989	0.671
7	0.831	0.849	0.804	0.785	0.841	0.828	0.790	0.810	0.843	0.393	0.985	0.714
8	0.936	0.737	0.713	0.821	0.846	0.870	0.816	0.633	0.827	0.452	0.992	0.684
9	0.907	0.806	0.808	0.775	0.763	0.770	0.863	0.734	0.872	0.531	0.984	0.707
10	0.954	0.724	0.812	0.826	0.883	0.816	0.861	0.632	0.852	0.574	0.993	0.573
11	0.909	0.744	0.900	0.870	0.716	0.703	0.920	0.889	0.847	0.734	0.985	0.428
12	0.971	0.767	0.751	0.770	0.884	0.717	0.933	0.879	0.788	0.682	0.983	0.704
13	0.964	0.693	0.887	0.804	0.772	0.772	0.873	0.815	0.848	0.599	**0.997**	0.560
14	0.965	0.655	0.811	0.773	0.859	0.812	0.908	0.850	0.634	0.510	0.984	0.637
15	0.975	0.749	0.923	0.912	0.812	0.732	0.892	0.822	0.789	0.498	0.977	0.571
16	0.952	0.817	0.817	0.808	0.644	0.687	0.850	0.778	0.872	0.554	0.990	0.601
17	0.931	0.752	0.731	0.842	0.890	0.784	0.883	0.907	0.809	0.626	0.984	0.606
18	0.949	0.780	0.874	0.741	0.821	0.705	0.852	0.759	0.890	0.603	0.982	0.581
19	0.907	0.868	0.775	0.907	0.887	0.697	0.838	0.746	0.619	0.346	0.976	0.574
20	0.908	0.608	0.876	0.824	0.820	0.595	0.804	0.815	0.791	0.499	0.991	**0.256**

Bolded number are the lowest and highest R value.

Table A12. MAE for percentage of CH4 in biogas for different training algorithms at different number of hidden neurons.

Hidden Neurons	Training Algorithm											
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	1.642	6.822	1.397	1.378	1.621	1.462	1.525	1.488	1.49	2.392	**1.337**	1.538
2	1.111	1.75	1.414	1.431	1.399	1.365	1.257	1.839	1.551	4.669	1.129	2.918
3	2.86	1.121	1.08	1.675	1.122	1.488	0.915	1.29	1.174	2.392	0.831	2.025
4	0.806	1.435	1.157	1.032	1.536	1.242	1.028	1.22	1.109	3.127	0.606	1.673
5	0.937	1.322	1.246	1.281	1.719	1.189	1.032	0.944	1.562	1.902	0.114	3.519
6	0.624	1.11	1.235	1.27	0.93	1.066	1.046	1.176	1.093	8.924	0.134	1.414
7	0.587	1.085	1.057	1.333	1.106	1.196	1.224	1.336	1.046	6.414	0.142	1.393
8	0.408	1.294	1.445	1.126	0.939	0.983	1.103	1.456	1.147	10.763	0.105	4.515
9	0.463	1.197	1.95	1.265	1.336	1.074	0.965	1.326	0.982	14.349	0.143	15.518
10	0.332	1.294	1.276	1.073	0.922	1.067	0.952	1.552	1.058	8.583	0.088	4.126
11	0.648	1.433	0.863	0.731	1.34	1.343	0.648	0.92	0.966	4.209	0.162	3.029
12	0.358	1.308	1.332	1.122	0.92	1.146	0.646	1.078	1.519	3.296	0.147	3.621
13	0.334	1.508	0.882	1.208	1.222	1.32	0.853	1.174	1.009	8.283	**0.072**	4.256
14	0.422	1.633	1.173	1.357	1.056	1.15	0.878	1.496	1.498	5.281	0.179	2.824
15	0.331	1.13	0.725	0.84	1.231	1.348	0.888	1.071	1.267	12.657	0.201	4.762
16	0.384	1.197	1.167	0.95	1.211	1.301	0.915	1.201	0.955	5.027	0.13	8.688
17	0.912	1.219	1.566	0.991	0.945	1.152	0.939	0.768	1.273	1.66	0.144	3.503
18	0.37	1.336	0.905	1.178	1.084	1.381	0.816	1.211	0.868	12.714	0.145	3.806
19	0.731	0.896	1.351	0.801	1.119	1.61	1.093	1.186	1.863	5.1	0.199	3.972
20	0.45	1.454	0.978	1.238	0.951	1.495	1.084	1.182	1.07	5.171	0.119	19.23

Bolded number are the lowest and highest MAE value.

Table A13. MAPE for percentage of CH4 in biogas for different training algorithms at different number of hidden neurons.

Hidden Neurons	Training Algorithm											
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	0.03	0.02	0.02	0.02	0.03	0.02	0.03	0.02	0.02	0.04	0.020	0.03
2	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.08	0.020	0.05
3	0.05	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.04	0.010	0.03
4	0.01	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.05	0.010	0.03
5	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.03	0.03	0.001	0.06
6	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.15	0.003	0.02
7	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.11	0.001	0.02
8	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.18	0.001	0.07
9	0.01	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.24	0.002	0.25
10	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.14	0.001	0.07
11	0.01	0.02	0.01	0.01	0.02	0.02	0.01	0.02	0.02	0.07	0.003	0.05
12	0.01	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.05	**0.001**	0.06
13	0.01	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.14	0.001	0.07
14	0.01	0.03	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.09	0.003	0.05
15	0.01	0.02	0.01	0.01	0.02	0.02	0.01	0.02	0.02	0.21	0.003	0.08
16	0.01	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.08	0.002	0.14
17	0.01	0.02	0.03	0.02	0.02	0.02	0.02	0.01	0.02	0.03	0.002	0.06
18	0.01	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.01	0.21	0.002	0.06
19	0.01	0.01	0.02	0.01	0.02	0.03	0.02	0.02	0.03	0.08	0.003	0.07
20	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.09	0.002	**0.32**

Bolded number are the lowest and highest MAPE value.

Table A14. MSE for percentage of CH4 in biogas for different training algorithms at different number of hidden neurons.

Hidden Neurons	Training Algorithm											
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	3.77	3.14	3.76	3.95	3.77	4.35	3.52	4.03	4.10	9.00	3.77	4.70
2	1.94	4.52	2.69	3.15	2.77	2.89	2.02	5.23	3.22	25.78	1.78	11.64
3	1.16	1.76	1.98	3.90	1.69	3.44	1.31	2.32	1.91	8.57	0.99	6.17
4	0.97	2.81	2.04	1.55	3.89	2.34	1.49	2.09	1.90	14.69	0.98	4.64
5	1.39	2.78	2.49	2.13	4.25	2.10	1.54	1.62	3.78	5.66	0.13	23.74
6	1.16	1.75	2.10	2.42	1.58	1.74	1.48	2.01	1.76	87.02	0.13	3.83
7	1.71	1.54	1.98	2.72	1.70	2.18	2.16	2.42	1.68	46.60	0.19	2.85
8	0.80	2.64	3.61	1.77	1.57	1.37	1.89	3.48	1.91	123.63	0.09	30.70
9	0.97	2.02	6.12	2.66	2.56	2.30	1.42	3.17	1.44	209.59	0.19	267.10
10	0.55	2.99	2.75	1.87	1.37	1.85	7.03	3.78	1.56	87.74	0.08	20.69
11	1.14	3.25	1.04	1.43	2.74	3.11	0.85	1.16	1.63	23.20	0.17	12.12
12	0.37	2.31	2.50	2.53	1.20	2.86	0.88	1.74	3.32	13.03	0.19	18.95
13	0.46	3.22	1.26	2.41	2.30	2.49	1.51	2.05	1.66	87.05	**0.03**	26.11
14	0.49	4.17	2.24	3.63	1.98	1.86	1.03	3.93	4.19	36.49	0.20	10.95
15	0.32	2.44	0.83	1.00	2.56	2.93	1.13	1.85	2.09	174.16	0.29	29.62
16	0.53	2.05	1.92	2.06	3.34	2.91	1.59	2.31	1.40	32.86	0.14	92.72
17	1.18	2.89	3.43	1.61	1.24	2.18	1.36	0.98	2.25	4.54	0.18	16.35
18	0.57	2.45	1.32	2.46	1.93	3.04	1.66	2.44	1.27	210.13	0.20	22.09
19	1.27	1.39	2.82	0.98	1.61	3.81	1.71	2.45	4.66	37.64	0.28	26.60
20	0.97	4.41	1.31	1.95	1.94	3.53	2.24	2.35	2.19	34.99	0.10	**398.13**

Bolded number are the lowest and highest MSE value.

Table A15. R value for methane yield in biogas for different training algorithms at different number of hidden neurons.

Hidden Neurons	Training Algorithm											
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	0.741	0.594	0.665	0.644	0.769	0.665	0.614	0.693	0.679	0.467	0.659	0.603
2	0.926	0.644	0.908	0.879	0.926	0.844	0.834	0.834	0.779	0.683	0.935	0.548
3	0.930	0.797	0.817	0.886	0.814	0.888	0.870	0.840	0.671	0.477	0.894	0.845
4	0.790	0.842	0.827	0.857	0.752	0.796	0.913	0.800	0.839	0.629	0.872	0.530
5	0.838	0.626	0.733	0.704	0.881	0.692	0.915	0.915	0.757	0.496	0.941	0.563
6	0.930	0.639	0.670	0.915	0.666	0.676	0.855	0.858	0.727	0.420	0.978	0.529
7	0.914	0.505	0.645	0.684	0.716	0.551	0.936	0.663	0.612	0.653	0.989	0.314
8	0.949	0.427	0.713	0.692	0.780	0.779	0.895	0.822	0.770	0.560	0.990	0.457
9	0.925	0.417	0.807	0.640	0.678	0.724	0.899	0.655	0.650	0.554	0.992	0.494
10	0.925	0.584	0.692	0.642	0.771	0.617	0.933	0.841	0.622	0.577	0.989	0.534
11	0.956	0.497	0.706	0.579	0.777	0.794	0.912	0.581	0.665	**0.267**	0.993	0.664
12	0.914	0.473	0.555	0.628	0.691	0.622	0.886	0.641	0.706	0.549	0.990	0.531
13	0.958	0.578	0.585	0.570	0.721	0.644	0.839	0.653	0.704	0.491	0.988	0.450
14	0.891	0.662	0.675	0.708	0.556	0.602	0.902	0.573	0.742	0.495	0.990	0.587
15	0.943	0.476	0.649	0.576	0.667	0.603	0.873	0.662	0.600	0.597	**0.996**	0.485
16	0.963	0.530	0.603	0.637	0.771	0.593	0.912	0.618	0.669	0.409	0.984	0.404
17	0.964	0.673	0.591	0.658	0.638	0.638	0.882	0.506	0.519	0.443	0.992	0.480
18	0.953	0.544	0.624	0.631	0.567	0.644	0.933	0.512	0.578	0.514	0.993	0.539
19	0.958	0.427	0.698	0.640	0.634	0.672	0.906	0.601	0.666	0.721	0.994	0.384
20	0.961	0.692	0.512	0.669	0.647	0.675	0.834	0.642	0.647	0.438	0.995	0.629

Bolded number are the lowest and highest R value.

Table A16. MAE for methane yield for different training algorithms at different number of hidden neurons.

Hidden Neurons	Training Algorithm											
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	0.015	**1.389**	0.042	0.020	0.024	0.020	0.016	0.061	0.044	0.046	0.033	0.077
2	0.048	0.031	0.016	0.065	0.097	0.016	0.011	0.082	0.064	0.042	0.033	0.073
3	0.086	0.068	0.027	0.080	0.024	0.115	0.012	0.012	0.051	0.079	0.028	0.021
4	0.014	0.050	0.020	0.073	0.071	0.128	0.010	0.055	0.028	0.106	0.021	0.052
5	0.038	0.025	0.024	0.025	0.012	0.078	0.009	0.040	0.023	0.034	0.007	0.052
6	0.008	0.072	0.023	0.037	0.061	0.068	0.012	0.018	0.108	0.119	0.004	0.116
7	0.009	0.048	0.033	0.017	0.023	0.078	0.008	0.061	0.092	0.030	0.004	0.072
8	0.007	0.028	0.071	0.034	0.027	0.034	0.011	0.160	0.103	0.095	0.003	0.057
9	0.009	0.086	0.068	0.099	0.077	0.021	0.010	0.130	0.053	0.083	0.002	0.103
10	0.008	0.047	0.094	0.025	0.119	0.065	0.008	0.035	0.081	0.042	**0.001**	0.058
11	0.007	0.059	0.033	0.070	0.033	0.182	0.010	0.109	0.018	0.146	0.002	0.052
12	0.008	0.067	0.081	0.058	0.119	0.060	0.013	0.155	0.123	0.100	0.001	0.058
13	0.005	0.103	0.059	0.018	0.023	0.094	0.015	0.097	0.097	0.132	0.003	0.052
14	0.011	0.075	0.023	0.071	0.057	0.055	0.011	0.068	0.163	0.062	0.002	0.149
15	0.006	0.030	0.091	0.033	0.160	0.032	0.012	0.108	0.141	0.066	0.002	0.089
16	0.004	0.104	0.061	0.086	0.019	0.060	0.012	0.057	0.175	0.124	0.002	0.047
17	0.009	0.050	0.087	0.100	0.141	0.104	0.013	0.083	0.054	0.053	0.003	0.092
18	0.006	0.105	0.032	0.177	0.091	0.045	0.008	0.186	0.084	0.066	0.003	0.083
19	0.006	0.167	0.057	0.050	0.039	0.074	0.011	0.142	0.056	0.047	0.002	0.078
20	0.005	0.103	0.062	0.089	0.107	0.025	0.012	0.043	0.121	0.049	0.001	0.043

Bolded number are the lowest and highest MAE value.

Table A17. MAPE methane yield for different training algorithms at different number of hidden neurons.

Hidden Neurons	Training Algorithm											
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	0.08	0.22	0.19	0.10	0.12	0.10	0.09	0.31	0.20	0.21	0.160	0.34
2	0.22	0.16	0.08	0.32	0.47	0.09	0.06	0.41	0.29	0.22	0.160	0.36
3	0.39	0.33	0.14	0.40	0.14	0.55	0.06	0.07	0.25	0.39	0.140	0.12
4	0.08	0.26	0.10	0.36	0.35	0.62	0.05	0.27	0.14	0.51	0.110	0.26
5	0.18	0.14	0.12	0.12	0.06	0.39	0.05	0.20	0.11	0.17	0.040	0.26
6	0.04	0.32	0.13	0.18	0.28	0.31	0.06	0.08	0.52	0.57	0.020	0.55
7	0.05	0.25	0.16	0.09	0.11	0.39	0.04	0.29	0.44	0.16	0.020	0.36
8	0.03	0.15	0.34	0.18	0.13	0.17	0.06	0.77	0.50	0.46	0.010	0.27
9	0.04	0.41	0.33	0.47	0.38	0.10	0.05	0.62	0.25	0.38	0.010	0.50
10	0.04	0.22	0.46	0.12	0.57	0.31	0.04	0.17	0.40	0.21	**0.006**	0.28
11	0.04	0.29	0.16	0.34	0.15	0.85	0.05	0.53	0.10	0.69	0.012	0.26
12	0.04	0.32	0.37	0.27	0.56	0.27	0.06	0.75	0.59	0.46	0.007	0.29
13	0.02	0.48	0.30	0.10	0.11	0.44	0.07	0.47	0.46	0.64	0.014	0.25
14	0.06	0.35	0.12	0.34	0.27	0.26	0.05	0.32	0.75	0.29	0.008	0.68
15	0.03	0.14	0.44	0.17	0.74	0.15	0.06	0.52	0.67	0.31	0.009	0.42
16	0.02	0.50	0.30	0.39	0.09	0.29	0.06	0.28	0.82	0.58	0.011	0.23
17	0.04	0.23	0.40	0.46	0.68	0.48	0.06	0.39	0.27	0.25	0.014	0.44
18	0.03	0.50	0.15	0.82	0.45	0.22	0.04	**0.88**	0.41	0.32	0.013	0.39
19	0.03	0.79	0.28	0.24	0.19	0.36	0.06	0.69	0.26	0.22	0.010	0.39
20	0.02	0.47	0.29	0.43	0.51	0.12	0.06	0.20	0.58	0.23	0.007	0.19

Bolded number are the lowest and highest MAPE value.

Table A18. MSE for methane yield for different training algorithms at different number of hidden neurons.

Hidden Neurons	Training Algorithm											
	LM	GDX	CGP	SCG	BFG	OSS	RP	CGB	CGF	GD	BR	GDm
1	0.0004	0.0023	0.0019	0.0008	0.0007	0.0008	0.0006	0.0043	0.0021	0.0023	0.0014	0.0066
2	0.0025	0.0015	0.0003	0.0045	0.0095	0.0005	0.0003	0.0072	0.0046	0.0023	0.0014	0.0061
3	0.0001	0.0050	0.0011	0.0069	0.0011	0.0135	0.0002	0.0003	0.0032	0.0073	0.0011	0.0010
4	0.0004	0.0030	0.0006	0.0055	0.0054	0.0166	0.0002	0.0034	0.0011	0.0126	0.0008	0.0036
5	0.0018	0.0011	0.0007	0.0011	0.0002	0.0066	0.0002	0.0018	0.0007	0.0020	0.0001	0.0032
6	0.0001	0.0057	0.0010	0.0015	0.0042	0.0053	0.0003	0.0005	0.0121	0.0190	2.32×10^{-05}	0.0158
7	0.0002	0.0030	0.0019	0.0006	0.0007	0.0070	0.0001	0.0047	0.0095	0.0013	3.22×10^{-05}	0.0072
8	0.0001	0.0012	0.0055	0.0015	0.0009	0.0015	0.0002	0.0261	0.0111	0.0108	2.16×10^{-05}	0.0041
9	0.0001	0.0103	0.0050	0.0119	0.0065	0.0007	0.0002	0.0177	0.0036	0.0076	1.45×10^{-05}	0.0125
10	0.0001	0.0027	0.0093	0.0013	0.0146	0.0064	0.0006	0.0015	0.0071	0.0028	2.27×10^{-05}	0.0038
11	0.0001	0.0047	0.0017	0.0070	0.0013	0.0346	0.0002	0.0132	0.0005	0.0259	1.39×10^{-05}	0.0037
12	0.0002	0.0070	0.0071	0.0039	0.0161	0.0041	0.0002	0.0246	0.0160	0.0109	2.06×10^{-05}	0.0040
13	0.0001	0.0120	0.0043	0.0007	0.0008	0.0099	0.0003	0.0101	0.0104	0.0197	2.21×10^{-05}	0.0043
14	0.0002	0.0084	0.0010	0.0059	0.0048	0.0038	0.0002	0.0071	0.0298	0.0057	2.11×10^{-05}	0.0336
15	0.0001	0.0014	0.0089	0.0017	0.0311	0.0013	0.0002	0.0123	0.0219	0.0057	7.39×10^{-06}	0.0096
16	0.0001	0.0169	0.0045	0.0115	0.0006	0.0044	0.0002	0.0041	0.0329	0.0197	3.30×10^{-05}	0.0031
17	0.0001	0.0031	0.0114	0.0162	0.0220	0.0172	0.0002	0.0087	0.0041	0.0038	1.62×10^{-05}	0.0104
18	0.0001	0.0164	0.0014	0.0404	0.0108	0.0039	0.0001	0.0407	0.0083	0.0060	1.36×10^{-05}	0.0086
19	0.0001	0.0358	0.0045	0.0041	0.0027	0.0058	0.0002	0.0211	0.0052	0.0035	1.10×10^{-05}	0.0080
20	0.0001	0.0155	0.0068	0.0104	0.0128	0.0012	0.0003	0.0023	0.0155	0.0040	8.63×10^{-06}	0.0043

Table A19. R, MAE, MAPE, MSE value for COD removal for different training algorithms at different number of hidden neurons (aerobic process).

| Hidden Neuron | Training Algorithm | | | | | | | |
| | LM | | | | BR | | | |
	R	MAE	MAPE	MSE	R	MAE	MAPE	MSE
1	0.923	4.401	0.061	31.3975	0.845	5.963	0.080	61.8890
2	0.938	3.420	0.048	26.3683	0.960	3.146	0.040	16.2590
3	**0.982**	1.920	0.021	**7.3187**	0.904	2.575	0.030	47.6270
4	0.887	3.365	0.038	58.1403	0.975	1.497	0.020	10.5095
5	0.955	2.851	0.036	27.3267	0.995	0.651	0.010	2.2123
6	0.935	2.827	0.032	44.2447	**0.997**	0.459	0.005	**1.0288**
7	0.909	3.131	0.036	60.1302	0.997	**0.321**	**0.004**	1.0767
8	0.963	**1.525**	**0.016**	19.0559	0.982	0.956	0.010	7.9566
9	0.966	1.954	0.022	16.4031	0.980	1.070	0.016	12.4715
10	0.943	2.503	0.028	27.9288	0.996	0.605	0.007	1.9727
11	0.967	2.754	0.034	15.7244	0.991	0.674	0.010	4.7788
12	0.916	4.015	0.044	46.5166	0.993	0.601	0.007	2.9810
13	0.875	4.688	0.059	91.5153	0.991	0.664	0.008	3.5952
14	0.939	2.823	0.034	27.3849	0.978	1.136	0.017	13.6653
15	0.922	4.054	0.053	30.8332	0.997	0.359	0.004	1.1977
16	0.954	2.748	0.035	19.0488	0.997	0.372	0.004	1.5340
17	0.913	3.774	0.042	47.5349	0.983	1.026	0.012	7.4674
18	0.866	3.264	0.036	74.7672	0.997	0.528	0.006	1.0595
19	0.898	4.005	0.046	54.9089	0.991	0.639	0.007	4.0884
20	0.929	2.818	0.03	33.8239	0.989	0.830	0.012	5.6022

Bolded number are the highest R, and lowest MAE, MAPE and MSE value.

Table A20. R, MAE, MAPE, MSE value for BOD removal for different training algorithms at different number of hidden neurons (aerobic process).

| Hidden Neuron | Training Algorithm | | | | | | | |
| | LM | | | | BR | | | |
	R	MAE	MAPE	MSE	R	MAE	MAPE	MSE
1	0.447	2.918	0.031	12.5023	0.502	2.727	0.030	11.8510
2	0.875	1.235	0.013	4.6747	0.674	2.565	0.030	9.3357
3	0.929	1.260	0.013	2.2124	0.948	0.829	0.010	1.6992
4	0.939	0.930	0.010	1.9114	0.920	1.230	0.010	2.3544
5	0.927	1.172	0.012	2.2400	0.956	0.833	0.010	1.3128
6	0.935	0.874	0.009	2.3717	0.981	0.459	0.005	0.6047
7	0.928	0.848	0.009	3.4287	0.984	0.406	0.004	0.5496
8	0.957	**0.578**	**0.006**	**1.3668**	0.993	0.324	**0.003**	0.2235
9	0.936	0.889	0.009	2.2842	**0.995**	**0.247**	0.003	**0.1734**
10	0.945	0.606	0.006	1.6660	0.968	0.730	0.008	1.2808
11	0.944	0.709	0.007	1.8058	0.990	0.329	0.004	0.3182
12	0.932	1.056	0.011	2.1750	0.991	0.387	0.004	0.3320
13	0.908	1.240	0.013	6.0826	0.986	0.378	0.004	0.5745
14	0.938	0.735	0.008	2.3268	0.990	0.321	0.003	0.3284
15	0.938	0.975	0.010	1.9248	0.985	0.261	0.003	0.4613
16	0.933	0.961	0.010	2.2338	0.976	0.362	0.004	0.8479
17	0.876	1.203	0.012	3.7892	0.981	0.418	0.005	0.6142
18	0.929	0.603	0.006	2.3706	0.964	0.814	0.009	1.1946
19	0.942	0.734	0.008	1.7587	0.984	0.463	0.005	0.5261
20	**0.960**	0.612	0.006	1.5370	0.986	0.338	0.004	0.5504

Bolded number are the highest R, and lowest MAE, MAPE and MSE value.

Table A21. R, MAE, MAPE, MSE value for TSS removal for different training algorithms at different number of hidden neurons (aerobic process).

Hidden Neuron	Training Algorithm							
	LM				BR			
	R	MAE	MAPE	MSE	R	MAE	MAPE	MSE
1	0.638	4.535	0.050	27.0476	0.637	4.629	0.050	27.6880
2	0.682	4.265	0.047	24.6278	0.833	3.284	0.040	14.1102
3	0.918	1.930	0.021	7.2376	0.918	1.867	0.020	7.5477
4	0.925	1.707	0.018	7.5925	0.982	0.712	0.010	1.6731
5	0.935	1.667	0.018	6.9112	0.981	0.720	0.010	1.7804
6	0.983	0.583	0.007	1.8520	**0.995**	**0.399**	**0.004**	**0.5560**
7	0.923	1.189	0.013	6.7818	0.991	0.408	0.005	0.9309
8	0.959	0.880	0.009	3.9917	0.984	0.503	0.006	1.4665
9	**0.992**	**0.382**	**0.004**	**0.7541**	0.978	0.576	0.007	2.1969
10	0.937	1.210	0.013	6.2074	0.974	0.944	0.011	2.4801
11	0.890	1.720	0.019	9.4141	0.984	0.525	0.006	1.7226
12	0.891	1.739	0.019	10.1281	0.97	0.784	0.009	3.1270
13	0.945	1.057	0.012	5.7302	0.977	0.594	0.007	2.6534
14	0.943	1.276	0.014	6.0531	0.985	0.524	0.006	1.7528
15	0.951	1.637	0.018	5.3287	0.981	0.459	0.005	2.0162
16	0.957	1.584	0.018	4.8823	0.965	0.687	0.008	3.4741
17	0.930	1.633	0.018	7.2215	0.99	0.482	0.005	1.0659
18	0.891	1.419	0.015	10.3910	0.975	0.915	0.010	2.3832
19	0.893	1.400	0.015	9.8483	0.984	0.569	0.006	1.6107
20	0.903	1.654	0.018	9.9505	0.983	0.496	0.006	1.5933

Bolded number are the highest R, and lowest MAE, MAPE and MSE value.

Appendix A.2. MATLAB CODE

```
clear all
load('finalone.mat')
x = Input';
t = Output';

for k=1:20 ;% for 1 hidden neurons to 20 hidden neurons
load('finalone.mat')
x = Input';
t = Output';
trainFcn = 'trainlm';% Levenberg-Marquardt backpropagation.

% Create a Fitting Network

for i=1:1000
hiddenLayerSize = (k);
net = fitnet(hiddenLayerSize,trainFcn);

% Choose Input and Output Pre/Post-Processing Functions
net.input.processFcns = {'removeconstantrows','mapminmax'};
net.output.processFcns = {'removeconstantrows','mapminmax'};

% Setup Division of Data for Training, Validation, Testing
net.divideFcn = 'dividerand'; % Divide data randomly
net.divideMode = 'sample'; % Divide up every sample
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
```

```
% Choose a Performance Function
net.performFcn = 'mse'; % Mean Squared Error

% Choose Plot Functions
net.plotFcns = {'plotperform','plottrainstate','ploterrhist', ...
    'plotregression', 'plotfit'};

% Train the Network
[net,tr] = train(net,x,t);

% Test the Network
y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y)

% Recalculate Training, Validation and Test Performance
trainTargets = t .* tr.trainMask{1};
valTargets = t .* tr.valMask{1};
testTargets = t .* tr.testMask{1};
trainPerformance = perform(net,trainTargets,y);
valPerformance = perform(net,valTargets,y);
testPerformance = perform(net,testTargets,y);
%extraction of weight and biases
w1(i)=net.IW(1,1);
w2(i)=net.LW(2,1);
b1(i)=net.b(1);
b2(i)=net.b(2);
Rval(:,i)=regression(t,y);
Predval(i)={y};
MSEtrain(i)=trainPerformance;
MSEval(i)=valPerformance;
MSEtest(i)=testPerformance;
MAE(i)={1/96*sum(abs(Output-y'))};
MAPE(i)={(1/96*(sum(rdivide(abs(Output-y'),Output))))};
MSE(i)={(sum((Output-y').^2))/24};

end

%R value
R=regression(t,y);

%extract predicted value
m=cell2mat(Predval);
p=cell2mat(MAE);
c=cell2mat(MAPE);
d=cell2mat(MSE);

%To export to excel
filename=['gdm','.xlsx'];
Sheet=sprintf('neuron%d',k);
col_header={'R value','','','Sum of R2value','Average if R2
    value','Best R2 value','','MAEvalue','','','','SUM of MAE','','MAPE value','','','','Sum
of MAPE','','MSE
    value','','','','Sum of MSE','','Predvalue','','','Position','','Expval'};
```

```
sumofr2={'=SUM(A2:C2)','','=MAX(D:D)'};
match1={'=MATCH(F2,D:D,0)'};
maeoffset={'=OFFSET($H$2,COLUMNS($H2:H2)-1+(ROWS($2:2)-
    1)*3,0)','','','=SUM(I2:K2)','=MIN(L2:L1001)'};
match2={'=MATCH(M2,L:L,0)'};
mapeoffset1={'=OFFSET($N$2,COLUMNS($N2:N2)-1+(ROWS($2:2)-
    1)*3,0)','','','=SUM(O2:Q2)','=MIN(R2:R1001)'};
match3={'=MATCH(S2,R:R,0)'};
mseoffset1={'=OFFSET($T$2,COLUMNS($T2:T2)-1+(ROWS($2:2)-
    1)*3,0)','','','=SUM(U2:W2)','=MIN(X2:X1001)'};
match4={'=MATCH(Y2,X:X,0)'};
position1={'=INT((ROW(E1)-1)/24)+1'};
xlswrite(filename,col_header,Sheet,'A1');
xlswrite(filename,sumofr2,Sheet,'D2');
xlswrite(filename,match1,Sheet,'F3');
xlswrite(filename,maeoffset,Sheet,'I2');
xlswrite(filename,match2,Sheet,'M3');
xlswrite(filename,mapeoffset1,Sheet,'O2');
xlswrite(filename,match3,Sheet,'S3');
xlswrite(filename,mseoffset1,Sheet,'U2');
xlswrite(filename,match4,Sheet,'Y3');
xlswrite(filename,position1,Sheet,'AC2');


xlswrite(filename,Rval',Sheet,'A2')
xlswrite(filename,p',Sheet,'H2')
xlswrite(filename,c',Sheet,'N2')
xlswrite(filename,d',Sheet,'T2')
xlswrite(filename,m',Sheet,'Z2')
xlswrite(filename,t',Sheet,'AE2')
save(['workspacefor' num2str(k)])
clear all

end
```

References

1. Department of Statistics Malaysia. Department of Statistics Malaysia Press Release. 2021. Available online: https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=72&bul_id=TDV1YU4yc1Z0dUVyZ0xPV0ptRlhWQT09&menu_id=Z0VTZGU1UHBUT1VJMFlpaXRRR0xpdz09 (accessed on 1 March 2022).
2. Malaysian Palm Oil Industry Council; Malaysian Palm Oil Council. 2020. Available online: https://mpoc.org.my/malaysian-palm-oil-industry/#:~{}:text=In2020%2CMalaysiaaccountedfor,fatsinthesameyear (accessed on 1 March 2022).
3. Poh, P.E.; Yong, W.-J.; Chong, M.F. Palm Oil Mill Effluent (POME) Characteristic in High Crop Season and the Applicability of High-Rate Anaerobic Bioreactors for the Treatment of POME. *Ind. Eng. Chem. Res.* **2010**, *49*, 11732–11740. [CrossRef]
4. Madaki, Y.S.; Seng, L.G. Palm oil mill effluent (pome) from malaysia palm oil mills: Waste or resource. *Int. J. Sci. Environ. Technol.* **2013**, *2*, 1138–1155.
5. Ge, D.; Yuan, H.; Xiao, J.; Zhu, N. Insight into the enhanced sludge dewaterability by tannic acid conditioning and pH. regulation. *Sci. Total Environ.* **2019**, *679*, 298–306. [CrossRef] [PubMed]
6. Ahmad, A.L.; Ismail, S.; Bhatia, S. Water recycling from palm oil mill effluent (POME) using membrane technology. *Desalination* **2003**, *157*, 87–95. [CrossRef]
7. Chong, J.; Chan, Y.; Chong, S.; Ho, Y.; Mohamad, M.; Tan, W.; Cheng, C.; Lim, J. Simulation and Optimisation of Integrated Anaerobic-Aerobic Bioreactor (IAAB) for the Treatment of Palm Oil Mill Effluent. *Processes* **2021**, *9*, 1124. [CrossRef]
8. Tan, H.M.; Poh, P.E.; Gouwanda, D. Resolving stability issue of thermophilic high-rate anaerobic palm oil mill effluent treatment via adaptive neuro-fuzzy inference system predictive model. *J. Clean. Prod.* **2018**, *198*, 797–805. [CrossRef]
9. Batstone, D.; Keller, J.; Angelidaki, I.; Kalyuzhnyi, S.; Pavlostathis, S.; Rozzi, A.; Sanders, W.; Siegrist, H.; Vavilin, V. Anaerobic digestion model No 1 (ADM1). *Water Sci. Technol.* **2002**, *45*, 65–73. [CrossRef]
10. Güçlü, D.; Dursun, S. Amelioration of Carbon Removal Prediction for an Activated Sludge Process using an Artificial Neural Network (ANN). *CLEAN–Soil. Air. Water* **2008**, *36*, 781–787. [CrossRef]

11. Yang, S.-S.; Yu, X.-L.; Ding, M.-Q.; He, L.; Cao, G.-L.; Zhao, L.; Tao, Y.; Pang, J.-W.; Bai, S.-W.; Ding, J.; et al. Simulating a combined lysis-cryptic and biological nitrogen removal system treating domestic wastewater at low C/N ratios using artificial neural network. *Water Res.* **2021**, *189*, 116576. [CrossRef]

12. Manu, D.S.; Thalla, A.K. Artificial intelligence models for predicting the performance of biological wastewater treatment plant in the removal of Kjeldahl Nitrogen from wastewater. *Appl. Water Sci.* **2017**, *7*, 3783–3791. [CrossRef]

13. Güçlü, D.; Yılmaz, N.; Ozkan-Yucel, U.G. Application of neural network prediction model to full-scale anaerobic sludge digestion. *J. Chem. Technol. Biotechnol.* **2011**, *86*, 691–698. [CrossRef]

14. Sathish, S.; Vivekanandan, S. Parametric optimization for floating drum anaerobic bio-digester using Response Surface Methodology and Artificial Neural Network. *Alex. Eng. J.* **2016**, *55*, 3297–3307. [CrossRef]

15. Park, Y.-S.; Lek, S. Artificial Neural Networks. In *Developments in Environmental Modelling*; Elsevier BV: Amsterdam, The Netherlands, 2016; pp. 123–140.

16. Nair, V.V.; Dhar, H.; Kumar, S.; Thalla, A.K.; Mukherjee, S.; Wong, J.W. Artificial neural network based modeling to evaluate methane yield from biogas in a laboratory-scale anaerobic bioreactor. *Bioresour. Technol.* **2016**, *217*, 90–99. [CrossRef] [PubMed]

17. Mougari, N.E.; Largeau, J.F.; Himrane, N.; Hachemi, M.; Tazerout, M. Application of artificial neural network and kinetic modeling for the prediction of biogas and methane production in anaerobic digestion of several organic wastes. *Int. J. Green Energy* **2021**, *18*, 1584–1596. [CrossRef]

18. Khalil, A.; Almasri, M.N.; Mc Kee, M.; Kaluarachchi, J.J. Applicability of statistical learning algorithms in groundwater quality modeling. *Water Resour. Res.* **2005**, *41*, 1–16. [CrossRef]

19. Lecun, Y.; Bottou, L.; Muller, K. Efficient BackProp. 2000; pp. 1–44. Available online: http://yann.lecun.com/exdb/publis/pdf/lecun-98b.pdf (accessed on 1 March 2022).

20. Lee, D.-H.; Kim, Y.-T.; Lee, S.-R. Shallow Landslide Susceptibility Models Based on Artificial Neural Networks Considering the Factor Selection Method and Various Non-Linear Activation Functions. *Remote Sens.* **2020**, *12*, 1194. [CrossRef]

21. Alver, A.; Kazan, Z. Prediction of full-scale filtration plant performance using artificial neural networks based on principal component analysis. *Sep. Purif. Technol.* **2020**, *230*, 115868. [CrossRef]

22. Roopnarain, A.; Rama, H.; Ndaba, B.; Bello-Akinosho, M.; Bamuza-Pemu, E.; Adeleke, R. Unravelling the anaerobic digestion 'black box': Biotechnological approaches for process optimization. *Renew. Sustain. Energy Rev.* **2021**, *152*, 111717. [CrossRef]

23. Merkel, W.; Krauth, K. Mass transfer of carbon dioxide in anaerobic reactors under dynamic substrate loading conditions. *Water Res.* **1999**, *33*, 2011–2020. [CrossRef]

24. Roopnarain, A.; Adeleke, R. Current status, hurdles and future prospects of biogas digestion technology in Africa. *Renew. Sustain. Energy Rev.* **2017**, *67*, 1162–1179. [CrossRef]

25. Sebti, A.; Boutra, B.; Trari, M.; Aoudjit, L.; Igoud, S. Application of Artificial Neural Network for Modeling Wastewater Treatment Process. In *Lecture Notes in Networks and Systems*; Springer Science and Business Media LLC: Berlin, Germany, 2019; pp. 143–154.

26. Chen, M.; Chen, C.-J.; Bao, B.-C.; Xu, Q. Multi-stable patterns coexisting in memristor synapse-coupled Hopfield neural network. In *Mem-Elements for Neuromorphic Circuits with Artificial Intelligence Applications*; Elsevier BV: Amsterdam, The Netherlands, 2021; pp. 439–459.

27. Miner, G.; Elder, J.; Fast, A.; Hill, T.; Nisbet, R.; Delen, D. *Practical Text Mining and Statistical Analysis for Non-Structured Text Data Applications*; Elsevier BV: Amsterdam, The Netherlands, 2012; pp. 1007–1016.

28. Thanki, R.; Borra, S. Application of Machine Learning Algorithms for Classification and Security of Diagnostic Images. In *Machine Learning in Bio-Signal Analysis and Diagnostic Imaging*; Elsevier BV: Amsterdam, The Netherlands, 2019; pp. 273–292.

29. Schmitt, F.; Banu, R.; Yeom, I.-T.; Do, K.-U. Development of artificial neural networks to predict membrane fouling in an anoxic-aerobic membrane bioreactor treating domestic wastewater. *Biochem. Eng. J.* **2018**, *133*, 47–58. [CrossRef]

30. Bekkari, N.; Zeddouri, A. Using artificial neural network for predicting and controlling the effluent chemical oxygen demand in wastewater treatment plant. *Manag. Environ. Qual. Int. J.* **2019**, *30*, 593–608. [CrossRef]

31. Madaeni, S.S.; Shiri, M.; Kurdian, A.R. Modeling, Optimization, and Control of Reverse Osmosis Water Treatment in Kazeroon Power Plant Using Neural Network. *Chem. Eng. Commun.* **2015**, *202*, 6–14. [CrossRef]

32. Jawad, J.; Hawari, A.H.; Zaidi, S.J. Artificial neural network modeling of wastewater treatment and desalination using membrane processes: A review. *Chem. Eng. J.* **2021**, *419*, 129540. [CrossRef]

33. Kaur, H.; Salaria, D.S. Bayesian Regularization based Neural Network Tool for Software Effort Estimation. *Glob. J. Comput. Sci. Technol.* **2013**, *13*, 1–7.

34. Burden, F.; Winkler, D. Bayesian Regularization of Neural Networks. *Methods Mol. Biol.* **2008**, *458*, 23–42. [CrossRef]

35. Bajpai, P. Basics of Anaerobic Digestion Process. In *Gigaseal Formation in Patch Clamping*; Springer Science and Business Media LLC: Berlin, Germany, 2017; pp. 7–12.

36. Buchanan, J.R.; Seabloom, R.W. *Aerobic Treatment of Wastewater and Aerobic Treatment Units*; University Curriculum Development for Decentralized Wastewater Management: Fayetteville, AR, USA, 2004; pp. 1–27. Available online: http://onsite.tennessee.edu/Aerobic%20Treatment%20&%20ATUs.pdf (accessed on 21 December 2021).

37. Li, J.-Y.; Chow, T.; Yu, Y.-L. The estimation theory and optimization algorithm for the number of hidden units in the higher-order feedforward neural network. In Proceedings of the ICNN'95-International Conference on Neural Networks; IEEE, Perth, Australia, 27 November–1 December 1995; pp. 1229–1233.

38. Ke, J.; Liu, X. Empirical Analysis of Optimal Hidden Neurons in Neural Network Modeling for Stock Prediction. In *2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application*; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2008; Volume 2, pp. 828–832.

39. Sheela, K.G.; Deepa, S.N. Review on Methods to Fix Number of Hidden Neurons in Neural Networks. *Math. Probl. Eng.* **2013**, *2013*, 425740. [CrossRef]

40. Hunter, D.; Yu, H.; Pukish, M.S., III; Kolbusz, J.; Wilamowski, B.M. Selection of Proper Neural Network Sizes and Architectures—A Comparative Study. *IEEE Trans. Ind. Inform.* **2012**, *8*, 228–240. [CrossRef]

41. Shibata, K.; Ikeda, Y. Effect of number of hidden neurons on learning in large-scale layered neural networks in 2009. In Proceedings of the ICROS-SICE International Joint Conference, Fukuok, Japan, 18–21 August 2009; pp. 5008–5013.

42. Xu, S.; Chen, L. A Novel Approach for Determining the Optimal Number of Hidden Layer Neurons for FNn_s and Its Application in Data Mining. 2008, pp. 683–686. Available online: https://eprints.utas.edu.au/6995/1/02-au-xu.pdf. (accessed on 1 March 2022).

43. Zhang, Z.; Ma, X.; Yang, Y. Bounds on the number of hidden neurons in three-layer binary neural networks. *Neural Netw.* **2003**, *16*, 995–1002. [CrossRef]

44. Fujita, O. Statistical estimation of the number of hidden units for feedforward neural networks. *Neural Netw.* **1998**, *11*, 851–859. [CrossRef]

45. Tamura, S.; Tateishi, M. Capabilities of a four-layered feedforward neural network: Four layers versus three. *IEEE Trans. Neural Netw.* **1997**, *8*, 251–255. [CrossRef] [PubMed]

46. Garson, D.G. Interpreting neural-network connection weights. *AI Expert* **1991**, *6*, 46–51.

47. Aleboyeh, A.; Kasiri, M.; Olya, M. Prediction of azo dye decolorization by UV/H2O2 using artificial neural networks. *Dye. Pigment.* **2008**, *77*, 288–294. [CrossRef]

48. Yetilmezsoy, K.; Demirel, S. Artificial neural network (ANN) approach for modeling of Pb(II) adsorption from aqueous solution by Antep pistachio (*Pistacia Vera* L.) shells. *J. Hazard. Mater.* **2008**, *153*, 1288–1300. [CrossRef]

49. Dibaba, O.R.; Lahiri, S.K.; T'jonck, S.; Dutta, A. Experimental and Artificial Neural Network Modeling of a Upflow Anaerobic Contactor (UAC) for Biogas Production from Vinasse. *Int. J. Chem. React. Eng.* **2016**, *14*, 1241–1254. [CrossRef]

50. Antwi, P.; Li, J.; Meng, J.; Deng, K.; Quashie, F.K.; Li, J.; Boadi, P.O. Feedforward neural network model estimating pollutant removal process within mesophilic upflow anaerobic sludge blanket bioreactor treating industrial starch processing wastewater. *Bioresour. Technol.* **2018**, *257*, 102–112. [CrossRef]

51. Jaroenpoj, S.; Yu, Q.J.; Ness, J. Development of Artificial Neural Network Models for Biogas Production from Co-Digestion of Leachate and Pineapple Peel. *Glob. Environ. Eng.* **2014**, *1*, 42–47. [CrossRef]

52. Xu, F.; Wang, Z.-W.; Li, Y. Predicting the methane yield of lignocellulosic biomass in mesophilic solid-state anaerobic digestion based on feedstock characteristics and process parameters. *Bioresour. Technol.* **2014**, *173*, 168–176. [CrossRef]

53. Ghani, W.; Idris, A. Preliminary study on biogas production of biogas from municipal solid waste (MSW) leachate. *J. Eng. Sci. Technol.* **2009**, *4*, 374–380.

54. Patel, P.; Modi, A.; Minipara, D.; Kumar, A. Microbial biosurfactants in management of organic waste. In *Sustainable Environmental Clean-Up*; Elsevier BV: Amsterdam, The Netherlands, 2021; pp. 211–230.

55. Basim, K. The effect of MLSS value on removal of COD and phosphorus control method of return activated sludge concentration. *J. Eng. Appl. Sci.* **2018**, *13*, 9730–9734.

56. Senapati, T.; Samanta, P.; Roy, R.; Sasmal, T.; Ghosh, A.R. Artificial neural network: An alternative approach for assessment of biochemical oxygen demand of the Damodar River, West Bengal, India. In *Intelligent Environmental Data Monitoring for Pollution Management*; Elsevier BV: Amsterdam, The Netherlands, 2020; pp. 231–240.

57. Olden, J.D.; Joy, M.K.; Death, R.G. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecol. Model.* **2004**, *178*, 389–397. [CrossRef]

58. Chan, Y.J. Foo Seng, H.; Chong, M.F.; Ng, D.; Lim, D.L.K. Pre-commercialized integrated anaerobic-aerobic bioreactor (iaab) for palm oil mill effluent (pome) treatment & biogas generation. *Environ. Health* **2020**, *40*, 57–66. [CrossRef]