Advances in Computational Design, *Vol. 5, No. 3 (2020) 323-347*
DOI: https://doi.org/10.12989/acd.2020.5.3.323

Modeling methods used in bioenergy production processes: A review

Hamza Akroum*[1], Dahbia Akroum-Amrouche[2] and Abderrezak Aibeche[1]

[1]Laboratoire d'Automatique Appliquée, Université M'Hamed Bougara de Boumerdès, 1 Av. de l'indépendance 35000 Boumerdès, Algeria
[2]Département de chimie, faculté des sciences, Université M'Hamed Bougara de Boumerdes Av. de l'independance 35000 Boumerdes, Algeria

(Received March 1, 2019, Revised February 28, 2020, Accepted March 5, 2020)

Abstract. The enhancements of bioenergy production effectiveness require the comprehensively experimental study of several parameters affecting these bioprocesses. The interpretation of the obtained experimental results and the estimation of optimum yield are extremely complicated such as misinterpreting the results of an experiment. The use of mathematical modeling and statistical experimental designs can consistently supply the predictions of the potential yield and the identification of defining parameters and also the understanding of key relationships between factors and responses. This paper summarizes several mathematical models used to achieve an adequate overall and maximal production yield and rate, to screen, to optimize, to identify, to describe and to provide useful information for the effect of several factors on bioenergy production processes. The usefulness, the validity and, the feasibility of each strategy for studying and optimizing the bioenergy-producing processes were discussed and confirmed by the good correlation between predicted and measured values.

Keywords: system modeling; experimental design methods; neural network design, identification, optimization, bioenergy

1. Introduction

The worldwide energy requirement has been rising exponentially, and the reserves of fossil fuels have been decreasing. The dependence on fossil fuels as our primary energy source contributes to global climate change, environmental degradation, health problems and contributes to the greenhouse effect. Environmental and energy concerns are major factors driving research on the green energy production and use of renewable fuels, an eco-friendly, sustainable and cost-effective energy sources that can replace a significant amount of the petroleum fuel, to reduce air pollution and greenhouse gas emissions (Pan *et al.* 2008, Perera 2017, Pandey *et al.* 2018).

To ensure the sustainability of bioenergy, it is necessary to confirm that the energy produced is greater than the energy consumed in its production. Life cycle assessment is a widely used method to quantify the environmental impacts and cost in bioenergy development (Lardon *et al*. 2009, Guerrero and Muñoz 2018). Biomass is considered a renewable resource because of its short life

*Corresponding author, Ph.D., E-mail: akroum@univ-boumerdes.dz

http://www.techno-press.org/?journal=acd&subpage=7 ISSN: 2383-8477 (Print), 2466-0523 (Online)

cycle when biomass grows; all carbon in biomass comes from the atmosphere and is liberated into the environment when it is burned, also, life-cycle assessments of biomass-derived biofuels found a significant net reduction in greenhouse gas emissions and environmental burdens. Therefore, biomass is thought of as a carbon-neutral fuel. The bioenergies, produced in optimal conditions based on the biomass used, are potential substitutes to fossil fuels (Patterson *et al*. 2013, Tonini *et al*. 2016, Chen *et al.* 2019). However, during the last few years, biofuel production, such as biogas, biomethanol, bioethanol, biodiesel, and biohydrogen, received considerable attention from many researchers in the world (Demirbas 2009, Kaparaju *et al.* 2009, Hasegawa *et al.* 2010, Wirth *et al.* 2015, Pérez-Sariñana *et al.* 2019). Bioenergy is considered to be an ideal energy alternative in the future. It can be produced using a variety of technologies and from various substrates such as wastewater and lignocellulosic materials, including agricultural, wood, and forest residue (Kumar *et al.* 2009, Sathish and Vivekanandan 2016). Bioenergy derived from the cultivation of algae has therefore been proposed as an alternative approach that does not impact on agriculture (Pittman *et al.* 2011, Onumaegbu *et al.* 2018, Nassef *et al.* 2019). Compared with the chemical generation process, the bioenergy production process is a more attractive approach because it is recyclable, environment-friendly and low cost (Huang *et al.* 2012).

Recently, a large number of bioprocesses and pre-treatment methods have been used and investigated for biofuels production under different operation conditions. Biomass can be converted into energy via thermochemical conversions, biochemical conversions, and extraction of oil from oil-bearing seeds (Van der Stelt *et al.* 2011, Canabarro *et al*. 2013, Hossain 2019). Among reaction conditions the composition of fermentation media, is a very important factor that determines the nutrimental and physic-chemical environment such as the substrate, reduction agent, bivalents cations, pH, light intensities and temperature for the entire cell biocatalysts in the reactor (Weuster-Botz 2000, Karthic *et al.* 2013, Adnan *et al.* 2014, Garimella *et al.* 2019).

Owing to the intricacy of bioprocesses several mathematical models are used to achieve an adequate overall and maximal production efficiency, to optimize, to identify, to describe and to provide useful information for the effect of several factors on bioenergy production processes. Analysis of the resulting models can aid substantially in avoiding such mistakes or in identifying errors or omissions in earlier thinking and interpretations. In this paper, the mathematical experiment designs used to study the factors that affect operations and effectiveness of bioenergy production, including linear, non-linear, experimental design, neural network design, and hybrid systems, in different bioenergy production pathways, are described and evaluated.

2. Experimental design and mathematical modeling methods

An experimental design is a process by which certain selected factors are intentionally varied in a controlled manner to get their effects on a response of interest, often followed by the analysis of the experimental results. According to the number of factors involved for investigation at a time, the experimental design can be classified into two categories: One-factor-at-a time design (single-factor design) and factorial design (multiple factor design).

Mathematical modeling methods improve the understanding of certain selected factors affecting the studied response. The most frequent empirical models fit the experimental data to take either a linear form or quadratic form. The selection of a suitable model is important for bioprocesses' mathematical modeling ((Luftig 1998, Czitrom 1999, Daniel *et al.* 2003, Frey *et al.* 2018).

Design of experiment (DOE) is a strategy for carrying out experiments and examining the influence of several input factors on the process output. Multivariate statistical approaches require the user to establish minimum and maximum values for each factor and define the experimental domain to be investigated during the optimization procedure. A combination of different levels of effective variables for achieving the optimum operation conditions is attained using response surface approaches (Zolgharneina *et al*. 2013, Frey *et al.* 2018, Chollom *et al.* 2020). Therefore, the use of experimental design methods is very attractive in the analytical sciences.

Various designs, investigate the relationship between the process variables or factors x_1, x_2, ..., x_k and a quality characteristic y of the product, to construct the empirical models describing the response and predicting performance measures over the factor design space, and also synchronize study of several factors and reduce the number of experiments (Adnan *et al.* 2014, Sathish and Vivekanandan, 2016, Onumaegbu *et al.* 2018, Nassef *et al.* 2019, Ishola *et al.* 2019, Garimella *et al.* 2019).

2.1 One factor at-a-time design

One-factor-at-a-time design OFAT is a usual experimental design where only one factor is studied in a certain time period while keeping the levels of other factors constant. The level of the factor to be investigated is then changed, over the desired range, to study its effects on the response. After the experimental results are obtained, certain graphs are generally constructed showing how the response is influenced by the factor studied (Wang and Wan 2009, Iqbal Syaichurrozi and Sumardiono 2014, Miñón-Fuentes and Aguilar-Juárez 2020) and can be analyzed to yield valid and objective conclusions using linear and non-linear modeling.

Michaelis–Menten, Monod, Luedeking–Piret, modified Gompertz and modified logistic models are mostly used in the literature because they have easily interpretable parameters. Kinetic constants obtained from these models can be used for the design and operation of the bioenergy producing process in bioreactors.

2.1.1 Linear regression

Y is a variable it can be defined in amount, in rate and yield of bioenergy produced, substrate consumed and micro-organisms produced. The best-fitting straight line through the points called a regression line is determined by the linear regression, Eq. (1):

$$Y' = ax + b \tag{1}$$

where Y' is the predicted response, a is the line slope, and b is the Y-intercept. Linear regression was widely used to describe the bioprocesses.

The Michaelis–Menten equation Eq. (2) and The Monod equation Eq. (3) have the same form and are based respectively on empirical and theoretical considerations. Both models were widely used to describe the effects of substrate concentration on the rates of substrate degradation and cell growth, respectively, in the bioenergy production processes. The Monod model Eq. (4) is based, particularly, on growth kinetic classic substrate-limited, in which k_d is the endogenous decay coefficient (Boni *et al*. 2013). The best way to analyze biological kinetic data is to fit the data directly to the Michaelis-Menten and Monod equations using nonlinear regression or linear regression by taking the inverse of these equations, the linearization of Eq (2) and Eq (3) gives Eq. (5) and Eq (6), respectively. The plotting of $1/v$ against $1/S$, gives a straight line with an intersection of $1/v_m$ and a slope of k_m/v_m, k_m indicates the substrate content necessary to reach 50% of the maximum specific rate of substrate degradation. The Luedeking–Piret model Eq. (7) (Gadhe

et al. 2014a) is used to describe the relationship between the bioenergy-producing microorganisms' growth rate and the product formation rate. The first term in the Eq. (7), i.e. α_i *dX/dt*, represents the rate of formation of growth-linked product *i*, implies that growing cells produce the product in the constant proportion of their growth. While the second term, i.e. $\beta_i X$, represents the non-growth linked the formation of product *i*, implies that microorganisms produce a product in the constant proportion of their concentration, irrespective of the growth phase.

$$v = \frac{v_m S}{k_m + S} \tag{2}$$

$$\mu = \frac{1}{X}\frac{dX}{dt} = \frac{\mu_m S}{k_S + S} \tag{3}$$

$$\mu = \frac{1}{X}\frac{dX}{dt} = \frac{\mu_m S}{k_S + S} - k_d \tag{4}$$

$$\frac{1}{v} = \frac{k_m}{v_m}\frac{1}{S} + \frac{1}{v_m} \tag{5}$$

$$\frac{1}{\mu} = \frac{k_s}{\mu_m}\frac{1}{S} + \frac{1}{\mu_m} \tag{6}$$

$$\frac{dp_i}{Xdt} = \alpha_i \frac{\mathrm{dX}}{\mathrm{dt}} + \beta_i X \tag{7}$$

where v ($g\ g^{-1}h^{-1}$) is the specific substrate degradation rate; v_m ($g\ g^{-1}\ h^{-1}$) is the maximum specific substrate degradation rate; k_m (g L^{-1}) is the dissociation constant; S ($g\ L^{-1}$) is the substrate concentration, μ is the specific growth rate; μ_m is the maximum specific growth rate; k_s ($g\ L^{-1}$) is the half-velocity constant equal to the substrate concentration corresponding to half of the maximum growth rate.

2.1.2 Nonlinear regression

Like the linear regression, the nonlinear regression relates a response *Y* to a vector predictor variables *x*, nonlinearly for building a purely empirical model. For example, if something is growing exponentially, which means growing at a steady rate, the relationship between *X* and *Y* is curved, to fit it, non-linear regression is required (Smyth 2002, Motulsky and Christopoulos 2013).

The modified Gompertz Eq. (8) (Zwietering *et al.* 1990) and modified logistic Eq. (9) are well known and extensively used in various aspects of biology, as well as the substrate degradation efficiency, the bacteria growth (Sevinç *et al.* 2012), the yeast growth (Phukoetphim, *et al.* 2017) and the bioenergy production like biohydrogen (Wang and Wan 2008, Sevinç *et al.* 2012, Bina, *et al*. 2019, Miñón-Fuentes and Aguilar-Juárez 2020), biogas (Iqbal Syaichurrozi and Sumardiono 2014, Deepanraj *et al.* 2015, Zhu *et al*, 2019), and bioethanol (Phukoetphim, *et al.* 2017).

$$P = P_{\max} \exp\left\{-\exp\left[\frac{R_{\max}.e}{P_{\max}}(\lambda - t) + 1\right]\right\} \tag{8}$$

$$P = \frac{P_{\max}}{1 + \exp\left[4R_{\max}(\lambda - 1)/P_{\max} + 2\right]} \tag{9}$$

where P is the expected value (*ml or g*), P_{max} is the maximum production potential (*ml or g*), R_m is the maximum production rate (*ml/h or g/l*), λ is the lag-phase time (*h*) and *e* equals 2.718.

Both the modified Gompertz model and the modified logistic model were statistically significant and could describe successfully the bioenergy production data obtained from each batch test and provide the kinetic parameters P, R_{max}, λ and P_{max}. Sevinç *et al.* (2012) reported that the cell growth data fitted well to the logistic model and the cumulative hydrogen production data fitted well to the Modified Gompertz Model. On the other hand, Phukoetphim *et al.* (2017) reported that the logistic model excellently predicted the biomass profile after the beginning of the exponential growth phase. However, it failed to estimate the biomass concentration during the initial stage of fermentation while the modified Gompertz model describes successfully the cell growth curves. Similarly, Deepanraj *et al.* (2015) reported that the modified Gompertz model produced perfect goodness of fit than the modified Logistic model.

2.2 Experimental design methods DOE

Today's science uses a statistical approach to experimental design due to the diversity and the difficulty to identify significant factors, to estimate a response function or to optimize a process. Therefore, the use of experimental design methods is very interesting in the analytical sciences because they allow a synchronized study of several factors and considerably reduce the number of experiments. Experimental design methods are faster to perform and more cost-effective than standard experimental design.

2.2.1 First-order models

The first-order design is the full and fractional factorial designs, due to their simplicity and relatively low cost, in which every factor is experimentally studied at only two levels. They can also be used to determine simple response surfaces that are linear with respect to all of the investigated factors. Only the first stage in a multivariate investigation, where a linear response surface is determined, will be mentioned. For a j-factor case, the response surface is given by the linear model (Richard *et al.* 2009, Tarley *et al.* 2009, Suen *et al.* 2013) as detailed below Eq. (10)

$$\mathrm{Y} = \mathrm{X}_0\alpha_0 + \mathrm{X}_1\alpha_1 + \ldots.\mathrm{X}_\mathrm{m}\alpha_m + \mathrm{E} \tag{10}$$

where Y is the vector of n observations, α_j is the vector of all j-factor interactions, X_j is the matrix of orthonormal contrast coefficients for α_j, and E is the vector of independent random errors.

Full Factorial Design (FD) and Fractional Factorial Design (FFD) $n^{(k-p)}$:

A full factorial design is very useful for preliminary studies or in initial optimization steps. This design allows free interaction with data, the ability to make comparisons, seeking similarities, differences, and trends. Fractional factorial designs $n^{(k-p)}$ are among the main important statistical contributions to the efficient investigation of the effects of numerous controllable factors on a response of interest (Richard *et al.* 2009). It was developed when the problem involves a large number of factors.

Fractional factorial experiments use known properties of the design to reduce selectively the size of an experiment while, at the same time, limiting the trade-off of critical information that

may be omitted if a thorough study of all possible combinations, of the levels of factors of interest, is not carried out. Fractional factorial design $n^{(k-p)}$ having N runs and k factors of n levels each with $N=n^{k-p}$. The design is represented by an N x k matrix D with entries 0; 1… n-1, where each row represents a run, and each column represents a factor. The generalized minimum aberration criterion to judge the optimality of the designs is used. For a design D having N runs and m factors and a $1/2^p$ fraction of the number of possible factor level combinations, the response surfaces are planar (Suen *et al.* 2013).

Fractional factorial designs enable the screening and identification of the main effective factors upon the response, which is estimated as the difference between both averages of measurements made at the high and low levels of that factor (Li *et al.* 2013). This design can obtain information concerning the major effects in relatively few numbers of runs to simplify the analysis and understand the results. The effects are tested to determine whether non-linear terms are required during the construction of the model. An $n^{(k-p)}$ fractional factorial design was also used as the experimental model to optimize the process parameters for the production of hydrogen. The objective of this design was to develop a model that would predict the hydrogen production efficiency as a function of various factors selected affecting this bioprocess.

Taguchi design:

Taguchi design is a fractional factorial design developed by Taguchi and Konishi, it allows the effects of many factors with two or more levels on a response, to be studied in a relatively small number of runs (Athreya *et al*. 2012, Wang *et al.* 2013, Usmanbaha *et al.* 2019). The values of the functions were converted into signal-to-noise ratios (SNRs), which are the log functions of the expected outcome that would serve as the objective of an optimization problem (Singh *et al.* 2018). The S/N ratio for each response is computed by the following three formulas for the larger-the-better response Eq. (11) the smaller-the-better response Eq. (12) and the nominal best response Eq. (13):

$$S/N = -10\log_{10}\left(\sum_{j=1}^{N}\left(1/Y_j^2/n\right)\right) \tag{11}$$

$$S/N = -10\log_{10}\left(\sum_{j=1}^{N}\left(Y_j^2/n\right)\right) \tag{12}$$

$$S/N = -10\log_{10}\left(\sum_{j=1}^{N}\left(Y_j^2/s_i^2\right)\right) \tag{13}$$

where y_j is the measured property, n is the number of samples in each test trial, s_i is the standard deviation

Plackett-Burman design:

For the screening and the identifying procedure, Plackett–Burman experimental design (Plackett and Burman, 1946) is adequate (Karthic *et al*. 2013, Paintsil *et al.* 2016, Thao Vi *et al.* 2017); it is based on the first-order model, linear approach Eq. (14):

$$Y_v = \alpha_0 + \sum \alpha_i X_i \qquad (i = 1..k) \tag{14}$$

where Y_v is the response or dependent variable expressed in bioenergy production as hydrogen

production potential ($ml\ l^{-1}$ medium), rate ($ml\ 1^{l-}\ h^{-1}$) and/or yield ($mol.\ mol^{-1}$ substrate), α_i is the linear regression coefficient, X_i is the level of the independent variable, and α_0 is the model intercept.

Plackett-Burman design (PBD) is a small-sized two-level factorial experimental design low and a high level coded as (-1) and (+1) respectively, programmed to identify critical physicochemical parameters from N number of variables in N+1 experiment, without recourse to the interaction effects between and among the variables. Since the sample size is traditionally small, the interaction effects are completely shrouded in the main effects. To identify the main important effects, the PBD only screens the design space. The selected parameters are further optimized by the means of an appropriate design technique of a response surface method (RSM) (Plackett and Burman 1946, Sivamani and Baskar 2015, Ekpenyong *et al.* 2017).

2.2.2 Quadratic model

A second-order model or quadratic polynomial model typically used in response surface methodology (RSM). The RSM is a tool consisting of mathematical and statistical methods used to define a relationship linking the response with the independent variables. The RSM method combines a two-level full or fractional factorial design with additional points (star points) and at least one point at the center of the experimental region, selected to find properties for example orthogonality or rotatability, to fit quadratic model. For statistical calculations, the relation between the coded and actual values is described as Eq. (15):

$$X_i = \left(A_i - A_0\right)/\Delta A \tag{15}$$

where X_i is a coded variable value; A_i is the actual variable value; A_0 is the actual value of A_i at the center point, and ΔA is the step change to the variable. The quadratic equation of the variables is given as, Eq. (16):

$$y = \beta_0 + \sum \beta_i X_i + \sum \beta_{ii} X_i^2 + \sum \beta_{ij} X_i X_j \tag{16}$$

where Y is the predicted response; β_0 is a constant; β_i is the linear coefficient; β_{ii} is the squared coefficient, and β_{ij} is the cross-product coefficient.

Differential calculations were then employed to predict the optimum values needed to acquire the maximum hydrogen yield (Sun *et al.* 2010). An analysis of variance (ANOVA) was conducted to test the significance of the fit to the second-order polynomial equation for the experimental data (Table 2). The p-value is used to determine the significance of each coefficient and the interaction degree between each independent variable. A greater F-value and a smaller p-value indicate that the independent variables are more significant. The terms with p-values less than 0.05 suggest that the model fit is statistically significant. The model terms with p-values greater than 0.10 indicate they are insignificant (Chaganti *et al.* 2012).

The coefficient of determination R^2 value indicated a close relationship between the experimental and predicted values, which suggests that this is a very reliable mathematical model for bioenergy production (Sun *et al.* 2010), the statically significance of this model was checked by F-test. The conduction of experiments replicated at the optimum conditions confirms the validity of the statistical experimental strategies by comparing the differences of the observed results and the predicted values. The accuracy of the empirical model was confirmed by several researchers (Thao Vi *et al.* 2017, Xingyong *et al.* 2019).

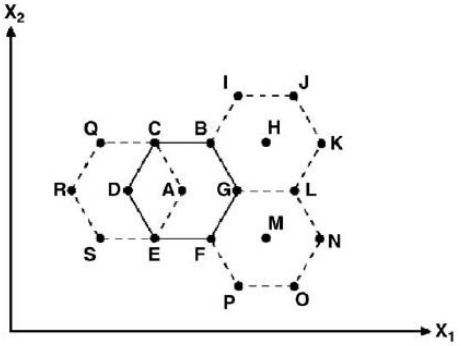

Fig. 1 Hexagonal Doehlert two factors design with three possible displacements in the experimental space

The RSM method was used to define the impact of independent first order (linear), second-order and combination mixed variable terms on the hydrogen production via more complex experimental designs such as Doehlert matrix (DM), central composite designs (CCD) and three-level designs such as the Box-Behnken design (BBD), and to optimize the responses using desirability function method. Predicted values obtained from each response surface are transformed into a dimensionless scale d_i. The desirability function scale ranges between $d = 0$ (for an undesirable response value) and d=1 (for a totally desirable response value). D is calculated combining the individual desirability values by applying the geometric mean, Eq. (17):

$$D = \left(d_1 \times d_2 \times .. d_m\right)^{1/m} \tag{17}$$

To determine the set of variable values that maximize it, an algorithm is applied to the D function. This function has been frequently used during the optimization of analytical systems, which involve several responses (Chaganti *et al.* 2012, Ferreira *et al.* 2004, Ferreira *et al.* 2007).

Doehlert matrices:

For two factors, the Doehlert design consists of one central point and six points forming a regular hexagon (i.e., situated on a circle). In three dimensions it can be viewed in different ways, depending on the plane projection selected. The number of experiments required (N) is given by Eq. (18):

$$N = k^2 + k + C_0 \tag{18}$$

where k is the number of factors and C_0 is the number of center points. Replicates at the central level of the factors are performed to validate the model using an estimate of experimental variance. The basic hexagon in Fig. 1 has six points lying on a circumference around the center point. The two-factor central composite design has eight points, also lying on a circumference surrounding its center point.

In Doehlert designs, the number of levels is not the same for all factors. In a two-factor case, for example, one factor is studied at five different levels, while the other is studied at only three levels. This property enables a free choice of the factors to be assigned to a large or small number of levels. Different criteria can be used to assign the factors. As a general rule, it is preferable to choose the factor with the stronger effect as the factor with five levels, to obtain the most system

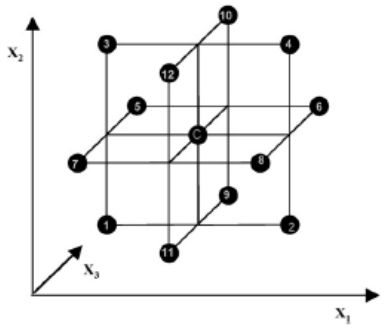

Fig. 2 Three-level full factorial design for three factors experiment

information. Also, Doehlert designs are more efficient in mapping space. The adjoining hexagons can fill space completely and efficiently, as the hexagons fill space without overlapping. Another advantage is its potential for sequencing, where experiments can be reused when the boundaries were not well chosen at first (Tarley *et al.* 2009).

Barekati-Goudarzi *et al.* (2016) used Doehlert experimental design to study the effect of the four factors and their interactions (Catalyst concentration, solvent ratio, reaction time and temperature) for microwave-assisted in-situ transesterification of Chinese tallow tree seeds (*Triadica sebifera L.*). The regression coefficient of determination and the absolute average deviation (ADD) calculated for product yield (89.19%) were 0.906% and 1.19%, respectively, indicating a good fit of the model.

Three-level full (FD) and fractional factorial design (FFD):

The three-level full factorial design was adopted by Wang and Wan (2011) to investigate the combined effects of temperatures and initial pH on fermentative hydrogen production by mixed cultures in batch tests. To obtain the optimal temperature and the optimal initial pH for fermentative hydrogen production the modified Ratkowsky model Eq. (19) was used to describe and to fit the combined effects of two factors on the substrate degradation efficiency, hydrogen yield and the average hydrogen production rate.

$$y = a.\left\{\left(T - T_{\min}\right).\left\{1 - \exp\left[b\left(T - T_{\max}\right)\right]\right\}.\left(pH - pH_{\min}\right).\left\{1 - \exp\left[c.\left(pH - pH_{\max}\right)\right]\right\}\right\}^2 \quad (19)$$

where y is the response variable, a, b and c are constants, T_{min} and T_{max} are the minimum and maximum temperatures at which y is zero, and pH_{min} and pH_{max} are the minimum and maximum initial pH at which y is zero. Subsequently, the variables at which the maximum response variable was obtained were estimated from Eq. (16). Also, the effect of two variables on the response variable was studied by plotting response surface plots and contour plots based on Eq. (16) with varying the two variables within the experimental design range. Gadhe *et al*. (2014b) evaluate the efficacy of the ultrasonication pretreatment method for complex food waste before anaerobic digestion for enhancement of H_2 yield (Y_{H2}) and rate (R_{H2}) using a statistical 3^2 full factorial design.

A $3^{(k-p)}$ fractional factorial design was used as the experimental model to optimize the process parameters for the production of H_2 to develop a model that would predict the H_2 yield as a function of k factors and to optimize them (Fig.2). Chaganti *et al.* (2012) develop a model that

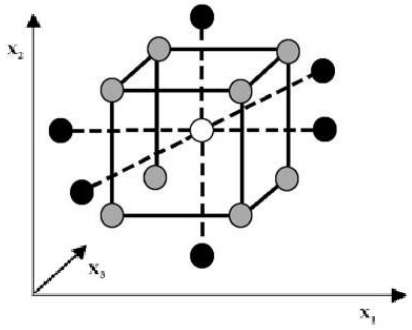

Fig. 3 Central composite design for three factors experiment

would predict the H_2 yield as a function of pH, the oleic acid (OA) concentration and the biomass concentration using $3^{(k-p)}$ fractional factorial designs. The authors reported that the model obtained to provide a useful approach for predicting optimum conditions for maximum H_2 production by inhibiting the H_2 consumers in mixed anaerobic cultures.

Central composite design(CCD) and Box- Behnken design(BBD):

A statistical designs CCD and BBD are an efficient tool for optimizing process parameters of bioenergy production and enhancing rate and yield of bioenergy production (Bouaid *et al.* 2009, Jha *et al*. 2017, Selvakumar *et al.* 2018, Garimella *et al.* 2019, Chollom *et al*. 2020).

A central composite design (CCD), is a five-level fractional factorial design developed by Box and Wilson (1951), was used to optimize the key factors that enhanced bioenergy production at the design center point. The number of experiments for a central composite design is given by Eq. (20) (Ferreira *et al.* 2007).

$$N = 2^k + 2k + C_0 \tag{20}$$

where k is the number of factors and C_o is the number of central points. The overall structure of a three-factor central composite design is shown in Fig.3. The levels of the independents variables are coded as (-1, 0, +1, -α and +α).

Box–Behnken design (BBD) is an autonomous, rotatable quadratic design without inserted factorial or fractional factorial points where the combinations of variables are at the midpoints of the edges of the variable space and at the center. Besides, Box–Behnken design permits evaluating the response function at middle levels and permits the determination of the system performance at any experimental point within the range studied via suitable design and analysis of experiments (Abbar 2019). For each considered factor, three equally spaced levels coded -1, 0 and +1 are taken by the design. The number of experiments (N) required for the development of BBD is defined as Eq.(21):

$$N = 2k(k-1) + C_0 \tag{21}$$

where k is the number of factors and C_o is the number of central points. The overall structure of a three-factor Box–Behnken design is shown in Fig. 4 (Ferreira *et al.* 2007).

The CCD designs provide high-quality predictions over the entire design space but require factor settings outside the range of the factors in the factorial part ($\pm\alpha$). The Box-Behnken design is rotatable or nearly rotatable but it contains regions of poor prediction quality.

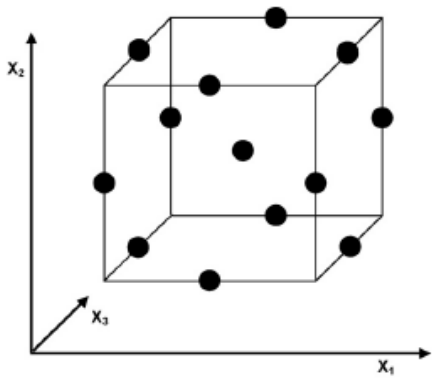

Fig. 4 Box–Behnken design for a three-factor experiment

Table 1 Structural comparison of CCD, and BBD designs for three factors

CCD				Box-Behnken			
Replication	X_1	X_2	X_3	Replication	X_1	X_2	X_3
1	-1	-1	-1	1	-1	-1	0
1	+1	-1	-1	1	-1	-1	0
1	-1	+1	-1	1	-1	+1	0
1	+1	+1	-1	1	+1	+1	0
1	-1	-1	+1	1	-1	0	-1
1	+1	-1	+1	1	+1	0	-1
1	-1	+1	+1	1	-1	0	+1
1	+1	+1	+1	1	+1	0	+1
1	-1.682	0	0	1	0	-1	-1
1	1.682	0	0	1	0	+1	-1
1	0	-1.682	0	1	0	-1	+1
1	0	1.682	0	1	0	+1	+1
1	0	0	-1.682	3	0	0	0
1	0	0	1.682				
6	0	0	0				
Total Runs = 20				Total Runs = 15			

For three factors, the central composite design and Box-Behnken design offer some advantages in requiring a fewer number of runs. For 4 or more factors, this advantage disappears. The CCD and BBD designs require, respectively, 5 and 3 levels for each factor. Table 1 summarizes the structural comparisons of CCD and BBD designs for three factors.

The α value for rotatability depends on the number of experimental runs (the number of points in the factorial portion) of the CCD, which is given in Eq. (22) (Behbahani *et al.* 2011):

$$\alpha = \left[N_f\right]^{\frac{1}{4}} \tag{22}$$

Table 2 The α value and the number of experiments required by CC and Box-Behnken Designs

Number of Factors	CCD		BBD
	α value	Number of experiments	Number of experiments
2	1.414	13	-
3	1.682	20	15
4	2.000	30	27
5	2.000	33	46
6	2.378	54	54

where N_f is the number of factorial runs of the design, $N_f=2^k$, k is the number of factors.

Table 2 compares the number of runs required for a given number of factors for Central Composite and Box-Behnken designs. A comparison between the Box-Behnken design and CCD has demonstrated that the Box-Behnken design is slightly more efficient and more economical design than the central composite design.

2.3 Neural network

A neural network (NN) can depict successfully the interactive effects among different factors in multivariate non-linear bioprocesses. Neural network design frequently reported in the literature include recurrent, Hopfield network, Kohonen, Boltzmann machine, support vector machines, radial basis function networks, the single-layer/multilayer perceptron and feed-forward backpropagation (Sathasivam *et al*. 2011, Hanrahan 2011, Bhattacharyya 2012).

In general, a feed-forward neural network with one hidden layer containing certain hidden neurons can give accurate approximations to many nonlinear functions. The inputs chosen are factors, while the outputs are response efficiencies, such as the production yield and the average production rate. The critical aspect is the choice of the number of neurons in the hidden layers. In the training process, the mean square error (MSE) between the experimental data and the corresponding predicted data is calculated and propagated backward through the network (Nagata *et al*. 2003, Wang and Wan 2009b).

The backpropagation (BP) algorithm adjusts the weights in each successive layer to reduce the error. This procedure is repeated until the error between the experimental data and the corresponding predicted data satisfies certain error criteria and achieves the desired result (Nagata *et al*. 2003, Wang and Wan 2009b). The BP neural network has many advantages, it can solve the nonlinear complex relationship problems of the research object, it can improve the computing speed of the algorithm, there is no fixed algorithm to calculate the relationship of network input-output, which is determined by the weights of layers connected (Li and Liu 2019).

The backpropagation neural network (BPNN) is one of the most significant and extensively used models for training of feedforward networks and has been frequently studied in bioenergy process (Prakasham *et al*. 2011, Mohamed *et al*. 2013, Mahanty *et al*. 2013, Karthic *et al*. 2013, Betiku and Taiwo 2015, Sewsynker-Sukai *et al* 2017, Sivamani *et al*. 2019).

2.4 Hybrid systems

The movement toward more intelligent systems requires consideration of alternative strategies

to help optimize the processes studied (Hanrahan 2011) for example, the combination between genetic algorithms (GA) and RSM (Wang and Wan 2009c, Fayyazi *et al.* 2015), (GA) and NN (Wang and Wan 2009c), fuzzy logic (FL) and NN (ANFIS) (Ishola *et al*. 2019), FL and RSM, FL and Particle Swarm Optimization (PSO) (Nassef *et al*. 2019), PSO and GA (Bertram *et al*. 2016), NN and Ant Colony Optimisation (ACO) (Beltramo *et al*. 2016, Sebayang *et al*. 2017, Silitonga *et al.*. 2019, Silitonga *et al*. 2019), NN-ACO-GA (Beltramo *et al*. 2019). The hybrid technology can impart the efficiency and accuracy needed (Whiteman and Kana 2014, Jha *et al*. 2017, Ardabili *et al*. 2018).

Fuzzy logic (FL) is an intelligent method, robust, quick, and cheap method; it is an extension of the binary logic. Unlike binary-valued logic 0 or logic 1, FL is a multivalued logic. Also, FL mimics the humanlike representation of events. For example in binary logic, if the temperature is above certain deterministic value we call it hot and it is cold if it is below this value. On the other hand, in FL the temperature variable can be represented, for example by linguistic notations such as very hot, hot, normal, cold, and very cold. The output of the fuzzy system is usually obtained after certain sequential steps; fuzzification, inference, and defuzzification (Nassef *et al*. 2019, Ardabili *et al.* 2019).

Genetic algorithm (GA) is a highly parallel, randomly searching algorithm that emulates evolution according to the Darwinian survival of the fittest principle. GA had a much higher modeling ability; it explores all regions of the solution space using a population of individuals. Each individual represents a set of independent variables (Wang and Wan 2009c, Hanrahan 2011, Ardabili *et al.* 2018).

Particle Swarm Optimization (PSO) is an algorithm that emulates the birds' movements. It is an optimization tool based on the use of several proposed solutions usually called particles. These particles constitute a swarm to move within the search space for exploring the location of the best solution. During the search process, every particle adjusts its "flying" direction and position based on its own best experience in the swarm to be towards the area that contains the global optimum (Wolf *et al*. 2008, Nassef *et al*. 2019, Kadi *et al.* 2019).

Ant colony optimization (ACO) describes one of the swarm intelligence techniques; it is a recently represented optimization algorithm capable of optimizing complex process parameters. The development of this method was inspired by insect colonies that possess an outstanding social structure. The basis of the algorithm is the structure of the ant's natural behavior looking for food. On their way to the food source and back to the nest, they leave a pheromone trail. The pheromone trail serves as an indirect means of communication between the ants, identifying the pathways to the food source. All the ants move at the same speed and spread pheromone at the same rate. The pheromone evaporates at a constant rate as well. Hence the shortest pathways will be mostly used and contain accordingly the highest concentration of pheromone. This principle of the "shortest path" is used by the ACO algorithm. The described methodology has already been applied successfully to the optimization of bioenergy production processes (Beltramo *et al*. 2016, Sebayang *et al*. 2017).

GA, PSO, and ACO have been widely employed in pure and hybrid form with successful results. The use of hybrid methods such as GA-ANN or GA-RSM leads to an improvement and optimization of the process of bioenergy production whereas the use of Multilayered perception (MLP) and Adaptive Neuro-Fuzzy inference system (ANFIS) methods leads to the highest correlation and the lowest error for prediction of the bioenergy production (Wang and Wan 2009c, Whiteman and Kana 2014, Jha *et al*.2017, Ardabili *et al*. 2018). PSO and GA, NN-ACO and NN-ACO-GA hybrids help to minimize the model dimension, to identify the significant process

variables, to improve the prediction performance of the ANN models and the optimization capacity to produce a high yield of bioenergy and to reduce the cost, time and effort associated with experimental techniques (Beltramo *et al*. 2015, Beltramo *et al*. 2016, Sebayang *et al*. 2017, Beltramo *et al*. 2019, Silitonga *et al*. 2019).

Noteworthy that, the Artificial Bee Colony (ABC) optimization has been utilized in pure form (Rostami *et al*. 2016, Garlapati *et al* 2017), to our knowledge, the hybrid form of ABC has not yet been utilized in the bioenergy production field.

The Artificial Bee Colony (ABC) algorithm is an optimization algorithm based on the intelligent behavior of honey bee swarm do when looking for the food source. In ABC, the main components are the colony of artificial bees which contains three groups of bees (employed bees, onlookers bees, and scouts bees), and food source. For every food source, there is only one employed bee (Karaboga and Basturk 2007, Rostami *et al*. 2016). Artificial Bee colony optimization has been utilized to optimize the variables to enhance the production of biodiesel from mountain almond (Prunus Scoparia) oil using an ultrasonic system (Rostami *et al*. 2016) and methyl butyrate through immobilized lipase-mediated transesterification (Garlapati *et al* 2017). The BCA's pitfall is that it is good for exploration but it lacks the exploitation factor. The capabilities of the bee colony can be enhanced using a hybrid representation (Castillo-Villar 2014).

2.5 Model validation criteria

To evaluate the adequacy and capabilities of the developed model and to assess the accurate prediction of each variable studied, different statistical indicators were used in the literature, such as coefficient of determination (R^2), mean square error (*MSE*), root mean square error (*RMSE*), mean absolute error (*MAE*), standard error of prediction (*SEP*) and absolute average deviation (*AAD*) (Wang and Wan 2009c, Sathish and Vivekanandan, 2016, Adepoju *et al*. 2018, Sivamani *et al.* 2019, Silitonga *et al.* 2019), (Table 3).

R^2 is the measure of how well the model describes the experimental data, in other words, it measures the alignment of the dataset to the regression line of the plot of data predicted against the corresponding experimental data to visualize the modeling ability of the model used. R^2 equals 0 when the values of the factors do not allow any prediction of the responses, and equals 1 when the model can perfectly predict the responses from the factors studied. For a model to be adjudged reliable, its R^2 should be at least 80% (Motulsky and Christopoulos 2013, Ishola *et al*. 2019).

MSE is defined as the expectation of the squared difference between an estimator y_i and its true value $y_{i,exp}$. *RMSE* depends on the scale of the dependent variable, it is the square root of *MSE*; both will control the models in the same way. *MAE* is the sum of absolute differences between true value $y_{i,exp}$ and predicted variables *yi*. It gives the average magnitude of forecast errors, while *RMSE* and *MSE* give more weight to the largest errors. *SEP* is the square root of an estimate of the unconditional mean squared error of prediction. *AAD* is used to address both the mean and the variability in a single measure. These statistical indicators can only be used for a relative comparison of forecasts for the same series through different models. The deviance measures are zero only if the values are identical; consequently, the most preferable model is the one that has the lowest value (Booth and Hobert 1998, Burati and Weed 2006, Jachner *et al.* 2007, Siu-Woon Ng 2011, Kato 2016). In the literature, Wang, and Wan (2009c), Sivamani *et al.* (2019) and Selvakumar *et al*. (2018) have used several model evaluation metrics simultaneously to evaluate the optimization efficiency of RSM and ANN modeling.

Table 3 Model evaluation criteria

R^{2*}	$R^2 = 1 - \sum_{i=1}^{n}\left(\frac{\left(y_{i,\exp} - y_{i,cal}\right)^2}{\left(y_{i,avg} - y_{i,\exp}\right)^2}\right)$
MSE^*	$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_{i,\exp}\right)^2$
$RMSE^*$	$RMSE = MSE^{\frac{1}{2}}$
SEP^*	$SEP = \frac{RMSE}{y_e}.100$
MAE^*	$MAE = \frac{1}{n}\sum_{i=1}^{n}\left\|y_i - y_{i,\exp}\right\|$
AAD^*	$AAD = \frac{1}{n}\left\{\sum_{i=1}^{n}\left(\frac{y_{i,\exp} - y_i}{y_{i,\exp}}\right)\right\}.100$

*where n is the number of experimental data, y_i is the predicted value obtained from the model, $y_{i,exp}$ is the actual value, $y_{i,avg}$ is the average of the actual values, y_e is the mean value of experimental data.

3. A comparative study between the significations of different DOE in the bioenergy production process

A comparison of the efficiency of CCD, BBD, and Doehlert designs indicates that both the FFD and CCD were applied to optimize the main factors which significantly affect the production process. Doehlert matrices and BBD are more efficient than CCD. CCD is much more efficient than FFD designs. The FFD is also expensive when the factor number is upper than 2 (Adnan *et al.* 2014, Oiwoh *et al.* 2018, Garimella *et al*. 2019).

Doehlert designs are more efficient in mapping space; adjoining hexagons can fill space completely and efficiently since the hexagons fill space without overlap. Another advantage is its potential for sequentially, where experiments can be re-used when the boundaries have not been well chosen at first (Ferreira *et al.* 2004, Massart *et al.* 1997, Barekati-Goudarzi *et al.* 2016).

BBD does not contain combinations of factors at their highest or lowest level (extreme conditions), which avoids dissatisfaction with experimental results. Conversely, they are not indicated for situations in which we would like to know the responses at the extremes, that is, at the vertices of the cube (Ferreira *et al.* 2007).

RSM is the most widely used method, ANN and ANFIS modeling could be presented as a better alternative tool for the prediction of process parameters in optimization. Talebian-Kiakalaieh *et al*. (2013) reported that RSM (R^2 = 0.9987) was slightly better than ANN (R^2 = 0.985) in data fitting and estimation capabilities. However, ANN prediction for reaction conversion at the optimum amount of variables was 88.40% which was better than 87.98% achieved by RSM. Ishola

B *et al*. 2019 reported that the R^2 for the RSM, ANN and ANFIS models were 0.97899, 0.98746 and 0.99435, respectively. The MSE values obtained for the three models were all low, supporting a good fit of the models. Similarly, it has been shown that ANFIS is superior to ANN.

NND is a superior and more accurate modeling technique compared with the RSM, as it represents the non-linearities in a much better way. It has been reported that the genetic algorithm based on a neural network had a higher optimizing ability than response surface methodology. Otherwise, without multiple-response optimization, several responses would have to be optimized separately (Ferreira *et al.* 2004, Ferreira *et al.* 2007, Wang and Wan 2011). Simultaneous optimization of multiple responses by the method of desirability function involves first building an appropriate model for each response and then trying to find a set of optimal operating conditions.

In Table 4, several parameters have been listed to compare the validity of mathematical models used for the production of biodiesel, biohydrogen, bioethanol, and biogas.

Table 4 Bioenergy production studies with different mathematical modeling methods

Bioenergy production	Biomass/ microbial biomass	Factors studied	experimental design	Production yield	Model accuracy indicators	References
Biodiesel	oleic acid	Time, Catalyst concentration Solvent ratio, T*	NND	81.8%	R^2 0,95	(Fauzi and Amin, 2013)
			CCD	77.7%	R^2 0.832	
	waste cooking oil	T*, Methanol /oil molar ratio, time, Catalyst-loading	NND	88.40%	R^2 0.985	(Talebian-Kiakalaieh *et al.* 2013)
			CCD	87.98%	R^2 0.9987	
	yellow oleander oil	Methanol/oil ratio, H_2SO_4, time, T*	NND	94.97%	R^2 0.9999	(Betiku and Ajala, 2014)
			BBD	95.25%	R^2 0.9947	
	rubber seed oil	molar ratio catalyst, reaction	NND	95.95%	R^2 0.9885 MSE 1.24	(Bharadwaj *et al.* 2018)
			CCD		R^2 0.8732	
	Palm Oil	Methanol/oil molar ratio, T*, catalyst content, time	NND	97.95%	R^2 0.9958 MSE 0.0010	(Thoai *et al.* 2018)
			CCD	99.64%	R^2 0.9953 MSE 0.0879	
	waste used oil	Reaction time, Catalyst, Methanol/oil ratio amount	NND	98.46%	R^2 0.9950 RMSE 0.697	(Adepoju *et al.* 2018)
			BBD	92.45%	R^2 0.9979 RMSE 0.708	
	Microalgae	pressure, number of passes, time	BBD	18%	R^2 0.6853	(Onumaegbu *et al*. 2018)
			FL-PSO	29.57%	MSE 2.060	(Nassef *et al.* 2019)
	Waste cooking oil	T*, catalyst loading, methanol/ oil, reaction time	NND	88.3%	R^2 0.99957 RMSE 0.570	(Anbessa and Karthikeyan, 2019)
			CCD		R^2 0.9976 RMSE 1.178	
	Sorrel oil	methanol/oil molar ratio, T*, catalyst weight, time	NND	99.42%	R^2 0.9875 MSE 0.2774	(Ishola *et al.* 2019)
			ANFIS	99.71%	R^2 0.9944 MSE 0.1210	
			CCD	98.61%	R^2 0.9790 MSE 0.4482	

Table 4 (Continued)

Bioethanol production	oil of paradise kernel (*Simarouba glauca*)	Oil-to-alcohol ratio, Time, T*	BBD	62%	RMSE 0.54 R^2 0.97 AAD 7.15 %	(Sivamani *et al.* 2019)
			NND		RMSE 0.29 R^2 0.99 AAD 5.91%	
	E. coli SS1, glycerol	pH, T*, Trace element, ON*, Substrate, Salt	FFD	15.72 ± 0.26 g/L	R^2 0.9133	(Adnan *et al.* 2014)
		pH, Salt, ON*, Substrate	CCD		R^2 0.9474	
	Bread fruit starch	Bread fruit starch hydrolyzate concentration, Time, pH	NND	4.21%	R^2 1 AAD 0.09%	(Betiku and Taiwo 2015)
			BBD	3.95%	R^2 0.9882 AAD 1.67%	
			CCD	9.39 ± 0.33 g/L	R^2 0.9691 AAD 27.09% MSE 1.1082	
	Manihot esculenta Crantz YTP1	pH, T*, agitation and time			RMSE 1.0526 R^2 0.9794	(Selvakumar *et al.* 2018)
			NND	9.39 ± 0.33 g/L	AAD 13.48% MSE 0.4324 RMSE 0.6575	
	mixed cultures	T*, pH, Glucose concentrations	RSM	289.8 ml/g of glucose	RMSE 16.6% SEP 38.4%	(Wang and Wan, 2009c)
			ANN	360.5 ml/g of glucose	RMSE 7.7% SEP 17.8%	
Bio-hydrogen production	*Enterobacter* species (MTCC 7104)	Xylose concentration, pH, and peptone concentration	CCD	1.94 mol/mol	R^2 0.961 RMSE 7.6%	(Karthic *et al.* 2013)
			NND		R^2 0.995 RMSE 2.7%	
	Upflow Anaerobic Sludge Blanket	HRT*, ICV*, T*	NND	0.92 mol/mol	R^2 0.99 RMSE 2.22%	(Jha *et al.* 2017)
			BBD	0.91mol /mole	R^2 0.90 RMSE 9.64%	
	Rhodobacter sphaeroides	Lactose, Uracil, biotin	BBD	6.8ml/30ml	R^2 0.9681	(Garimella *et al.* 2019)
			CCD	7.6ml/30ml	R^2 0.8848	
Biogas	Agricultural waste, mixed cultures	T*, pH, Substrate concentration, Agitation time	CCD	537 ml CH_4/g Vs	R^2 0.991 MAE 1.98%	(Sathish and Vivekanandan, 2016)
			NND		R^2 0.998 MAE 1.01%	
	mixed cultures	Organic Loading Rate, HRT*, pH	BBD	5955.4mL/d ± 225.3	R^2 0.978	(Chollom *et al.* 2020)
			CCD	4636.31mL/d ± 439.81	R^2 0.961	

*T: Temperature, ON: Organic Nitrogen, HRT: Hydraulic Retention Time, ICV: Immobilized Cell Volume.

4. Comparative study between OFAT and DOE

Many researchers often carry out one-factor-at-time experiments. While, when studying more than one factor, the use of statistical experimental designs is a more profitable and more effective

way allowing to vary several factors simultaneously and to determine precisely and systematically the interaction between factors and the impact of each one on the response. Moreover, the prediction of the response and the optimization of the processes become more efficient because the experimental information is located in a larger region of the whole factor space (Czitrom 1999, Frey *et al.* 2018).

In a one-factor-at-a-time (OFAT) design experiment, at one time all factors are kept fixed, while only one factor is changed to be studied (Czitrom 1999, Daniel *et al*. 2003). In the OFAT, the effects of each factor and model parameters $y=f(x)$ are estimated more accurately at the cost of a high number of experiences, which involves a waste of time and money. It is also necessary to repeat the experience at least twice. Furthermore, the interactions between factors are not estimable.

Frey *et al*. (2018), optimize simultaneously saccharification and fermentation for biobutanol production from a pretreated oil palm empty fruit bunch by *Clostridium acetobutylicum ATCC 824*. Approximately, 2.47 *g/L* and 3.97 *g/L* of biobutanol concentration were obtained using one factor at a time (OFAT) and central composite design (CCD), respectively. The RSM approach explores the interaction effects of studied factors towards biobutanol production and improves this as well.

5. Conclusions

In this paper, different modeling methods and experimental designs including linear, non-linear, OFAT, RSM, fuzzy logic and neural network designs used to screen, identify, investigate and optimize the different chemical and physical factors in bioenergy production processes were summarized. The usefulness, validity, and feasibility of each strategy were discussed to acquiring the maximum information about empirical models.

In the OFAT, the effects of each factor and model parameters are estimated more accurately at the cost of a high number of experiences, which involves a waste of time and money. It is also necessary to repeat the experience at least twice. Furthermore, the interactions between factors are not estimable.

The Plakett Burman design screens the important variables affecting the bioenergy production as well as their significance levels but does not consider the interaction effects among the variables. The FFD is less efficient and more expensive than BBD and CCD. The interactions between factors can be esteemed using CCD design, Box-Behnken design, fuzzy logic design, and neural network design. The hybrid technology (like RSM-GA, NND-GA, ANFIS, and ANFIS-GA) can provide the effectiveness and accuracy required for research.

A comparison of the different modeling techniques was carried out to help researchers choose the appropriate modeling technique to simplify the interpretation of the experimental results, to acquire the maximum amount of information with a minimum of experience, and to optimize the response studied.

References

Abbar, A.H. (2019), "Effect of electrolysis parameters on the specific surface area of nickel powder: Optimization using Box-Behnken design", *Int. J. Electrochem. Sci*., **14**, 662-678. https://doi.org/10.20964/2019.01.60.

Adepoju, T.F., Udoetuk, E.N., Olatunbosun, B.E., Mayen, I.A. and Babalola, R. (2018), "Evaluation of the effectiveness of the optimization procedure with methanolysis of waste oil as case study", *South African J. Chem. Eng.,* **25**, 169-177. https://doi.org/10.1016/j.sajce.2018.05.002.

Adnan, N.A.A., Suhaimi, S.N., Abd-Aziz, S., Hassan, M.A. and Phang, L.Y. (2014), "Optimization of bioethanol production from glycerol by Escherichia coli SS1", *Renewable Energy*, **66**, 625-633. https://doi.org/10.1016/j.renene.2013.12.032.

Anbessa, T. T. and Karthikeyan, S. (2019), "Optimization and mathematical modeling of biodiesel production using homogenous catalyst from waste cooking oil", *J. Eng. Adv. Technol.*, **9**(1), 1733-1739. https://doi.org/10.35940/ijeat.F9005.109119.

Ardabili, S. F., Najafi, B. and Shamshirband, S. (2019), "Fuzzy logic method for the prediction of cetane number using carbon number, double bounds, iodic, and saponification values of biodiesel fuels AQ3", *Environ. Progress Sustainable Energy*, **38**, 584-599. https://doi.org/10.1002/ep.12960.

Ardabili, S.F., Najafi, B., Shamshirband, S., Bidgoli, B. M., Deo, R.C. and Chau, K.W. (2018), "Computational intelligence approach for modeling hydrogen production: a review", *Eng. Appl. Comput. Fluid Mech.*, **12**(1), 438-458. https://doi.org/10.1080/19942060.2018.1452296.

Athreya, S. and Venkatesh, Y.D. (2012), "Application of taguchi method for optimization of process parameters in: improving the surface roughness of lathe facing operation", *Referred J. Eng. Sci.,* **1**(3), 13-19.

Barekati-Goudarzi, M., Boldor, D. and Nde D.B. (2016), "In-situ transesterification of seeds of invasive Chinese tallow trees (Triadica sebifera L.) in a microwave batch system (GREEN^3) using hexane as co-solvent: Biodiesel production and process optimization", *Bioresource Technol.*, **201**, 97-104. https://doi.org/10.1016/j.biortech.2015.11.028.

Behbahani, M., Moghaddam, M.R. A. and Arami, M. (2011), "Techno-economical evaluation of fluoride removal by electrocoagulation process: Optimization through response surface methodology", *Desalination*, **271,** 209-218. https://doi.org/10.1016/j.desal.2010.12.033.

Beltramo, T., Klocke, M. and Hitzmann, B. (2019), "Prediction of the biogas production using GA and ACO input features selection method for ANN model", *Info. Process. Agriculture*, **6**(3), 349-356. https://doi.org/10.1016/j.inpa.2019.01.002.

Beltramo, T., Ranzan, C., Hinrichs, J. and Hitzmann, B. (2016), "Artificial neural network prediction of the biogas flow rate optimised with an ant colony algorithm", *Biosyst. Eng.*, **143**, 68-78. https://doi.org/10.1016/j.biosystemseng.2016.01.006.

Bertram, A.M., Zhang, Q. and S.C. (2016) "A novel particle swarm and genetic algorithm hybrid method for diesel engine performance optimization", *J. Engine Res*., **17**(7), 732-747. https://doi.org/10.1177/1468087415611031.

Betiku, E. and Ajala, S.O. (2014), "Modeling and optimization of Thevetia peruviana (yellow oleander) oil biodiesel synthesis via Musa paradisiacal (plantain) peels as heterogeneous base catalyst: A case of artificial neural network vs. response surface methodology", *Industrial Crops Products*, **53**, 314-322. https://doi.org/10.1016/j.indcrop.2013.12.046.

Betiku, E. and Taiwo, A.E. (2015), "Modeling and optimization of bioethanol production from breadfruit starch hydrolyzate vis-a-vis response surface methodology and artificial neural network", *Renew Energy*., 74:87-94. https://doi.org/10.1016/j.renene.2014.07.054.

Bharadwaj, A. V. S. L. S., Niju, S., Begum, K.M.M. and Anantharaman N. (2018), "Optimization and modeling of biodiesel production using fluorite as a heterogeneous catalyst", *Energy Sources, Part A: Recovery, utilization, and environmental effects*, **41**(11), 1-17. https://doi.org/10.1080/15567036.2018.1549165.

Bhattacharyya, S. (2012), *Neural Networks: Evolution, Topologies, Learning algorithms and Applications, in Cross-Disciplinary Applications of Artificial Intelligence and Pattern Recognition: Advancing Technologies*, IGI Global edition, USA.

Bina, B., Amin, M. M., Pourzamani, H., Fatehizadeh, A., Ghasemian, M., Mahdavi, M. and Taheri, E. (2019), "Biohydrogen production from alkaline wastewater: The stoichiometric reactions, modeling, and electron equivalent", *MethodsX*, **6**, 1496-1505. https://doi.org/10.1016/j.mex.2019.06.013.

Boni, M.R., Sbaffoni, S., Tuccinardi, L. and Viotti, P. (2013), "Development and calibration of a model for biohydrogen production from organic waste", *Waste Manage.,* **33,** 1128-1135. https://doi.org/10.1016/j.wasman.2013.01.019.

Booth, J. G. and Hobert, J. P. (1998), "Standard errors of prediction in generalized linear mixed models", *J. American Statistical Assoc.*, **93**(441), 262-272.

Bouaid, A., Martinez, M. and Aracil, J., (2009), "Production of biodiesel from bioethanol and Brassica carinata oil: Oxidation stability study", *Bioresource Technol.,* **100**, 2234-2239. https://doi.org/110.1016/j.biortech.2008.10.045.

Box, G.E.P. and Wilson, K.B. (1951), "On the Experimental Attainment of Optimum Conditions", *J. Royal Statistical Society B*, **13**, 1-45, United Kingdom.

Burati, J.L. and Weed, R.M. (2006), "Accuracy and precision of typical quality measures. transportation research record", *J. Transport. Res. Board*, **1946**(1), 82-91.

Canabarro, N., Soares, J.F., Anchieta, C.G., Kelling, C.S. and Mazutti, M.A. (2013), "Thermochemical processes for biofuels production from biomass", *Sustainable Chem. Processes*, **1**(1), 22. https://doi.org/10.1186/2043-7129-1-22.

Castillo-Villar, K.K. (2014), "Metaheuristic algorithms applied to bioenergy supply chain problems: theory, review, challenges, and future", *Energies,* **7**, 7640-7672. https://doi.org/10.3390/en7117640.

Chaganti, S.R., Kim, D.H. and Lalman, J.A. and Shewa, W.A. (2012), "Statistical optimization of factors affecting biohydrogen production from xylose fermentation using inhibited mixed anaerobic cultures", *J. Hydrogen Energy,* **37,** 11710 -11718. https://doi.org/10.1016/j.ijhydene.2012.05.036.

Chen, W.H., Lee, K. T. and Ong H. C. (2019), "Biofuel and bioenergy technology", *Energies*, **12**(2), 290. https://doi.org/10.3390/en12020290.

Chollom, M.N., Rathilal, S., Swalaha, F.M., Bakare, B.F. and Tetteh, E. K. (2020), "Comparison of response surface methods for the optimization of an upflow anaerobic sludge blanket for the treatment of slaughterhouse wastewater", *Environ. Eng. Res.*, **25**(1) 114-122. https://doi.org/10.4491/eer.2018.366.

Czitrom, V. (1999), "One-Factor-at-a-Time versus designed experiments", *The American Statistician*, **53**(2), 126-131. https://doi.org/10.1080/00031305.1999.10474445.

Deepanraj, B., Sivasubramanian, V. and Jayaraj, S. (2015), "Experimental and kinetic study on anaerobic digestion of food waste: The effect of total solids and pH", *J. Renewable Sustainable Energy,* **7**, 063104. https://doi.org/10.1063/1.4935559 .

Demirbas, M.F. (2009), "Biorefineries for biofuel upgrading: a critical review", *Appl Energy*, **86**, S151-61. https://doi.org/10.1016/j.apenergy.2009.04.043.

Ekpenyong, G., Antai1, S.P., Asitok, A.D. and Ekpo, B.O. (2017), "Plackett-Burman Design and Response Surface Optimization of Medium Trace Nutrients for Glycolipopeptide Biosurfactant Production Maurice", *Iranian Biomedical J.*, **21**(4), 249-260. https://doi.org/10.18869/acadpub.ibj.21.4.249.

Fayyazi, E, Ghobadian., B, Najafi., G, Hosseinzadeh., B, Mamat., R. and Hosseinzadeh, J. (2015), "An ultrasound-assisted system for the optimization of biodiesel production from chicken fat oil using a genetic algorithm and response surface methodology", U*ltrason Sonochem*., **26**, 312-320. https://doi.org/10.1016/j.ultsonch.2015.03.007.

Ferreira, S.L.C., Bruns, R.E., Ferreira, H.S., Matos, G.D., David, J.M., Brand˜ao, G.C., da Silva, E.G.P., Portugal, L.A., dos Reis, P.S., Souza a, A.S. and dos Santos, W.N.L. (2007), "Box-Behnken design: An alternative for the optimization of analytical methods", *Analytica Chimica Acta*, **597**, 179-186. https://doi.org/10.1016/j.aca.2007.07.011.

Ferreira, S.L.C., dos Santos, W.N.L., Quintella, Ben´ıcio. Neto C.M. and Bosque-Sendra J.M. (2004), "Doehlert matrix: A chemometric tool for analytical chemistry", *Talanta*, **63**, 1061-1067. https://doi.org/10.1016/j.talanta.2004.01.015.

Frey, D. D., Engelhardt, F. and Greitzer, E.M. (2003) "A role for 'one-factor-at-a-time' experimentation in parameter design", *Res. Eng. Design,* **14**, 65-74. https://doi.org/10.1007/s00163-002-0026-9.

Gadhe, A., Sonawane, S.S. and Varma, M.N. (2014a), "Kinetic analysis of biohydrogen production from complex dairy wastewater under optimized condition", *J. Hydrogen Energy*, **39**, 1306-1314. https://doi.org/10.1016/j.ijhydene.2013.11.022.

Gadhe, A., Sonawane, S.S. and Varma, M.N. (2014b), "Ultrasonic pretreatment for an enhancement of biohydrogen production from complex food waste", *J. Hydrogen Energy,* **39**, 7721-7729. https://doi.org/10.1016/j.ijhydene.2014.03.105.

Garlapati, V.K., Vundavilli, P.R. and Banerjee, R. (2017), "Optimization of flavour ester production through artificial bee colony algorithm: ABC optimization approach for flavour ester production", *4th IEEE ICIIP 2017 Fourth International Conference on Image Information Processing,* Shimla, India, December. https://doi.org/10.1109/iciip.2017.8313694.

Guerrero, A.B. and Muñoz, E. (2018), "Life cycle assessment of second generation ethanol derived from banana agricultural waste: environmental impacts and energy balance", *J. Clean Prod.*, **174**,710-717. https://doi.org/10.1016/j.jclepro.2017.10.298.

Gunst, R.F. and Robert, L. (2009), "Fractional factorial design", *Computational Statistics*, **1**(2), 234-244.

Hanrahan, G. (2011), *Artificial Neural networks in Biological and Environmental Analysis*, *Analytical Chemistry Series,* **182**, Duke University, Los Angeles, California.

Hasegawa, F., Yokoyama, S. and Imou, K. (2010), "Methanol or ethanol produced from woody biomass: Which is more advantageous?", *Bioresource Technol.*, **101**, S109-S111. https://doi.org/10.1016/j.biortech.2009.05.008.

Hossain, S.M.Z. (2019), "A Tutorial review on biochemical conversion of microalgae biomass into biofuel.", *Chem. Eng. Technol.*, https://doi.org/10.1002/ceat.201800605.

Huang, D., Zhou, H. and Lin, L. (2012), "Biodiesel: An alternative to conventional fuel", *Energy procedia,* **16**, 1874-1885. https://doi.org/10.1016/j.egypro.2012.01.287.

Ishola, N.B., Okeleye, A.A., Osunleke, A.S. and Betiku, E. (2019), "Process modeling and optimization of sorrel biodiesel synthesis using barium hydroxide as a base heterogeneous catalyst: appraisal of response surface methodology, neural network and neuro-fuzzy system", *Neural Comput. Appl.*, **31**, 4929-4943. https://doi.org/10.1007/s00521-018-03989-7.

Jachner, S., Boogaart, K. G.v.d. and Petzoldt, T. (2007), "Statistical Methods for the Qualitative Assessment of Dynamic Models with Time Delay (R Package qualV)", *J. Statistical Software*, **22**(8), 1-30. https://doi.org/10.18637/jss.v022.i08.

Jha, P., Kana, E. and Schmidt, S. (2017), "Can artificial neural network and response surface methodology reliably predict hydrogen production and COD removal in an UASB bioreactor?", *J. Hydrogen Energy*, **42**(30), 18875-18883. https://doi.org/10.1016/j.ijhydene.2017.06.063.

Kadi, M.A., Akkouche, N., Awad, S., Loubar, K. and Tazerout, M. (2019), "Kinetic study of transesterification using particle swarm optimization method", *Heliyon*, **5**(8), e02146. https://doi.org/10.1016/j.heliyon.2019.e02146 .

Kaparaju, P**.**, Serrano, M**.**, Thomsen, A.B**.**, Kongjan, P. and Angelidaki, I. (2009), "Bioethanol, biohydrogen, and biogas production from wheat straw in a biorefinery concept", *Bioresour Technol.,* **100**(9), 2562-2568. https://doi.org/10.1016/j.biortech.2008.11.011.

Karaboga, D. and Basturk, B. (2007), "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm", *J. Global Optimization*, **39**(3), 459-471. https://doi.org/10.1007/s10898-007-9149-x.

Karthic, P., Joseph, S., Arun, N., Kumaravel, S. (2013), "Optimization of biohydrogen production by Enterobacter species using artificial neural network and response surface methodology", *J. Renewable Sustainable Energy*, **5**(3), 033104, 1-13. https://doi.org/ 10.1063/1.4803746.

Kato, T. (2016), *Prediction of Photovoltaic Power Generation Output and Network Operation, in Integration of Distributed Energy Resources in Power Systems*, 77-108, Academic Press, Japan.

Kumar, P., Barrett, D.M., Delwiche, M. J. and Stroeve, P. (2009), "Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production", *Ind. Eng. Chem. Res.,* **48**(8), 3713-3729. https://doi.org/10.1021/ie801542g.

Lardon, L., Hélias, A., Sialve, B., Steyer, J.P. and Bernar, O. (2009), "Life-Cycle assessment of biodiesel production from microalgae", *Environ. Sci. Technol.,* **43**(17), 6475-6481. https://doi.org/10.1021/es900705j .

Li, D. and Liu, S. (2019), *Water Quality Evaluation in Water Quality Monitoring and Management: Basis,*

Technology and Case Studies, 113-159, Academic Press, London. https://doi.org/10.1016/B978-0-12-811330-1.00004-1.

Li, P., Xu, L., Mou, Y., Shan, T., Mao, Z., Lu, S., Peng, Y. and Zhou, L. (2012), "Medium optimization for exopolysaccharide production in liquid culture of *endophytic ungus berkleasmium* sp. dzf12", *Int. J. Mol. Sci*., **13**(9), 11411-11426. https://doi.org/10.3390/ijms130911411.

Luftig, J.T. and Jordan, V.S. (1998), *Design of Experiments in Quality Engineering*, McGraw-Hill, NY, USA.

Mahanty, B., Zafar, M. and Park, H.S. (2013), "Characterization of codigestion of industrial sludges for biogas production by artificial neural network and statistical regression models", *Environ Technol*., **34**, 2145-2153. https://doi.org/10.1080/09593330.2013.819022.

Massart, D.L., Vandeginste, B.G.M., Buydens, L.M.C., de Jong, S., Lewi, P.J. and Smeyers-Verbeke, J. (1997), *Handbook of Chemometrics and Qualimetrics*, Part A, **20A**, Elsevier, Amsterdam.

Miñón-Fuentes, R. and Aguilar-Juárez, O. (2020). "Hydrogen production from coffee pulp by dark fermentation", *Water Scie. Technol.*, **80**(9), 1692-1701. https://doi.org/10.2166/wst.2019.416.

Mohamed, M.S., Tan, J.S., Mohamad, R., Mokhtar, M.N. and Ariff A.B (2013), "Comparative analyses of response surface methodology and artificial neural network on medium optimization for Tetraselmis sp. FTC209 grown under mixotrophic condition", *Sci World J*., **2013**, 1-14. https://doi.org/10.1155/2013/948940 .

Motulsky, H. and Christopoulos, A. (2003), *Fitting Models to Biological Data using Linear and Nonlinear Regression : A Practical Guide to Curve Fitting*, GraphPad Software, Inc San Diego CA, USA.

Nagata, Y. and Chu, K.H. (2003), "Optimization of a fermentation medium using neural networks and genetic algorithms", *Biotechnol Lett.*, **25**, 1837-1842. https://doi.org/10.1023/A:1026225526558.

Nassef, A. M., Sayed, E. T., Rezka, H., Abdelkareem, M. A., Rodriguez, C. and Olabi, A.G. (2019), "Fuzzy-modeling with Particle Swarm Optimization for enhancing the production of biodiesel from Microalga", *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, **41**(17), 2094-2103. https://doi.org/10.1080/15567036.2018.1549171.

Oiwoh, O., Ayodele, B.V., Amenaghawon, N.A. and Okieimen, C.O. (2018), "Optimization of bioethanol production from simultaneous saccharification and fermentation of pineapple peels using Saccharomyces cerevisiae", *J. Appl. Sci. Environ. Manage*., **22**(1) 54-59. https://doi.org/10.4314/jasem.v21i7.5.

Onumaegbu, C., Alaswad, A., Rodriguez., C. and Olabi, A. (2018), "Optimization of pre-treatment process parameters to generate biodiesel from Microalga", *Energies*, **11**(4), 806. https://doi.org/10.3390/en11040806.

Paintsil, A., Armah, F.A. and Yanful, E.K. (2016), "Assessment of the transesterification stage of biodiesel, production i: application of a plackett–burman design to select the process variables", *Waste Biomass Valor*, **8**(2), 473-481. https://doi.org/10.1007/s12649-016-9583-4.

Pan, C.M., Fan, Y.T., Xing, Y., Hou, H.W. and Zhang, M.L. (2008), "Statistical optimization of process parameters on biohydrogen production from glucose by Clostridium sp. Fanp2", *Bioresource Technol.,* **99**, 3146-3154. https://doi.org/10.1016/j.biortech.2007.05.055.

Pandey, R. K., Chand, K. and Tewari, L. (2018), "Solid state fermentation and crude cellulase based bioconversion of potential bamboo biomass to reducing sugar for bioenergy production", *J. Food Agriculture,* **98**(12), 4411-4419. https://doi.org/10.1002/jsfa.8963.

Patterson, T., Esteves, S., Dinsdale, R., Guwy, A. and Maddy, J. (2013), "Life cycle assessment of biohydrogen and biomethane production and utilization as a vehicle fuel", *Bioresource Technol.,* **131**, 235-245. https://doi.org/10.1016/j.biortech.2012.12.109.

Perera, F. (2017), "Pollution from fossil-fuel combustion is the leading environmental threat to global pediatric health and equity: solutions exist", *J. Environ. Res. Public Health,* **15**(1), 16. https://doi.org/10.3390/ijerph15010016 .

Pérez-Sariñana, B.Y., Díaz-González, A., León-Rodriguez, A.D., Saldaña-Trinidad, S., Pérez-Luna, Y.D.C., Guerrero-Fajardo, C.A.O. and Sebastian, P.J. (2019), "Methane production from coffee crop residues", *Rom Biotechnol Lett.,* **24**(4), 669-675. https://doi.org/10.25083/rbl/24.4/669.675.

Phukoetphim, N., Salakkam, A., Laopaiboon, P. and Laopaiboon, L. (2017), "Kinetic models for batch

ethanol production from sweet sorghum juice under normal and high gravity fermentations: Logistic and modified Gompertz models", *J. Biotechnol.*, **243**, 69-75. https://doi.org/10.1016/j.jbiotec.2016.12.012.

Pittman, J. K., Dean, A. P. and Osundeko, O. (2011), "The potential of sustainable algal biofuel production using wastewater resources", *Bioresource Technol.,* **102**, 17-25. https://doi.org/10.1016/j.biortech.2010.06.035.

Plackett, R.L. and Burman, J.P. (1946), "The Design of optimum multifactorial experiments", *Biometrika*, **33**(4), 305-325. https://doi.org/10.1093/biomet/33.4.305.

Prakasham, R.S., Sathish, T. and Brahmaiah, P. (2011), "Imperative role of neural networks coupled genetic algorithm on optimization of biohydrogen yield", *Int J. Hydrog. Energy*, **36** 4), 4332-4339. https://doi.org/10.1016/j.ijhydene.2011.01.031.

Razali, N.A.A.M., Ibrahim ID, M. F., Bahrin, E.K. and Abd-Aziz, S. (2018), "Optimisation of simultaneous saccharification and fermentation (ssf) for biobutanol production using pretreated oil palm empty fruit bunch", *Molecules*, **23**, 1944. https://doi.org/10.3390/molecules23081944.

Rostami, S., Samani, B. H. and Saeidi, K. (2016) "Optimization of biodiesel production from prunus scoparia using artificial bee colony algorithm", *JREE*, **3**(1), 52-58. https://doi.org/10.30501/jree.2016.70078.

Sathasivam, S., Hamadneh, N. and Choon, O.H. (2011), "Comparing neural networks: hopfield network and rbf network", *Appl. Math. Sci.*, **5**(69), 3439-3452.

Sathish, S. and Vivekanandan, S. (2016), "Parametric optimization for floating drum anaerobic bio-digester using response surface methodology and artificial neural network", *Alexandria Eng. J.,* **55**, 3297-3307. https://doi.org/10.1016/j.aej.2016.08.010.

Sebayang, A.H., Masjuki, H.H., Ong, H.C., Dharma, S., Silitonga, A.S., Kusumo, F. and Milano, J. (2017), "Optimization of bioethanol production from sorghum grains using artificial neural networks integrated with ant colony", *Industrial Crops Products*, **97**, 146-155. http://dx.doi.org/10.1016/j.indcrop.2016.11.064.

Selvakumar, P., Kavitha, S. and Sivashanmugam, P. (2018), "Optimization of process parameters for efficient bioconversion of thermo-chemo pretreated manihot esculenta Crantz YTP1 stem to ethanol", *Waste Biomass Valorization*, https://doi.org/10.1007/s12649-018-0244-7.

Sevinç, P., Gündüz, U., Eroglu, I. and Yücel, M. (2012), "Kinetic analysis of photosynthetic growth, hydrogen production and dual substrate utilization by *Rhodobacter capsulatus*", *J. Hydrogen Energy,* **37**(21), 16430-16436. http://dx.doi.org/10.1016/j.ijhydene.2012.02.176.

Sewsynker-Sukai, Y., Faloye, F. and Kana, E.B.G. (2017), "Artificial neural networks: an efficient tool for modelling and optimization of biofuel production (a mini review)", *Biotechnol. Biotechnol. Equipment*, **31**(2), 221-235. https://doi.org/10.1080/13102818.2016.1269616.

Shalaby, E.A. (2015), "A Review of Selected Non-Edible Biomass Sources as Feedstock for Biodiesel Production"*, Biofuels - Status and Perspective*, chapter 1, 1-20, InTech, London, United Kingdom.

Silitonga, A., Mahlia, T., Shamsuddin, A., Ong, H., Milano, J., Kusumo, F. and Rahman, S. (2019), "Optimization of cerbera manghas biodiesel production using artificial neural networks integrated with ant colony optimization", *Energies*, **12**(20), 3811. https://doi.org/10.3390/en12203811.

Singh, G., Mohapatra, S.K., Ragit, S.S. and Kundu, K. (2018), "Optimization of biodiesel production from grape seed oil using Taguchi's orthogonal array", *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, **40(**18), 2144-2153. https://doi.org/10.1080/15567036.2018.1495778

Siu-Woon Ng, E. (2011) "MSE < Variance? A pitfall in calculating the mean square error", *Model Assisted Statistics Appl.*, **6**, 369-371. https://doi.org/10.3233/MAS-2011-0195.

Sivamani, S. and Baskar, R. (2015), "Optimization of Bioethanol Production from Cassava Peel Using Statistical Experimental Design", *Environ. Progress Sustainable Energy,* **34**(2), 567-574. https://doi.org/10.1002/ep.11984.

Sivamani, S., Selvakumar, S., Rajendran, K. and Muthusamy, S. (2019), "Artificial neural network–genetic algorithm based optimization of biodiesel production from Simarouba glauca", *Biofuels,* **10**(3), 393-401. https://doi.org/10.1080/17597269.2018.1432267. https://doi.org/10.1080/17597269.2018.1432267.

Smyth, G.K. (2002), "Nonlinear Regression", *Encyclopedia of Environmetrics*, John Wiley Sons, Ltd,

Chichester, **3**, 1405-1411. https://doi.org/10.1002/9780470057339.van017.

Suen, C.y., Das, A. and Midha, C.K. (2013), "Optimal fractional factorial designs and their construction", *J. Statistical Planning Inference,* **143**, 1828-1834. https://doi.org/10.1016/j.jspi.2013.05.004.

Sun, Q., Xiao, W., Xi, D., Shi, J., Yan, X. and Zhou, Z. (2010), "Statistical optimization of biohydrogen production from sucrose by a co-culture of Clostridium acidisoli and *Rhodobacter sphaeroides*", *J. Hydrogen Energy,* **35**, 4076-4084. https://doi.org/10.1016/j.ijhydene.2010.01.145.

Talebian-Kiakalaieh, A., Amin, N. A. S., Zarei, A. and Noshadi, I. (2013), "Transesterification of waste cooking oil by heteropoly acid (HPA) catalyst: Optimization and kinetic model", *Appl. Energy*, **102**, 283–292. https://doi.org/10.1016/j.apenergy.2012.07.018.

Tarley, C.R.T., Silveira, G., Lopes dos Santos, W. N., Matos, G. D., Paranhos da Silva, E. G., Bezerra, M. A., Miró, M. and Ferreira S.L.C. (2009), "Chemometric tools in electroanalytical chemistry: Methods for optimization based on factorial design and response surface methodology", *Microchem. J.,* **92**, 58-67. https://doi.org/10.1016/j.microc.2009.02.002.

Thao Vi, L.V., Salakkam, A. and Reungsang, A. (2017), "Optimization of key factors affecting bio-hydrogen production from sweet potato starch", *Energy Procedia*, **138**, 973-978. https://doi.org/10.1016/j.egypro.2017.10.092.

Thoai, D. N., Tongurai, C., Prasertsit, K. and Kumar, A. (2018) "Predictive Capability Evaluation of RSM and ANN in Modeling and Optimization of Biodiesel Production from Palm (Elaeisguineensis) Oil", *J. Appl. Eng. Res.*, **13**(10), 7529-7540.

Tonini, D., Hamelin, L., Alvarado-Morales, M. and Astrup, T.F. (2016), "GHG emission factors for bioelectricity, biomethane, and bioethanol quantified for 24 biomass substrates with consequential life-cycle assessment", *Bioresour, Technol.,* **208**,123-133. https://doi.org/10.1016/j.biortech.2016.02.052.

Usmanbaha, N., Jariyaboon, R., Reungsang, A., Kongjan, P. and Chu, C.Y. (2019), "Optimization of batch dark fermentation of chlorella sp. using mixed-cultures for simultaneous hydrogen and butyric acid production", *Energies*, **12**(13), 2529. https://doi.org/10.3390/en12132529.

Van der Stelt, M.J.C., Gerhauser, H., Kiel, J.H.A. and Ptasinski, K.J. (2011), "A Biomass upgrading by torrefaction for the production of biofuels", *A Rev. Biomass Bioenergy,* **35**, 3748-3762. https://doi.org/10.1016/j.biombioe.2011.06.023.

Wang, J. and Wan W. (2011), "Combined effects of temperature and pH on biohydrogen production by anaerobic digested sludge", *Biomass Bioenergy,* **35**, 3896-3901. https://doi.org/10.1016/j.biombioe.2011.06.016.

Wang, J. and Wan, W. (2009a), "Experimental design methods for fermentative hydrogen production: a review". *J. Hydrogen Energy,* **34**, 235-44, https://doi.org/10.1016/j.ijhydene.2008.10.010.

Wang, J. and Wan, W. (2009b), "Application of desirability function based on neural network for optimizing biohydrogen production process", *J. Hydrogen Energy*, **34**, 1253-1259. https://doi.org/10.1016/j.ijhydene.2008.11.055.

Wang, J. and Wan, W. (2009c), "Optimization of fermentative hydrogen production process using genetic algorithm based on neural network and response surface methodology", *J. Hydrogen Energy*, **34**(1), 255-261. https://doi.org/10.1016/j.ijhydene.2008.10.010.

Wang, J.L. and Wan, W. (2008), "The effect of substrate concentration on biohydrogen production by using kinetic models", *Sci. China Ser. B Chem.*, **51**(11), 1110-1117. https://doi.org/10.1007/s11426-008-0104-6.

Wang, K.S., Chen, J.H., Huang, Y.H. and Huang, S.L. (2013), "Integrated Taguchi method and response surface methodology to confirm hydrogen production by anaerobic fermentation of cow manure", *J. Hydrogen. Energy*, **38**, 45-53. https://doi.org/10.1016/j.ijhydene.2012.03.155.

Weuster-Botz, D. (2000), "Experimental design for fermentation media development: Statistical design or global random search?", *J. Biosci. Bioeng.*, **90**(5), 473-83. https://doi.org/10.1016/s1389-1723(01)80027-x.

Whiteman, J. and Kana, E.G. (2014), "Comparative assessment of the artificial neural network and response surface modeling efficiencies for biohydrogen production on sugar cane molasses", *BioEnergy Research*, **7**(1), 295-305. https://doi.org/10.1007/s12155-013-9375-7.

Wirth, R., Lakatos, G., Maróti, G., Bagi, Z., Minárovics, J., Nagy, K., Kondorosi, É., Rákhely, G. and

Kovács, K.L. (2015), "Exploitation of algal-bacterial associations in a two-stage biohydrogen and biogas generation process", *Biotechnol. Biofuels,* **8**(59), 1-14. https://doi.org/10.1186/s13068-015-0243-x .

Wolf, C., McLoone, S., & Bongards, M. (2008), "Biogas plant optimization using genetic algorithms and particle swarm optimization", *IET Irish Signals and Systems Conference* (ISSC 2008), NUI, Galway, Ireland, June. https://doi.org/10.1049/cp:20080670.

Xingyong, L., Yubao, C., Xu, Z., Juhua, W., Wenhua, Z., Shijie, L., Jing, L. amd Xingling, Z. (2019), "Optimization of Operating Parameters in C8–C16 iso-Alkanes Selectivity with Response Surface Methodology", *J. Biobased Mater. Bioenergy*, **13**(1), 11-18(8). https://doi.org/10.1166/jbmb.2019.1815.

Yono, B., Syaichurrozi, I. and Sumardiono, S. (2014), "Kinetic Model of Biogas Yield Production from Vinasse at Various Initial pH: Comparison between Modified Gompertz Model and First Order Kinetic Model", *Res. J. Appl. Sci., Eng. Technol.*, **7**(13), 2798-2805. http://dx.doi.org/10.19026/rjaset.7.602.

Zhu, H., Yang, J. and Xiaowei, C. (2019), "Application of Modified Gompertz Model to Study on Biogas production from middle temperature co-digestion of pig manure and dead pigs", *4th International Conference on Advances in Energy and Environment Research ICAEER 2019, E3S Web of Conferences* **118** (03022), 1-7, Shanghai, China, August. https://doi.org/10.1051/e3sconf/201911803022.

Zolgharneina, J., Shahmoradia, A. and Ghasemi, J.B.J. (2013), "Comparative study of Box–Behnken, central composite, and Doehlert matrix for multivariate optimization of Pb (II) adsorption onto *Robinia* tree leaves", *Chemometrics*, **27**(1-2), 12-20. https://doi.org/10.1002/cem.2487.

Zwietering, M.H., Jongenburger, I., Rombouts, F.M. and van't Riet, K. (1990), "Modelling of the bacterial growth curve", *Appl Environ Microbiol*, **56**(6),1875-1881.

TK