OPTIMIZED BIODIESEL PRODUCTION FROM *C. INOPHYLLUM* BIO-OIL USING KRIGING AND ANN PREDICTIVE MODELS

by

Hariram VENKATESAN[a*], Godwin John JOHNROSE[b], Saravanan ARUMUGAM[c], Vinoth Kumar MARI[a], Balachandar MOORTHY[a], Johnson Samuel STALIN[a], and Kondeti Lakshmi VASUDEV[d]

[a]Department of Mechanical Engineering, Hindustan Institute of Technology and Science, Chennai, Tamil Nadu, India
[b]Department of Automobile Engineering, Hindustan Institute of Technology and Science, Chennai, Tamil Nadu, India
[c]Department of Chemical Engineering, Hindustan Institute of Technology and Science, Chennai, Tamil Nadu, India
[d]Department of Mechanical Engineering, Amrita School of Engineering, Amrita Vishwa Vidhyapeetham, Chennai, Tamil Nadu, India

Original scientific paper
https://doi.org/10.2298/TSCI211127032V

This work aimed at optimizing the two-stage transesterification efficiency of the production of C. inophyllum biodiesel using artificial neural network and Kriging predictive models. Response surface methodology was used to develop the central rotatable composite design of 27 trial experimental runs with variations in the input process parameters like methanol to oil molar ratio, potassium hydroxide catalyst loading, and reaction time. A multi-layered non-linear regressive artificial neural network model with feed-forward propagation and a numerical surrogate Kriging model was used to predict the C. inophyllum biodiesel yield. The efficacy of the developed model was verified using analysis of variance by comparing its coefficient of determination and the mean relative percentage deviation values. The optimized C. inophyllum biodiesel as 98.1% is derived with 0.94 v/v of methanol to oil molar ratio, 0.98 wt.% of potassium hydroxide catalyst loading, and 80 minutes reaction time with 70 ℃ constant reaction temperature as predicted by Kriging model. The optimized parameters were also verified experimentally.

Keywords*: transesterification, reaction time, Molar ratio, biodiesel yield, optimization*

Introduction

Sustained development of a country's economy depends on the transportation sector which needs energy resources. As fuel energy requirement is one of the primary responsible factors for the advance of any nation, the world has witnessed an upsurge in the demand for fuel. This led to continued depletion of fossil-based fuels. On the other hand, environmental

*Corresponding author, e-mail: connect2hariram@gmail.com

concerns are also discussed over the release of emissions like CO, SO, and hydrocarbons at greater concentrations. These factors played a major role in ascertaining an alternate source of energy which are bio-degradable, non-toxic, and environmentally friendly. The economic viability of the produced bio-oil from vegetable feedstocks is lower as there is a trade-off that exists between the edible and non-edible oil sources [1]. Presently, India is reported to import around 47% of edible oil from oil-rich countries even though it is the fifth largest bio-oil seed producing country in the world with an annual production of 9.73% of oil seeds. Bio-oil from non-edible vegetable sources is one such alternative that can be transformed into usable energy source through scientific methods like heating, thermal cracking, pyrolysis, *etc*., and the outcomes are commonly termed as *Biodiesel*. Neem (*Azadirachta indica*), Jatropha (*Jatropha curcas*), Pongamia (*Millettia pinnata*), Cotton Seed (*Gossypium herbaceum*), *etc*., are presently being used for producing biodiesel in India [2]. Transesterification is one such methodology through which vegetable-based bio-oil is transformed into biodiesel through a series of reversible chemical chain reactions [3].

The transesterification efficiency mainly depends upon its input variable factors like methanol to oil molar ratio, catalyst loading and reaction time. Maintaining one factor as constant and varying the other parameters is an optimizing strategy to study its effect on the yield of biodiesel but the inter-relationship between the input parameters could not be assessed by this method. To obtain the interaction effect among multiple parameters, researchers have reported use of techniques like design of experiments, Taguchi, soft computational approach like artificial neural networks (ANN), adaptive neuro-fuzzy interference system, response surface methodology (RSM), and Kriging to optimize the yield of biodiesel [4].

Arumugam and Ponnusami [5] outlined on the production of biodiesel from *C. inophyllum* oil through eco-friendly techniques. Biocatalyst was used in the transesterification process to stimulate the reaction rate along with supercritical methanol. Hariram *et al*. [6] optimized the transesterification process adopted to produce biodiesel from *C. inophyllum*. The analysis of variance (ANOVA) indicator proved that the developed model was more suitable with a superior accuracy and precision. Venkanna and Venkatarmana [7] used a three-stage transesterification process to transform the bio-oil of *C. inophyllum* (Honne oil) into its biodiesel. A biodiesel yield of 89% was achieved by optimizing the variable input parameters such as, reaction time, catalyst loading and molar ratio. Ayoola *et al*. [8] optimized the input variable parameters used for the transesterification process of crude palm kernel oil using ANN and RSM. The investigation outcome proved that the RSM had a better predicting capability of the biodiesel yield between 87% and 99% when potassium hydroxide (KOH) was used as catalyst.

Akshay and Siddharth [9] predicted the optimized biodiesel production factors for the algal oil using ANN and RSM. The three process parameters namely, molar ratio, catalyst loading, and reaction time were varied based on the box-behnken design of experimental trials. The ANN prediction was in good agreement with the experimental outcomes along with an appreciable value of coefficient of determination (R^2) as 0.9918 in comparison with 0.9627 for the RSM. Betiku *et al*. [10] optimized the biodiesel production process from the *Vitellaria paradox* (Shea tree nut butter) which possesses a higher free fatty acid (FFA) content. The experimental trials were modelled based on the central composite rotatable design (CCRD) using RSM technique. The ANN coupled with genetic algorithm (GA) and RSM methodology are employed to predict and optimize the biodiesel yield by varying the base-catalysed transesterification parameters like KOH catalyst loading, methanol to oil molar ratio, and reaction time. The efficacy of the developed model was assessed using ANOVA by observing

the R^2 value which was above 0.9923 and 0.9981 for RSM and ANN, respectively. Benjamin *et al*. [11] focussed on the effect of RSM and ANFIS in producing the optimized biodiesel from *Thevetia peruvianta* seed oil (Yellow oleander). The first stage reduced the FFA content to less than 1% by using methanol and ferric sulphate. The second stage being the transesterification process was accomplished using sodium hydroxide and methanol. The efficacy of the developed models viz., ANN and ANFIS were evidenced through R^2 values of ANOVA. Biodiesel yield of 99.8% was obtained by the experimental procedure which was in good agreement with the yield prediction of ANFIS. The extracted biodiesel was subjected to fourier transform infrared and gas chromatography mass spectrometer analysis for its confirmation.

Many researchers [4, 6, 10] have carried out studies on the optimization strategies of biodiesel yield by varying the input operating parameters using soft computing tools but advanced techniques like Kriging is not used. Kriging model of optimization is a common practise in areas such as mining, geo-statistical applications involving spatially and temporally correlated data, aerospace, structural engineering, *etc*. Few literatures based on the Kriging model is discussed here. Timothy *et al*. [12] used Kriging models for global approximation in solving aero spike nozzle design problem. Second order Response Surface models and Kriging models were developed for the weight, thrust and gross lift-off weight factors. The Kriging models outperformed the response surface models in predicting accurate results. Kaymaz [13] compared the classical RSM with Kriging based RSM for solving the structural reliability problems. The effects of the Kriging parameters based on β-computation and the fitting behaviour were investigated and found that Kriging method gave a better prediction in comparison with the classical RSM on the parameters that were adjusted appropriately. Jeong *et al*. [14] implemented the Kriging based GA to the aerodynamic design problems. Kriging model was developed between the objective function and the design variables to study the influence of each design variable and its interactions to the objective function by functional ANOVA.

The ANN model was used to optimize the process parameter to maximize biodiesel yield. Based on the literature review, it is observed that advanced techniques like Kriging model is not used till date to accurately optimize process parameters and maximize biodiesel yield. Therefore, the current work focuses on using the Kriging model and the outcomes were compared with the optimization results obtained using ANN model. The efficacies of the ANN and Kriging models on the prediction of biodiesel yield of the *C. inophyllum* biodiesel were verified with experimental outcomes. Experiments are conducted involving a two-stage transesterification process using sulphuric acid, KOH, and methanol. Statistical ANOVA data were used to estimate the efficiency of the developed ANN model and Kriging model using parameters such as coefficient of correlation, R, R^2, root mean square error (RSME), mean square error, standard error percentage, and mean relative percent deviation (MRPD) values. The inter-dependence among the input parameters such as, molar ratio, catalyst loading and reaction time at constant reaction temperature were also assessed in this study.

Materials and methods

Calophyllum inophyllum seed sources

The *C. inophyllum* is a mid-sized, slow growing evergreen tree having its nativity in the tropical Asia and it belongs to the kingdom of Plantae. Its family, order and genus classifications are *Calophyllaceae*, *Malpighiales,* and *Calophyllum*, respectively. Usually, it grows

up to 9 to 13 meters with perennial flowering. The fruit possessed by this tree is of 2.5-3.5 cm size is the source of bio-oil. The fruit yielding capability of the tree is throughout the year which makes it more favourable for its usage as biodiesel feed stock. The ripened seed of the *C. inophyllum* plant possess about 60% to 75% of bio-oil which is mainly composed of 68% unsaturated fatty acids and 32% of saturated fatty acid. The feedstock of *C. inophyllum*, its seeds and the fruits along with the extracted bio-oil are shown in fig. 1.

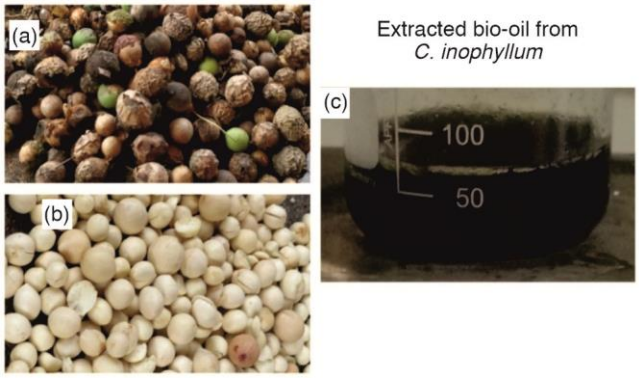

Figure 1. The *C. inophyllum* bio-oil; (a) feedstock, (b) seed kernels, and (c) bio-oil

Biodiesel feedstock and chemical reagents

The seeds of *C. inophyllum* were collected from the matured trees planted along the Puducherry-Tindivanam highway of Tamil Nadu, India. Around 13 kilograms of naturally ripened seeds are transported to the Antonin Lavoisier Laboratory, Hindustan Institute of Technology and Science, Padur, Chennai, Tamil Nadu, India. The outer shell of the seeds are detached carefully, and its inner fruits are isolated for its usage as the feed stock for biodiesel production. Analytical grade sulphuric acid is procured from the Alpha Chemika, Mumbai, India. The KOH and sodium sulphate are purchased from Girnar Chemical Industries, Vepery, Chennai, India. Laboratory grade methanol of around 96% and n-hexane are procured from the Hydrova Chemical and Systems, Urapakkam, Chennai, Tamil Nadu, India.

Calophyllum inophyllum – Bio-oil extraction

Bio-oil extraction from the *C. inophyllum* seeds was accomplished by a combination of cold pressing and Soxhlet extraction method. The obtained seeds are incubated at a temperature of 70 °C for 15 hours to remove moisture. Firstly, the seeds of ripened *C. inophyllum* fruit are subjected to hydraulic cold pressing through batch process at a rate of 1.5 kg per batch which produced around 4.6 L of bio-oil. The remains of the batch process are milled, crushed, and transferred to the next step involving Soxhlet apparatus. The milled *C. inophyllum* seeds are filled in the 5 L Soxhlet apparatus at a rate of 800 g per cycle. The bottom of the Soxhlet apparatus is filled with 3.2 L of n-hexane solvent which is circulated inside the solvent extracting system at a temperature of 62 °C thereby extracting the remaining bio-oil by this approach. By this combinational method, 6.75 L of bio-oil is extracted with an extracting efficiency of around 51.92%.

C. inophyllum bio-oil – Pre-treatment process (acid catalysed esterification)

Figure 2 shows the *C. inophyllum* – biodiesel preparation.

Base catalyzed transesterification of pre-treated C. inophyllum bio-oil

The extracted *C. inophyllum* bio-oil is found to possess an acid value of 3.2 by phenolphthalein titration method. Therefore, it is necessary to reduce the FFA content of the bio-

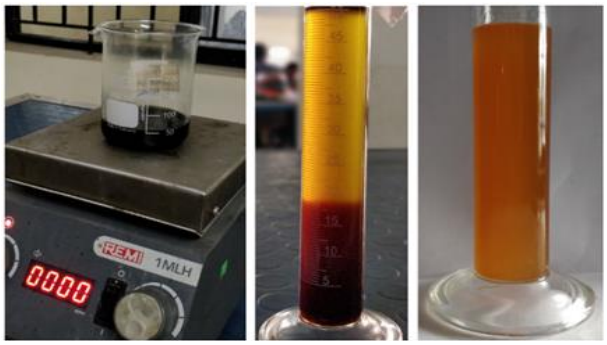

Figure 2. The *C. inophyllum* – Biodiesel preparation

oil as it could affect the transesterification reaction resulting in soapy sludge formation. For this purpose, acid catalysed pre-treatment is deployed to reduce the acid value. Appropriate quantity of *C. inophyllum* bio-oil and methanol is thoroughly mixed in the 1000 ml round necked Erlenmeyer flask. The flask containing the mixture is placed on a magnetic stirrer equipped with heating arrangement and it is maintained at 70 °C at 450 rpm agitation speed for 30 minutes. The 2% of H_2SO_4 is added to this mixture to catalyse the reaction at a faster rate and a reaction period of 45 minutes is allowed. Further, the resultant mixture is a two-layered content with FFA reduced oil as the top layer and denser fatty acid at the bottom layer which is formed due to its density variations. A separating funnel is used to isolate the pre-treated *C. inophyllum* bio-oil for further usage. Surplus methanol present in the pre-treated oil is removed by heating the oil up to 75 °C for 30 minutes. By this method, nearly 87% of pre-treated bio-oil is obtained. As explained earlier, methanol and KOH reagents are used in the base catalysed transesterification reaction. The acid esterified *C. inophyllum* bio-oil is taken in a 1000 ml flat bottomed Erlenmeyer flask. Pre-determined quantity (as per the experimental design) of KOH and laboratory grade methanol is thoroughly mixed at 55 °C to initiate the formation of potassium methoxide solution. A known quantity of potassium methoxide solution and pre-treated *C. inophyllum* bio-oil are thoroughly mixed in the Erlenmeyer flask and it is maintained at an encoded temperature with 600 rpm of agitation speed using the magnetic stirrer with extraction arrangement. After the specified reaction period, a ring formation appears separating the mono-alkyl ester and glycerol in the reaction chamber. The yield of *C. inophyllum* biodiesel is then transferred into the separating funnel and at a stagnation time of 180 minutes separated the biodiesel and glycerol at bottom layer due to its density differences. A needle valve arrangement removes the glycerol formation at the bottom of the separating funnel which also comprises excess methanol and catalyst reagents [14]. The obtained *C. inophyllum* is washed three times with double distilled water and finally with sodium sulphate to remove presence of moisture. The yield of *C. inophyllum* biodiesel is calculated using:

$$C.inophyllum \text{ biodiesel yield} = \frac{\text{Weight of } C.inophyllum \text{ biodiesel produced}}{\text{Weight of pre-treated } C.inophyllum \text{ used}} \times 100 \qquad (1)$$

Experimental model development

The RSM approach available in the Design Expert 12.0 software was utilized to develop the experimental model for this investigation. The major parameters which affect the *C. inophyllum* biodiesel yield are catalyst loading, methanol to oil molar ratio, reaction time, and reaction temperature and these are also mentioned by the earlier researchers like Hariram *et al*. [15] and Selvabala *et al*. [16] in their studies. In this optimization study, the reaction temperature was maintained constant as 70 °C, and all the other significant input parameters are varied.

A central rotatable composite design is developed with 27 experimental trials with varying degrees of catalyst loading, methanol to oil molar ratio, and reaction duration as shown in tab. 1. The input variable parameters are confined within the limits *i.e*., catalyst loading (0.87-1.12 wt.%), methanol to oil molar ratio (0.73-1.12), and reaction time (60-100 minutes) at a constant reaction temperature of 70 °C. The orthogonality of the experimental design is ensured by assessing the axial point and low, medium, and high levels of the input parameters. The ANOVA is deployed to evaluate the developed model by estimating the multiple regression approach to fit the quadratic regression model.

Table 1. Experimental design – Range of input parameters

Input parameters	Units	Decremental optimal	Optimal	Incremental optimal
KOH – catalyst loading	wt.%	0.87	0.98	1.12
Methanol to oil molar ratio	v/v	0.73	0.94	1.12
Reaction time	minute	60	80	100

Description of ANN and Kriging model

Sample data for fitting in ANN and Kriging models to estimate the bio-diesel yield is obtained from a 27 point orthogonal array. The sample data and corresponding response values are summarized in tab. 2. The data is scaled against the baseline design to protect the proprietary nature of the data. For two reasons, 27 point orthogonal array is chosen over the traditional central composite design as it mentions only 15 sample points for three design variables. Firstly, as each analysis needs a computational time of around 25-30 minutes, 25 iterations can be conveniently completed in a few minutes with the existing high performance computing facilities. Finally, it is observed that central composite designs frequently lead to singularities in the correlation matrix as shown in eq. (5) while performing the maximum likelihood estimation because of the location and spacing of the sample points in the design space.

The ANN model

The ANN is a soft computational tool which operate based on the information processing structure. It gathers data in terms of input, output, and hidden layers and inter-relate them as neurons. A multi-layered perceptron with forward feed-back propagation is used in the modelling of ANN. The non-linear regression and the need for longer cross validation is avoided by adoption of the Bayesian regularized algorithm thus formulating a simpler statistical regression. The topographical architecture used for the prediction of *C. inophyllum* biodiesel yield through ANN approach is shown in fig. 3. The input variable parameters are categorized as three input neurons *i.e*., catalyst loading, methanol to oil molar ratio, and reaction time, and *C. inophyllum* biodiesel yield is considered as the output neuron.

Kriging model

A surrogate model using Kriging is developed numerically to evaluate the prediction capability of the *C. inophyllum* biodiesel yield. Notations are used to develop the empirical relationship like methanol to oil molar ratio, m, catalyst loading, r, *C. inophyllum* biodiesel yield, Z, and constant mean of function, μ, In the present study, optimization of bio-diesel yield, Z, is performed by developing a surrogate using ordinary Kriging model using eqs. (2)-(5). The surrogate is generated based on the three variables, namely, molar ratio, m, catalyst concentration, c, and reaction duration, r. Ordinary Kriging is a spatial prediction method

which works on the assumptions specified by Cressie [17] and Wackernagel [18] and were derived using the equations 2 to 5.

Table 2. The CCRD design – Comparison of ANN and Kriging model with Experimental results

S.No	Molar ratio [v/v]	Catalyst loading [wt.%]	Reaction time [Minute]	C. inophyllum Bio-diesel yield (%)				
				Experiment [wt.%]	Kriging model [wt.%]	Error [%]	ANN predicted value [wt.%]	Error [%]
1	0.73	0.87	60	90.8	90.4	–0.44	91.2	0.44
2	0.73	0.87	80	91.6	91.88	0.30	91.6	–0.31
3	0.73	0.87	100	92.1	92.08	–0.02	92.2	0.02
4	0.73	0.98	60	91.7	91.47	–0.25	91.4	0.25
5	0.73	0.98	80	92.8	92.99	0.20	92.8	–0.20
6	0.73	0.98	100	93.4	93.27	–0.13	93.4	0.14
7	0.73	1.12	60	92.1	92.3	0.21	92.2	–0.66
8	0.73	1.12	80	93.8	93.7	–0.10	93.6	–0.58
9	0.73	1.12	100	95.1	94.71	–0.41	95.1	0.41
10	0.94	0.87	60	94.9	94.7	–0.21	93.9	0.21
11	0.94	0.87	80	95.7	95.6	–0.10	95.9	–0.82
12	0.94	0.87	100	97.1	96.98	–0.12	97.3	0.12
13	0.94	0.98	60	93.1	93.25	0.16	93.5	–1.84
14	0.94	0.98	80	98.1	98.1	0	97.8	0.41
15	0.94	0.98	100	97.1	97.18	0.08	97.2	–1.44
16	0.94	1.12	60	95.1	94.84	–0.27	95	1.31
17	0.94	1.12	80	97.1	96.79	–0.31	96.9	0.32
18	0.94	1.12	100	97.3	97.46	0.16	97.5	–0.16
19	1.12	0.87	60	95.2	95.25	0.05	95.5	–0.05
20	1.12	0.87	80	97.3	97.29	–0.01	97.4	0.01
21	1.12	0.87	100	98	98.1	0.10	98.1	–0.05
22	1.12	0.98	60	94.8	94.67	–0.13	94.7	0.14
23	1.12	0.98	80	96.9	96.78	–0.12	97	0.12
24	1.12	0.98	100	98	98.1	0.10	98	0.68
25	1.12	1.12	60	93.7	93.72	0.02	93.7	–0.02
26	1.12	1.12	80	94.8	94.9	0.10	95	–1.19
27	1.12	1.12	100	96.9	96.88	–0.02	97.2	0.02

The global, constant mean $\mu \in \mathbb{R}$ of the random function $Z(x)$ is unknown. The data comes from an intrinsically stationary random function $Z(x)$ with known variogram function $\gamma(\mathrm{h})$.

$$\gamma(h) = \frac{1}{2} Var\left[Z(x+h) - Z(h)\right] = \frac{1}{2}\mathbb{E}\left\{\left[Z(x+h) - Z(h)\right]^2\right\} \tag{2}$$

The ordinary Kriging predictor is:

$$Z_{\mathrm{w}}^{*}(x_0) = \sum_{i=1}^{n} w_i Z(x_i) = w^T Z \tag{3}$$

where $w=(w_1,\ldots w_n)^T \in \mathbb{R}$ provides the unknown weights corresponding with the influence of the variable $Z(x_i)$ in the computation of $Z_w^*(x_0)$. To ensure consistency of the linear predictor $Z_w^*(x_0)$, the sum of the weights is set equal to 1:

$$\sum_{i=1}^{n} w_i = 1 \tag{4}$$

Therefore, the expected error vanishes:

$$\mathbb{E}\left[Z_w^*(x_0)-Z(x_0)\right]=\mathbb{E}\left[\sum_{i=1}^{n} w_i Z(x_i)-Z(x_0)\sum_{i=1}^{n} w_i\right]=\sum_{i=1}^{n} w_i \mathbb{E}\left[Z(x_i)-Z(x_0)\right]=0$$

The error variance σ_E^2 of $Z_w^*(x_0)$ by substituting variogram function $\Upsilon(h)$ of $Z(x)$ is:

$$\sigma_E^2 = Var\left[Z_w^*(x_0)-Z(x_0)\right]=\mathbb{E}\left\{\left[Z_w^*(x_0)-Z(x_0)\right]^2\right\}=w^T\left(2\gamma_0-\Gamma w\right)\geq 0 \tag{5}$$

where γ_0 is semivariances and is equal to:

$$\gamma_0=\left[\gamma(x_1-x_0),\ldots,\gamma(x_n-x_0)\right]^T \in \mathbb{R}^n$$

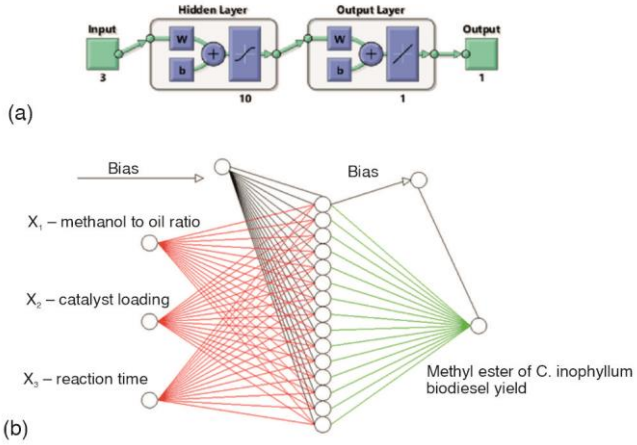

Figure 3. The ANN – topological architecture to predict *C. inophyllum* biodiesel yield

The predicting capability of ANN and Kriging models on assessing the *C. inophyllum* biodiesel yield is compared by varying the input variable parameters such as, catalyst loading between 0.87-1.12 wt.%, methanol to oil molar ratio between 18:1 and 22:1, and reaction duration between 60 and 100 minutes as tabulated in tab. 2. The interdependence between the input variable parameters on the yield of *C. inophyllum* biodiesel is also studied. The generation of child population were based in eqs. (6)-(12) and the random number, α, was between 0 and 1. The crossover and mutation was identified in the kriging model development.

If $\alpha \leq \bar{\alpha}$ then generate two new configurations using crossover:

$$\delta_i=(2\gamma_i)^{\frac{1}{\eta+1}}\ (\text{if } \gamma_i<0.5),\ \delta_i=1/\left[2(1-\gamma_i)\right]^{\frac{1}{\eta+1}}\ (\text{if } \gamma_i\geq 0.5) \tag{6}$$

$$l = l + 1, \quad x_i^{(l)} = \frac{1}{2}\left\{\left(1-\delta_i\right)x_i^{(q_1)} + \left(1+\delta_i\right)x_i^{(q_2)}\right\} \text{ exit step 4 if } l = p \tag{7}$$

$$l = l + 1, \quad x_i^{(l)} = \frac{1}{2}\left\{\left(1+\delta_i\right)x_i^{(q_1)} + \left(1-\delta_i\right)x_i^{(q_2)}\right\} \text{ exit step 4 if } l = p \tag{8}$$

where γ_i (i = 1,2,3) random numbers generated each time. Clearly, if there are three variables, *i.e.*, N = 6, one would need 6 values of γ_i for these two configurations. This step is repeated until the number of configurations generated is equal to the size p.

The mutation of the input variable parameters was identified based on eqs. (9) and (10). If $\alpha \le \bar{\alpha}$ then generate a new configuration using:

$$\delta_i = \left(2\gamma_i\right)^{\frac{1}{\eta+1}} - 1 \left(\text{if } \gamma_i < 0.5\right), \ \delta_i = 1 - \left[2\left(1-\gamma_i\right)\right]^{\frac{1}{\eta+1}} \left(\text{if } \gamma_i \ge 0.5\right) \tag{9}$$

$$l = l + 1, \quad x_i^{(l)} = x_i^{(q_1)} + \delta\left(x_{Ui} - x_{Li}\right) \text{ exit step 4 if } l = p \tag{10}$$

where $\bar{\alpha} = 0.8$ and $\eta = 20$ is used in this study. At the end of step 4, p new configurations is called child population.

The final step involved in computing the biodiesel yield $Z^{(j)}$ of each of these p new configurations and construct a set of 2p configurations by appending the new set in Step 4 to the set in Step 1. Reorder these 2p configurations as done in Step 2 and consider the first p configurations and remove the rest. With this new p configurations go to Step 3 and proceed for next iteration till the biodiesel yield of the first configuration does not decrease further with an assumed tolerance. This will then be the best configuration. The GA parameters used in this study are given in tab. 3.

Table 3. The GA parameters used in the present work

Number of variables	3
Population size	100
Crossover rate	0.8
Mutation rate	0.2
Number of generations	100
Number of evaluations	10000

Results and discussion

Efficacy of the developed models

The predictive model is developed using the empirical and regression relations of the ANN and Kriging approach to evaluate the biodiesel yield of *C. inophyllum* comprehensively. The ANOVA is used to understand the efficacy of the developed model through its statistical indicators. Both the developed model are compared, and the efficacy is studied using:

$$R = \frac{\sum_{i=1}^{N}\left(x_{pi} - x_{p,\,\text{ave}}\right)\left(x_{ai} - x_{a,\,\text{ave}}\right)}{\sqrt{\left[\sum_{i=1}^{w}\left(x_{pi} - x_{p,\,\text{ave}}\right)^2\right]\left[\sum_{i=1}^{n}\left(x_{ai} - x_{a,\,\text{ave}}\right)^2\right]}} \tag{11}$$

$$MSE = \frac{1}{n}\sum_{i=n}^{n}\left(x_{pi} - x_{ai}\right)^2 \tag{12}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=n}^{n}\left(x_{pi} - x_{ai}\right)^2} \tag{13}$$

$$MRPD = \frac{1}{n}\left[\sum_{i=1}^{n}\left(\frac{x_{ai} - x_{pi}}{x_{ai}}\right)\right] \times 100 \tag{14}$$

where x_{pi} and x_{ai} are the experimental and predicted data, n – the number of experimental runs, $x_{p,\text{ave}}$ and $x_{a,\text{ave}}$ are the average values of experimental and predicted outputs.

The ANN model

The ANN are the computer enabled codes which operates based on the artificial intelligence. The segregation of input data such as training, validation and testing increases the efficacy of the ANN model. The ANN establishes the inter-dependence between the dependent and independent variable operating parameters. Considering their non-linearity, 10 hidden neurons are deployed in this ANN model to predict the *C. inophyllum* biodiesel yield. As mentioned earlier, 70% of the available experimental data is used to train the ANN-GA, and the remaining 30% is used in the validation and testing of the output functions. Topographical layers are used to arrive at an optimum result of the hidden neurons. The efficacy of the ANN model is evidenced through ANOVA approach in which the R value is determined as 0.97279. The R^2 and adjusted R^2 are observed as 0.98744 and 0.99645, respectively. The prediction capability and precision accuracy of the developed ANN model is superior with its lower MRPD value as 0.22457%. The confidence level of the ANOVA is found to be 94.5% which supports the importance of ANN model in the estimation of *C. inophyllum* biodiesel yield. The detailed regression analysis on ANN prediction is shown in fig. 4.

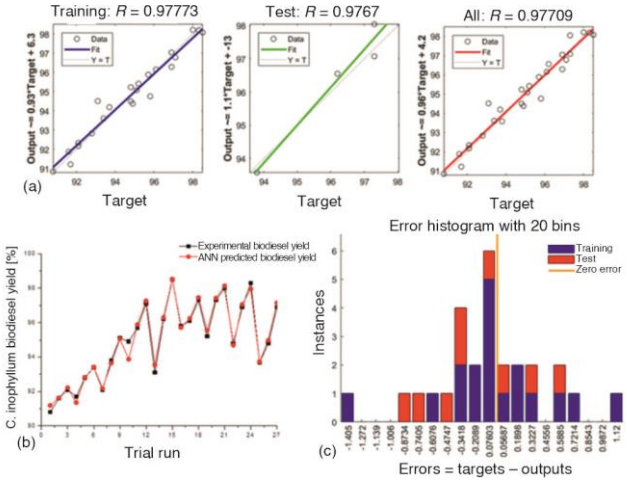

Figure 4. Regression analysis on ANN prediction

Mathematically, a neuron can be expressed as shown in eq. (15) and the output/input layer's transfer function as in eqs. (16) and (17):

$$x_i = \sum_{i=1}^{m} y_1 w_{ji}^{in} + \beta_j^{in} \tag{15}$$

$$f(x) = \text{tans}ig(x) \tag{16}$$

$$\text{tans}ig(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{17}$$

where w_{ji}^{in} is the linking factor between β_j^{in} and j^{th} neuron, x_i is the node input and y_i is the input neuron.

Kriging model

As Kriging models interpolate the sample data, 27 additional, randomly selected validation points are used to verify the accuracy of the Kriging models. Moreover, the accuracy of the response surface models is also examined using these 27 points. Error is defined as the difference between the actual response from the computer analysis, *y*, and the predicted value, $\hat{y}$ from either the ANN model or the Kriging model. The maximum absolute percent error, the average absolute percent error, and the RMSE for these 27 validation points are summarized in tab. 4. As listed in the table, it is evident that Kriging models has a lower maximum absolute error, lower average absolute error and lower RMSE values for bio-diesel yield than the ANN model. Finally, it appears that both models predict well, whereas Kriging model offers a better accuracy because of the lower RMSE values.

Table 4. Error analysis of ANN and Kriging models

	Bio-diesel yield (Z), Kriging model	Bio-diesel yield (Z), ANN model
Max |% error|	0.44	1.05
Average |% error|	0.02	0.19
RMSE	0.19	0.28

The effect of variations of catalyst concentration, reaction time, and molar ratio on the response biodiesel yield are plotted in Figures 5, 6 and 7. As can be seen in these Figures, initially the response biodiesel yield increases and attains the peak and then, it decreases. This process of increase, reaching a peak and then decreasing is evident as the catalyst concentration increases. It is also clear that, the peak value gradually reduces with increase in catalyst concentration beyond certain value (c=1). This happens as the catalyst concentration up to (c=1) helps in faster reactions of triglycerides and avoid the formation of sludge. If the concentration of KOH increases beyond (c=1), then the process of transesterification deteriorates and prevents the formation of fatty acid methyl esters resulting in reduced biodiesel yield. To study the outcome of independent process variables on dependent response biodiesel yield, three different 3D plots are drawn by selecting two process independent parameters at its minimum and maximum values along with the yield while keeping the third parameter constant.

Effect of variable operating parameters in Kriging model

Effect of molar ratio and reaction duration at
various catalyst concentration on biodiesel yield

The percentage of biodiesel yield of the mix increases with the increase in the amount of catalyst as stated by Hariram *et al*. [19]. Furthermore, the percentage of yield changes its trend as the amount of catalyst reaches a certain limit which is 1% in this study, fig. 5. This may be due to the reason that addition of too much catalyst results in faster reaction of triglycerides which resulted in increased biodiesel yield. Initially, blend yield increases

due to the increase in the conversion rate of fatty acid with increase in reaction time and after reaching maximum value, it decreases. The results observed are in line with earlier finding by Narula *et al*. 2017 [20].

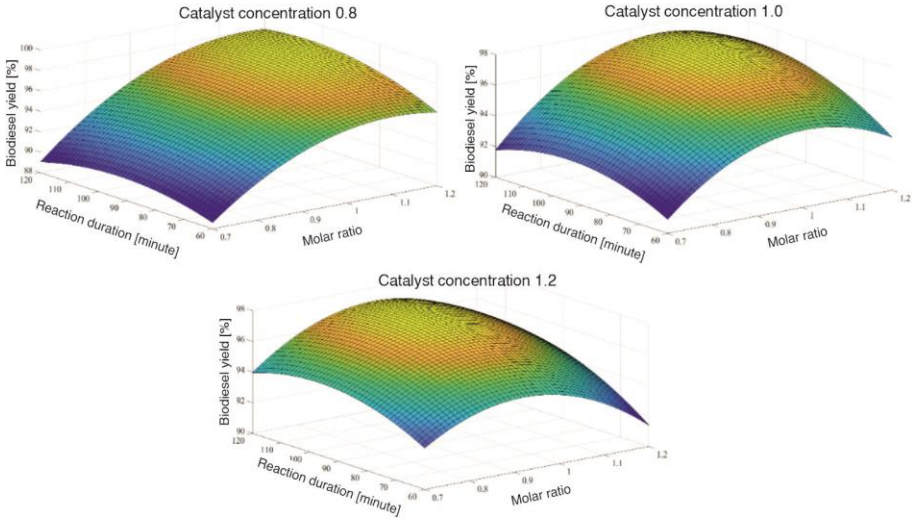

Figure 5. Influence of molar ratio and reaction duration at various catalyst loading

Effect of reaction duration and catalyst concentration at various molar ratio on biodiesel yield

The percentage of molar ratio also has a proportional effect on the response yield of the blend as a percentage of yield decreases with the rising quantity of methanol to oil ratio within the mixture. The observed results are also reported by Narula *et al*. [20] and Verma *et al*. [21].

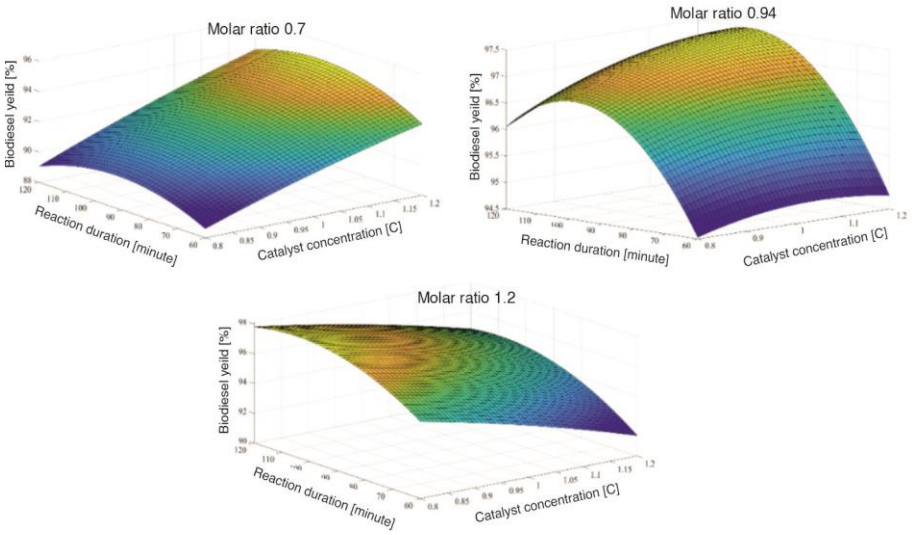

Figure 6. Influence of reaction duration and catalyst loading at various molar ratio

Effect of catalyst loading and molar ratio at
various reaction duration on biodiesel yield

Reaction duration has a proportional relation to the response biodiesel yield of the mix. The percentage of response yield initially rises due to the accessibility of adequate time to finish the process, and afterwards, the percentage of response decreases with duration. These findings are also supported by Hariram *et al*. [22] and Vasudev *et al*. [23].

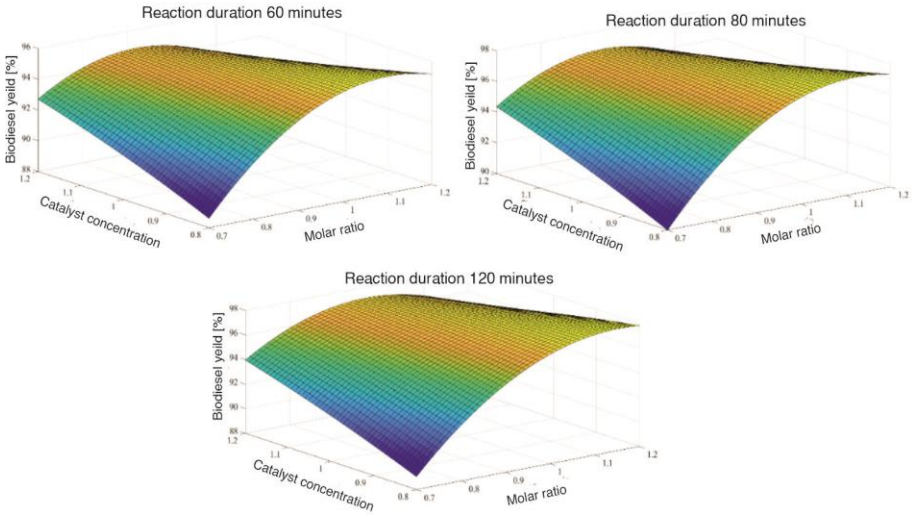

Figure 7. Influence of catalyst loading and molar ratio at various reaction duration

Optimization problem results

The statistical outcomes of both Kriging and ANN model is very close to each other and is in good agreement with the earlier study of Zhao *et al*. [24]. Also, the experimental results of the *C. inophyllum* biodiesel yield is nearer to the predicted outcomes of ANN and Kriging as shown in tab. 2. With respect to ANN, the co-efficient of correlation, *R*, value with respect to training, testing and combinational data are 0.97773, 0.9764, and 0.97709, respectively, which implicated the accuracy and interdependency of the developed model. The GA disclosed a significant variation of the fitness value up to 14^{th} generation of iterations which is inconsequential thereafter due to the minimal mutation and negligible cross-over. The continued iteration up to 25^{th} order generation confirmed the non-participatory effect which later resulted in the optimized *C. inophyllum* biodiesel yield. The synergetic effect of the input process parameters is obtained using the developed Kriging model with its surface plots. Favourable limits for the optimized *C. inophyllum* biodiesel yield is derived using methanol to oil molar ratio between 0.73 and 1.12 v/v, KOH catalyst loading between 0.87 wt.% and 1.12 wt.%, and reaction duration between 60 minutes and 100 minutes. The biodiesel yield significantly increased at improved concentrations of methanol higher than the pre-determined levels.

The optimal combination of input parameters in comparison with ANN and Kriging model is found to be 0.94 v/v of methanol to oil molar ratio, 0.98 wt.% of KOH catalyst loading and 80 minutes reaction duration which yielded 98.1% of *C. inophyllum* biodiesel. The outcomes of the transesterification reaction is in good agreement with the earlier findings of Hariram *et al*. [15], and Selvabala *et al*. [16]. The effectiveness of Kriging and ANN models

over other optimization tools are reported by many researchers. Finally, it can be evidenced that Kriging and ANN tools can be successfully used to assess and predict the transesterification efficiency of the *C. inophyllum* bio-oil biodiesel yield. A comparison of ANN and Kriging predictive models with experimental biodiesel yield is graphically represented in fig. 9.

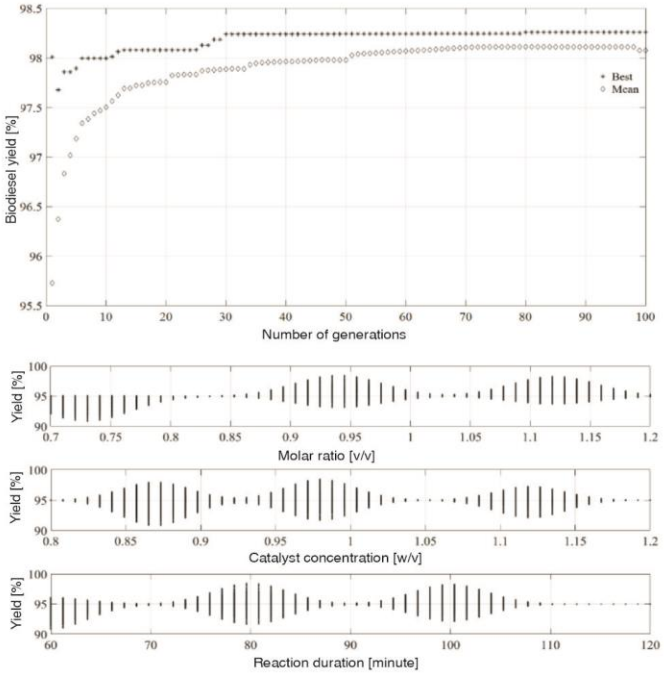

Figure 8. Kriging model prediction of *C. inophyllum* biodiesel yield – a comparison

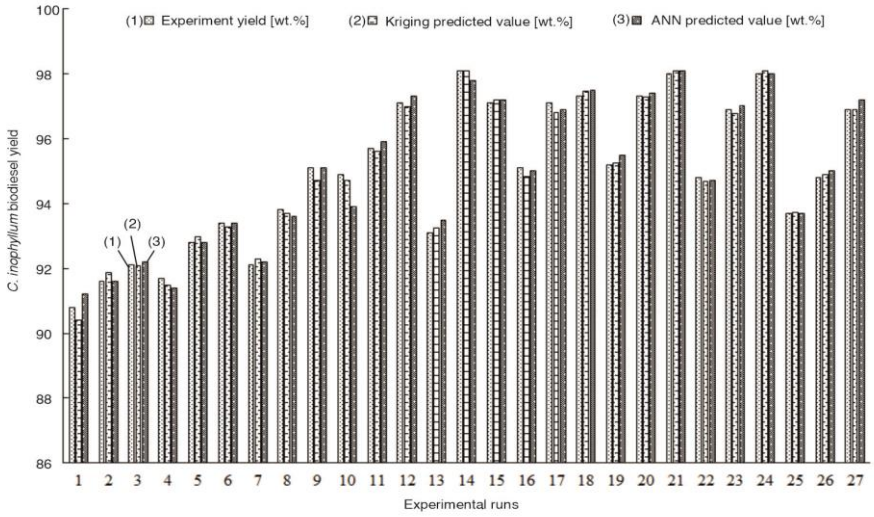

Figure 9. Comparison of ANN and Kriging model prediction with experimental results

Conclusion

The present work mainly focused on the comparison between ANN model coupled with Bayesian regularized algorithm and Kriging model developed to study the transesterification process of *C.inophyllum*. The GA has been used to solve the optimization problem with given bounds on the process variables. Both ANN model and Kriging model were used to find the optimum biodiesel yield. The prediction error analysis has been carried out between the ANN and Kriging model with experimental data. The maximum error of 0.44% in Kriging model and 1.05% in ANN model are noticed. Similarly, the average error in kriging model is 0.02% and in ANN model is 0.19% and finally, the RSME in kriging is 0.19% and in ANN model 0.28%. The optimization results using both ANN and Kriging models converged to the same values of process variables with only difference in the calculation of biodiesel yield. The values of the optimized process variables, namely, molar ratio is 0.94, catalyst concentration is 0.98 and reaction duration is 80 minutes using both models. The biodiesel yield is 98.1% using Kriging model with no difference when compared with experimental data, whereas it differs by 0.41% when ANN model has been used. Therefore, it can be concluded that Kriging model can be adopted to predict the yield of biodiesel through transesterification process.

References

[1] Demirbas, A., Importance of Biodiesel as Transportation Fuel, *Energy Policy, 35* (2007), 9, pp. 4661-4670
[2] Venkatesan, H., *et al*., Cotton Seed Biodiesel as Alternative Fuel: Production and Its Characterization Analysis Using Spectroscopic Studies, *International Journal of Renewable Energy Research, 7* (2017), 3, pp. 1333-1339
[3] Cakirca, E. E., *et al*., Catalytic Activity of CaO-Based Catalyst in Transesterification of Microalgae Oil with Methanol, *Energy Environ, 30* (2018), 1, pp. 176-187
[4] Betiku, E., Ajala, S. O, Modelling 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 and Products, 53* (2014), Feb., pp. 314-1-322
[5] Arumugam, A., Ponnusami, V., Biodiesel Production from Calophyllum Inophyllum oil a Potential Non-Edible Feedstock: An Overview, *Renewable Energy, 131* (2019), Feb., pp. 459-471
[6] Hariram, V., *et al*., Performance Assessment of Artificial Neural Network on the Prediction of Calophyllum Inophyllum Biodiesel Through Two-Stage Transesterification, *Energy Sources, Part A: Recovery, Utilization and Environmental Effects, 43* (2021), 9, pp. 1060-1072
[7] Venkanna, B. K., Venkataramana Reddy, C., Biodiesel Production and Optimization from Calophyllum Inophyllum Linn Oil (Honne Oil) – A Three Stage Method, *Bio Resource Technology, 100* (2009), 21, pp. 122-125
[8] Ayoola, F. K., *et al*., Response Surface Methodology and Artificial Neural Network Analysis of Crude Palm Kernel Oil Biodiesel Production, *Chemical Data Collection, 28* (2020), 100478
[9] Garg, A., Siddharth, J., Process Parameter Optimization of Biodiesel Production from Algal Oil by Response Surface Methodology and Artificial Neural Networks, *Fuel, 277* (2020), 118254
[10] Betiku, E., *et al*., Performance Evaluation of Artificial Neural Network Coupled with Generic Algorithm and Response Surface Methodology in Modelling and Optimization of Biodiesel Production Process Parameters from Shea Tree (Vitellaria Paradoxa) Nut Butter, *Rene. Energy, 76* (2015), C, pp. 408-417
[11] Benjamin, O. O., *et al*., Optimization of Biodiesel Production from Thevetia Peruviana Seed Oil by Adaptive Neuro-Fuzzy Interference System Coupled with Genetic Algorithm and Response Surface Methodology, *Energy Conversion and Management, 132* (2017), Jan.,pp. 231-240
[12] Timothy, W., *et al*., Kriging Models for Global Approximation in Simulation-Based Multidisciplinary Design Optimization, *AIAA Journal, 39* (2001), 12, pp. 2233-2241
[13] Kaymaz, I., Application of Kriging Method to Structural Reliability Problems, *Structural safety, 27* (2005), 2, pp. 33-151

[14] Jeong, S., *et al*., Efficient Optimization Design Method Using Kriging Model, *Journal of Aircraft, 42* (2005), 2, pp. 413-420
[15] Hariram, V., Vasanthaseelan, S., Optimization of Base Catalyzed Transesterification and Characterization of Brassica Napus (Canola Seed) for the Production of Biodiesel, *International Journal of ChemTech Research, 8* (2015), 9, pp. 418-423
[16] Selvabala, V. S., *et al*., Two-Step Biodiesel Production from Calophyllum Inophyllum Oil: Optimization of Modified B-Zeolite Catalyzed Pre-Treatment, *Bio Resource Technology, 102* (2011), 2, pp. 1066-1072
[17] Cressie, N., *Statistics for Spatial Data*, Wiley, New York, USA. 1993
[18] Wackernagel, H., Geo-Statistical Models and Kriging, *IFAC Proceedings, 36* (2003), 16, pp. 543-548
[19] Hariram, V *et al*., Spectrometric Analysis of Algal Biodiesel as a Fuel Derived through Base-Catalysed Transesterification, *International Journal of Ambient Energy, 40* (2019), 2, pp. 195-202
[20] Narula, V., *et al*., Process Parameter Optimization of Low Temperature Transesterification of Algae-Jatropha Curcas Oil Blend, *Energy, 119* (2017), Jan., pp. 983-988
[21] Verma, P., Sharma, M, P., Review of Process Parameters for Biodiesel Production from Different Feedstocks, *Renew Sustain Energy Rev, 62* (2016), Sept., pp. 1063-1071
[22] Hariram, V., *et al*., Reduction of Exhaust Emission Using a Nano-Metallic Enriched Lemongrass Biodiesel Blend, *Energy Sources, Part A: Recovery, Utilization and Environmental Effects, 39* (2017), 21, pp. 2065-2071
[23] Vasudev, K. L., *et al*., A Modular and Integrated Optimization Model for UVs, *Defence Sci. J, 66* (2016), 1, pp. 71-80
[24] Zhao, X., *et al*., Optimization of Transesterification of Beef Tallow for Biodiesel Production Catalyzed by Solid Catalysts, *Nongye Gongcheng Xuebao/Transactions Chinese Soc. Agric Eng; 29* (2013), 17, pp. 196-203

Paper submitted: November 27, 2021
Paper revised: January 25, 2022
Paper accepted: February 1, 2022