Hybrid nero-fuzzy methods for estimation of ultrasound and mechanically stirring influences on biodiesel synthesis through transesterification

Baharak Sajjadi a, Perumal Asaithambi a, Abdul Aziz Abdul Raman a,⇑, Shaliza Ibrahim b

⇑Corresponding author.
E-mail address: azizraman@um.edu.my (A.A.A. Raman).

Article info
Article history:
Received 3 August 2016
Received in revised form 20 January 2017
Accepted 26 January 2017
Available online 8 February 2017

Keywords:
Biodiesel
Transesterification
Ultrasound
Neuro-fuzzy inference system
ANFIS-PSO
ANTIS-GA
ANTIS-DE

Abstract
Accurate estimation of transesterification efficiency is needed for designing operational conditions and obtaining the maximum yield of biodiesel production. Accordingly, the objective of this study is to analyse the applicability of three different hybrid soft computing techniques for the prediction of transesterification yield under ultrasound irradiation as a novel and under mechanical stirring as a traditional method of biodiesel synthesis. The models include ANFIS-PSO (Adaptive Neuro-Fuzzy Inference System linked with Particle Swarm Optimization), ANFIS-GA (Adaptive Neuro-Fuzzy Inference System linked with Genetic Algorithm) and ANFIS-DE (Adaptive Neuro-Fuzzy Inference System linked with Differential Evolution). Independent variables including reaction temperature, reaction time, reactant concentrations, catalyst loading and power input were considered as the network inputs while the reaction yield was considered as the network output. The obtained simulation results were then analysed using Kolmogrov-Smirnov method as well as root mean-square error (RMSE) and coefficient of determination (R^2). The analyses confirmed the validity of the proposed models. It was found that although ANFIS-PSO had better performance in training phase, it generated the weakest results in testing phase. Meanwhile, ANFIS-DE provided the best statistical characteristics compared to the other methods for estimating the transesterification yield under either ultrasound irradiation or mechanically stirring.

1. Introduction
Currently, the world’s energy demand is considerably (about 80%) met through non-renewable resources. However, considerations over environmental problems and depletion of fuel resources caused by widespread use of fossil fuels has resulted in significant depletion of fossil fuel resources and considerable impact on economies of many countries [1]. Therefore, there should be an economically competitive and renewable energy source with low environmental impact.

1.1. Biodiesel synthesis difficulties
Among different renewable energy sources, biodiesel has been recognized as a feasible and potential alternative to petroleum based fuel. Fatty Acid Methyl Ester (FAME), also known as biodiesel, is one of the best substitutes for diesel fuel. Chemically, biodiesel is a mixture of alkyl esters with long-chain fatty acids. It is traditionally produced by transesterification of nontoxic and biological resources such as edible or non-edible oils [2]. Transesterification is a three-step reaction in which triglyceride reacts with an alcohol to form esters as the main products and glycerol as the by-product. Di-Glycerides (DG) and Mono-Glycerides (MG) are also the other by-products (Eq. (1)).

\[
\begin{align*}
\text{CH}_2\text{-COOR}_1 & \quad \text{R}^1\text{COOR'} & \quad \text{CH}_2\text{-OH}^* \\
\text{CH}_2\text{-COOR}_2 & \quad 3\text{R'}\text{OH} & \quad \text{R}^2\text{COOR'} + \quad \text{CH}_2\text{-OH}^* \\
\text{CH}_2\text{-COOR}_3 & \quad \text{R'}\text{COOR'} & \quad \text{CH}_2\text{-OH}^* \\
\text{Triglyceride(oil)} & \quad \text{Alcohol} & \quad \text{Alkyl ester} & \quad \text{Glycerol}
\end{align*}
\]

Immiscible nature of the reactant (oil and alcohol) and poor mass transfer between them are the key drawbacks associated with transesterification to produce biodiesel [3] resulting in low conversion rates, long reaction time, significant costs of feedstock and

http://dx.doi.org/10.1016/j.measurement.2017.01.044
0263-2241/© 2017 Elsevier Ltd. All rights reserved.
downstream separations. Accordingly, the cost of biodiesel production is almost twice the cost of producing mineral diesel. Improving the biodiesel producing reaction through acceleration of mass transfer among the reactants can be a solution to eliminate a number of the mentioned challenges above [4]. Traditionally, biodiesel is synthesized by mechanical mixing in a batch stirred vessel. Mixing in such equipment is done by a shaft and impeller which is highly time, energy and cost consuming. This type of mixing is in macro level and not efficient [5]. Employing new techniques that generate fine microemulsion in the system, including transesterification reaction, is a key to overcome such limitations. Among the available techniques, ultrasound irradiation based method can be considered as an effective and energy efficient method to intensify transesterification. Mixing phenomenon under ultrasound energy is in micro level. This method has been extensively used to enhance the conversion of the product and reaction yield compared to traditional method which uses mechanically stirring [6].

1.2. Advantages of biodiesel production under ultrasound energy

Ultrasound waves are sinusoidal mechanistic waves that transfer energy. They are categorised into three major categories considering their frequencies: 20–100 kHz, 100 kHz to 2 MHz and above 2.5 MHz [7]. When Ultrasound waves are applied in a liquid, acoustic waves comprise a cycle of expansion (rarefaction) and compression phases. The expansion and contraction phenomena of the transfer medium creates empty regions which are named cavitation bubbles [8]. Ultrasound waves with the frequency within 20–100 kHz are in conventional range. Cavitation bubbles are significantly generated in this category. In the second range, bubbles generation becomes dramatically difficult and at frequency above 2.5 MHz, cavitation does not happen [7]. In systems containing chemical reactions cavitation bubbles act as micro reactors with small volumes. In other words, micro fine cavitation bubbles filled with the reactants vapour resulting in sufficient mass transfer within these bubbles. The cavitation bubbles grow and collapse subsequently [9]. The collapse of acoustic cavities interrupts the phase boundary and generates micro jet streams and shock waves that impinge one liquid into another, assisting the system in mixing of immiscible reactants. As a result, a fine micro emulsion is generated and the mass transfer rate increases subsequently. Therefore, cavitation assists the process physically by turbulence and chemically through generation of free ions. Acoustic streaming is the other effect of ultrasound waves which is defined as any flow generated by the force arising from the dissipation of acoustic energy flux and Reynolds stress inducing gradients in momentum flux [10]. Cavitation and acoustic streaming as the main effects of low frequency ultrasound (20–100 kHz) can improve multiphase transfer processes, increase movement and mixing of fluid and prevent of samples clogging. Accordingly, low frequency ultrasound is mostly applied on systems containing chemical reactions such as transesterification. Ultrasound energy provides the activation energy for initiating the transesterification and the mechanical energy for mixing and processing. Compared to biodiesel synthesis using mechanical stirring, ultrasound energy provides the advantages of shorter reaction duration, less energy and alcohol consumption compared to conventional mechanical stirring [11].

1.3. New applicable methods for analysis of biodiesel synthesis

In order to analyse the capability of this synthesis method in industries and compare it with other methods, estimation of reaction yield in different conditions is of importance. Response Surface Methodology (RSM) is an alternative which has been widely used in statistically analysing the relationships between several explanatory variables and the response variables in different systems including transesterification. RSM models the output(s) of a system according to its input variables obtained from predesigned experiments. This method is based on the fit of a mathematical model to the results of a pre-designed set of experiments. RSM is a subsection of a fundamental tool entitled DOE (designing the experiments). The basic concept of DOE is to consider the effect of all significant parameters simultaneously and then analyse their importance through a mathematical model. RSM then develops regression models, evaluating the individual and interactive effects of different variables and optimizing the process. The optimization through DOE assisted by RSM method follows six consequent steps [12] including (i) identification of independent variables, (ii) identification of experimental design strategy, (iii) implementation of experiments and observing results, (iv) fitting the model equation to obtained results, (v) obtaining response graphs and verification of the model through Analysis of Variances (ANOVA), (vi) determination of optimal conditions. Most of the published works in transesterification area have been analysed by RSM method [13,14]. However, RSM can only analyse the interactive or individual effects of operating parameters through descriptive models and cannot predict what will happen if the operating conditions are changed. In industrial scale, it is of importance to predict the result (Herein reaction yield) and to consider the necessary changes to keep the reaction yield at its highest value. In recent decade, soft computing methods have showed great ability for estimating and forecasting the final results of different processes [15]. The routine procedure to employ the soft computing methods is to train the planned network using the collected input and output data [16]. These powerful and robust methods can bring accurate and satisfying results in different area. Fuzzy systems are powerful computing techniques due to the concepts of fuzzy set theory, fuzzy reasoning and fuzzy if-then rules. A robust neuro-fuzzy network, which has a potential to obtain the benefits of both neural network and fuzzy system in a single framework is obtained by integrating fuzzy systems with neural networks. Adaptive Neuro Fuzzy Inference System (ANFIS) is a popular framework of the neuro-fuzzy systems that uses the Takagi-Sugeno inference system [17]. Although ANFIS has high learning and prediction capabilities, its conventional training algorithm is gradient descent, which is difficult to be calculated in each step and the use of chain rules may cause a local minimum [18]. Furthermore, the number of the generated fuzzy rules increases exponentially in modelling problems with a large number of inputs [19]. Therefore, ANFIS is not suitable in such cases. ANFIS training parameters can be improved by integrating ANFIS with another potential algorithm to overcome these limitations and for the purpose of rapid convergence. In industrial scale, the operating conditions should be analysed fast enough and its effect on reaction should be predicted simultaneously. Despite decades of manufacturing, there is no well-established online method for continuous monitoring and control of chemical reactions and their yields. The difficulty of developing such methods arises from the complexity of chemical reactions, complexity of yield analysis and the applicability and accuracy of the methods/models. However, recently, a technology entitled ACOMP/CI (Automatic Continuous Online Monitoring of Polymerization reactions combined with a Control Interface) based on the idea of on-line modelling a chemical reaction, has been invented and developed at Tulane (US patent 6052184 and US patent 6653150). ACOMP continuously and automatically analyse and control all relevant characteristics of polymerization reactions. The benefits of such a fully automate the manufacturing of chemical processes can be summarised as: (i) yield increase, (ii) reduced energy use, (iii) improved product quality, (vi) enhanced sustainability, (v) increased worker safety. It also in beneficial in terms of (vi) energy impact: reduced electricity usage for cooling/heating, equipment utilization and more efficient use of initial feedstock. Accordingly, it is of importance to develop technologies based on
on-line analysis and controlling the chemical reactions i.e. transesterification for biodiesel production. In order to make biodiesel production process efficient, all effects should be analysed online and the corresponding and suitable change should be applied to the system without any delay. Accordingly, the objective of this manuscript is to analyse the applicability of three different networks for predicting the transesterification yield under different operating conditions and compare their prediction accuracy. Toward this objective, the capability of three ANFIS-based hybrid soft computing techniques (ANFIS linked with PSO, ANFIS linked with GA and ANFIS linked with DE) for estimation of transesterification yield in biodiesel synthesis was examined. The key operating parameters including temperature, alcohol concentration, catalyst loading, power input and reaction time were considered as the network inputs while the reaction yield was considered as the network output. In experimental part, a low frequency/high power ultrasonicator as a modern technique and mechanical stirring along with conventional heating as the traditional technique were employed for biodiesel synthesis. It should be noted that, the individual effect of each operating factor as well as their interactive effects on transesterification yield were analysed in details in the author’s previous work using sensitivity analysis. The detailed analysis, comparison of these two methods and optimization results have been discussed in the author’s previous work [20].

2. Methodology

2.1. Materials

Bleached, deodorized and refined palm oil was used as a triglyceride source. Its iodine (51.8 ppm), acid (0.151 mg/g) and moisture content (400 x 10⁻⁶ ppm) were determined through automatic titrations (Metrohm; Titrando models 737, 808 Karl Fischer coulometer). Pure KOH, catalyst, methanol (99.99%) were also used as the catalyst, second reactant and the biodiesel standards (GLC-10 Fame Mix (100 mg)).

2.2. Design of Experiments (DOE)

In current study, CCD (Central Composite Design) combined with RSM was employed to design and arrange the experiment. CCD is a first-order design augmented by additional center and axial points to allow estimation of the tuning parameters of a second-order model. In short, one central point, 10 (=2N) axial points and 32 (=2N) factorial points were applied to combine 5 (N) independent variables as the model inputs. Five operating parameters (N) including reaction temperature (medium temperature, °C), reactant concentrations (methanol to oil ratio), catalyst concentrations (weight percent to oil), reaction time and power input (ultrasound power for mixing reactants, Watt) were identified and used as the independent variables for biodiesel synthesis. Two sets of experiments were prepared for biodiesel synthesis using sonorization and mechanical stirring, respectively. A total of 50 experiments (2 x N + 2N + 8 (central point)), total 100, were performed for each category (ultrasound and mechanically mixing). Note that, 8 replications were at the center point with temperature of 57°C; MeOH:Oil of 9:1; catalyst concentration of 1.5 wt.%, ultrasound power of 350 W; mixing intensity of 6000RPM and reaction time of 40 min to determine the experimental error, as shown in Table 1.

2.3. Biodiesel synthesis

In this study, biodiesel was synthesised by batch transesterification. The reactor was a stainless steel vessel equipped with a thermocouple and a water bath jacket. Initially, palm oil of predetermined weight in the reaction vessel was heated through the water bath jacket. Then, the pre-heated solution of potassium hydroxide and methanol was transferred into the reactor. The sonication was then immediately started by a ultrasonic processor, (Frequency: 24 kHz, Maximum Power: 400 W, UP400S Hielscher Ultrasound). After a fixed duration, the final mixture was decanted to a separator funnel to be separated into two layers gravitationally. After 48 h, the bottom-layer glycerol was removed and remained biodiesel was neutralized with hydrochloric acid solution. The final product was then washed three times with distilled water to remove traces of methanol, hydrochloric acid and catalyst at a ratio of 1:1. Lastly, it was stirred at the temperature of 70 °C for 120 min. The process of collecting data is presented in Fig. 1. For the second set of experiment, all steps repeated except that, instead of sonicator, a stirred tank equipped with an overhead stirrer (13516 IKA Eurostar 60 Digital) was used for mixing the reactants.

2.4. Determination of biodiesel content

The composition of biodiesel (also known as FAME: Fatty Acid Methyl Ester) of each sample was analysed using a gas chromatography. The system was equipped with a capillary column (film thickness: 15 μm; column length: 60 m; inner diameter: 0.250 mm) and flame ionized detector (FID). Nitrogen, helium and hydrogen were used as the make-up gas (flow rate of 25 mL/min), carrier gas (flow rate: 1 mL/min) and fuel gas, respectively. The combination of cis-9-Octeic Methyl ester, Methyl Palmitate (C17H34O2), Methyl Linolate (C19H34O2), Methyl Linoleate (C19H34O2), and Methyl Stearate (C17H34O2) was used as the internal standard while n-hexane (C6H14) was employed as the solvent. The oven temperature was fixed at 50 °C and then gradually increased to 175, 230 and 280 °C at a rate of 25, 4 and 5 °C/min. The temperature of the injector and detector was fixed at 250 and 280 °C respectively. The calibration and operating conditions were in accordance with the EN14115 and ASTM D6584 standards [21,22]. Conversion of triglycerides to FAEE was calculated by:

\[
\text{Conversion (\%)} = \frac{\sum T_G_i - \sum T_G_t}{\sum T_G_t} \times 100
\]

where \(T_G_i\) and \(T_G_t\) denote the initial mass of triglyceride at initial time (\(t = 0\)) and remaining mass of triglycerides at the end of reaction (\(t = t\)), respectively. Table 1 presents two sets of experimental designs along with the subsequent responses (transesterification yield) under ultrasound irradiation and mechanical stirring respectively.

3. Soft computing approaches

The approach in this study was based on the merging of PSO, GA, DE with ANFIS and PSO and BP with RBFN.

3.1. ANFIS

Fig. 2 illustrates an ANFIS construction including five inputs (temperature, alcohol to oil molar ratio, KOH concentration, mixing intensity and time). The present study was conducted with fuzzy logic rules of Takagi and Sugeno’s, which has the following form:

If \(x_1\) is \(A_{11}\) and ... \(x_m\) is \(A_{1m}\), then \(y_i = f_i\) \((x_1, ..., x_m)\)

Herein \(x_1, ..., x_m\) denote the input variables, \(A_{11}, ..., A_{1m}\) are the membership function corresponds to each input variable for the ith rule and \(y_i\) is the output of the ith rule. \(m\) and \(i\) are also the number of...
In the present work, three bell-shaped membership functions are used. The consequent part \( f_i \) of Takagi and Sugeno’s rules has the following form:

\[
 f_i(x_1, \ldots, x_m) = \sum_{j=1}^{m} q_{ij} x_j + r_i
\]

where \( q_{ij} \) and \( r_i \) are the weighting value of the \( j \)th input onto the \( i \)th rule output, the bias of the \( i \)th output and the number of rules, respectively.

In the present work, three bell-shaped membership functions with the minimum value of 0 and maximum value of 1, were chosen for each input.

\[
 A_{ij}(x_j) = \left[ 1 + \left( \frac{x_j - c_{ij}}{a_{ij}} \right)^2 \right]^{-1}
\]

where \( a_{ij}, b_{ij}, \) and \( c_{ij} \) are the parameters of the MF, which are affected in shape of MFs.

The bell-shaped membership is a function of parameters \( a, b \) (mostly positive) and \( c \).

By receiving the input values from the 1st membership layer, the weights of each MFs are assessed by the second layer. This layer represents the fuzzy sets of the respective input variables.

The signals received from the 1st layer are multiplied in the 2nd layer and then sends them out as follows:

\[
 w_i = \mu(l)_{i-1} \cdot \mu(l)_i
\]

The output of each node demonstrates the firing strength of a weight or rule.

The 3rd layer which is known as the rule layer performs the precondition matching of the fuzzy rules. Accordingly, each node of this layer computes the weights which are normalized. Each node also estimates the ratio between the rules firing strength and the sum of all rules firing strengths using the below equation:
Normalized firing strengths or normalized weights are the outputs of the 3rd layer.

The output values resulting from the inference of rules are prepared in the 4th layer or defuzzification layer prepares, in which, each node is adaptive with the node function of:

\[
O^4_i = w_i \cdot f = w_i \cdot \left( \sum_{j=1}^{m} q_{ij} x_j + r_i \right)
\]

where \(\{p_i, q_i, r_i, s, t\}\) demonstrates the consequent parameters.

Finally, all the outputs of the fourth layer are collected in the output layer (fifth layer), where the fuzzy classification results are transferred into a binary (crisp). Then, a single node in the last layer calculates an overall output from all receiving signals as:

\[
O^5 = \sum_{i} w_i \cdot f = \frac{\sum_{i} w_i \cdot f}{\sum_{i} w_i}
\]

### 3.2. Training algorithms

The training algorithm of an ANFIS algorithm is to determine the parameters \((a_i, b_i, c_i)\) and \((q_{ij}, r_{ij})\) so that the error between the actual and predicted data should be minimized [23,24].

#### 3.2.1. Training ANFIS algorithm with Particle Swarm Optimization (PSO)

Particle swarm optimization [25] is an evolutionary computational technique that deals with the optimization of continuous/discontinuous function over a number of iterations. As a population-based search technique, each particle in a population (namely swarm) refers to a potential solution. Particles randomly distribute and change their location till the optimal state or equilibrium is obtained or computation restrictions arise. In short, for optimization of \(M\) variables, PSO considers a population (swarm) of \(N\) particles in which a position is randomly dedicated to each particle in the multidimensional space. Accordingly, each particle’s position represents a solution for the optimization problem. \(x\) refer to a particle’s position and \(v\) denoted particle flight velocity over a
solution space, respectively. A scoring function in the swarm scores each individual $x$ that obtains the accuracy of fitting. The position and velocity of a particle are determined and updated by the stochastic and deterministic update rules, as follows:

$$x_i(t) = x_i(t - 1) + v_i(t)$$

$$v_i(t) = \omega v_i(t - 1) + q_1(x_{pbest_i} - x_i(t)) + q_2(x_{gbest} - x_i(t))$$

In these equations, $q_1$ ($=r_1c_1$) and $q_2$ ($=r_2c_2$) are random variables, with $r_1$, $r_2$, $U(0,1)$ and $x$ denoting inertia weight. $c_1$ and $c_2$ denote the weights of the stochastic rule which accelerate a particle toward the personal best ($p_{best}$) and global best ($g_{best}$), respectively [26]. The best values are those with the minimum fitness values which are defined as mean square error (MSE). Large values indicate an abrupt movement of particles toward the target areas whereas small values indicate that the particles roam far from the target areas. The PSO algorithm is sensitive to the values of the inertia weight and acceleration coefficients. In this study, the value of 2.0 was used for both $c_1$ and $c_2$ which are the original parameters of particle swarm. These values satisfy the heuristic in $\omega > \frac{1}{2}(c_1 + c_2) - 1$, ensuring convergent trajectories [27,28].

Besides, a balance between the local and global explorations can be obtained by a suitable correction of inertia $x$, resulting in a reduction in the number of iterations.

In the present study, an inertia weight approach (IWA) as an inertia correction function was employed at which the inertia weight was corrected using the below equation.

$$\omega = 0_{max} - \frac{0_{max} - 0_{min}}{Itr_{max}} Itr$$

In this equation, $Itr$, $Itr_{max}$, $x_{max}$ and $x_{min}$ refer to the current number of iteration, maximum number of iteration, initial inertia weights and final inertia weights, respectively [29]. The ANFIS-PSO combination is presented in Fig. 3.

3.3 Genetic Algorithm (GA)

In genetic algorithm [30], the evolutionary process of optimization starts from a population of randomly generated solutions (chromosomes) and the process is repeatedly and loosely applied on the population of candidate solutions to achieve the best solution. A new population called generation is bred in every iteration, the fitness of each individual in the population in each generation is assessed. Accordingly, the algorithm generates new design solutions from a population of potential solutions and removes the solutions which have an inferior fitness. The number of above-average solutions thus increases after each iteration. The evolutionary process continues until a stopping criterion which is often a specific value for fitness, is met. Generally, selection, crossover and mutation are three steps of each iteration. In the first step (selection), a portion of the population is selected according to their fitness to breed a new generation. In ‘crossover’, the information (gene) of a pair of “parent” solutions is combined in such a way to produce a better new solution, called “child” solution. In ‘mutation’, the generation of the new population is complete by toggling the randomly selected gene of the randomly selected child [31].

Population size, mutation rate and cross over rate are the crucial factors in the genetic algorithm performance. Generally, “small” population size could bring unsatisfying solutions while the algorithm could consume more computation time to reach a solution in “large” population size. Too high mutation rate may lead to loss of satisfying solutions whereas; genetic drift may arise from a very small mutation rate. Additionally, a very high crossover may lead to premature convergence. Accordingly, employing the best setting could lead to find reasonable answers [31]. In this study, a wide range of population size (starting from 200) was tested and the fitness values were considered for each population size. The objective was to maximise the fitness value based on which the optimal population size was 512 and the number of chromosomes was 25. More details about the parameters used per run of GA in this study, were summarised in Table 2. It should be noted that the found mutation rate, crossover rate and inversion rate were reported in percentage values. Root Mean Square Error (RMSE) was employed as fitness function error type. The ANFIS-GA combination involving the main steps in GA assisted selection of the three parameters is depicted in Fig. 4.

3.4 Fuzzy Adaptive Differential Evolution Algorithm (ANFIS-DE)

In ANFIS-DE, the mutation factor $F$ and the crossover rate CR of DE are controlled by using a Fuzzy Logic Control (FLC); whereby the control parameters are updated in a dynamic adaptive manner through a fuzzy logic-based system [32]. The basic idea of ANFIS-DE combination is depicted in Fig. 5. ANFIS-DE has first been developed to overcome stagnation and premature convergence as the two main problems in DE. These problems often arise in any optimization method when the control parameters of these techniques are not well estimated initially and population’s information are not available [33]. The main steps of ANFIS-DE are summarised in the next parts.
3.4.1. Initialization

The process starts with the standard equation by transferring the initial population into the candidate solutions. 

\[ x_{t_0} = x_{min,j} + a_j(x_{max,j} - x_{min,j}) \]  

(13)

where \( x_{max} \) and \( x_{min} \) refer to the upper and lower bound of the parameter values. These parameters are required to specify the domain from which the values, \( x_{t_0,j} \), \( j \in \{1, \ldots, D\} \) of each \( N_p \) vectors in this initial population are selected. A random number generator \( a_{i,j} \) is also employed to provide a uniformly distributed random number ranging \([0, 1)\) [33].

3.4.2. Mutation

The standard DE/rand/1 mechanism is employed to generate the mutant vector \( v_{t} = \{v_{t,1}, v_{t,2}, \ldots, v_{t,D}\} \). This operation is used by differentiating members selected randomly, \( r_1 \) and \( r_2 \in \{1, 2, \ldots, N_p\} \), of the current population. A difference value \((x_{t,r_1} - x_{t,r_2})\) is then multiplied by the parameter namely scaling factor, \( F \) to qualify the reinforcement of the differential variation [33]. The mutation equation in the standard form is defined as:

\[ v_{t+1}^{i,j} = x_{t,r_1} + F \cdot (x_{t,r_2} - x_{t,r_3}) \quad 1 \leq i \neq r_1 \neq r_2 \neq r_3 \leq N_p \]  

(14)

3.4.3. Crossover

In crossover step, the trial vector \( U_{t} = \{u_{t,1}, u_{t,2}, \ldots, u_{t,D}\} \) is delivered by the perturbation operation called binomial crossover (bin), as illustrated in Eq. (14). The mixing process is controlled by using a user-specified real parameter value called crossover rate or crossover probability, \( CR \). Afterwards, a component of the donor vector \( v_{t}^{i,j} \) is inherited with probability \( CR \) for the offspring \( u_{t}^{i,j} \), and with probability \( 1 - CR \) from the target vector \( x_{t}^{i,j} \).
\[ \text{DE/}x/y/\text{bin } u^t_{ij} = \begin{cases} u^t_{ij} & (\beta_{ij} \leq CR) \text{ or } (j = j_{\text{rand}}) \\ x^t_{ij} & \text{otherwise} \end{cases} \quad j = 1, 2, \ldots, D \] (15)

Herein, \(\beta_{ij}\) and \(j_{\text{rand}}\) refer to a real number, uniformly generated ranging \([0, 1]\) and an integer generated randomly in the range of \([1, D]\) respectively. \(j_{\text{rand}}\) ensures that target vector \(X_i(t)\) will differ from its corresponding trial vector \(U_i(t)\) by at least one component [34].

3.4.4. Selection

Then, one-to-one greedy selection is employed to figure out whether the trial vector \(U_i(t)\) is more efficient than its corresponding target vector \(X_i(t)\), which determines if it becomes a member in the population \(X_i(t + 1)\) of the next generation. This evaluation process is usually accomplished on the basis of both individuals’ fitness functions [33].

\[ X^t_{i+1} = \begin{cases} U^t_i & \text{if } f(U^t_i) < f(X^t_i) \\ X^t_i & \text{otherwise} \end{cases} \] (16)

4. Statistical performance evaluation

To analyse the efficiency of the mentioned hybrid models, three statistical indicators were chosen:

![Performance graphs of ANFIS-PSO, ANFIS-GA, and ANFIS-DE in training and testing phase for estimating the transesterification under ultrasound irradiation.](image)

**Fig. 6.** Performance of (a) ANFIS-PSO, (b) ANFIS-GA and (c) ANFIS-DE in training and testing phase for estimating the transesterification under ultrasound irradiation.
(1) Root-Mean-Square Error (RMSE)

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}},
\]

(17)

(2) Pearson correlation coefficient (R²)

\[
R^2 = \frac{\left[\sum_{i=1}^{n}(O_i - \bar{O}) \cdot (P_i - \bar{P})\right]^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2 \cdot \sum_{i=1}^{n}(P_i - \bar{P})^2}
\]

(18)

In these equations, \(P_i\) and \(O_i\) refer to the actual (experimental) and predicted values, respectively. \(n\) denotes the total number of the experimental data which were 50 for each set of experiment. It should also be noted that, the experimental data set was divided in training and testing set several times randomly to find optimal results and the average results were considered and reported.

5. Results and discussion

In order to evaluate the reliability of the developed soft computing techniques and compare their performance, the test data were hidden during the training step. Accordingly, the obtained
transesterification yield data were categorised into training and testing data sets. The data (67%) was used for training the phase while the second set of the data (34%) was utilized for the testing phase. It should be noted that the central point was designed to detect and reduce the mistake or uncertainty of the results provided from different experiments but under similar operational conditions. As observed, the operational conditions of the test number 42–50 were exactly the same. However, it was not possible to train the mathematical algorithms under similar conditions but with different results! Accordingly, the average of the outputs was considered for these tests. The transesterification yield estimated by ANFIS-based models is represented in Figs. 6 and 7 for ultrasound irradiation and mechanical stirring respectively. Moreover, RBFN-based models have also been reported to be able to predict the behaviour of many systems well [35]. Accordingly, in this study, the applicability and accuracy of the data predicted by RBFN-PSO and RBFN-BP were also analysed. However, the results were not as accurate as those predicted by ANFIS based models.

### Table 3
Comparison the impacts of PSO, GA and DE in ANFIS algorithm.

<table>
<thead>
<tr>
<th>Parameter/function</th>
<th>PSO</th>
<th>GA</th>
<th>DE</th>
</tr>
</thead>
<tbody>
<tr>
<td>General features</td>
<td>Multi-agent, probabilistic search</td>
<td>Multi-agent, probabilistic search</td>
<td>Multi-agent, probabilistic search</td>
</tr>
<tr>
<td>Individual operators</td>
<td>Personal best position history, velocity, inertia</td>
<td>Mutation</td>
<td>Mutation</td>
</tr>
<tr>
<td>Social operators</td>
<td>Neighborhood best position history</td>
<td>Selection, crossover</td>
<td>Cross over and selection</td>
</tr>
<tr>
<td>Particle's variables (tracked by the population manager)</td>
<td>Position and velocity</td>
<td>Position</td>
<td>The component of each vector and the mutant vector are tracked</td>
</tr>
<tr>
<td># of algorithmic parameters (basic)</td>
<td>w, cn, cp, k, m, position and velocity range</td>
<td>pm, pc, selection par, m, position range</td>
<td>F, CR, NP</td>
</tr>
<tr>
<td>Population diversity</td>
<td>Mainly via local neighborhood</td>
<td>Somehow tunable via pc/pm ratio and selection schema</td>
<td>Somehow tunable via F. F, known as mutation scaling factor, is typically chosen from the range (0,2]. Larger values in higher diversity in the generated population and lower values cause faster convergence</td>
</tr>
<tr>
<td>Global/local search balance</td>
<td>Tunable with w (w → global search; w → local search)</td>
<td>Somehow tunable via pc/pm ratio and selection schema</td>
<td>Somehow tunable via F.</td>
</tr>
<tr>
<td>Original search space</td>
<td>Continuous (but some work on discrete nowadays)</td>
<td>Discrete (but continuous ok nowadays)</td>
<td>Continuous</td>
</tr>
</tbody>
</table>

Note: In PSO and GA algorithms, the particle's variable are tracked, while in DE algorithm the vector's component are tracked.

### Table 4
Performance indices of various approaches for the estimation of the transesterification yield Using Root-Mean-Square Error & Pearson correlation coefficient.

<table>
<thead>
<tr>
<th>Method</th>
<th>Training</th>
<th>Testing</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Error (RMSE)</td>
<td>Pearson coefficient (R²)</td>
<td>Error (RMSE)</td>
<td>Pearson coefficient (R²)</td>
</tr>
<tr>
<td>Ultrasound irradiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANFIS-PSO</td>
<td>0.423</td>
<td>0.977</td>
<td>1.284</td>
<td>0.847</td>
</tr>
<tr>
<td>ANFIS-GA</td>
<td>0.911</td>
<td>0.886</td>
<td>3.209</td>
<td>0.436</td>
</tr>
<tr>
<td>ANFIS-DE</td>
<td>0.816</td>
<td>0.909</td>
<td>1.097</td>
<td>0.841</td>
</tr>
<tr>
<td>Mechanically stirring</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANFIS-PSO</td>
<td>0.242</td>
<td>0.996</td>
<td>1.794</td>
<td>0.956</td>
</tr>
<tr>
<td>ANFIS-GA</td>
<td>0.780</td>
<td>0.963</td>
<td>1.285</td>
<td>0.978</td>
</tr>
<tr>
<td>ANFIS-DE</td>
<td>0.743</td>
<td>0.967</td>
<td>1.293</td>
<td>0.978</td>
</tr>
</tbody>
</table>

Bold values point out to the most accurate predictions.

### Table 5
Performance indices of various approaches for the estimation of the transesterification yield using Kolmogrov-Smirnov Analysis.

<table>
<thead>
<tr>
<th>Method</th>
<th>Training</th>
<th>Testing</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D-Statistic</td>
<td>P-value</td>
<td>D-Statistic</td>
<td>P-value</td>
</tr>
<tr>
<td>Ultrasound irradiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANFIS-PSO</td>
<td>0.097</td>
<td>0.999</td>
<td>0.308</td>
<td>0.588</td>
</tr>
<tr>
<td>ANFIS-GA</td>
<td>0.100</td>
<td>0.998</td>
<td>0.385</td>
<td>0.100</td>
</tr>
<tr>
<td>ANFIS-DE</td>
<td>0.133</td>
<td>0.958</td>
<td>0.385</td>
<td>0.300</td>
</tr>
<tr>
<td>Mechanically stirring</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANFIS-PSO</td>
<td>0.100</td>
<td>0.999</td>
<td>0.538</td>
<td>0.044</td>
</tr>
<tr>
<td>ANFIS-GA</td>
<td>0.129</td>
<td>0.963</td>
<td>0.385</td>
<td>0.300</td>
</tr>
<tr>
<td>ANFIS-DE</td>
<td>0.100</td>
<td>0.999</td>
<td>0.231</td>
<td>0.898</td>
</tr>
</tbody>
</table>

Bold values point out to the most accurate predictions.
Therefore, the results were not reported. A comparison between the training algorithms in RBFN-based predictions indicates that even though the performance of both RBFN-BP and RBFN-PSO is slightly similar, the $R^2$ value for RBFN-PSO is considerably greater than that for RBFN-BP, indicating better performance of RBFN-PSO. Compared to BP, statistical results show that PSO algorithm is superior in finding the optimal weights for the RBFN neuro identifier and in term of robustness. However, compared to RBFN-based modes, ANFIS-based models presented higher prediction accuracy. In this study, only the results obtained by ANFIS-based models (ANFIS-PSO, ANFIS-GA and ANFIS-DE) were reported. The differences between these methods in the training of ANFIS algorithm are summarised in Table 3. Generally, PSO benefits from a stochastic optimization technique, which is well adapted to the optimization of nonlinear functions in multidimensional space. In ANFIS-PSO, ANFIS provides the search space while PSO serves to find the optimal solution by comparing objective function from time to time. The difference between the actual data and the model output can be reduced to a minimum value by iterating the PSO algorithm \cite{23,36,37}. The main advantages of the PSO are that the
algorithm has a high degree of stability and does not rely on the derivative nature of objective function and can reach to the best solution by tuning the membership functions. In this study, ANFIS-PSO could reach a satisfying result and among the RBFN-PSO and RBFN-BP, the former provided stronger results. Genetic algorithms (GAs) are search algorithms that are capable of locating optimal solutions of a large-scale optimization problems among a population of individuals that evolve according to a set of rules such as selection, crossover and mutation. GA is very similar to PSO; but unlike PSO, GA involves evolutionary operators such as crossover and mutation. GA may perform better for more complex and discrete search problems. Genetic Algorithm and Evolutionary Strategies (ES) are the basics of Differential evolution (DE). Accordingly, DE has some similarities with GA, but the correspondences are different. The genotype in DE is some form of real-valued vector. The ideas of mutation/crossover operations in DE are substantially different from those in GA. The mutation and crossover make use of the difference between two or more vectors in the

![Graphs showing cumulative relative frequency](image-url)
population to create a new vector. As a result, in DE algorithm, GA changes the structure of individuals through mutation and crossover operators, while ES achieves self-adaptation by a geometric manipulation of individuals. It has been reported that DE algorithm generates more robust results compared to the GA [38,39]. The reason is that the vectors can be considered to form a “cloud” that explores the high-value areas of the solution space quite effectively. It can be concluded that, ANFIS is a robust method which has high adaptability with strong nonlinear data. However the merging of the ANFIS with other soft computing methods enhanced the ANFIS capabilities more. It is confirmed by statistical indicators. The performance of ANFIS-PSO, ANFIS-DE and ANFIS-GA in training and testing phases of transesterification under ultrasound energy as well as mechanically stirring have been reported in Figs. 6 and 7. According to the figures, all models, “especially in terms of transesterification under mechanically stirring” have been well trained. The reason relates to the nature of the process under mechanically stirring, which is more systematic compared to the transesterification under ultrasound energy. In other words, cavitation phenomena are activated under ultrasound energy of which their effects cannot be easily considered. These phenomena have been well trained. The reason relates to the nature of the process under mechanically stirring, which is more systematic compared to the transesterification under ultrasound energy. In other words, cavitation phenomena are activated under ultrasound energy of which their effects cannot be easily considered. These phenomena affect the range of results to some extent. Similar level of scattering can be observed in testing phase, which is more highlighted in case of transesterification under ultrasound energy. 

In order to better analysis, root-mean-square error (RMSE) and coefficient of determination ($R^2$) were used to evaluate the differences between the predicted and real values for the soft computing techniques and the results were reported in Table 4. Based on the results, although ANFIS-PSO presented the highest accuracy in training phase, it generated the weakest results in testing of the output. ANFIS-DE had the best performance in testing phase and generated more reliable results compared to the other two techniques. These observations were consistent in both cases of transesterification under ultrasound energy and mechanically stirring. However, RMSE and correlation coefficient as the lone methods of analysis, cannot give reliable criteria of the accuracy and applicability of the predictive models. In other words, the models cannot only be compared based on the RSM values. In order to better judge the performance of each model in a trustful way, the distributions of individual error from each model were investigated and compared with a non-parametric test. Toward this objective, Kolmogorov-Smirnov test (K-S test) was used. The K-S test determines if two datasets differ significantly. This test has the advantage of making no assumption about the distribution of the data and the determination is considered through two parameters: D statistic and P-value. Kolmogorov-Smirnov’s D statistic is the highest deviation occurring between the two curves. P-values report if the numbers differ significantly. The null hypothesis is rejected if P is “small”. In this study, two series of analysis were accomplished using the K-S test: (i) based on the distribution of an output parameter (transesterification yield), (ii) based on the distribution of the residuals. The graphs related to the distribution of output parameter have not been presented however, the results were summarised in Table 5.

As per results, all models were well trained using the data of transesterification either under ultrasound energy or mechanically stirring. The maximum deviations were in a narrow range (0.097–0.133) with the corresponding P-values between 0.958 and 0.999. These values proved that the models were all well trained and the null hypothesis could not be rejected (P-value > 0.05). In other words, the distribution of both data sets (Actual and predictions) was similar and did not have any differences in the training phase. On the other side and as expected, the testing results had more deviations compared to the training ones. Although all three models had similar accuracy in predictions of the results under ultrasound energy, the results under mechanically stirring were quite different. In consistent with RMSE and $R^2$ values, although ANFIS-PSO had the highest accuracy in training phase, the minimum deviation between experimental and predicted results was observed using ANFIS-DE algorithm. Accordingly and based on P-values, ANFIS-DE had the best prediction in this case while the null hypothesis was accepted in case of ANFIS-PSO. Since a model with poor residuals but a similar reaction yield distribution can get favorable results. Thus, it is important to compare the distributions of the residuals (obtained on the same training/testing datasets). Accordingly, the residuals of each model (training vs model, testing vs model) were obtained and compared with those of the other models through the non-parametric K-S test. In this way we can answer to the question: “Are the model performances (in terms of residuals) significantly different from each other?". The K-S graphs for the residuals of transesterification predictions in both training and testing phases under ultrasound energy and mechanically stirring were presented in Figs. 8 and 9, respectively. Moreover, the D and P values of each prediction were reported in Table 6. As per results, the binary comparison of the residual distributions depicted that null hypothesis cannot be rejected in training phase of reaction yield under ultrasound energy. However, the risk to reject the null hypothesis while it is true is only 21.98% and 12.24% at the binary comparison of ANFIS-PSO/ANFIS-GA and ANFIS-PSO/ANFIS-DE; but the value is 99.78% at the binary comparison of ANFIS-GA/ANFIS-DE. In other words, the residuals obtained by predictions using ANFIS-GA and ANFIS-DE follows similar distribution. This difference is more visible in testing phase at which the null hypothesis is rejected except for the binary comparison of ANFIS-GA/ANFIS-DE. In other words, the distribution of residuals predicted by ANFIS-PSO does not follow those obtained by ANFIS-GA and ANFIS-DE. While the two latter, follow similar distribution. Distribution of residuals in case of transesterification under mechanical mixing confirm the similar results. In this case, the hypothesis is rejected for the two comparisons of ANFIS-PSO/ANFIS-GA and ANFIS-PSO/ANFIS-DE, confirming the distribution of residuals obtained from ANFIS-PSO differs from the other two.

<table>
<thead>
<tr>
<th>Method</th>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D-Statistic</td>
<td>P-value</td>
</tr>
<tr>
<td>Ultrasound irradiation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANFIS-PSO/ANFIS-GA</td>
<td>0.276</td>
<td>0.220</td>
</tr>
<tr>
<td>ANFIS-PSO/ANFIS-DE</td>
<td>0.310</td>
<td>0.122</td>
</tr>
<tr>
<td>ANFIS-GA/ANFIS-DE</td>
<td>0.103</td>
<td>0.998</td>
</tr>
<tr>
<td>Mechanically stirring</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANFIS-PSO/ANFIS-GA</td>
<td>0.552</td>
<td>0.000</td>
</tr>
<tr>
<td>ANFIS-PSO/ANFIS-DE</td>
<td>0.345</td>
<td>0.064</td>
</tr>
<tr>
<td>ANFIS-GA/ANFIS-DE</td>
<td>0.241</td>
<td>0.367</td>
</tr>
</tbody>
</table>

Bold values point out to the most accurate predictions.
models. In testing phase, although the null hypothesis is not rejected in neither of the models, the risk to reject the null hypothesis while it is true is $8.75$ for the residuals obtained for binary comparison of ANFIS-PSO/ANFIS-GA and ANFIS-PSO/ANFIS-DE but it is $10$ for those obtained for ANFIS-GA/ANFIS-DE. Consistent with the previous results, obtained based on the output variable, the ANFIS-PSO could not be used as an online tool for predicting the biodiesel yield.

6. Conclusion

The applicability of three hybrid neural networks, ANFIS-PSO, ANFIS-DE and ANFIS-GA to estimate transesterification yield for biodiesel synthesis was analysed in this study. Two different methods of synthesis were employed: transesterification under (i) ultrasound irradiation and (ii) mechanical stirring along with conventional heating. Moreover, operating factors and reaction yield were considered as the independent variables (or input factors) and the output variable, respectively. Two statistical indicators, RMSE and $R^2$ along with Kolmogorov-Smirnov method were also used to evaluate the performances of the introduced models. The simulation results demonstrated that ANFIS-GA and ANFIS-DE models could be successfully applied to estimate transesterification yield. However, the best prediction was obtained by ANFIS-DE. Generally, in hybrid networks, finding the ANFIS optimal structure would have taken very short time, facilitating their applications as online tools. In order to outperform the ANFIS-DE, future works include the investigation of DE variants. In terms of ANFIS-PSO, although it presented the highest accuracy in training phase, the weakest results in testing of the output was obtained. Accordingly, the results obtained by ANFIS-PSO were not reliable based on K-S Analysis. In this case modification of PSO algorithm can be an alternative to improve the ANFIS-PSO results. Accordingly, the results obtained by ANFIS-PSO were not reliable based on K-S Analysis. In this case, modification of PSO algorithm can be an alternative to improve the ANFIS-PSO results. For example, for removing worst particle in population and replace it by new particle; it is important that how to determine worst particle and how to generate new particle for current population. Generally, in order to understand where the algorithm suffers, more works should focus on topology of communication, parameter adjustment, initial distribution of particles and methods to deal with stagnation.

Acknowledgement

The authors are grateful to the University of Malaya High Impact Research Grant (HIR-MOHE-D00038-16001) from the Ministry of Higher Education Malaysia and University of Malaya IPPP (PG115-2012B) which financially supported this work.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at [http://dx.doi.org/10.1016/j.measurement.2017.01.044](http://dx.doi.org/10.1016/j.measurement.2017.01.044).

References

[10] B. Sajjadi, A.A.A. Raman, S. Ibrahim, A comparative fluid flow characterisation test with the previous results, obtained based on the output variable, respectively. Two statistical indicators, RMSE and $R^2$ along with Kolmogorov-Smirnov method were also used to evaluate the performances of the introduced models. The simulation results demonstrated that ANFIS-GA and ANFIS-DE models could be successfully applied to estimate transesterification yield. However, the best prediction was obtained by ANFIS-DE. Generally, in hybrid networks, finding the ANFIS optimal structure would have taken very short time, facilitating their applications as online tools. In order to outperform the ANFIS-DE, future works include the investigation of DE variants. In terms of ANFIS-PSO, although it presented the highest accuracy in training phase, the weakest results in testing of the output was obtained. Accordingly, the results obtained by ANFIS-PSO were not reliable based on K-S Analysis. In this case modification of PSO algorithm can be an alternative to improve the ANFIS-PSO results. Accord-