Research Article

Comparison of Neural Networks. An Estimation Model in Yield of Monoglycerides from Biodiesel by-Product

Dolores M. E. Álvarez[1], Nancy F. Bálsamo[1], Mario R. Modesti[2] and Mónica E. Crivello[1,*]

[1]*Centro de Investigación y Tecnología Química-Universidad Tecnológica Nacional-CONICET-Facultad Regional Córdoba, Maestro Marcelo López esq. Cruz Roja Argentina, Ciudad Universitaria, CP 5016, Córdoba, Argentina.*
[2]*Centro de Investigación en Informática para la Ingeniería (CIII), Universidad Tecnológica Nacional - Facultad Regional Córdoba, Córdoba, Argentina.*

___

Abstract

Biodiesel is generally manufactured by transesterification, obtaining glycerol as a by-product. The transesterification of methyl stearate selectively produced monoglycerides, for glycerol valuation. Mixed oxides containing lithium catalysed the reaction. The purpose of this work was to develop and compare mathematical models obtained through artificial neural networks (ANN), capable for characterising the relationship between the mole percent conversion of methyl stearate and the yield of the products mono-, di- and triglycerides. The lowest mean squared error (MSE), the highest correlation coefficient (R), similarity in the evolution of validation and simulation errors and absence of data overlearning were considered to select the best model. Three ANNs with backpropagation structures were compared. They evidenced high correspondence between the estimated product yield values and the interpolated experimental ones. The ANN containing 35 neurons with sigmoid transfer function in the hidden layer and a linear neuron in the output one was the simplest. Consequently, the 5, 15 and 60 neurons were also explored in the hidden layer. The ANN structured with an intermediate number of neurons (35) achieved the most adequate MSE, considering mono- and diglyceride products (0.011193, 0.000489). The development of these models contributes to the dynamic estimation of the process.

Keywords: Artificial Neural Network; Monoglycerides; Yield;

___

1. Introduction

The use of renewable energy sources is being proposed as an important way of development, due to the environmental benefits involved. Thus, there is growing interest in obtaining biofuels [1]. In this context, biodiesel has gained significant attention as a non-toxic, biodegradable and renewable alternative to petroleum-based fuels. This biofuel is usually manufactured by transesterification of oils with methanol or ethanol, obtaining glycerol as a by-product [2].

The increase in biodiesel production is generating a glycerol surplus. Accordingly, it is desirable to convert this low-cost secondary product into value-added chemicals or materials. The synthesis of monoglycerides (MG) by transesterification of fatty acid methyl esters is an attractive option to revalue glycerol. MG are emulsifiers widely used in food and pharmaceutical industries. They improve the long-term stability of food emulsions and facilitate the displacement of protein from oil surfaces [3]. The transesterification reaction can be catalysed by solid bases such as mixed oxides derived from layered double hydroxides (LDH) modified with metal ions [4, 5].

LDH are a group of materials that have attracted great scientific and industrial interest due to the possibility of customized nanodesign, control of accessibility to active sites, and a wide range of applications. At present, many structural, textural and compositional modifications of these materials are known, which allows a fine adjustment and therefore, the control of catalytic reactivity [6-7]. Lithium has positioned itself as one of the catalysts in a wide range of applications. One of its possible uses is as a catalyst with basic properties, incorporated in solid materials [8].

Current industrial processes for MG production, in the presence of inorganic catalysts at high temperatures (>200 °C), reach a selectivity around 50%. This reaction is known as glycerolysis and produces a mixture of MG and diglycerides (DG), and sometimes triglycerides (TG). MG are the most important of them because they contribute to the emulsifying property of the obtained mixture product. For this reason, their highest yield was sought. Thus, the possibility of estimating MG yield is significant in the chemical process. However, many of the technologies offered still need further development to make them cost-effective and operationally feasible for incorporation into biorefineries [9, 2].

Since more than three million tons of glycerol are expected to be introduced on the market in the immediate future [10] as a consequence of the biodiesel industry, the possibility of having a mathematical model to improve the MG obtaining process is of great interest.

Artificial neural networks (ANN) are powerful mathematical tools used to model and simulate different processes of non-linear characteristics. They can estimate the practical relationships between the independent variables (input) of the process and the dependent ones (output), using data which were obtained experimentally. There is no

___

criterion that can be applied to its constitution [11]. However, there are ANN architectures with more appropriate characteristics according to the problem to be solved.

The ANN backpropagation algorithm is based on the least squares error calculation and uses an associated error gradient function, used to update the weights of the connections between neurons, as shown in Eq. 1:

$$w_{ij}(t+1) = w_{ij}(t) - \eta \, \partial E/(\partial w_{ij}) \qquad (1)$$

Where *wij* represents the weights of the connections between the neurons *j* (of the hidden layer) and *i* (of the output layer); *t*, the time and *η*, the learning rate. The gradient is obtained from the generalized delta rule. It modifies the weights in order to locate the global minimum on the error surface in weight space. A parameter called *momentum* adjusts the network to ignore the minimum local values on the surface [12].

Both the learning rate *η* and the *momentum* coefficient are optimized heuristically in relation to the number of neurons in the hidden layer [13]. The training ends when a specific objective is achieved, according to the operator pre-established criteria or when the response is closer to the desired one, that is, the minimum error is reached. Once the network is trained, the application continues with its validation. This is the ability to give satisfactory answers to inputs that the system has never seen in its training phase [12]. This procedure allows the estimation of the model generalization capacity.

A preliminary simulation to evaluate the generalization capacity of the model is executed. This process has no effects on training but provides an independent approach to the performance of the network. The complexity of the networks can reduce their capacity for generalization (overlearning). It is possible that an increasing number of neurons in the hidden layer also increase the mean squared error (MSE) [14]. In this situation, if the validation MSE increases while the training MSE decreases, an ANN overlearning may occur.

The application of ANNs in science and engineering is increasing rapidly; they have been used to model and predict biodiesel characteristics and production processes [15, 16]. Chakraborty and Sahu [17] employed ANN for modelling the transesterification process of biodiesel synthesis from waste goat tallow. Rocabruno-Valdés et al. [18] presented models based on ANN to predict the density, dynamic viscosity, and cetane number of biodiesel. Moradi et al. [19] studied the transesterification of soybean oil to biodiesel using KOH in different process conditions and applied ANN to estimate the biodiesel yield.

The purpose of this work is to develop and evaluate different ANN structures in order to characterise the relationship between the mole percent conversion of the methyl stearate reagent and the yield of the obtained products. In the first phase, three different ANN structures were evaluated. One of them was selected to value its performance, modifying the number of neurons in its hidden level. In this way, the impact produced by the different number of hidden neurons in the selected model was analysed.

2. Experimental

2.1. Catalyst synthesis
The LDH consist of sheets of Mg-Al hydroxides and a third metal cation that, depending on its charge, replaces Mg or Al.

There are also charge compensating anions located in the interlayer completing the characteristic structure.

They have the general formula $[M(II)_{1-x} M(III)_x(OH)_2]^{x+}(A^{n-}_{x/n}).m \, H_2O$, where $M(II)$ and $M(III)$ represent the di- and trivalent positive ions and A^{n-}, the negative ones. The Mg and Al layer structure was modified. The theoretical molar ratio, $(M^{2+}+M^{1+}):M^{3+}$, was 3:1. The LDH were synthesized with the addition of Li^{1+} as a third metal. The incorporation of Li was carried out replacing 15% of Mg moles. The synthesis consisted of the co-precipitation method. Two solutions were prepared, one containing the nitrates of Mg, Al and Li and the other, Na_2CO_3. Both solutions were dropped slowly at a constant rate of 60 mL h^{-1} to obtain the desired structure. A third solution of NaOH was added to keep the pH constant at the value of 10 ± 0.2. When the drip was finished, the system continued stirring for 4 h and then it kept aging for 8 h. Finally, the resulting precipitate was washed with distilled water until reaching a pH = 7. The purpose of the distilled water wash was to extract all the ions that are not part of the structure. The obtained solid was dried at 90 °C for 12 h. The corresponding mixed oxides were obtained by thermal decomposition of LDH at 450 °C for 9 h.

2.2. Process description
The transesterification reaction was performed in a batch reactor at 220 °C. A stream of nitrogen was circulated to remove the methanol generated, which was recovered by condensation. A glycerol/methyl stearate molar ratio of 6:1 was employed. The catalyst (3 wt%) was added. The reaction samples were taken every 30 min for 7 h.

The reaction products were identified and quantified by gas chromatography using Perkin Elmer Clarus 500 equipment with flame ionization detector. Quantification was carried out using the standards 1-steoryl-rac-glycerol (Sigma, >99% GC), Glyceryl 1,3-distearate (Sigma, >99% GC) and Glyceryl tristearate (Sigma, ≥99% GC). Silylated samples of products and reactants with a known amount of tricaprin (Glyceryl tridecanoate, Sigma, >99% GC), added as an internal standard, were analysed.

The relative response factors were also calculated for a quick follow-up of the conversion. In this way, the mole percents of stearate conversion and selectivity of the products MG, DG and TG were obtained. The yield was calculated as the multiplication between reactant conversion and the selectivity of each product.

The synthesis method and the characterization of the catalysts, as well as a preliminary catalytic testing in the transesterification reaction, have been exhaustively discussed in a previous publication [20].

2.3. Model development
Three ANN backpropagation models were proposed, ANN1, ANN2 and ANN3 (Fig. 1, a-c). The first one (ANN1, Fig. 1-a) consisted of 35 sigmoid neurons in the hidden layer followed by a linear layer. The second one (ANN2, Fig. 1-b) was composed of a cascade structure with 25 neurons in the hidden layer and an output layer, both with sigmoid transfer function. The third model (ANN3, Fig. 1-c), in contrast with the previous ones, was composed of two hidden layers of 25 and 10 neurons respectively. The first layer was structured with a sigmoid transfer function and the second, with radial basis one. As in ANN1, the transfer function for the output layer was linear.

The structure of the employed transfer function influences the neural network learning rate. The sigmoid transfer function (Fig. 2-a, Eq. 2) used in the hidden layers is

continuous and derivable. This allows ANN to adapt to the non-linearity of the process [13]. The radial basis function is Gaussian and also continuous, which provides adaptation characteristics different to those proposed by the sigmoid function (Fig. 2-b, Eq. 3). The linear transfer function allows the simulation without discontinuities [11] (Fig. 2-c, Eq. 4).

Fig. 1. Scheme of the artificial neural networks: (a) ANN1, (b) ANN2 and (c) ANN3.

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

$$f(x) = a \, e^{-b \, x} \tag{3}$$

$$f(x) = a \, x \tag{4}$$

Fig. 2. Transfer functions: (a) sigmoidal, (b) radial basis, (c) linear.

The input matrix of the ANNs was composed of the mole percent conversion of methyl stearate, plus first and second derivatives (3x420 elements). The output included the mole percent yield, corresponding to the production of MG, DG. These data constitute three matrices of 1x420 elements. All these values correspond to the process and are not shown in this work.

The experimental data were taken at 30 min intervals. They were interpolated applying differentiable curves defined in portions by means of polynomials of a lower order (splines) to obtain a continuous trend in the 7 h study period. The interpolation method was chosen to provide the model with the ability to determine the parameter values that were not obtained experimentally. Due to the non-linearity of the process, the first and second derivatives of the mole percent conversion were incorporated as training information of the ANNs. This allowed a fast convergence in the training phase.

The training application divided the input and output matrix into three parts, 70% was used for network training, 15% for validating, and the remaining percentage for evaluating its generalization capacity. The first and second

derivatives of the mole percent conversion were a substantial contribution for the network convergence to an acceptable error. The performance of the ANN models were measured by MSE (Eq. 5) and correlation coefficient (R) between the network estimated values and the experimental ones. The similarity of the validation and test errors and the inexistence of overlearning were also considered.

$$MSE = \frac{1}{n} \sum_{j=1}^{p} (Y_j^t - Y_j)^2 \tag{5}$$

Where p is the number of training patterns; Y, the estimated values and Y^t, the experimental ones.

According to the analysis of the indexes for the proposed neuronal structures (mentioned above and analysed in the results and discussion section), ANN1 (Fig. 1-a) was selected for evaluation. Its configuration was 3-35-1 (3 inputs, 35 neurons at hidden layer and 1 output). In order to find the optimum number of neurons in the hidden layer, four different 3-*x*-1-architectures (*x* increases from 5 to 60, including the 35 provided in the original ANN1 structure) were evaluated. The same evaluation criteria used for the preliminary ANN analysis were applied.
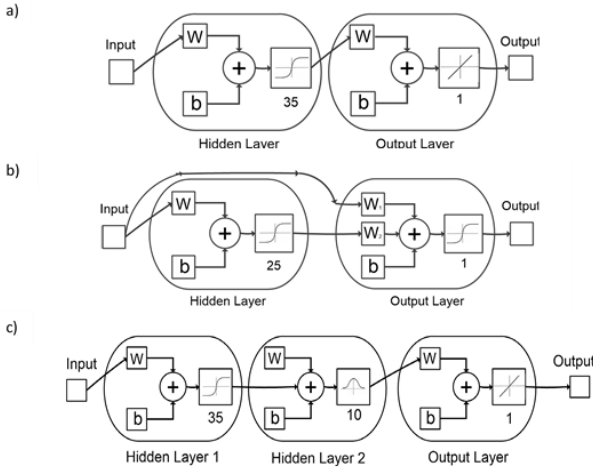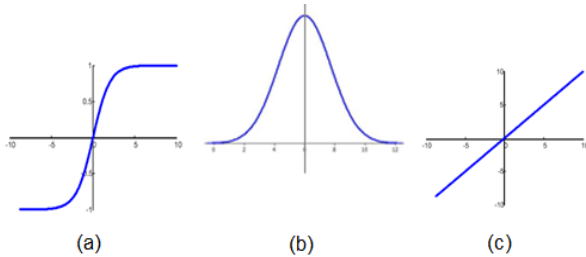
3. Results and discussion

Figs. 3 and 4 compare experimental performance with the one estimated by the three proposed neural models for MG (Fig. 3, a-c) and DG (Fig. 4, a-c). TG models were avoided because of their low mole percent yield. As can be seen, the estimated values have a very good approximation in relation to the interpolated experimental ones for MG and DG. The R values for each network are summarized in Tab. 1.
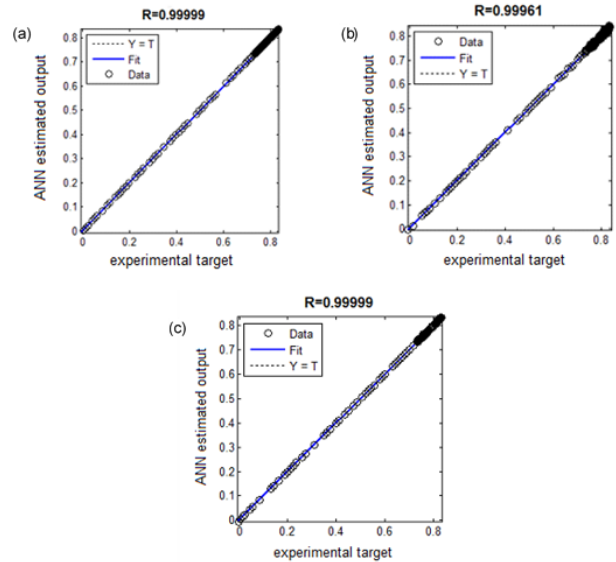

Fig. 3. Correlation between the experimental and estimated values of the models: (a) ANN1, (b) ANN2 and (c) ANN3 for the validation data of MG product.

Table 1. R values for different ANN architectures.

Network	R	
	MG	DG
ANN1	0.99999	0.99770
ANN2	0.99961	0.99948
ANN3	0.99999	0.99983

For MG it was observed that ANN1 and ANN3 had better performance (Fig. 3, a-c). For DG, instead, the models did not have the same precision as MG. However, the best response

is corresponded to ANN2 and ANN3 (Fig. 4, a-c). This may be a consequence of the model structures; ANN1 and ANN3 have 35 neurons in the hidden layer in comparison with ANN2, which includes 25. In general, by increasing the number of neurons in the hidden layer, it is possible to improve the approach capability [14].
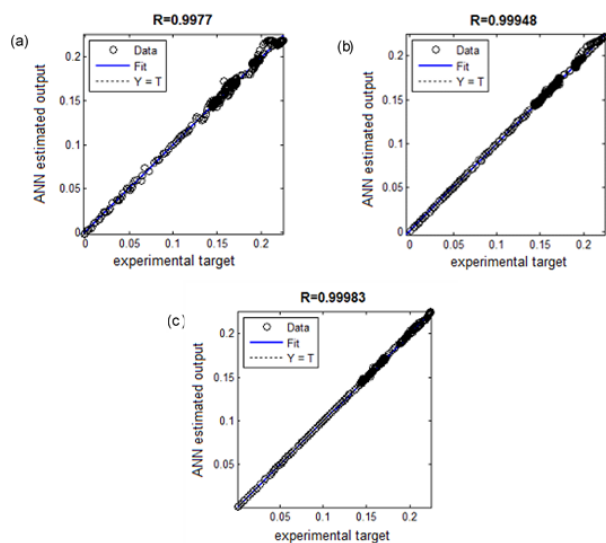

Fig. 4. Correlation between the experimental and estimated values of the models: (a) ANN1, (b) ANN2 and (c) ANN3 for the validation data of DG product.

In the case of MG, the most important product in comparison with DG, all the proposed ANNs had similar performance. It was observed that the increment in complexity does not lead to substantial improvements in the approach (Tab. 1). So, the simplest neural structure, ANN1 (configuration 3-35-1, Fig. 1-a), was selected [9, 11, 15]. Consequently, it was chosen for evaluation, modifying the number of neurons in its hidden level.

Four different ANN 3-x-1-architectures (x: 5, 15, 35 and 60 neurons) were considered. Fig. 5 and 6 show the evolution of the MSE during the training, validation and testing of the ANNs in relation to the number of validation checks for MG and DG, respectively. It was verified that the evolution of the validation and simulation errors was similar in all the cases, differing in the number of required iterations and the final error reached.
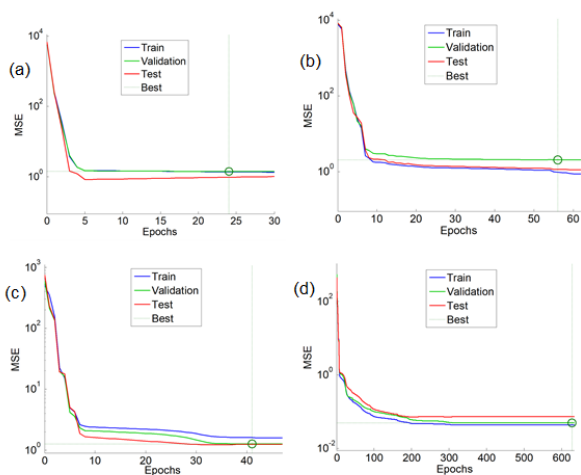

Fig. 5. Evolution of MSE during training, validation and testing of ANN1 with (a) 5 neurons, (b) 15 neurons, (c) 35 neurons and (d) 60 neurons in the hidden layer for MG.
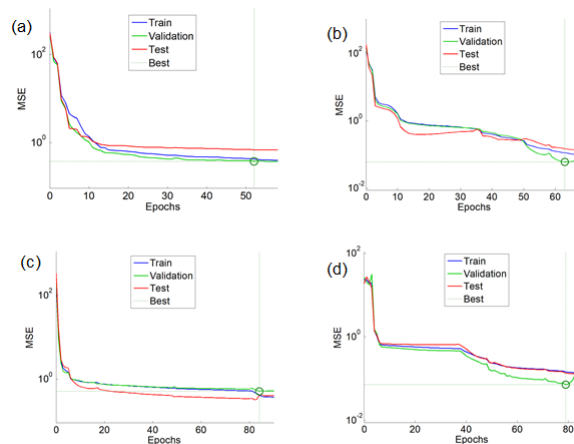

Fig. 6. Evolution of MSE during training, validation and testing of ANN1 with (a) 5 neurons, (b) 15 neurons, (c) 35 neurons and (d) 60 neurons in the hidden layer for DG.

Tab. 2 displays the MSE and R values reached, corresponding to the experimental data and the estimated ones of the evaluated models for MG and DG. These errors indicate the fitting rate of the ANNs. Fig. 7 shows the evolution of MSE according to the number of neurons in their hidden layers (5, 15, 35 and 60), during the ANN testing for MG and DG. It can be seen that the MSE and R values generated by changing the number of neurons in the hidden layer produced acceptable results with respect to the experimental ones (Tab. 2). In general, for both products a decrease in MSE is observed when the number of neurons increases. This behaviour is also reflected when comparing R values. While the number of neurons increases, R values rise too.

Table 2. MSE and R values for different numbers of neurons of the ANN1

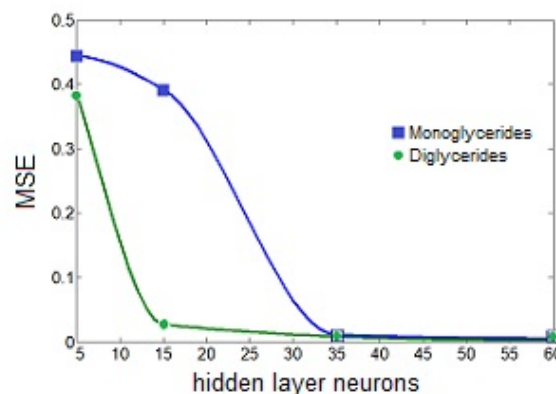| ANN	MG		DG	
	MSE	R	MSE	R
3-5-1	0.443736	0.999422	0.384768	0.994108
3-15-1	0.392565	0.999503	0.027687	0.999464
3-35-1	0.011193	0.999977	0.000489	0.999990
3-60-1	0.005524	0.999989	0.002319	0.999952

Fig. 7. MSE trend according to the number of neurons in the hidden layer for MG and DG products.

In particular, it was verified that the errors achieved by ANNs composed of 35 and 60 neurons in the hidden layer were similar to each other and smaller than those of 5 and 15 neurons for MG (Tab. 2, Fig. 7). It was observed that the

smallest error occurred in the network composed of 60 neurons in the hidden layer (0.005524). The MSE observed in the 35 neuron network was 0.011193.

Moreover, according to Tab. 2, in all cases the MSE values for DG were lower than those for MG. The best value for DG was obtained for 35 neurons. In this case, MSE and R values between the experimental and estimated responses were 0.000489 and 0.999990, respectively.

Although none of the networks proposed for MG (Fig. 5, a-d) or DG (Fig. 6, a-d) showed overlearning in the training phase, when the number of neurons in the hidden layer increases, the MSE could also rise [14]. Consequently, if there are two ANNs with similar estimation capacity, the simplest one is desirable [12]. From the current evaluation and considering that MG is the product of interest in the reaction, the ANN model composed of 35 neurons is accepted as the best option for representing the relationship between the inputs and outputs of the system.

The criterion used for the selection of the number of hidden neurons is the same as that assumed by Abdul Rahman et al. [11] and Moradi et al. [16]. They concluded that the ANN backpropagation model that showed the best performance to explain the relationship between the input and output variables of the reactions studied in their respective works included an intermediate number of neurons in the hidden layer.

4. Conclusions

In this work, different ANN backpropagation structures were designed and evaluated by comparing their performances using the correlation coefficient (R) and MSE to select the best model. All of them could successfully reproduce the relationship between the mole percent of the reagent conversion, methyl stearate, and the yield of the products MG and DG. Mixed oxides derived from layered double hydroxides (LDH) containing lithium were used as catalyst. However, the most suitable network was a two-layer structure with 35 neurons in the sigmoidal hidden layer and a linear neuron in the output layer.

In addition, the impact produced by different numbers of neurons (5, 15, 35 and 60) in the hidden layer was analysed for the selected ANN. It was found that the best accuracy was related to the 3-35-1 final model.

It can be affirmed that ANN models successfully represent the reaction carried out, which allows a good estimation of the non-linearity of the process. Therefore, it is possible to use this methodology to obtain mathematical models of these experiments in the simulation phase. The development of these models is of great interest since they contribute to the dynamic estimation of the process.

5. Acknowledgments

References

[1] D.M. Alonso, J.Q. Bon and J.A. Dumesic, Green Chem. 12 1493 (2010).
[2] J.C. Juan, D.A. Kartika, T.Y. Wu and T.Y. Hin, Bioresour. Technol. 102, 452 (2011).
[3] D. Singh, P. Patidar, A. Ganesh and S. Mahajani, Ind. Eng. Chem. Res. 52 (42), 14776 (2013).
[4] P.G. Belelli, C.A. Ferretti, C.R. Apesteguía, R.M. Ferullo and J.I. Di Cosimo, J. Catal. 323, 132 (2015).
[5] N. Balsamo, S. Mendieta, M. Oliva, G. Eimer and M. Crivello, Procedia Mat. Sci. 1, 506 (2012).
[6] R. Chimentao, S. Abelló, F. Medina, J. Llorca, J. Sueiras, Y. Cesteros and P. Salagre, J. Catal. 252, 249 (2007).
[7] E. Angelescu, O. Pavel, R. Birjega, M. Florea and R. Zavoianu, Appl. Catal. A: General 341, 50 (2008).
[8] C. Ferretti, C. Apesteguía and J. Di Cosimo, Appl. Catal. A: General 399, 146 (2011).
[9] F. Yang and M. Hannaand, Biotech. for Biofuels, 5, 1 (2012).
[10] Ö.D. Bozkurt, F.M. Tunç, N. Bağlar, S. Çelebi, İ.D. Günbaş and A. Uzun, Fuel Process. Technol. 138, 780 (2015).
[11] F. Elfghi, Chem. Eng. Res. Design, 113, 264 (2016).
[12] Z. Nagy, Chem. Eng. J., 127, 95 (2007).
[13] K. Desai, S. Survase, P. Saudagar, S. Lele and R. Singhal, Biochem. Eng. J. 41, 266 (2008).
[14] M. B. Abdul Rahman, N. Chaibakhsh, M. Basri, A. B. Salleh and R.N.Z.R. Abdul Rahman, Appl. Biochem. Biotechnol. 158, 722 (2009).
[15] K. Shahbaz, S. Baroutian, F. S. Mjalli,, M. A. Hashim and I. M. AlNashef, Chemom. Intell. Lab. Syst. 118, 193 (2012).
[16] P. Shivakumar, B.R. Srinivasa Pai and R. Shrinivasa, Appl. Energy 88, 2344 (2011).
[17] R. Chakraborty, Appl. Energy 114, 827 (2014).
[18] C.I. Rocabruno-Valdés, L.F. Ramírez-Verduzco and J.A. Hernández, Fuel 147, 9 (2015).
[19] G.R. Moradi, S. Dehghani, F. Khosravian and A. Arjmandzadeh, Renew. Energy 50, 915 (2013).
[20] N. Bálsamo, K. Sapag, M. Oliva, G. Pecchi, G. Eimer and M. Crivello, Catal. Today. 279 (2), 209 (2017).