Revista Mexicana de Ingeniería Química

A COMPARATIVE STUDY BETWEEN MULTIVARIATE CALIBRATION AND ARTIFICIAL NEURAL NETWORK IN QUANTIFICATION OF SOYBEAN BIODIESEL
UN ESTUDIO COMPARATIVO ENTRE LA CALIBRACIÓN MULTIVARIANA Y LA RED NEURAL ARTIFICIAL EN LA CUANTIFICACIÓN DEL BIODIESEL DE SOJA

C.Z. Brusamarello[1]*, M. Di Domenico[1], C. Da Silva[2], F. de Castilhos[3]

[1]*Federal University of Technology - Paraná, Francisco Beltrão 85601-970, PR- Brazil.*
[2]*The State University of Maringá, Umuarama 87506-370, PR- Brazil.*
[3]*Federal University of Santa Maria, Santa Maria 97105-900, RS-Brazil.*

Abstract

Biodiesel is an alternative fuel that can be obtained by the transesterification of vegetable oils. Spectrophotometric methods can be used for the quantification of mixtures of oils, associated with chemometric tools, enabling the analysis of overlapping signals. The aim of this work was to apply the multivariate calibration with PLS (partial least squares) and artificial neural network (ANN), to estimate the concentration of esters in the transesterification of soybean oil using molecular absorption spectrophotometry as analytical technique. The absorbance measurements were performed in a spectrophotometer UV/VIS. Synthetic solutions were prepared with standards of the five major compounds of soybean biodiesel and the real samples were obtained by the reaction of transesterification of soybean oil with two types of catalysts: NaOH and enzymatic method using Lipozyme® IM (Novozymes). According to the results, all components of the reaction medium absorb in the wavelength range of 190-280 nm. The reactions of the basic catalysis reached conversions close to 100%, whereas the enzymatic reactions reached lower conversion values. For both methods, the calibration and validation groups were composed, respectively, by the synthetic and the real samples. Results showed that the concentrations of esters estimated by the ANN model in the real samples are more accurate (R^2 of 0.93), showing the great ability of the ANN in estimating the data.
Keywords: biofuels, chromatography, enzymology, spectroscopy.

Resumen

El biodiesel es un combustible alternativo que puede obtenerse mediante la transesterificación de aceites vegetales. Se pueden utilizar métodos espectrofotométricos para la cuantificación de mezclas de oleos, asociados con herramientas quimiométricas, que permiten el análisis de señales superpuestas. El objetivo de este trabajo fue aplicar la calibración multivariable con PLS (mínimos cuadrados parciales) y la red neuronal artificial (ANN), para estimar la concentración de ésteres en la transesterificación del aceite de soja utilizando espectrofotometría de absorción molecular como técnica analítica. Las mediciones de absorbancia se realizaron en un espectrofotómetro UV/VIS. Se prepararon soluciones sintéticas con estándares de los cinco compuestos principales del biodiesel de soja y las muestras reales se obtuvieron mediante la reacción de transesterificación de aceite de soja con dos tipos de catalizadores: NaOH y método enzimático utilizando Lipozyme® IM (Novozymes). De acuerdo con los resultados, todos los componentes del medio de reacción se absorben en el rango de longitud de onda de 190-280 nm. Las reacciones de catálisis básica alcanzaron conversiones cercanas al 100%, mientras que las reacciones enzimáticas alcanzaron valores de conversión más bajos. Para ambos métodos, los grupos de calibración y validación fueron compuestos, respectivamente, por muestras sintéticas y reales. Los resultados mostraron que las concentraciones de ésteres estimadas por el modelo ANN en las muestras reales son más precisas (R^2 de 0.93), lo que muestra la gran capacidad de la ANN para estimar los dados.
Palabras clave: biocombustibles, cromatografía, enzimología, espectroscopia.

1 Introduction

Biodiesel is a non-toxic, biodegradable and renewable diesel fuel that can be used alone or in blends with petroleum diesel fuels. Biodiesel has many advantages when compared to diesel fuels: it has a higher cetane number, contains no aromatics, almost no sulphur and 10-12% oxygen by weight (Haagenson *et al*., 2014; Özener *et al*., 2014; Jiménez-Pérez *et al*., 2015).

According to ANP (Brazilian National Agency for Petroleum, Natural Gas and Biofuels), biodiesel

accepts low percentage of monoglycerides (1.00 max, % m/m), diglycerides (0.25 max, % m/m), triglycerides (0.25 max %,m/m), alcohol (0.50 max, % m/m), free glycerine (0.02 max % m/m) and total glycerine (0.38 max % m/m) (ANP 255 authorization). Such low values of these compounds demand an accurate quantitative analysis, since errors could strongly affect the fuel specification.

Accordingly, the specification of this biofuel is fundamental in ensuring its quality and specificity. The instrumental analysis such as gas chromatography, for biodiesel specification, generates peaks hindering their interpretation (Baptista *et al*., 2008). An alternative to biodiesel specification is the use of spectrophotometric methods combined with processing data by chemometric methods, in order to build the calibration models. This technique has been increasingly adopted for the determination of compounds present in multicomponent samples due to its rapidity, low cost and reduced amount of sample required for analysis (Silva *et al*. 2015). Spectroscopic techniques associated with chemometric tools appear as one of the most appropriate to evaluate the quality of fuel quickly, accurately and non-destructively. These methodologies have been used primarily combined with spectroscopic techniques for in-line or in-situ monitoring, for the non-destructive analysis of samples.

Ultraviolet-Visible (UV-Vis) spectroscopy has been used to determine the content of heptane-diluted biodiesel/diesel blends in a large concentration range. For biodiesel, the UV spectra show two intense absorption bands along the 230-260 nm wavelength (Silva *et al*. 2015). The variation in aromatic composition in the diesel may falsify the results, since they show an intense absorbance peak in this spectral region. Nevertheless, the UV-Vis spectroscopy provides easier and faster results in routine analyses, due to the reduced time and lower reagent consumption. Despite these methods, research studies continually attempt to develop alternative methods, combining low cost, fast and accurate results. Among efforts to develop a methodology capable of meeting all necessary requirements, the UV-Vis spectroscopy may be a potential tool to quantify the biodiesel content in biodiesel/diesel blends.

Shimamoto & Tubino (2016) proposed an alternative method for the quantitative analysis of biodiesel in diesel-biodiesel blends. This method is based on UV-Vis spectroscopy and applies the univariate calibration. Authors concluded that the

alliance between the UV-Vis technique and the PLS provides the quantification of biodiesel in blends adulterated with vegetable oil. The UV-Vis PLS model can recognize the difference between the biodiesel and the vegetable oil.

The multivariate calibration is effectively used in cases with the problem of overlapping analytical signals and simultaneous determinations. It is produced a model, based on all available information, that can make a connection between the entire analytical signal and the ownership interest (concentration in many cases). Works using this technique to predict biodiesel constituents can be found in the literature (Vakh *et al*., 2014).

Another method that has found increasing applications for multicomponent determination, in the last few years, is the artificial neural network (ANN). This technique is capable of handling incomplete data and can deal with nonlinear problems. Once trained, the ANN model can perform predictions and generalizations at high speeds (Rajendra *et al*., 2009). The ANN model approach is effective for spectrophotometric analysis because it can be used to solve analytical problems. Several comparative studies on these two techniques have been conducted using various data sets (Jamrógiewicz, 2012; Wu *et al*., 2013).

Rocabruno-Valdés *et al*. (2015) developed ANN models to predict the density, dynamic viscosity and the cetane number of biodiesel, using temperature, composition of methyl esters, number of carbon atoms and number of hydrogen atoms as input variables. Correlation coefficients of 0.9195-0.9940 were obtained by comparing the experimental and calculated values, while a mean squared error (MSE) of 1.842×10^3 was obtained in the validation stage.

Lopes *et al*., (2019) applied four methods for direct classification (decision tree classifier, K-nearest neighbors, support vector machine and ANN) to optimize and compare the biodiesel samples according to their compliance to viscosity, density, oxidative stability and iodine value, having as input the composition of fatty acid methyl esters. A comparison between these methods of direct classification and empirical equations distinguished positively the direct classification methods in the problem addressed.

Hosseini & Pierantozzi (2019) studied the molecular thermodynamics and the ANN modeling of surface tensions of several fatty acid esters and biodiesels. The selected ANN architecture was a two hidden layer network with nine neurons each, with three input parameters. The model confirmed

its validity for the surface tension of fatty acids and biodiesel, giving an excellent result in the database of the pure fluids containing 137 points.

Thus, the aim of this paper was to apply the PLS and the ANN models to estimate the concentration of esters in the transesterification of soybean oil, using molecular absorption spectrophotometry as analytical technique. Synthetic solutions were prepared with standards of the five major compounds of soybean biodiesel and real samples were obtained by the reaction of transesterification of soybean oil with two different catalysts: NaOH and enzymatic method with Lipozyme® IM (Novozymes). The lack of data in the literature related to enzymatic biodiesel samples justifies the purpose of this work.

2 Materials and methods

2.1 Chemicals

The synthetic samples for the gas chromatographic (GC) analyzes were prepared from the authentic standards of ethyl esters of soybean biodiesel (ethyl palmitate, ethyl linoleate, ethyl linolenate, ethyl oleate, ethyl stearate) and methyl palmitate as internal standard. All reactants were purchased from Sigma Aldrich.

The real samples were obtained from the transesterification reactions, using as substrates the commercial soybean oil (Soya), without any prior treatment, and ethanol (Merck 99.9% purity). In the alkaline catalysis reactions, it was used sodium hydroxide (Quimex, 97% purity) as catalyst, along with sulfuric acid (Quimex, 10%), sodium chloride (Quimex), heptane (Quimex, 99.9% purity) and sodium sulfate anhydrous (Nuclear). In addition, for the enzymatic catalysis method, it was used Lipozyme® IM (Novozymes) as catalyst and n-hexane as solvent (Merck 99.9% purity). Finally, the samples for the uptake measurement were prepared using heptane (Merck 99.9% purity) as solvent and also ethyl alcohol (Merck 99.9% purity).

2.2 Sample preparation

Synthetic samples were prepared using authentic standards of ethyl esters of soybean from biodiesel in a solution of n-heptane and ethyl alcohol in the volumetric ratio of 1:1. The solutions were prepared at different concentrations of ethyl esters, and different proportions between them, as presented in Table 1.

According to Ndiaye *et al.* (2006) the fatty acid content of a commercial soybean oil is 11.2% (± 0.08) palmitic acid, 3.45% (± 0.01) stearic acid, 23.41% (± 0.27) oleic acid, 54.21% (± 0.30) linoleic acid and 6.75% (± 0.01) linolenic acid.

The real samples were obtained by the transesterification reaction of soybean oil in ethanol as described in detail by Oliveira *et al.* (2004). For the alkali catalyst, the reaction temperature was held at 70 °C, with an oil:ethanol molar ratio of 1:9 and 0.5% (w/w) of NaOH, during 20 min (reaction time). For the enzymatic catalyst, the reaction temperature was held at 35 °C, with an oil:ethanol molar ratio of 1:3, 10% (w/w) of distilled water, 20% (w/w) of enzyme and a ratio of solvent (*n*-hexane) to oil of 1:40 (w/w), during 8 h (reaction time).

2.3 Measurements

Absorbance measurements were performed in a spectrophotometer (Agilent Model 8453), in the wavelength region between 190-800 nm, using a quartz cuvette (Helmann) of 3.5 ml and the moisture of heptane/ethanol of 1:1 (v/v) as blank solution.

The original concentration of the samples was determined by gas chromatography as described by Silva *et al.* (2007).

Table 1. Proportions in mass of ethyl esters used in the preparation of synthetic solutions.

Ester	Fraction-1 (%) w/w	Fraction-2 (%) w/w	Fraction-3 (%) w/w	Fraction-4 (%) w/w
ethyl palmitate	10	10	10	10
ethyl stearate	6	6	6	5
ethyl oleate	20	23	23	20
ethyl linolenate	9	11	9	7
ethyl linoleate	55	50	52	58

Table 2. Samples of calibration and validation for the three compositions used in this work.

Setting	Calibration	Number of Calibration Sample	Validation	Number of Validation Sample
1	S1 to S26	26	R1 to R22	22
2	S1 to S26;R6;R10;R14; R15;R17;R18;R21	33	R1 to R5; R7 to R9;R12;R13;R16; R19;R20;R22	15
3	S1 to S26;R6;R10;R14; R15;R17;R18;R20;R21	34	R1 to R5; R7 to R9;R12;R13; R16;R19;R22	14

R - Real sample; S - Synthetic sample

Samples for the quantification of ethyl esters of soybean biodiesel were prepared at concentrations of 2000 mg.L^{-1}, using methyl palmitate as internal standard at a concentration of 500 mg.L^{-1}. The concentration of the original samples obtained by gas chromatography varied between 21.70-100 g.mL^{-1}.

2.4 Chemometric and ANN models

The ANN model approach used in this work was a MultiLayer Perceptron type with a bias and an inner layer with four inputs and one output. After several preliminary tests, the absorbance measurements of wavelengths at 202, 203, 204, 211 nm were used as inputs, being the ester content used as output of the reaction. The hyperbolic tangent activation function was applied in both inner and output layers. The input and output data were normalized within the range of [-1,1]. Also, the range of 3 to 18 internal units was investigated in order to define the best neural network architecture.

From the conventional chemometric tool, it was made a principal component analysis (PCA) with the software Unscrambler 6.11, where it was investigated the range of up to 15 principal components from the calibration group. After choosing the number of principal components, a model PLS1 (one output) was developed from the calibration group. Finally, the model validation was performed with the validation group.

For both models applied in this work, it was used the first derivative method. The derivative spectrophotometry is an analytical technique of great utility for extracting both qualitative and quantitative information from spectra composed of unresolved bands. For a single-peak spectrum, the first derivative is a plot of the absorption gradient $dA/d\lambda$ as a function of the wavelength and features a maximum and a minimum in the vertical distance between these, which is the amplitude. The parameter amplitude is proportional to the analyte concentration; theoretically, $dA/d\lambda$ is zero at λ_{max} for the band in the normal spectrum (Mark and Workman Jr, 2003).

The Root Mean Square Error of Prediction (RMSEP) from an independent test set was used as the measurement of the ANN and PLS models performance, as described by Eq. 1:

$$RMSEP = \sqrt{\frac{\sum(Y_{PRE} - Y_{REF})^2}{N_{TEST}}} \tag{1}$$

where N_{test} is the number of samples in the test set, Y_{ref} is the actual value and Y_{pre} is the predicted value provided by the ANN and PLS models. All computations were performed with Fortran 90 Software.

Validation and calibration groups were built by means of the preparation of 48 samples, consisting by real and synthetic samples, as presented in Table 2. The detailed compositions of samples can be find at the appendix, in Tables 5 and 6.

3 Results and discussion

3.1 Absorption spectrum

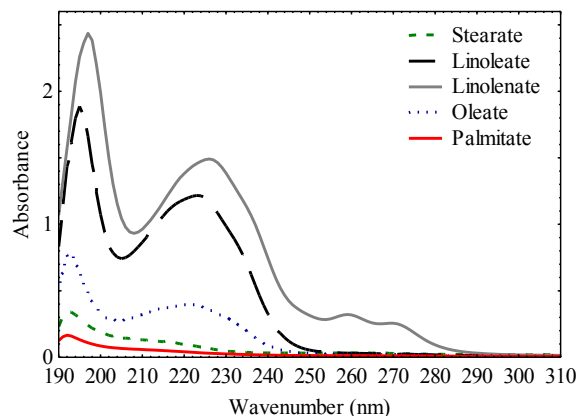Fig. 1 depicts the absorption spectrum of soybean ethyl esters (biodiesel).

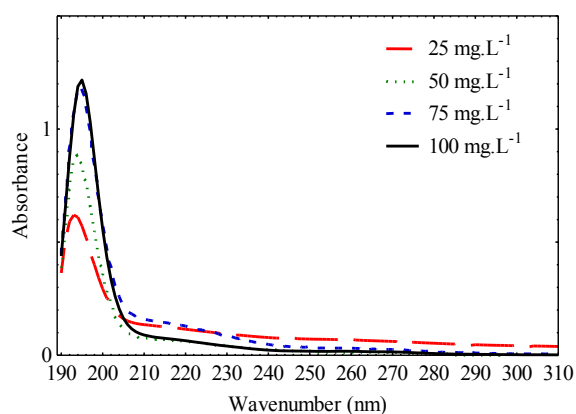Fig. 1. Absorption spectra of soybean ethyl esters.

Fig. 2. Absorption spectra of synthetic solutions in the Fraction-1 of components at different concentrations.

According to this figure, all components of the biodiesel absorb in the same wavelength region. The components of this reaction medium absorb in the range of 190-280 nm, as pointed out in the work of Dantas *et al.* (2011). This result indicates the difficulty in interpreting the obtained data due to the overlap of peaks. Such behavior justifies the use of chemometric tools for the analysis of soybean oil transesterification.

Fig. 2 report the absorption spectrum of the synthetic solutions at different concentrations (25-100 mg.L^{-1}) using Fraction-1 of components, previously given in Table 1. It is visible that the solutions absorbance is proportional to the concentration, i.e., the higher the concentration, more of a particular wavelength is absorbed.

Fig. 3 depicts the spectrum of synthetic solutions for the concentration of 75 mg.L^{-1} in the four prepared fractions, previously given by Table 1. This concentration was selected due to the absorption

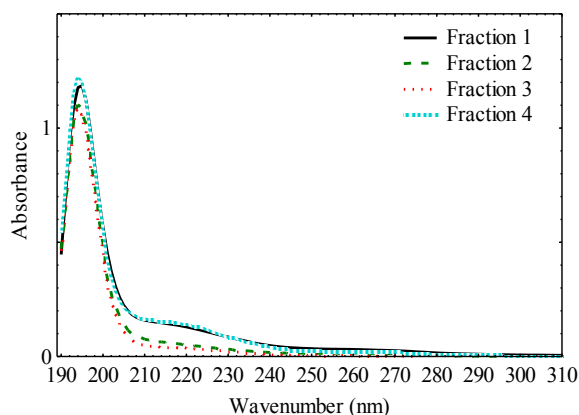UV/VIS signal proximity with the concentration of

Fig. 3. Absorption spectrum of synthetic solutions at a concentration of 75 mg.L^{-1} in the four prepared fractions.

100 mg.L^{-1}. These results demonstrate that different fractions in synthetic solutions result in insignificant changes in the absorption spectrum.

Fig. 4 reports the absorption spectrum of the real samples, obtained by the transesterification process with (a) the basic catalyst and (b) the enzymatic catalyst, at different reaction times. Firstly, it can be seen a considerable difference between the reaction times by using the two catalysts. This difference can be related to the higher conversion achieved with the basic catalyst: according to the gas chromatography analysis, after only 2 min of reaction, 93% of the ethyl esters were detected for this sample.

According to Fig. 4a, the absorption curves for the reaction with the basic catalyst are quite similar, even at different reaction times. The small difference between the curves is due to the high achieved yield, as previously described. In addition, according to Fig. 4b, for the reactions with the enzymatic catalyst, despite the low conversions for shorter reaction times, the absorbance curves did not show major differences.

3.2 *Performance of PLS and ANN method*

In Tables 3 and 4 are reported, respectively, the RMSEP values for the PLS and ANN models, with the normal range and the derivative spectrum. It seems that using the derivative spectrum, there is a substantial improvement in finding a more reliable model, for both PLS and ANN methodologies. As the RMSEP value approaches 0 (zero), there is an increase in the reliability of the model in predicting the concentrations of the samples.

Table 3. RMSEP values in the used settings with PLS method.

RMSEP	Setting 1	Setting 2	Setting 3
Normal spectrum	33.55	31.75	16.37
Derivative spectrum	31.84	11.17	9.86

Table 4. RMSEP values used in the settings with ANN method.

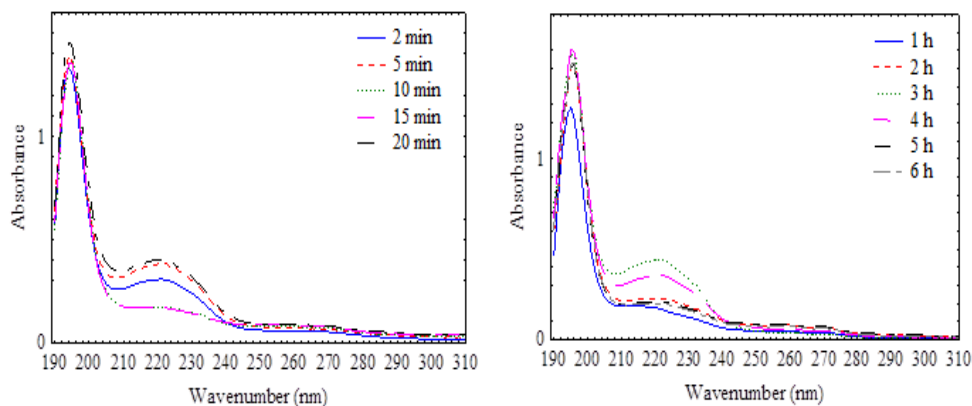| RMSEP	Setting 1	Setting 2	Setting 3
Normal spectrum	29.92	10.65	13.54
Derivative spectrum	13.88	9.13	8.77

Fig. 4. (a) Absorption spectrum of real samples obtained by transesterification process with basic catalyst at different reaction times. (b) Absorption spectrum of real samples obtained by transesterification process with enzymatic catalyst at different reaction times.
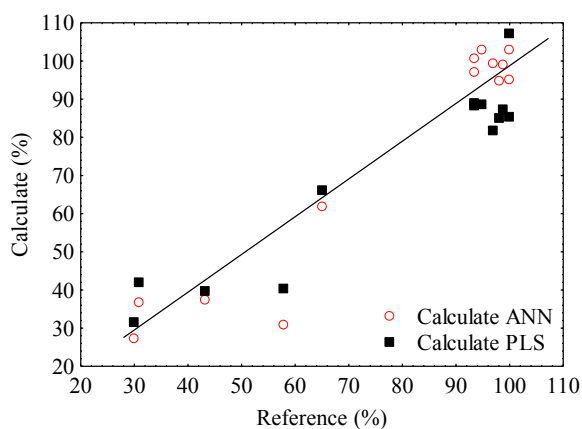

Fig. 5. ANN versus PLS method.

An essential point related to the different compositions of the calibration and validation groups is the number of samples used for analysis. The higher the number of samples used, the more representative are the results. The difficulty to achieve better results can be related to the lower conversions of the real samples. In addition, loss of monoglycerides, diglycerides and even esters may have occurred in the samples during their handling.

Fig. 5 reports a comparison between the two methods used in this research to predict the concentrations of biodiesel. It is compared the best sets of validation of both methods, Setting 3, with derivative spectrum. This spectrum can be used to reduce the analysis discrepancies, to resolve overlapping bands in qualitative analysis, and most importantly, to reduce the effects of interference from scattering, matrix or other absorbing compounds in quantitative analysis.

Therefore, another important result is that the best composition group is Setting 3, for both applied models.

The best obtained RMSEP results were 8.77 for the ANN method and 9.86 for the PLS method, which are close values. The numerical difference is not apparent; however, according to the predictions showed in Fig. 5, there are differences in the predictions for each model, particularly for the enzymatic biodiesel samples. Also, differences are especially found at lower conversion samples. The PLS method found difficulties to predict the samples close to the real values; in this case, the ANN model presented better results. Another characteristic is that the ANN has the ability to solve nonlinear problems; thus, better adjustments can be achieved with this model.

The results in this research reported data using enzymatic samples, and, for this reason, it is difficult to perform a comparison with other works found in the literature. Although, important discussions can be analyzed, especially about the RMSEP values achieved by other authors (Richard *et al.*, 2013) where the RMSEP obtained was lower than 5%. Also, a different spectroscopy method was used and concentrations of biodiesel above 60% of conversion were used. It is clear that the number of samples used in the calibration tests is very important to achieve good validation results. It is possible to see in Fig. 5 that the samples above 60% of conversion had good predictions by using the ANN and PLS models.

Conclusions

In this research it was found that the composition and the number of samples in the calibration group are important in the performance of the studied models, since the calibration group contains the information in which the method use to "learn".

The derivative spectrum was more appropriate than the normal spectrum for both ANN and PLS models. One reason is that the derivative spectrum provides more "rich" information through the curve slope, not affected by the baseline.

The multivariate calibration introduced in this research showed to be a good tool for the feasibility analysis with low cost and short time. It was noted the importance and the difficulty in the preparation of synthetic and real samples, since small experimental errors result in accumulation of estimate errors in the models.

The ANN model shows to be better in estimating the ester content in the biodiesel samples, with a R^2 of 0.93, in comparison to the R^2 of 0.91 for the PLS method. In addition, this research was the first attempt to use data from real and enzymatic samples of biodiesel for applications of chemometric tools and artificial neural networks.

In addition, this research was the first attempt to combine data from real and enzymatic samples of biodiesel, applying chemometric tools and artificial neural networks for their prediction.

Nomenclature

ANN	Artificial Neural Network
PLS	Partial Least Squares
RMSEP	Root Mean Square Error of Prediction
N_{test}	Number of samples in the test set
Y_{ref}	The actual value
Y_{pre}	The predicted value provided by ANN and PLS models

References

Balabin, R. M., Lomakina, E. I., & Safieva, R. Z. (2011). Neural network (ANN) approach to biodiesel analysis: Analysis of biodiesel density, kinematic viscosity, methanol and water contents using near infrared (NIR) spectroscopy. *Fuel 90*, 2007-2015. https://doi.org/10.1016/j.fuel.2010.11.038

Balabin, R. M., Safieva, R. Z., & Lomakina, E. I. (2010). Gasoline classification using near infrared (NIR) spectroscopy data: Comparison of multivariate techniques. *Analytica Chimica Acta 671*, 27-35. https://doi.org/10.1016/j.aca.2010.05.013

Baptista, P., Felizardo, P., Menezes, J. C., & Neiva Correia, M. J. (2008). Multivariate near infrared spectroscopy models for predicting the iodine value, CFPP, kinematic viscosity at 40 °C and density at 15 °C of biodiesel. *Talanta 77*, 144-151. https://doi.org/10.1016/j.talanta.2008.06.001

Dantas, M. B., Albuquerque, A. R., Barros, A. K., Rodrigues Filho, M. G., Antoniosi Filho, N. R., Sinfrônio, F. S. M., ... Souza, A. G. (2011). Evaluation of the oxidative stability of corn biodiesel. *Fuel 90*, 773-778. https://doi.org/10.1016/j.fuel.2010.09.014

de Fátima Bezerra de Lira, L., de Vasconcelos, F. V. C., Pereira, C. F., Paim, a. P. S., Stragevitch, L., & Pimentel, M. F. (2010). Prediction of properties of diesel/biodiesel blends by infrared spectroscopy and multivariate calibration. *Fuel 89*, 405-409. https://doi.org/10.1016/j.fuel.2009.05.028

Haagenson, D. M., Perleberg, J. R., & Wiesenborn, D. P. (2014). Fractionation of canola biodiesel sediment for quantification of steryl glucosides with HPLC/ELSD. *JAOCS, Journal of the American Oil Chemists? Society 91*, 497-502. https://doi.org/10.1007/s11746-013-2386-x

Hoekman, S. K., Broch, A., Robbins, C., Ceniceros, E., & Natarajan, M. (2012). Review of biodiesel composition, properties, and specifications. *Renewable and Sustainable Energy Reviews 16*, 143-169. https://doi.org/10.1016/j.rser.2011.07.143

Hosseini, S. M., & Pierantozzi, M. (2019). Molecular thermodynamic modeling of surface tensions of some fatty acid esters and biodiesels. *Journal of Molecular Liquids 281*, 431-443. https://doi.org/10.1016/j.molliq.2019.02.105

Jamrógiewicz, M. (2012). Application of the near-infrared spectroscopy in the pharmaceutical technology. *Journal of Pharmaceutical and Biomedical Analysis 66*, 1-10. https://doi.org/10.1016/j.jpba.2012.03.009

Jiménez-Pérez, J. L., López-Gamboa, G., Cruz-Orea, A., & Correa-Pacheco, Z. N. (2015). Thermal parameters study of biodiesel containing au nanoparticles using photothermal techniques. *Revista Mexicana de Ingeniería Química 14*, 481-487.

Killner, M. H. M., Rohwedder, J. J. R., & Pasquini, C. (2011). A PLS regression model using NIR spectroscopy for on-line monitoring of the biodiesel production reaction. *Fuel 90*, 3268-3273. https://doi.org/10.1016/j.fuel.2011.06.025

Lopes, M. V., Barradas Filho, A. O., Barros, A. K., Viegas, I. M. A., Silva, L. C. O., Marques, E. P., & Marques, A. L. B. (2019). Attesting compliance of biodiesel quality using composition data and classification methods.

Neural Computing and Applications 31, 539-551. https://doi.org/10.1007/s00521-017-3087-4

Mark, H., & Workman Jr, J. (2003). Derivatives in Spectroscopy - Part III. *Spectroscopy 18*, 106-111.

Ndiaye, P. M., Franceschi, E., Oliveira, D., Dariva, C., Tavares, F. W., & Oliveira, J. V. (2006). Phase behavior of soybean oil, castor oil and their fatty acid ethyl esters in carbon dioxide at high pressures. *Journal of Supercritical Fluids 37*, 29-37. https://doi.org/10.1016/j.supflu.2005.08.002

Oliveira, D., Di Luccio, M., Faccio, C., Rosa, C. D., Bender, J. P., Lipke, N., ... de Oliveira, J. V. (2004). Optimization of enzymatic production of biodiesel from castor oil in organic solvent medium. *Applied Biochemistry and Biotechnology 113-116*, 771-780. https://doi.org/10.1385/ABAB:115:1-3:0771

Özener, O., Yüksek, L., Ergenç, A. T., & Özkan, M. (2014). Effects of soybean biodiesel on a DI diesel engine performance, emission and combustion characteristics. *Fuel 115*, 875-883. https://doi.org/10.1016/j.fuel.2012.10.081

Rajendra, M., Jena, P. C., & Raheman, H. (2009). Prediction of optimized pretreatment process parameters for biodiesel production using ANN and GA. *Fuel 88*, 868-875. https://doi.org/10.1016/j.fuel.2008.12.008

Richard, R., Dubreuil, B., Thiebaud-Roux, S., & Prat, L. (2013). On-line monitoring of the transesterification reaction carried out in microreactors using near infrared spectroscopy. *Fuel 104*, 318-325. https://doi.org/10.1016/j.fuel.2012.07.054

Rocabruno-Valdés, C. I., Ramírez-Verduzco, L. F., & Hernández, J. A. (2015). Artificial neural network models to predict density, dynamic viscosity, and cetane number of biodiesel. *Fuel 147*, 9-17. https://doi.org/10.1016/j.fuel.2015.01.024

Rojas, F. S., Ojeda, C. B., & Pavon, J. M. (1988). Derivative ultraviolet-visible region absorption spectrophotometry and its analytical applications. *Talanta 35*, 753-761. https://doi.org/10.1016/0039-9140(88)80179-6

Shimamoto, G. G., & Tubino, M. (2016). Alternative methods to quantify biodiesel in standard diesel-biodiesel blends and samples adulterated with vegetable oil through UV-Visible spectroscopy. *Fuel 186*, 199-203. https://doi.org/10.1016/j.fuel.2016.08.076

Silva, C., Weschenfelder, T. a., Rovani, S., Corazza, F. C., Corazza, M. L., Dariva, C., & Oliveira, J. V. (2007). Continuous production of fatty acid ethyl esters from soybean oil in compressed ethanol. *Industrial & Engineering Chemistry Research 46*, 5304-5309. https://doi.org/10.1021/ie070310r

Silva, M. A. A., Correa, R. a, Tavares, M. G. D. O., & Antoniosi Filho, N. R. (2015). A new spectrophotometric method for determination of biodiesel content in biodiesel/diesel blends. *Fuel 143*, 16-20. https://doi.org/http://dx.doi.org/10.1016/j.fuel.2014.10.048

Vakh, C. S., Bulatov, A. V, Shishov, A. Y., Zabrodin, A. V, & Moskvin, L. N. (2014). Determination of silicon, phosphorus, iron and aluminum in biodiesel by multicommutated stepwise injection analysis with ?lassical least squares method. *Fuel 135*, 198-204. https://doi.org/http://dx.doi.org/10.1016/j.fuel.2014.06.059

Wu, J., Luo, W., Wang, X., Cheng, Q., Sun, C., & Li, H. (2013). A new application of WT-ANN method to control the preparation process of metformin hydrochloride tablets by near infrared spectroscopy compared to PLS. *Journal of Pharmaceutical and Biomedical Analysis 80*, 186-191. https://doi.org/http://dx.doi.org/10.1016/j.jpba.2013.03.004

Zagonel, G. (2004). Multivariate monitoring of soybean oil ethanolysis by FTIR. *Talanta 63*, 1021-1025. https://doi.org/10.1016/j.talanta.2004.01.008

Zhang, W. B. (2012). Review on analysis of biodiesel with infrared spectroscopy. *Renewable and Sustainable Energy Reviews 16*, 6048-6058. https://doi.org/10.1016/j.rser.2012.07.003

Appendix

Table 5. Concentration of real samples obtained by gas chromatography.

Sample	Concentration (mg/L)
R1	30
R2	65
R3	100
R4	93.41
R5	94.62
R6	100
R7	100
R8	96.87
R9	93.38
R10	96.23
R12	98.16
R13	98.75
R14	94.88
R15	36.74
R16	43.15
R17	36.15
R18	56.65
R19	57.93
R20	30.94
R21	21.7
R22	35.63

Table 6. Concentration of synthetic samples.

Sample	Concentration (mg/L)
S1	10
S2	15
S3	20
S4	25
S5	35
S6	40
S7	50
S8	55
S9	60
S10	70
S11	75
S12	80
S13	85
S14	90
S15	95
S16	105
S17	110
S18	115
S19	75
S20	75
S21	75
S22	75
S23	100
S24	100
S25	100
S26	100