Liquid–Liquid Equilibrium Data for Reactional Systems of Ethanolysis at 298.3 K

Marcelo Lanza,[†] Waldomiro Borges Neto,[‡] Eduardo Batista,[†] Ronei J. Poppi,[‡] and Antonio J. A. Meirelles*,[†]

ExTrAE, Department of Food Engineering, Faculty of Food Engineering, P.O. Box 6121, State University of Campinas – UNICAMP, 13083-862, Campinas, SP, Brazil, and LAQQA, Department of Analytical Chemistry, Chemistry Institute, P.O. Box 6154, State University of Campinas – UNICAMP, 13083-862, Campinas, SP, Brazil

The objective of this work was to determine liquid–liquid equilibrium data at (298.3 ± 0.2) K for systems of interest in biodiesel production, such systems being composed of vegetable oils + anhydrous ethanol + hexane. The measurements were performed using near-infrared spectroscopy (NIR) for quantification of the phases in equilibrium. The following vegetables oils were investigated: pretreated cottonseed oil, corn oil, canola oil, refined soybean oil, and degummed soybean oil. Initially, with the purpose of validating the NIR methodology, equilibrium data for the system composed of pretreated cottonseed oil (neutral) + commercial linoleic acid + anhydrous ethanol were determined at (298.2 ± 0.1) K by a conventional method based on acid–base titration and solvent evaporation and by the NIR method. The relative errors between the phase compositions determined by both analytical methods were 2.1 % for the alcoholic phase and 2.0 % for the oil phase. The mass balance errors for all investigated systems varied in the range from 0.03 % to 0.16 %, which indicates the good quality of the experimental data and the good performance of the NIR method. The experimental data were correlated using the NRTL model with an average global deviation of 0.70 %.

Introduction

Brazil is a country with a considerable production of commodities from biomass sources such as ethanol and vegetable oils. Recently, research on edible oil deacidification by liquid–liquid extraction[1–3] and on oil reactions using alternative solvents[4–7] has highlighted the importance of an appropriate solvent selection for the good performance of such processes. The conventional production of biodiesel by the alcoholic transesterification of vegetable oils with methanol or ethanol requires acid or base catalysts, leading to the formation of an undesirable byproduct and a loss of yield.[8,9]

On the other hand, the transesterification can be catalyzed by enzymes,[10,11] but such processes normally require the use of organic media in which the substrates (oil and ethanol) exhibit relatively high diffusion coefficients.[6,7] A typical system for biodiesel production using an enzymatic catalyst is composed of oil + ethanol + hexane.[6] Due to the limited mutual solubility of oil and ethanol, knowledge of the phase behavior for systems containing solvents is of the utmost importance for the reactive process because this can influence the reaction path as well as the subsequent purification steps of the biodiesel production. For these reasons, it is important to predict the occurrence of one or more liquid phases in the reaction environment and also to be able to estimate the compositions of the phases in equilibrium, as shown in the research of Zhou et al.[12] and Čerče et al.[13]

The scope of the present paper was to determine the liquid–liquid equilibrium data for fatty systems of interest in biodiesel production, using near-infrared spectroscopy (NIR) coupled with chemometric analysis as the analytical method for quantification of the components in each phase. Near-infrared spectroscopy supplies qualitative data about the samples that can be transformed into quantitative information using chemometric methods based on statistical and mathematical procedures.[14] Reports on the use of NIR spectroscopy in the analysis of vegetable oils can be found in the literature,[14–17] but no prior research was found using such a technique in the determination of liquid–liquid equilibrium data for fatty systems. For this reason, the first part of the present work involves the validation of this analytical methodology, measuring the equilibrium data for a similar system (cottonseed oil + linoleic acid + ethanol) by the new technique and by a conventional method, and the second part reports the determination of equilibrium data for systems of interest in the production of biodiesel.

Experimental Section

Materials. The solvents used in this work were anhydrous ethanol from Merck (Germany), with a purity of 99.9 %, and hexane, also from Merck, with purity greater than 96 %. Refined corn oil, refined canola oil, and refined soybean oil were from Cargill (Mairinque/SP, Brazil). Degummed soybean oil was kindly supplied by Cargill (Mairinque/SP, Brazil), having been submitted to the degumming process at the refinery after the oil extraction and solvent stripping processes. The pretreated cottonseed oil (neutral oil) was also submitted to a prior treatment (deacidification) in the refinery (kindly supplied by Maeda, Itumbiara/GO, Brazil), both being qualified as semi-processed oils. The other fatty reagent used in this study was commercial linoleic acid (Fluka, Steinheim, Germany).

All fatty reagents used in this work were analyzed by gas chromatography of the fatty acid methyl esters to determine the fatty acid composition, according to the official method (1-62) of the AOCS.[18] Prior to the chromatographic analysis, the fatty samples were prepared in the form of fatty acid methyl esters according to the official method (2-66) of the AOCS.[19] The refined soybean oil, degummed soybean oil, canola oil, and corn oil were submitted to the chromatographic analysis using

---

* Corresponding author. Fax: +55 19 3521 4027. E-mail: tomze@fea.unicamp.br.

[†] Laboratory of Extraction, Applied Thermodynamics and Equilibrium, Faculty of Food Engineering.

[‡] Chemometrics Laboratory in Analytical Chemistry, Chemistry Institute.

10.1021/je700120d CCC: $40.75 © 2008 American Chemical Society
Published on Web 10/20/2007

Table 1. Fatty Acid Compositions of the Vegetable Oils

									tie line 1		tie line 2	
			$\overline{M}$	pretreated cotton[b]	refined corn	refined canola	degum. soybean	refined soybean	AP[c]	OP[c]	AP[c]	OP[c]
fatty acids	symbol	C*x*:*y*[a]	g·mol^{-1}	100 *w*	100 *w*	100 *w*	100 *w*	100 *w*	100 *w*	100 *w*	100 *w*	100 *w*
myristic	M	C14:0[a]	228.38	0.75		0.07	0.08	0.09	0.21	0.09	0.21	0.09
palmitic	P	C16:0	256.43	22.79	12.71	4.72	11.21	11.18	19.41	11.35	19.64	11.64
palmitoleic	Po	C16:1	254.42	0.93	0.17	0.25	0.09	0.09	0.15	0.09	0.15	0.12
stearic	S	C18:0	284.49	2.35	2.25	2.55	3.61	4.13	6.87	4.35	6.92	4.46
oleic	O	C18:1	282.47	16.04	34.96	62.56	23.63	25.62	37.75	26.43	37.98	26.92
linoleic	Li	C18:2	280.45	56.41	47.61	20.13	53.42	50.88	31.55	49.93	31.08	49.25
linoelaidic[d]		C18:2T[d]			0.14			0.40	0.28	0.41	0.28	0.40
linolenic	Le	C18:3	278.44	0.16	0.83	7.08	6.77	4.97	1.28	4.71	1.23	4.50
translinolenic[d]		C18:3T[d]				0.41		1.15	0.33	1.11	0.32	1.06
arachidic	A	C20:0	312.54	0.26	0.54	0.62	0.36	0.45	0.70	0.47	0.71	0.48
gadoleic	Ga	C20:1	310.52	0.12	0.38	1.04	0.24	0.31	0.33	0.25	0.35	0.25
behenic	Be	C22:0	340.59	0.19	0.19	0.28	0.44	0.54	0.85	0.59	0.85	0.61
lignoceric	Lg	C24:0	368.65		0.22	0.17	0.15	0.19	0.29	0.22	0.28	0.22
nervonic	Ne	C24:1	366.63			0.12						

[a] C*x*:*y*, *x* = number of carbons and *y* = number of double bonds. [b] Composition taken from Rodrigues et al.[20] [c] Oil residues of the alcoholic (AP) and oil (OP) phases of tie lines 1 and 2 for the refined soybean oil + anhydrous ethanol + hexane system (see Tables 2 and 7). [d] Trans isomers.

Table 2. Probable Triacylglycerol Compositions of the Vegetable Oils

								tie line 1		tie line 2	
		$\overline{M}$	pretreated cotton	refined corn	refined canola	degum. soybean	refined soybean	AP[c]	OP[c]	AP[c]	OP[c]
main TAG[a]	group	g·mol^{-1}	100 *w*	100 *w*	100 *w*	100 *w*	100 *w*	100 *w*	100 *w*	100 *w*	100 *w*
MLiP	48:2[b]	803.29	0.73								
POP	50:1	833.36	2.90	1.83	0.50	0.94	1.01	5.15	1.08	5.32	1.16
PLiP	50:2	831.34	10.44	2.56		2.17	2.08	4.54	2.10	4.60	2.20
PPoLi	50:3	829.33	1.42								
MLiLi	50:4	827.31	0.92								
POS	52:1	861.41	0.58	0.62	0.50	0.57	0.71	3.41	0.77	3.51	0.84
POO	52:2	859.40	3.98	5.95	6.27	3.43	3.91	12.72	4.19	12.99	4.43
POLi	52:3	857.38	14.39	14.13	4.33	9.94	10.30	16.92	10.60	16.96	10.96
POLe	52:4	855.36			2.32						
PLiLi	52:4	855.36	25.66	10.06		12.39	11.44	8.02	11.23	7.87	11.26
PLiLe	52:5	853.35		0.47	0.59	2.91	2.52	0.77	2.39	0.74	2.34
PoLiLi	52:5	853.35	0.98								
SOS	54:1	889.46					0.20	0.91	0.22	0.94	0.24
PLiA	54:2	887.45	0.53								
SOO	54:2	887.45		1.20	3.33	1.00	1.31	4.15	1.43	4.23	1.51
SOLi	54:3	885.43	1.84			4.35	5.27	10.46	5.65	10.53	5.87
OOO	54:3	885.43		6.65	27.17						
OOLi	54:4	883.42	6.23	18.88	25.43	12.53	13.79	14.79	14.31	14.65	14.52
OLiLi	54:5	881.40	13.47	24.03		22.03	21.86	11.17	21.76	10.86	21.49
OOLe	54:5	881.40			17.21						
OLiLe	54:6	879.38			6.78						
LiLiLi	54:6	879.38	15.93	11.66		20.35	18.11	3.96	17.23	3.77	16.48
LiLiLe	54:7	877.37		0.57		6.13	5.07	0.47	4.65	0.44	4.31
OLeLe	54:7	877.37			2.05						
LiLeLe	54:8	875.35				0.74	0.57	0.02	0.51	0.02	0.46
OOA	56:2	915.50			0.84		0.32	0.77	0.35	0.78	0.36
OOGa	56:3	913.49			1.67						
OLiA	56:3	913.49		0.67			0.48	0.71	0.48	0.72	0.50
OLiGa	56:4	911.47		0.72	1.01						
LiLiA	56:4	911.47				0.52	0.63	0.45	0.59	0.46	0.59
OLiBe	58:3	941.54					0.42	0.61	0.46	0.61	0.48

[a] Groups with a total triacylglycerol (TAG) composition lower than 0.5 % were ignored, except for refined soybean oil and oil residues of tie lines 1 and 2. [b] *x*:*y*, *x* = number of carbons (except carbons of glycerol), *y* = number of double bonds. [c] Oil residues of the alcoholic (AP) and oil (OP) phases of tie lines 1 and 2 for the refined soybean oil + anhydrous ethanol + hexane system (see Tables 1 and 7).

a CGC Agilent 6850 Series GC capillary gas chromatograph system under the following experimental conditions: DB-23 Agilent capillary column (50 % cyanopropyl–methylpolysiloxane), 0.25 μm, 60 m × 0.25 mm i.d.; helium as the carrier gas at a rate of 1.0 mL·min^{-1}; linear velocity of 24 cm·s^{-1}; injection temperature of 523.2 K; column temperature of 383.2 K for 5 min, (383.2 to 488.2) K (rate of 5 K·min^{-1}), 488.2 K for 24 min; detection temperature of 553.2 K; and injection volume of 1.0 μL. The fatty acid methyl esters were identified by comparison with external standards purchased from Nu Check Prep (Elysian/MN, USA). Quantification was accomplished by internal normalization.

The fatty acid compositions of the pretreated cottonseed oil and the commercial linoleic acid were obtained from Rodrigues et al.[20] because these reagents were the same as those used in that work. The corresponding chromatographic conditions can be found in that paper. The fatty acid compositions of the vegetable oils are presented in Table 1, from which the probable

Table 3. Errors in the Quantification of the Prediction Standards for the System Containing Pretreated Cottonseed Oil (1) + Commercial Linoleic Acid (6) + Anhydrous Ethanol (7) at (298.2 ± 0.1) K Using the Conventional Analytical Method and the NIR Method

		alcoholic phase			oil phase			
method		oil (1)	acid (6)	ethanol (7)	oil (1)	acid (6)	ethanol (7)	E_R/%
conventional	E_R/%	1.60	0.60	1.59	1.98	5.98	1.07	2.14
NIR spectroscopy	E_R/%	1.91	0.27	1.73	0.74	2.21	1.05	1.32
	rmsep/%	0.24	0.13	0.23	0.60	0.15	0.66	-
	rmsecv/%	0.27	0.36	0.10	0.45	0.52	0.09	-
	R^2	0.99976	0.99990	0.99891	0.99857	0.99700	0.99824	-

Table 4. Quantification of the Mass Fractions of the Phases for the System Containing Pretreated Cottonseed Oil (1) + Commercial Linoleic Acid (6) + Anhydrous Ethanol (7) at (298.2 ± 0.1) K Using the Conventional Analytical Method and the NIR Method

	overall composition			conventional method			NIR spectroscopy method			
	100 w_1	100 w_6	100 w_7	100 w_1	100 w_6	100 w_7	100 w_1	100 w_6	100 w_7	$E_R{}^a$/%
alcoholic phase	49.65	0.00	50.35	6.69	0.00	93.31	6.43	0.00	93.57	1.39
	47.93	2.00	50.07	7.86	2.42	89.72	7.81	2.25	89.94	2.64
	46.24	3.75	50.01	9.76	4.57	85.67	10.04	4.42	85.54	2.10
	44.99	5.01	50.00	11.80	5.60	82.60	11.17	5.95	82.88	3.98
	42.50	7.50	50.00	15.49	8.74	75.77	15.20	8.41	76.39	2.16
	39.99	10.01	50.00	22.81	10.80	66.39	22.62	10.78	66.60	0.44
						$E_A{}^b$/%	0.28	0.17	0.29	-
									$E_R{}^a$/%	2.12
oil phase	49.65	0.00	50.35	85.46	0.00	14.54	86.23	0.00	13.77	2.07
	47.93	2.00	50.07	81.95	1.79	16.26	82.63	1.67	15.70	3.66
	46.24	3.75	50.01	77.40	3.42	19.18	77.71	3.31	18.98	1.55
	44.99	5.01	50.00	75.74	4.31	19.95	75.46	4.40	20.14	1.14
	42.50	7.50	50.00	68.19	6.82	24.99	67.84	6.52	25.64	2.50
	39.99	10.01	50.00	60.45	9.12	30.43	60.04	9.02	30.94	1.15
						$E_A{}^b$/%	0.47	0.12	0.48	-
									$E_R{}^a$/%	2.01

[a] Average relative error (E_R) calculated according to eq 1 using compositions determined by both analytical methods. [b] Average absolute error $= E_A = \sum_{i=1}^{n} |w_{i,\mathrm{NIR}} - w_{i,\mathrm{CONV}}| / n$, where i = component and n = number of tie lines.

triacylglycerol compositions of the vegetable oils could be determined (Table 2) using the algorithm suggested by Antoniosi Filho et al.[21]

To calculate the probable triacylglycerol compositions, the quantities of trans isomers (see Table 1) were added with their respective cis isomers. In Table 2, the main triacylglycerol represents the component with the greatest composition in the isomer set with *x* carbons and *y* double bonds. The results shown in Tables 1 and 2 allow us to calculate the average molar masses of the pretreated cottonseed oil, corn oil, canola oil, refined soybean oil, and degummed soybean oil. The values obtained were (861.09, 872.29, 880.44, 873.14, and 872.29) g·mol^{-1}, respectively. From the fatty acid compositions of the oils shown in Table 1, the average molar masses of the free fatty acids in each oil were estimated. The values for the average molar masses of the free fatty acids obtained were 274.46 g·mol^{-1} for the pretreated cottonseed oil, 278.36 g·mol^{-1} for corn oil, 280.76 g·mol^{-1} for canola oil, 278.61 g·mol^{-1} for refined soybean oil, and 278.41 g·mol^{-1} for degummed soybean oil. For the commercial linoleic acid, the average molar mass was 279.18 g·mol^{-1}, this value being taken from Rodrigues et al.[20]

Furthermore, the free fatty acid contents of the vegetable oils were determined by titration according to official method 2201 of IUPAC[22] using an automatic burette (Metrohm, model Dosimat 715, Herisan, Switzerland). The free fatty acid contents were 0.14 % for the pretreated cottonseed oil, 0.10 % for refined corn oil, 0.09 % for refined canola oil, 0.11 % for refined soybean oil, and 1.65 % for degummed soybean oil.

Apparatus and Procedures. Model fatty systems containing pretreated cottonseed oil + commercial linoleic acid were prepared by the addition of known quantities of commercial linoleic acid to the refined oil. The model fatty systems were mixed with anhydrous ethanol in the oil/solvent mass ratio of 1:1 at (298.2 ± 0.1) K. Liquid–liquid equilibrium data were determined using 50 mL polypropylene centrifuge tubes (Corning, Lowell/MA, USA). The components were weighed on an analytical balance with a precision of 0.0001 g (Adam, model AAA 250L, Milton Keynes, United Kingdom). The tubes were vigorously stirred for at least 20 min at room temperature (quite close to 298 K), centrifuged for 10 min at 4500*g* at (298.2 ± 1.5) K (Jouan Centrifuge, model BR4i, Saint-Herblain, France),

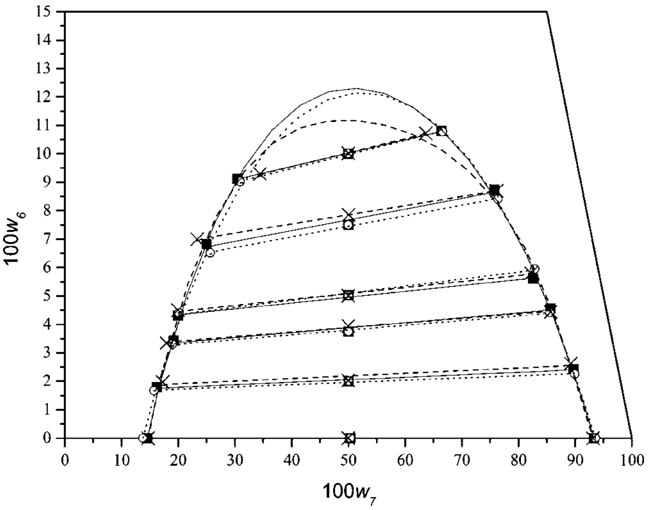

Figure 1. Liquid–liquid equilibrium for the system containing pretreated cottonseed oil (1) + commercial linoleic acid (6) + anhydrous ethanol (7) at (298.2 ± 0.1) K: ■, this work, LLE data for the conventional method; —, tie lines for the conventional method; —, binodal curve for the conventional method, fitting of a fourth-order polynomial equation; ○, this work, LLE data for the NIR method; ···, tie lines for the NIR method; ···, binodal curve for the NIR method; ×, LLE data taken from ref 20; ---, tie lines for the data taken from ref 20; ---, binodal curve for the data taken from ref 20.

and left to rest for (16 to 24) h in a thermostatic bath at (298.2 ± 0.1) K (Cole Parmer, model 12101-55, Chicago, USA). This contact time was fixed based on a previous study that showed that phase equilibrium was attained after 2 h of rest.[23]

After this treatment, the two phases became clear with a well-defined interface, and the composition of both phases was measured using two methods of quantitative analyses—a conventional analytical method and an analytical method based on near-infrared spectroscopy (NIR).

Subsequently, the liquid–liquid equilibrium data for the model systems containing vegetable oil + anhydrous ethanol + hexane were measured at (298.3 ± 0.2) K. The following oils were tested: pretreated cottonseed oil, refined corn oil, refined canola oil, refined soybean oil, and degummed soybean oil. The liquid–liquid equilibrium data were determined using equilibrium glass cells (50 mL) similar to those used by Silva et al.[24] Known quantities of each component, weighed on an analytical balance with a precision of 0.0001 g (Adam, model AAA 250L, Milton Keynes, United Kingdom), were added directly inside the equilibrium glass cells. The cell temperature was controlled using a thermostatic bath (Cole Parmer, model 12101-55, Chicago, USA). Thermometers (Incoterm, Porto Alegre/RS, Brazil) with 0.1 K subdivisions were used for monitoring the cell temperature. The mixture was stirred vigorously with a magnetic stirrer (IKA WERKE, model RH-KT/C, Staufen, Germany) for 20 min and left to rest for (16 to 24) h. This led to the formation of two clear, transparent phases, with a well-defined interface. For these systems, the composition of both phases was measured using only the NIR method.

As will be discussed later, the thermodynamic modeling developed in the present manuscript considered the system [vegetable oil + ethanol + hexane] to be a pseudoternary one, replacing the oil by a single triacylglycerol with the average molar mass of the oil. To check the validity of this approach, an overall composition for the pseudoternary system with refined soybean oil was selected, and the corresponding liquid–liquid equilibrium data were measured twice. After the analysis of the phases by the NIR method, samples of both phases were evaporated in a vacuum oven (pressure = 75.0 mmHg, temperature = 323.2 K) until complete removal of the solvents, ethanol and hexane. The oil residues of both phases were analyzed by gas chromatography of the fatty acid methyl esters, according to the procedure already described.

Conventional Analytical Method. The fatty acid content was determined by titration, according to the official method 2201 of IUPAC,[22] using an automatic burette (Metrohm, model Dosimat 715, Herisan, Switzerland). The total solvent composition was determined by evaporation at 333.2 K in an oven with air circulation and renewal (Marconi, model MA 035/3, Piracicaba/SP, Brazil) for 24 h. The drying time necessary to completely remove the solvent from the oil or alcoholic phases was determined by successively weighing the samples every 60 min until constant mass. The oil composition was determined by difference. In this method, all measurements were performed at least in triplicate. The uncertainties of the compositions varied within the following ranges: (0.08 to 0.67) % for cottonseed oil, (0.02 to 0.10) % for linoleic acid, and (0.01 to 0.58) % for ethanol, the lowest figures being obtained for the lowest compositions.

Analytical Method Based on Near-Infrared Spectroscopy (NIR). In the present work, near-infrared spectroscopy was applied for the quantitative determination of the compounds of the liquid–liquid equilibrium systems. Partial least-squares (PLS) regression was used for the calibration and prediction of the proposed method, employed to obtain relevant information from the complex NIR spectra.[25]

Instrumentation and Software. The near-infrared spectra were measured using a Thermo Nicolet infrared spectrometer (Nexus, model 670 FT-IR Esp, Madison, USA) connected to an air purging system, with a heated cuvette holder (Ventacon, model HNIR-1-1, Winchester, United Kingdom) and controlled temperature at (299.2 ± 0.5) K for a universal power supply (Ventacon, model D-2). Near-infrared absorbance spectra were registered in the range from (4400 to 11000) cm^{-1}, making 64 scans with 4 cm^{-1} resolution using a quartz cell (Hellma, model 110-QX, Müllheim/Baden, Germany) with a 1 mm optical path and 1 cm^{-1} spectra resolution. The background spectra (64 scans at 4 cm^{-1} resolution) were recorded daily, whereas for the samples, the 64 scans were taken with two repetitions to reduce instrument noise. The calibration methods were developed by PLS regression, available in PLS_Toolbox version 4.0, from Eigenvector Research (Wenatchee, WA, USA) for use with MATLAB 7.0 (The MathWorks, Natick/MA, USA).

Near-Infrared Method. To develop calibration models appropriate for each system, sets of mixtures with compositions close to the binodal curve were selected. Each binodal curve was divided into two parts, the oil phase and the alcoholic phase, and for each part, 30 pseudoternary mixtures were obtained by gravimetry and used as calibration standards. On the basis of these sets and the near-infrared spectra, the PLS regression generated the calibration models. A further set of about 10 pseudoternary mixtures for each part of the binodal curve was also obtained by gravimetry and used as a prediction set to check the accuracy of the calibration models. In the case of the system cottonseed oil + linoleic acid + ethanol, the compositions of the calibration and prediction sets were selected using the binodal curve reported by Rodrigues et al.[20] In the case of the systems containing vegetable oil + ethanol + hexane, the compositions were selected using a binodal curve previously determined by the cloud-point method. For the cloud-point method, individual conical flasks containing vegetable oil and ethanol in the composition range from 15 % to 90 % in mass of ethanol were used. Mixtures in this composition range generate two liquid phases. Using an automatic burette (Metrohm, model Dosimat 715, Herisan, Switzerland), hexane was added to the conical flasks with vigorous agitation, until the appearance of a single homogeneous liquid phase. The corresponding binodal curve was calculated for the amount of each component added.

The accuracy of the models was evaluated by the relative error (E_R), obtained from the difference between the predicted value and the actual value (gravimetry), the coefficient of determination (R^2), the linear plot between the predicted values and the actual ones, and the root-mean-square error (rmse). These values express the accuracy of the models and the proximity between the values predicted by the model ($w_{i,\text{pred}}$) and the actual values ($w_{i,\text{actual}}$) obtained using the reference method. The errors were defined as:

$$E_R = \sum_{i=1}^{n} \frac{|w_{i,\text{pred}} - w_{i,\text{actual}}|}{w_{i,\text{actual}}} \tag{1}$$

$$\text{rmse} = \sqrt{\frac{\sum_{i=1}^{n} (w_{i,\text{pred}} - w_{i,\text{actual}})^2}{n}} \tag{2}$$

where n is the total number of samples and the subscript i is the number of the sample.

The rmsecv (root-mean-square error of cross-validation) is the rmse calculated using the cross-validation samples, according to the cross-validation procedure known as "leave one out".[26,27] The cross-validation procedure and the corresponding rmsecv values were used for selecting the number of PLS factors considered in the construction of the calibration models. The rmsep (root-mean-square error of prediction) is the rmse calculated using the prediction standards, a special set of standards not included in the construction of the calibration models.

After determining and checking the calibration models for each phase, oil or alcoholic, of each system, these models could be used to calculate the compositions of the samples obtained in the equilibrium experiments. For comparative purposes and as a further check of the NIR method, the compositions of the equilibrium samples of the system cottonseed oil + linoleic acid + ethanol were also determined by the conventional method.

Most of the systems investigated in the present work were treated as pseudoternary ones, containing oil + hexane (or linoleic acid) + ethanol. The only exception was the system containing degummed soybean oil, which was treated as a pseudoquaternary one, because this semiprocessed oil contains a significant free fatty acid content (1.65 % in mass). For this reason, in the case of degummed soybean oil, eight calibration models were obtained, one for each pseudocomponent (oil, fatty acids, ethanol, or hexane) in each phase (oil or alcoholic). In the case of the NIR method, the uncertainties of the compositions varied within the following ranges: (0.04 to 0.60) % for oil, (0.04 to 0.15) % for linoleic acid, (0.02 to 0.66) % for ethanol, and (0.05 to 0.14) % for hexane, the lowest figures being obtained for the lowest compositions.

Theoretical Calculations

Calculation of the Errors in the Mass Balance of the Phases. To test the validity of the results of the equilibrium experiments, the procedure developed by Marcilla et al.,[28] and already applied to fatty systems by Rodrigues et al.,[3,20] was followed. This procedure consists of calculating the masses for both liquid phases and comparing their sum with the actual value for total mass used in the experiment, thus obtaining a relative error for each point of the overall mixture. According to this approach, i independent component balances can be written, with i being each component of the system, given by

$$M^{\mathrm{OC}} w_i^{\mathrm{OC}} = M^{\mathrm{AP}} w_i^{\mathrm{AP}} + M^{\mathrm{OP}} w_i^{\mathrm{OP}} \tag{3}$$

where M^{OC} is the amount of the overall composition (initial mixture); M^{AP} and M^{OP} are the amounts of the alcoholic and oil phases, respectively; w_i^{OC} is the mass fraction of component i in the initial mixture; and w_i^{AP} and w_i^{OP} are the mass fractions of component i in the alcoholic and oil phases, respectively. With these i equations, it is possible to calculate the values for M^{AP} and M^{OP} from the experimental values w_i^{AP} and w_i^{OP} using a least-squares fitting. If M is the matrix formed by the values for w_i^{OC}; B is the transformation matrix (formed by the values for w_i^{AP} and w_i^{OP}); and P is the matrix formed by the mass of each phase (M^{AP} and M^{OP}), the previous system can be written as

$$M = B \cdot P \tag{4}$$

Mathematical calculations transform eq 4 into the following expression:

$$P = (B^{\mathrm{T}} B)^{-1} B^{\mathrm{T}} M \tag{5}$$

where B^{T} is the transposed matrix of B and $(B^{\mathrm{T}}B)^{-1}$ is the inverse matrix of $(B^{\mathrm{T}}B)$. In this way, the values for M^{AP} and M^{OP} (matrix P), which minimize the errors of the previous system, have been calculated. The sum of M^{AP} and M^{OP} can be compared to M^{OC} to estimate an overall mass balance deviation. The relative errors between the sum ($M^{\mathrm{AP}} + M^{\mathrm{OP}}$) and M^{OC}, calculated according to ($|(M^{\mathrm{AP}} + M^{\mathrm{OP}}) - M^{\mathrm{OC}}|/M^{\mathrm{OC}}$), were always lower than 0.2 %, indicating the good quality of the experimental data.[28]

Thermodynamic Modeling. The experimental data determined for the model systems were used to adjust the parameters of the NRTL model. The mass fraction was used as a composition unit due to the large difference in molar mass of the components in the systems.[3,20,23,29–32] Rodrigues et al.[29] showed the activity coefficient equations and the isoactivity criterion, expressed in terms of mass fraction, according to the NRTL model. The isoactivity criterion of phase equilibrium developed on a molar fraction basis can be expressed in a mass fraction unity as follows:

$$(\gamma_i x_i)^{\mathrm{AP}} = (\gamma_i x_i)^{\mathrm{OP}} \tag{6}$$

$$(\gamma_i^w w_i)^{\mathrm{AP}} = (\gamma_i^w w_i)^{\mathrm{OP}} \tag{7}$$

where

$$\gamma_i^w = \frac{\gamma_i}{\overline{M}_i \sum_j^n \left(\dfrac{w_j}{\overline{M}_j} \right)} \tag{8}$$

γ_i is the activity coefficient of pseudocomponent i estimated according to the original NRTL equation; γ_i^w is the corresponding activity coefficient expressed on the mass fraction scale; and $\overline{M}_i$ is the molar mass of pseudocomponent i. Fitting was done by treating the model system vegetable oil + anhydrous ethanol + hexane as a pseudoternary one and the model system degummed soybean oil + linoleic acid + anhydrous ethanol + hexane as a pseudoquaternary one. It should be noted that the free fatty acid content was only significant in the case of degummed soybean oil, this being the reason the corresponding system was considered as a pseudoquaternary one.

For the fitting process, the vegetable oil was treated as a single triacylglycerol with the average molar mass of the oil. The same approach was extended to the linoleic acid. This approach assumes that the different triacylglycerols present in the vegetable oil behave in a very similar way in the liquid–liquid system under analysis. In this case, such compounds can be adequately replaced by a pseudocomponent having the corresponding average physical–chemical properties. The NRTL model has three parameters for each binary mixture: A_{ij} and A_{ji} are parameters representing the energy interactions between compounds i and j, and α_{ij} (= a_{ji}) is a nonrandomness parameter that derives from the local composition assumption.[33] With the exception of those parameters taken from the literature, in the present work, all three parameters corresponding to each binary mixture were adjusted to the experimental data, including the α_{ij} values. The adjustment of the α_{ij} values allowed a better correlation of the equilibrium data and is an approach already used in the literature.[34–37]

Estimation of the parameters was based on minimization of the objective function of compositions, eq 9, following the procedure developed by Stragevitch and d'Avila.[38]

Table 5. Errors in the Quantification of the Prediction Standards for the Systems Containing Vegetable Oil (1, 2, 3, and 4) + Anhydrous Ethanol (7) + Hexane (8) at (298.3 ± 0.2) K

		alcoholic phase			oil phase			
oil (x)[a]		oil (x)[a]	ethanol (7)	hexane (8)	oil (x)	ethanol (7)	hexane (8)	E_R/%
pretreated cottonseed (1)	E_R/%	1.27	0.22	0.82	0.18	0.60	0.77	0.64
	rmsep/%	0.11	0.17	0.09	0.14	0.10	0.07	-
	rmsecv/%	0.07	0.12	0.06	0.08	0.05	0.07	-
	R^2	0.99994	0.99996	0.99991	0.99998	0.99992	0.99997	-
corn (2)	E_R/%	1.14	0.06	1.12	0.07	0.11	0.84	0.56
	rmsep/%	0.09	0.06	0.13	0.05	0.03	0.06	-
	rmsecv/%	0.07	0.09	0.05	0.08	0.05	0.09	-
	R^2	0.99987	0.99999	0.99975	0.99998	0.99999	0.99990	-
canola (3)	E_R/%	1.98	0.07	0.76	0.06	0.11	0.39	0.56
	rmsep/%	0.10	0.06	0.06	0.04	0.02	0.05	-
	rmsecv/%	0.11	0.09	0.07	0.05	0.03	0.05	-
	R^2	0.99997	0.99999	0.99996	0.99999	0.99999	0.99998	-
refined soybean (4)	E_R/%	2.45	0.20	1.34	0.10	0.15	0.84	0.84
	rmsep/%	0.18	0.17	0.14	0.07	0.03	0.07	-
	rmsecv/%	0.11	0.12	0.17	0.18	0.10	0.14	-
	R^2	0.99999	0.99994	0.99960	0.99999	0.99999	0.99995	-

[a] x = reference number of the oil used.

Table 6. Errors in the Quantification of the Prediction Standards for the System Containing Degummed Soybean Oil (5) + Linoleic Acid (6) + Anhydrous Ethanol (7) + Hexane (8) at (298.3 ± 0.2) K

	alcoholic phase				oil phase				
	oil (5)	acid (6)	ethanol (7)	hexane (8)	oil (5)	acid (6)	ethanol (7)	hexane (8)	E_R/%
E_R/%	2.64	3.31	0.33	0.35	0.12	0.24	0.39	0.98	1.04
rmsep/%	0.25	0.08	0.23	0.05	0.10	0.04	0.08	0.08	-
rmsecv/%	0.04	0.04	0.04	0.05	0.06	0.05	0.05	0.04	-
R^2	0.99998	0.99924	0.99999	0.99995	0.99996	0.99965	0.99996	0.99995	-

$$\mathrm{OF}(w) = \sum_{m}^{D}\sum_{n}^{N}\sum_{i}^{K-1}\left[\left(\frac{w_{i,n,m}^{\mathrm{AP,exptl}} - w_{i,n,m}^{\mathrm{AP,calcd}}}{\sigma_{w_{i,n,m}^{\mathrm{AP}}}}\right)^2 + \left(\frac{w_{i,n,m}^{\mathrm{OP,exptl}} - w_{i,n,m}^{\mathrm{OP,calcd}}}{\sigma_{w_{i,n,m}^{\mathrm{OP}}}}\right)^2\right] \quad (9)$$

where D is the total number of data groups; N is the total number of tie lines; K is the total number of pseudocomponents in the data group; w is the mass fraction; the subscripts i, n, and m are the component, tie line, and group number, respectively; and the superscripts AP and OP stand for the alcoholic and oil phases, respectively; exptl and calcd refer to experimental and calculated compositions. $\sigma_{w_{i,n,m}^{\mathrm{AP}}}$ and $\sigma_{w_{i,n,m}^{\mathrm{OP}}}$ are the standard deviations observed in the compositions of the two liquid phases. The parameter estimation procedure involves flash calculations for the middle point compositions of the experimental tie lines and the minimization of the objective function given by eq. 9. The average deviations between the experimental and calculated compositions in both phases were calculated according to eq 10.

$$\Delta w = \sqrt{\frac{\sum_{n}^{N}\sum_{i}^{K}\left[\left(w_{i,n}^{\mathrm{AP,exptl}} - w_{i,n}^{\mathrm{AP,calcd}}\right)^2 + \left(w_{i,n}^{\mathrm{OP,exptl}} - w_{i,n}^{\mathrm{OP,calcd}}\right)^2\right]}{2NK}} \quad (10)$$

Results and Discussion

In the present work, the pseudocomponents used in the experimental data and in the fitting of the NRTL parameters received the following notations: pretreated cottonseed oil (1), corn oil (2), canola oil (3), refined soybean oil (4), degummed soybean oil (5), linoleic acid (6), anhydrous ethanol (7), and hexane (8).

Pretreated Cottonseed Oil + Commercial Linoleic Acid + Anhydrous Ethanol System. Table 3 gives the results predicted by the NIR method, including the errors between the compositions calculated by the calibration models and those calculated by gravimetry during the preparation of the prediction set. The errors of the conventional method are also shown. As can be seen, the calibration models exhibited high coefficients of determination (R^2) and low values for the root-mean-square errors of prediction and cross-validation (rmsep and rmsecv). In the case of the oil phase, the errors shown by the NIR method were always lower than the corresponding values shown by the conventional method. In the case of the alcoholic phase, the errors showed similar values for both methods. Considering the deviations in both phases, an overall average relative error of 1.3 % was obtained in the case of the NIR method and of 2.1 % in the conventional one.

Table 4 shows the equilibrium data for the cottonseed oil + linoleic acid + ethanol system according to both analytical methods and the corresponding average errors for the compositions. The overall average relative errors between the compositions obtained by both methods (Table 4) were similar to those observed when the compositions determined by each method were compared to the composition values of the prediction set (Table 3). The average absolute error between the compositions determined by both analytical methods (Table 4) were within the uncertainty ranges estimated for those methods.

As can be seen in Figure 1, the tie lines obtained for both methods showed similar inclinations, and the regions of two-phase coexistence were also very similar. Figure 1 also shows the liquid–liquid data taken from Rodrigues et al.[20] As can be seen, the three data sets are very similar, with larger composition differences only near the plait point.

Despite the deviations between the phase compositions determined by both methods, the mass balance errors calculated according to the procedure suggested by Marcilla et al.[28] were

Table 7. Liquid–Liquid Equilibrium Data for the Pseudoternary Systems Containing Pretreated Cottonseed (1), Corn (2), Canola (3), or Refined Soybean Oil (4) + Anhydrous Ethanol (7) + Hexane (8) at (298.3 ± 0.2) K

	overall composition			alcoholic phase			oil phase		
oil (x)[a]	100 w_x	100 w_7	100 w_8	100 w_x	100 w_7	100 w_8	100 w_x	100 w_7	100 w_8
pretreated cottonseed (1)	49.93	50.07	0.00	6.70	93.30	0.00	83.59	16.41	0.00
	47.71	50.11	2.18	7.72	90.13	2.15	80.56	17.41	2.03
	45.82	49.96	4.22	8.58	87.52	3.90	77.57	18.40	4.03
	43.94	49.95	6.11	9.87	84.37	5.76	74.28	19.68	6.04
	41.97	50.08	7.95	11.34	81.10	7.56	70.91	21.05	8.04
	39.87	50.05	10.08	13.07	77.40	9.53	67.04	22.82	10.14
	37.89	50.00	12.11	15.69	72.77	11.54	62.68	25.03	12.29
	35.98	50.07	13.95	18.62	67.91	13.47	57.76	28.06	14.18
	34.14	49.92	15.94	24.33	60.15	15.52	50.27	33.59	16.14
corn (2)	49.83	50.17	0.00	6.24	93.76	0.00	85.15	14.85	0.00
	48.03	49.98	1.99	7.00	91.07	1.93	82.18	15.85	1.97
	45.64	50.30	4.06	7.63	88.53	3.84	79.00	16.70	4.30
	43.71	50.34	5.95	8.89	85.44	5.67	75.75	17.84	6.41
	41.75	50.24	8.01	9.82	82.54	7.64	73.07	18.38	8.55
	39.59	50.22	10.19	11.47	78.90	9.63	68.83	20.52	10.65
	37.69	50.16	12.15	13.51	74.99	11.50	64.73	22.58	12.69
	35.65	50.23	14.12	15.95	70.50	13.55	60.23	25.08	14.69
	33.74	50.17	16.09	19.71	64.79	15.50	54.56	28.71	16.73
canola (3)	49.91	50.09	0.00	4.33	95.67	0.00	86.84	13.16	0.00
	47.73	50.02	2.25	5.35	92.48	2.17	82.72	15.11	2.17
	45.83	50.11	4.06	5.86	90.64	3.50	80.87	14.95	4.18
	43.79	50.13	6.08	7.46	86.90	5.64	77.03	16.88	6.09
	41.70	50.10	8.20	8.70	83.67	7.63	73.75	17.85	8.40
	39.66	50.16	10.18	9.86	80.63	9.51	69.86	19.51	10.63
	37.89	50.06	12.05	11.21	77.53	11.26	66.94	20.60	12.46
	35.82	49.87	14.31	13.65	72.77	13.58	62.29	22.97	14.74
	33.90	49.82	16.28	16.54	68.02	15.44	57.32	25.94	16.74
	32.09	49.87	18.04	20.70	61.88	17.42	52.20	29.34	18.46
refined soybean (4)	50.04	49.96	0.00	6.12	93.88	0.00	84.40	15.60	0.00
	47.99	50.13	1.88	7.07	91.35	1.58	81.80	16.40	1.80
	45.94	49.85	4.21	8.07	88.09	3.84	78.62	17.13	4.25
	44.03	50.03	5.94	9.21	85.38	5.41	75.43	18.41	6.16
	41.83	50.25	7.92	10.43	82.30	7.27	72.22	19.66	8.12
	39.84	50.24	9.92	12.06	78.75	9.19	68.86	21.07	10.07
	37.79	50.15	12.06	14.29	74.70	11.01	64.53	23.20	12.27
	36.01	50.03	13.96	16.93	69.97	13.10	60.04	25.70	14.26
	35.08[b]	49.91	15.01	18.59	66.53	14.88	56.17	28.26	15.57
	35.01[c]	49.96	15.03	18.56	66.48	14.96	56.14	28.14	15.72
	34.09	50.01	15.90	18.89	65.39	15.72	55.37	28.17	16.46
	33.00	50.11	16.89	21.63	61.83	16.54	51.47	31.02	17.51

[a] x = reference number of the oil used. [b] Tie line 1, for which the fatty acid compositions of the oil residues are given in Table 1, and the probable triacylglycerol compositions are given in Table 2. [c] Tie line 2, for which the fatty acid compositions of the oil residues are given in Table 1, and the probable triacylglycerol compositions are given in Table 2.

low in both cases. In the case of the NIR method, the average mass balance error was 0.06 %, and for the conventional method, the corresponding deviation was 0.10 %. Such low values indicate the good quality of the experimental data, slightly better in the case of the spectroscopy methodology.

Systems of Interest for Biodiesel Production. Tables 5 and 6 show the results obtained by the NIR spectroscopy method for the pseudoternary systems containing vegetable oils (pretreated cottonseed, corn, canola, and refined soybean) + anhydrous ethanol + hexane and for the pseudoquaternary system of degummed soybean oil + linoleic acid + anhydrous ethanol + hexane at (298.3 ± 0.2) K, respectively, including the errors between the compositions calculated by the calibration models and those calculated by gravimetry during the preparation of the prediction set.

As can be seen, the calibration models exhibited high coefficients of determination (R^2) and low values for the root-mean-square errors of prediction and cross-validation (rmsep and rmsecv). In almost all cases, the overall average error obtained was lower than 1.0 %. Tables 7 and 8 present the overall experimental composition of the mixtures and the corresponding tie lines for the system containing vegetable oil (pretreated cottonseed, corn, canola, and refined soybean) + anhydrous ethanol + hexane and for the degummed soybean oil + linoleic acid + anhydrous ethanol + hexane system at (298.3 ± 0.2) K, respectively. The experimental data measured for the model systems were used to fit the NRTL parameters to the pseudoternary and pseudoquaternary systems, as shown in Table 9. The parameters for the interaction between pretreated cottonseed oil (1) and ethanol (7) and linoleic acid (6) and ethanol (7) at 298.2 K were taken from Rodrigues et al.,[20] and the parameters for the interaction between corn oil (2) and ethanol (7) at 298.2 K were taken from Gonçalves et al.[31] For the system containing canola oil, Batista et al.[32] presented values for the parameters for the interaction between canola oil (1) and ethanol (7) at two temperatures, (293.2 and 303.2) K, so an average value of these parameters was used as an initial estimation for the fit at 298.3 K. For the systems containing refined and degummed soybean oils, the same sets of parameters were used for the interactions between oil and ethanol and oil and hexane. The parameters for the interactions between linoleic acid and hexane, and soybean oil and linoleic acid, were fitted using the experimental data for the degummed oil. This approach is possible because the fatty acid compositions of these oils are very similar, as can be seen in Table 1. The

Table 8. Liquid–Liquid Equilibrium Data for the Pseudoquaternary System Containing Degummed Soybean Oil (5) + Linoleic Acid (6) + Anhydrous Ethanol (7) + Hexane (8) at (298.3 ± 0.2) K

overall composition				alcoholic phase				oil phase			
100 w_5	100 w_6	100 w_7	100 w_8	100 w_5	100 w_6	100 w_7	100 w_8	100 w_5	100 w_6	100 w_7	100 w_8
49.12	0.82	50.06	0.00	8.31	0.15	91.54	0.00	82.25	1.36	16.39	0.00
47.02	0.79	49.92	2.27	9.10	0.17	88.77	1.96	78.79	1.31	17.42	2.48
45.10	0.75	49.93	4.22	9.99	0.18	85.99	3.84	75.85	1.26	18.36	4.53
43.00	0.72	50.02	6.26	11.13	0.20	82.88	5.79	72.43	1.20	19.76	6.61
41.21	0.69	49.81	8.29	12.49	0.22	79.51	7.78	68.89	1.14	21.20	8.77
39.30	0.66	49.66	10.38	14.17	0.25	75.79	9.79	65.05	1.07	23.05	10.83
37.46	0.63	49.86	12.05	16.00	0.28	72.16	11.56	61.56	1.01	24.94	12.49
35.27	0.59	49.94	14.20	19.31	0.33	66.62	13.74	56.10	0.93	28.23	14.74
33.40	0.56	50.00	16.04	24.79	0.42	59.16	15.63	48.84	0.82	33.81	16.53

Table 9. NRTL Parameters

pair *ij*	A_{ij}/K	A_{ji}/K	α_{ij}
17[a]	317.97	1410.31	0.52262
18	−472.36	208.84	0.47000
27[b]	−166.14	1620.9	0.40115
28	−605.88	328.22	0.26858
37[c]	33.495	1541.9	0.47000
38	−461.78	360.68	0.47000
47[c]	229.02	1465.6	0.52262
48	−529.26	319.60	0.39803
56	−2592.3	6.0130	0.69960
57[d]	229.02	1465.6	0.52262
58	−529.26	319.60	0.39803
67[a]	5000.0	−2020.22	0.21917
68[d]	3405.4	−787.35	0.21917
78	473.16	99.522	0.36793

[a] Parameters taken from Rodrigues et al.[20] [b] Parameters taken from Gonçalves et al.[31] [c] Pretreated cottonseed oil (1), corn oil (2), canola oil (3), refined soybean oil (4), anhydrous ethanol (7), and hexane (8). [d] Degummed soybean oil (5), linoleic acid (6), anhydrous ethanol (7), and hexane (8).

Table 10. Mean Deviations in the Phase Compositions of the Systems Studied

system	100 Δ*w*
pretreated cottonseed oil + anhydrous ethanol + hexane	0.80
corn oil + anhydrous ethanol + hexane	0.64
canola oil + anhydrous ethanol + hexane	0.48
refined soybean oil + anhydrous ethanol + hexane	0.64
degummed soybean oil + linoleic acid + anhydrous ethanol + hexane	0.86
average overall deviation	0.70

average deviations between the experimental and calculated compositions in both phases were calculated according to eq 10 and are shown in Table 10.

Figure 2 shows the experimental points, tie lines, and binodal curve calculated using the NRTL model for the system containing pretreated cottonseed oil + anhydrous ethanol + hexane at (298.3 ± 0.2) K. According to Figure 2, the selected thermodynamic model was able to accurately describe the phase compositions, as can be further confirmed by the corresponding low deviation (0.8 %) presented in Table 10. The mass balance error for this system, evaluated according to the procedure suggested by Marcilla et al.,[28] was 0.13 %, a value which indicates the good quality of the experimental data.

Figures 3, 4, and 5 show the experimental points, tie lines, and binodal curves calculated using the NRTL model for the systems containing corn oil, canola oil, and refined soybean oil, respectively. As can be seen in Figures 3, 4, and 5 and in Table 10, the NRTL model was also able to accurately describe the phase compositions for these three systems.

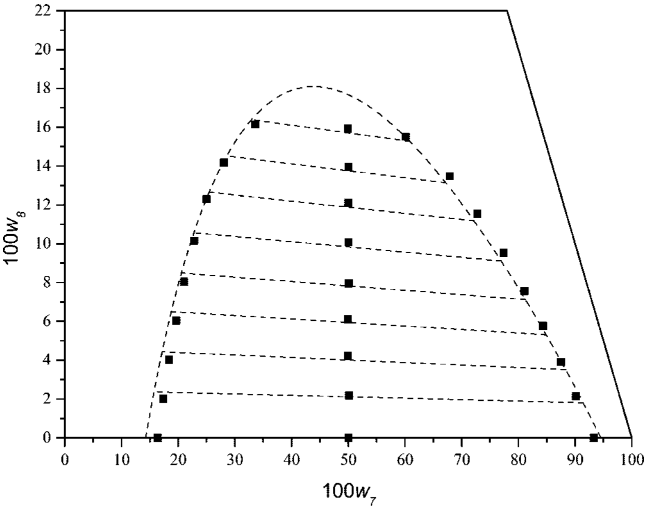

Figure 2. Liquid–liquid equilibrium for the system containing pretreated cottonseed oil (1) + anhydrous ethanol (7) + hexane (8) at (298.3 ± 0.2) K: ■, experimental; ---, NRTL model.

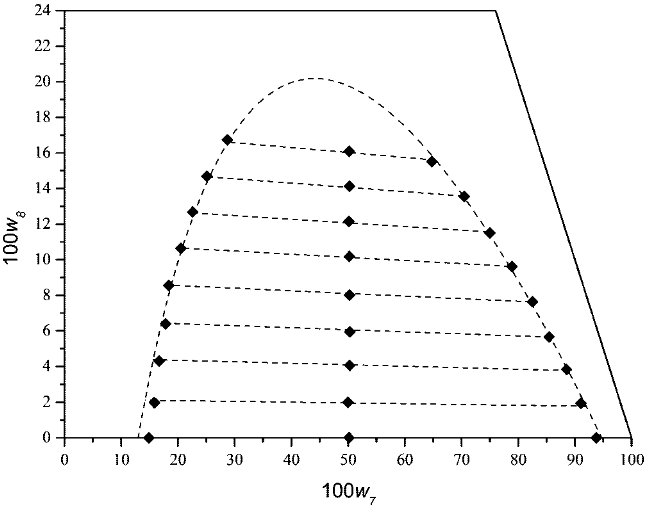

Figure 3. Liquid–liquid equilibrium for the system containing corn oil (2) + anhydrous ethanol (7) + hexane (8) at (298.3 ± 0.2) K: ◆, experimental; ---, NRTL model.

The evaluation of the mass balance errors indicated that the average error was 0.05 % for the system containing corn oil, 0.14 % for canola oil, and 0.16 % for refined soybean oil. These results suggest the good quality of the experimental data.

Table 1 shows the fatty acid compositions of the oil residues obtained after evaporation of the alcoholic and oil phases for two tie lines of the ternary system with refined soybean oil. These tie lines are also indicated in the data given in Table 7 and shown in Figure 5. As the overall

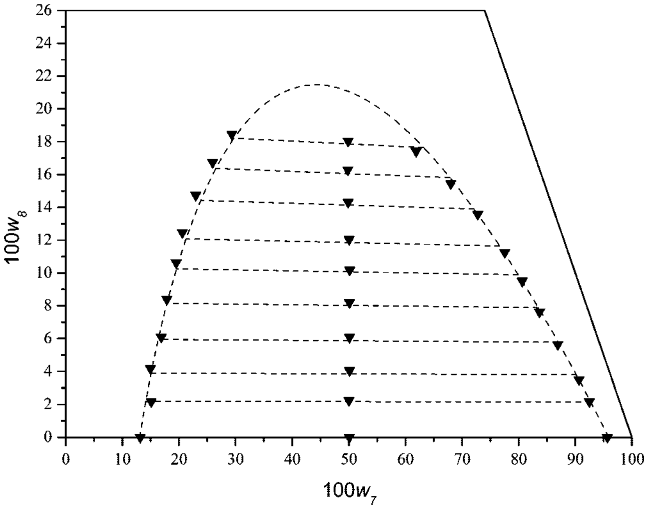

Figure 4. Liquid–liquid equilibrium for the system containing canola oil (3) + anhydrous ethanol (7) + hexane (8) at (298.3 ± 0.2) K: ▼, experimental; ---, NRTL model.

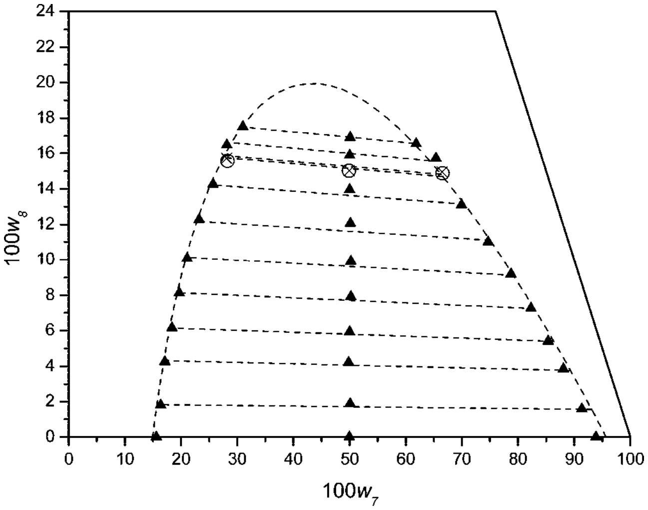

Figure 5. Liquid–liquid equilibrium for the system containing refined soybean oil (4) + anhydrous ethanol (7) + hexane (8) at (298.3 ± 0.2) K: ▲, experimental; ---, NRTL model; ○, tie line 1; and × , tie line 2 (see Tables 1, 2, and 7).

compositions were almost the same for both tie lines, the contents of oil, hexane, and ethanol in each corresponding phase were very similar. The same behavior was observed for the fatty acid compositions of the corresponding oil residues (Table 1): both alcoholic phases showed very similar fatty acid compositions, and the same occurred with the oil phases. Nevertheless, the fatty acid compositions of the alcoholic and oil phases of a specific tie line, 1 and 2, were not identical. Both oil phases showed a fatty acid composition very close to the values determined for the refined oil, but the alcoholic phases contained comparatively larger amounts of palmitic (C16:0), stearic (C18:0), and oleic (C18:1) acids and lower amounts of linoleic (C18:2) and linolenic (C18:3) acids.

On the basis of these fatty acid compositions, the probable triacylglycerol compositions of the oil residues were estimated according to the algorithm suggested by Antoniosi Filho et al.[21] The corresponding values are given in Table 2. As can be seen in this table, the triacylglycerol compositions of both oil phase residues are close to the refined soybean oil composition. But in the case of the alcoholic phases, triacylglycerols containing at least two fatty acids among palmitic, stearic, and oleic acids have mass fractions larger than the corresponding values in the refined oil. The opposite behavior is observed for those triacylglycerols containing at least two fatty acids among linoleic and linolenic acids. It is worthy to emphasize that these triacylglycerol compositions should be considered as estimated values. Although the fatty acid compositions were measured by gas chromatography, the corresponding triacylglycerol compositions were submitted to additional uncertainties associated to the statistical estimation procedure developed by Antoniosi Filho et al.[21]

Despite the differences in the fatty acid composition of the tie line phases, the average molar masses of the corresponding oil residues are very similar: 868.33 g·mol^{-1} and 868.20 g·mol^{-1} for the oil in the alcoholic phases of tie lines 1 and 2, respectively, and 873.13 g·mol^{-1} and 872.97 g·mol^{-1} for the oil in the oil phases of tie lines 1 and 2, respectively. These results indicate that the difference in the average oil molar masses in both phases was lower than 0.6 % for tie line 1 as well as for tie line 2. In fact, considering the oil residues of the alcoholic phases, the larger content of triacylglycerols with palmitic acid, a fatty component with 16 carbon atoms, was approximately balanced by the larger amounts of triacylglycerols with oleic and stearic acids, fatty compounds with 18 carbon atoms, so that the average molar masses of the oil residues in both phases tended to be very close to each other. Although some triacylglycerols were selectively distributed between both phases, this distribution was restricted to a narrow range of compositions, so that the pseudocomponent approach can be considered as approximately valid.

As explained in the Experimental Section, NIR calibration models were obtained for each of the oils, and the data reported in Tables 7 and 8 were calculated using those specific models. To check if the calibration models obtained for specific oils could be used in the quantification of an equilibrium system containing another oil, the calibration models obtained for the system with canola oil were also used to estimate the equilibrium data for the system with corn oil. The comparison between these estimated data and the data for corn oil reported in Table 7 indicated that the average relative deviation between both sets of data (eq 1) was 0.73 % for the alcoholic phase and 4.8 % for the oil phase. The average absolute deviations between both sets of data (see footnote of Table 4) were 0.11 % for the alcoholic phase and 0.51 % for the oil phase. The average mass balance error calculated using the estimated data was 0.31 %, larger than the error obtained for the corn oil data reported in Table 7 (0.05 %). Such results indicate that if a lower precision in the quantification of the equilibrium data is allowed calibration models obtained for specific oils can be used for a system containing another vegetable oil.

Figure 6 shows the equilibrium diagram for the system containing degummed soybean oil + linoleic acid + anhydrous ethanol + hexane. In order to represent such pseudoquaternary data in rectangular coordinates, degummed soybean oil and linoleic acid were considered as a mixed solute.

As can be seen in Figure 6, the NRTL model also correctly describes the experimental data for the pseudoquaternary system, with an average deviation of 0.86 % (Table 10). The very low value of the mass balance error for the system (0.03 %) confirms the excellent quality of the experimental data.

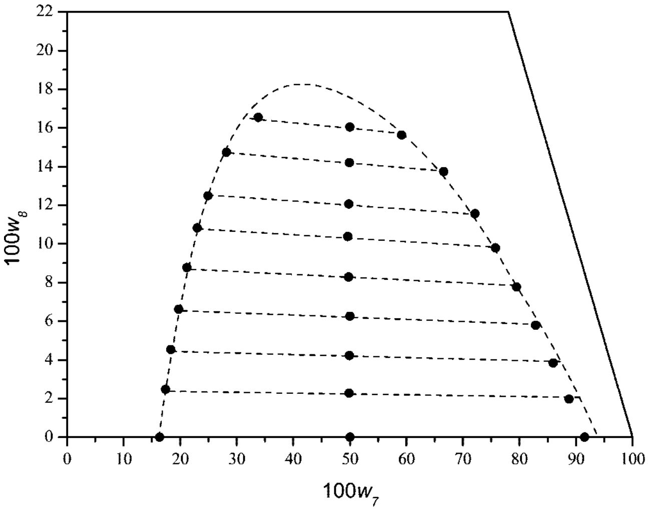

Figure 6. Liquid–liquid equilibrium for the system containing degummed soybean oil (5) + linoleic acid (6) + anhydrous ethanol (7) + hexane (8) at (298.3 ± 0.2) K: ●, experimental; ---, NRTL model.

Conclusions

The results presented in this paper show that the NIR method performed well in the determination of equilibrium data and can guarantee reliable results. According to the results obtained for the system containing pretreated cottonseed oil + commercial linoleic acid + anhydrous ethanol, the NIR method is highly accurate, with the further advantage of being less time intensive than other analytical methods. The low uncertainties of the experimental data and the low errors obtained in the mass balance indicate the good quality of the equilibrium data for the systems of interest for biodiesel production, containing different vegetable oils (cottonseed, corn, canola, or soybean oil) + anhydrous ethanol + hexane.

Literature Cited

(1) Gonçalves, C. B.; Pessôa Filho, P. A.; Meirelles, A. J. A. Partition of Nutraceutical Compounds in Deacidification of Palm Oil by Solvent Extraction. *J. Food Eng.* **2007**, *81*, 21–26.

(2) Rodrigues, C. E. C.; Onoyama, M. M.; Meirelles, A. J. A. Optimization of the Rice Bran Oil Deacidification Process by Liquid-Liquid Extraction. *J. Food Eng.* **2006**, *73*, 370–378.

(3) Rodrigues, C. E. C.; Silva, F. A.; Marsaioli, A., Jr.; Meirelles, A. J. A. Deacidification of Brazil Nut and Macadamia Nut Oils by Solvent Extraction: Liquid–Liquid Equilibrium Data at 298.2 K. *J. Chem. Eng. Data* **2005**, *50*, 517–523.

(4) Encinar, J. M.; González, J. F.; Rodríguez, J. J.; Tejedor, A. Biodiesel Fuels from Vegetable Oils: Transesterification of *Cynaracardunculus* L. Oils with Ethanol. *Energy Fuels* **2002**, *16*, 443–450.

(5) Rendón, X.; López-Munguía, A.; Castillo, E. Solvent Engineering Applied to Lipase-Catalized Glycerolysis of Triolein. *J. Am. Oil Chem. Soc.* **2001**, *78*, 1061–1066.

(6) Oliveira, D.; Oliveira, J. V. Enzymatic Alcoholysis of Palm Kernel Oil in n-Hexane and in $SCCO_2$. *J. Supercrit. Fluids* **2001**, *19*, 141–148.

(7) Iso, M.; Chen, B.; Eguchi, M.; Kudo, T.; Shrestha, S. Production of Biodiesel Fuel from Triglycerides and Alcohol Using Immobilized Lipase. *J. Mol. Catal. B: Enzym.* **2001**, *16*, 53–58.

(8) Di Serio, M.; Ledda, M.; Cozzolino, M.; Minutillo, G.; Tesser, R.; Santacesaria, E. Transesterification of Soybean Oil to Biodiesel by Using Heterogeneous Basic Catalysts. *Ind. Eng. Chem. Res.* **2006**, *45*, 3009–3014.

(9) Lotero, E.; Liu, Y; Lopez, D. E.; Suwannakarn, K.; Bruce, D. A.; Goodwin, J. G., Jr. Synthesis of Biodiesel via Acid Catalysis. *Ind. Eng. Chem. Res.* **2005**, *44*, 5353–5363.

(10) Noureddini, H.; Gao, X.; Philkana, R. S. Immobilized Pseudomonas Cepacia Lipase for Biodiesel Fuel Production from Soybean Oil. *Bioresour. Technol.* **2005**, *96*, 769–777.

(11) Shah, S.; Sharma, S.; Gupta, M. N. Biodiesel Preparation by Lipase-Catalyzed Transesterification of *Jatropha* Oil. *Energy Fuels* **2004**, *18*, 154–159.

(12) Zhou, H.; Lu, H.; Liang, B. Solubility of Multicomponent Systems in the Biodiesel Production by Transesterification of *Jatropha curcas* L. Oil with Methanol. *J. Chem. Eng. Data* **2006**, *51*, 1130–1135.

(13) Čerče, T.; Peter, S.; Weidner, E. Biodiesel-Transesterification of Biological Oils with Liquid Catalysts: Thermodynamic Properties of Oil–Methanol–Amine Mixtures. *Ind. Eng. Chem. Res.* **2005**, *44*, 9535–9541.

(14) Barthus, R. C.; Poppi, R. J. Multivariate Quality Control Applied to Detect the Soybean Oil Oxidation Using Fourier Transform Infrared Spectroscopy. *Spectrosc. Lett.* **2002**, *35*, 729–739.

(15) Gonzaga, F. B.; Pasquini, C.; Rodrigues, C. E. C.; Meirelles, A. J. A. Comparison of Near-Infrared Emission Spectroscopy and the Rancimat Method for the Determination of Oxidative Stability. *Eur. J. Lipid Sci. Technol.* **2007**, *109*, 61–65.

(16) Cozzolino, D.; Murray, I.; Chree, A.; Scaife, J. R. Multivariate Determination of Free Fatty Acids and Moisture in Fish Oils by Partial Least-Squares Regression and Near Infrared Spectroscopy. *LWT-Food Sci. Technol.* **2005**, *38*, 821–828.

(17) Downey, G.; McIntyre, P.; Davies, A. N. Geographic Classification of Extra Virgin Olive Oils from the Eastern Mediterranean by Chemometric Analysis of Visible and Near Infrared Spectroscopic Data. *Appl. Spectrosc.* **2003**, *57*, 158–163.

(18) AOCS. *Official Methods and Recommended Practices of the American Oil Chemists' Society*, 3rd ed.; AOCS Press: Champaign, 1988; *Vols. 1–2*.

(19) AOCS. *Official Methods and Recommended Practices of the American Oil Chemists' Society*, 5th ed.; AOCS Press: Champaign, 1998.

(20) Rodrigues, C. E. C.; Reipert, E. C. D.; Souza, A. F.; Pessôa Filho, P. A.; Meirelles, A. J. A. Equilibrium Data for Systems Composed by Cottonseed Oil + Commercial Linoleic Acid + Ethanol + Water + Tocopherols at 298.2K. *Fluid Phase Equilib.* **2005**, *238*, 193–203.

(21) Antoniosi Filho, N. R.; Mendes, O. L.; Lanças, F. M. Computer Prediction of Triacylglycerol Composition of Vegetable Oils by HRGC. *Chromatographia* **1995**, *40*, 557–562.

(22) *IUPAC Standard Methods for the Analysis of Oils, Fats and Derivatives*, 6th ed.,Part 1 (sections I and II); Paquot, C., Ed.; Pergamon Press: NY, 1979.

(23) Rodrigues, C. E. C.; Antoniassi, R.; Meirelles, A. J. A. Equilibrium Data for the System Rice Bran Oil + Fatty Acids + Ethanol + Water at 298.2 K. *J. Chem. Eng. Data* **2003**, *48*, 367–373.

(24) Silva, L. H. M.; Coimbra, J. S.; Meirelles, A. J. A. Equilibrium Phase Behavior of Poly(Ethylene Glycol) + Potassium Phosphate + Water Two Phase Systems at Various pH and Temperatures. *J. Chem. Eng. Data* **1997**, *42*, 398–401.

(25) Geladi, P.; Kowalski, B. R. Partial Least Squares Regression: A Tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17.

(26) Brereton, R. G. Introduction to Multivariate Calibration in Analytical Chemistry. *Analyst* **2000**, *125*, 2125–2154.

(27) Otto, M. *Chemometrics: Statistics and Computer Application in Analytical Chemistry*; Wiley: New York, 1999.

(28) Marcilla, A.; Ruiz, F.; García, A. N. Liquid-Liquid-Solid Equilibria of the Quaternary System Water-Ethanol-Acetone-Sodium Chloride at 25 °C. *Fluid Phase Equilib.* **1995**, *112*, 273–289.

(29) Rodrigues, C. E. C.; Pessôa Filho, P. A.; Meirelles, A. J. A. Phase Equilibria for the System Rice Bran Oil + Fatty Acids + Ethanol + Water + γ-Oryzanol + Tocols. *Fluid Phase Equilib.* **2004**, *216*, 271–283.

(30) Gonçalves, C. B.; Meirelles, A. J. A. Liquid-Liquid Equilibrium Data for the System Palm Oil + Fatty Acids + Ethanol + Water at 318.2 K. *Fluid Phase Equilib.* **2004**, *221*, 139–150.

(31) Gonçalves, C. B.; Batista, E.; Meirelles, A. J. A. Liquid-Liquid Equilibrium Data for the System Corn Oil + Oleic Acid + Ethanol + Water at 298.15 K. *J. Chem. Eng. Data* **2002**, *47*, 416–420.

(32) Batista, E.; Monnerat, S.; Kato, K.; Stragevitch, L.; Meirelles, A. J. A. Liquid-Liquid Equilibrium for Systems of Canola Oil, Oleic Acid, and Short-Chain Alcohols. *J. Chem. Eng. Data* **1999**, *44*, 1360–1364.

(33) Pereira, M. A. P.; Aznar, M. (Liquid + liquid) Equilibrium of (Water + 2-Propanol + 1-Butanol + Salt) Systems at $T = 313.15$ K and $T = 353.15$ K: Experimental Data and Correlation. *J. Chem. Thermodyn.* **2006**, *38*, 672–677.

(34) Zhao, J.; Jiang, X.-C.; Li, C.-X.; Wang, Z.-H. Vapor Pressure Measurement for Binary and Ternary Systems Containing a Phosphoric Ionic Liquid. *Fluid Phase Equilib.* **2006**, *247*, 190–198.

(35) Rodrigues, C. E. C.; Filipini, A.; Meirelles, A. J. A. Phase Equilibrium for Systems Composed by High Unsaturated Vegetable Oils + Linoleic Acid + Ethanol + Water at 298.2 K. *J. Chem. Eng. Data* **2006**, *51*, 15–21.

(36) Demirel, Y.; Paksoy, H. O. Calculations of Thermodynamic Derivative Properties from the NRTL and UNIQUAC Models. *Thermochim. Acta* **1997**, *303*, 129–136.
(37) Lee, M.-J.; Chang, Y.-K.; Lin, H.-m. Separation of 4-Morpholinepropanesulfonic Acid from Its Aqueous Solution. *Ind. Eng. Chem. Res.* **1997**, *36*, 5399–5402.
(38) Stragevitch, L.; d'Ávila, S. G. Application of a Generalized Maximum Likelihood Method in the Reduction of Multicomponent Liquid-Liquid Equilibrium Data. *Braz. J. Chem. Eng.* **1997**, *14*, 41–52.

Received for review March 07, 2007. Accepted September 19, 2007. The authors wish to acknowledge FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, 04/09591-2 and 05/53095-2), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, 303649/2004-6), and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for their financial support.

JE700120D