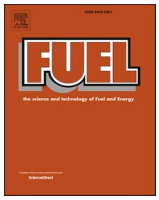
Fuel 319 (2022) 123794

Contents lists available at ScienceDirect

Fuel

journal homepage: www.elsevier.com/locate/fuel



Full Length Article

Near Infrared Spectroscopy: A useful technique for inline monitoring of the enzyme catalyzed biosynthesis of third-generation biodiesel from waste cooking oil

Josu López-Fernández [a], Desirèe Moya [b], María Dolors Benaiges [a], Francisco Valero [a,*], Manel Alcalà [b]

[a] Department of Chemical, Biological and Environmental Engineering, School of Engineering, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain
[b] Department of Chemistry, Faculty of Sciences, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

ARTICLE INFO

Keywords:
Biodiesel
Biocatalysis
Near infrared spectroscopy
Lipase
Komagataella phaffii
Inline monitoring

ABSTRACT

Near Infrared (NIR) spectroscopy was used for inline monitoring the enzymatic production of third-generation biodiesel from waste cooking oil (WCO) —with fatty acid profile similar to olive oil— under the principles of Continued Process Verification (CPV) and Process Analytical Technology (PAT). For this purpose, covalently immobilized mature *Rhizopus oryzae* lipase (rROL) was used to transesterify WCO with methanol and ethanol, firstly in 10 mL vials under orbital stirring and then it was successfully scaled up to a mechanically stirred 50 mL laboratory reactor specially designed for use of a NIR probe. Biocatalyst half-life and productivity after ten reaction cycles with methanol and ethanol were assessed. Slightly higher operational stability with methanol (337.5 h, 54 batches) vs with ethanol (146.7 h, 44 batches) was observed, but decreased productivity as a result of the increased reaction times used with the former (219 vs 327 μmol min^{-1}). The NIR spectroscopy results were highly correlated with those of gas chromatography (GC) used as reference. Thus, the root mean-square standard error of prediction (RMSEP) was 2.0% for methanol and 2.1% for ethanol. Therefore, NIR spectroscopy, which allows data acquisition in real time, is suitable for inline monitoring of enzymatic production of biodiesel.

1. Introduction

Circular economy is gaining momentum for reducing the environmental footprint of mankind and overcoming the widely adopted system grounded on linear resource use and waste generation. Using alternative and renewable sources of energy such as biogas, 2,5-dimethylfuran, hydrogen or biodiesel for transportation has been proposed to alleviate the current heavy dependence on fossil fuels [1–3]. The latter is a mixture of mono-alkyl esters of long-chain fatty acids that can be directly used by existing combustion engines and is arousing increasing interest by virtue of its being renewable and biodegradable, accelerating the carbon cycle and reducing net emissions of greenhouse gases. Biodiesel is obtained largely by transesterifying triacylglycerides with short-chain alcohols (ethanol and methanol, mainly). The resulting product is classified as first-, second- or third-generation biodiesel according to the oil substrate used. Due to the ethical dilemma of food vs fuel, public institutions are taking steps to shift production from first- to second- and third-generation biodiesel. The former is produced from non-edible oils such as radish and apricot kernel oils [4] while third-generation biodiesel is produced from substrates like waste cooking oil (WCO) in a manner that complies with the principles of circular economy [5,6].

At present, biodiesel is industrially produced mainly by chemical transesterification of first-generation substrates in the presence of a homogeneous basic catalyst. Using second- or third-generation substrates requires their pre-treatment to remove free fatty acids (FFAs), which they typically contain in high proportions [7,8]. In fact, unless properly neutralized, FFAs favor saponification and give biodiesel of decreased quality requiring further downstream work. Enzymatic transesterification with lipases (glycerol ester hydrolases, E.C.3.1.1.3) has emerged as a solution to avoid unwanted side reactions such as saponification and hence additional downstream processing. Although alternatives such as using acids and heterogeneous solid catalysts have been proposed, most are unfeasible or result in low reaction rates [9,10]. Using lipases for transesterification has additional advantages, such as more modest energy requirements from milder reaction conditions

* Corresponding author.
E-mail address: francisco.valero@uab.cat (F. Valero).

https://doi.org/10.1016/j.fuel.2022.123794
Received 22 December 2021; Received in revised form 4 March 2022; Accepted 5 March 2022
Available online 12 March 2022

—lower reaction temperatures, close to 30 °C, are required instead of the higher temperatures, usually over 50 °C, needed for chemical catalysts [11]—, a simplified flowsheet for purification [12] and the need for no organic solvents, which makes processes compliant with the principles of green chemistry [13,14].

Rhizopus oryzae lipase (ROL) has been extensively studied as a biocatalyst for biodiesel production on the grounds of its 1,3-regiospecificity preventing formation of glycerol during the transesterification reaction [15] —this alcohol is deemed a residue owing to its worldwide large scale production [16]. Instead, ROL leads to the formation of 2-monoacylglycerol, which is a by-product with a high added value by virtue of its capacity to increase biodiesel quality and its being useful for the pharmaceutical and food industries [17–19]. Although ROL is a stable biocatalyst, it is generally used in immobilized form to improve its industrial suitability. Immobilizing the enzyme allows it to be reused and usually increases its stability. ROL for biodiesel production has so far been immobilized in various ways including covalent binding, adsorption or the use of whole cells [15].

During chemically or enzymatically catalyzed transesterification for biodiesel production, the reaction medium contains a mixture of fatty acid methyl and ethyl esters (FAM/EE), FFA and tri-, di- and monoacylglycerols. Production performance is usually assessed by gas chromatography (GC), using an offline and sluggish analytical procedure [20,21]. Moreover, low volatile substances can damage capillary columns and other elements of the chromatographic system, so a number of alternative techniques including high-performance liquid chromatography (HPLC) [22] and proton nuclear magnetic resonance (^{1}H NMR) spectroscopy [23] are being increasingly used instead. Chromatographic techniques, however, are expensive and time-consuming; also, they cannot operate inline, which prevents timely decision-making during the process and can increase production costs by requiring stops —or even batch reprocessing if the target quality is not achieved [24]. The principles of Continued Process Verification (CPV) have made bioprocess automatization and Process Analytical Technology (PAT) increasingly attractive; also, they have promoted the use of analytical methods allowing real-time quality assessment, adoption of corrective measures and a better understanding of bioprocesses [25,26]. Such methods require in- or online monitoring of the process, and hence using, for instance, spectroscopic techniques (Raman, fluorescence, UV–Vis, IR, NIR) in combination with multivariate calibration [27].

NIR spectroscopy is an extensively studied analytical technique based on the interaction of matter and light radiation in the wavelength region from 780 to 2500 nm that affords multicomponent, fast, reliable, inexpensive and non-destructive analysis [20]. Besides, it avoids the need for sample withdrawal when used inline, waste production, and the need for complex pre-treatments of samples with solvents or other chemicals, all of which makes it a safe, clean, energy-saving choice fully compliant with the principles of green chemistry [28]. NIR spectra are complex and possess broad overlapping bands that require special mathematical procedures to accurately interpret spectra and understand the results, such as principal component analysis (PCA) or partial least-squares (PLS) regression [29]. NIR spectroscopy has so far been successfully used by the biodiesel industry to assess the quality or properties of biofuel/diesel blends [29,30], and also for inline monitoring of chemically catalyzed transesterification reactions [20,31,32].

In this work, the mature sequence of ROL (rROL) heterologously produced in the methylotrophic yeast *Komagataella phaffii* (*Pichia pastoris*) was immobilized onto a suitable support in order to catalyze the transesterification of waste cooking oil (WCO) in a solvent-free system with ethanol or methanol as acyl-acceptor. The main novelty of this work is the scale up of the transesterification reaction to a laboratory-scale reactor specially designed for use of an inline near infrared spectroscopy probe for monitoring the enzyme catalyzed transesterification reaction. The results were compared with those obtained by gas chromatography as reference in order to confirm the suitability of the NIR technique for accurate real-time monitoring of transesterification under the principles of CPV and PAT while avoiding the environmental and economic costs of withdrawing samples during the process. Unlike previous works on the NIR based monitoring of biodiesel production, the fact of performing a stepwise addition of alcohol and the employment of immobilized enzyme in this case might generate background noise in the spectra, making the NIR monitoring more challenging. Besides, the biocatalyst operational stability and mechanical strength were evaluated and compared with previously reported data.

2. Materials and methods

2.1. Materials

Waste cooking oil (WCO) was obtained from a local public waste management company and centrifuged prior to use. Polymethacrylate matrix support D6308 was kindly donated by Purolite® (King of Prussia, PA, USA). The colorimetric kit (11821729) for the enzymatic assay was supplied by Roche (Mannheim, Deutschland). Heptane, ethanol and methanol were purchased from Panreac (Barcelona, Spain). Standards of methyl/ethyl palmitate, methyl/ethyl stearate, methyl/ethyl oleate, methyl/ethyl linoleate and methyl linolenate, and all unstated reagents, were obtained from Sigma–Aldrich (St. Louis, MO, USA).

2.2. Lipase heterologous production

The mature sequence of *Rhizopus oryzae* lipase (rROL) was heterologously produced in the methylotrophic yeast *Komagataella phaffii* by the Bioprocess Engineering and Applied Biocatalysis Group of the Universitat Autònoma de Barcelona (Barcelona, Spain) as described elsewhere [33]. After fermentation, the culture broth was centrifuged, microfiltered, ultrafiltered and freeze-dried to remove biomass and concentrate the enzyme [34].

2.3. Lipase activity measurement

Enzymatic activity was determined by using the Roche lipase colorimetric kit in 200 mM Tris-HCl buffer at pH 7.25 at 30 °C on a Cary 300 spectrophotometer from Varian (Mulgrave, VIC, Australia) operating at 580 nm. All measurements were made in triplicate following an already published procedure in which one unit of lipolytic activity was defined as the amount of lipase necessary to hydrolyze 1 μmol of ester bond per minute under assay conditions [35].

2.4. Support functionalization and lipase immobilization

Polymethacrylate-matrix Purolite® D6308 with epoxide and octadecyl surface groups (EO) was functionalized as described elsewhere [10]. Briefly, epoxide functional groups were pre-treated by incubating 1 g of support with 1 M ethylenediamine solution at pH 10 under orbital stirring at 60 °C for 4 h. Then, the support was rinsed with distilled water by vacuum filtration and incubated in a 2.5% w/v glutaraldehyde solution in phosphate buffer at pH 8 on a roller at room temperature for 2 h. Finally, the glutaraldehyde-treated support was rinsed for incubation with a 3500 AU mL^{-1} lipase solution at 4 °C for 42 h. The biocatalyst thus obtained (EO-rROL) was dried on silica gel and stored at –20 °C until use.

The specific activity of the biocatalyst was calculated as the difference between those in the final blank and supernatant solutions divided by the final dry weight of enzyme.

2.5. Transesterification reactions

Transesterification was done in closed 10 mL vials containing 8 g of WCO at 30 °C under orbital stirring at 350 rpm [14]. The acyl-acceptor was added in 5 (ethanol) or 10 pulses (methanol) by splitting the stoichiometric volume of alcohol (2:1 alcohol:oil mixture) during the

Table 1
Percent fatty acid composition profile for sunflower, olive and waste cooking oil (WCO).

Acid	Sunflower oil	Olive oil	WCO
Stearic	2.8	2.3	2.46 ± 0.04
Oleic	28.0	66.4	70.78 ± 0.38
Linoleic	62.2	16.4	10.21 ± 0.01
Linolenic	0.16	1.6	2.45 ± 0.02

reaction in order to reach the theoretical maximum yield (200 min reaction for ethanol and 375 min for methanol) [17]. All substrates and biocatalysts were previously equilibrated for water activity by using a saturated KOH solution ($a_w = 0.093$) for a minimum of 16 h. The reaction was scaled up to a 50 mL HME-R mini reactor from Scharlab (Sentmenat, Barcelona, Spain) as described above but using a final oil mass of 40 g and mechanical stirring at 700 rpm.

The minireactor lid and tank were specially designed in collaboration with Scharlab (Supplementary Material 1). Thus, the lid was custom-made to include an inlet that allowed the NIR probe to be inserted in the reaction medium, and a second, slightly curved inlet for the stirrer, so that both could reach the reaction medium without colliding. Besides, the stirrer was modified by using an additional propeller —the two were "marine propellers"— to facilitate generation of a turbulent regime in the reactor.

Operational stability was assessed by allowing the biocatalyst to settle in the bottom of the reactor. Then, depleted reaction medium was removed and all components were prepared for the next run [17]. Relative yields of consecutive transesterification reactions cycles were used as described elsewhere to fit the results to a first-order exponential decay equation (Eq. (1)) and a two-component first-order exponential decay equation (Eq. (2)), using the software Sigma Plot v 14 [36–38].

$$Y(\%)_t = 100e^{-kt} \tag{1}$$

$$Y(\%)_t = 100e^{-k_1 t} + ce^{-k_2 t} \tag{2}$$

where k, k_1 and k_2 are deactivation coefficients.

2.6. Gas chromatography (GC) analysis

Fatty acid ethyl esters (FAEEs) and fatty acid methyl esters (FAMEs) were quantified on a model 7890A gas chromatograph from Agilent (Santa Clara, CA, USA) equipped with a 19095N-123 capillary column and an autosampler [39]. Relative standard deviations (RSD) never exceeded 3%.

2.7. Fatty acid composition

The fatty acid composition of WCO was established by adapting a previously reported procedure of chemically catalyzed transesterification [40]. Briefly, 40 g of oil was heated in an HME-R minireactor commercially available from Scharlab (i.e., with the original tank and lid) at 65 °C under mechanical stirring at 350 rpm. Then, KOH (1 % w/w) and methanol (6:1 alcohol/oil mixture) were added and the reaction was allowed to develop for 6 h with periodic sample withdrawal for GC analysis and calculation of the relative fatty acid composition.

2.8. Acquisition and processing of near infrared (NIR) spectra

The transesterification reactions were monitored by recording NIR spectra at 5 min intervals in each reaction cycle [28]. Spectra were acquired in the transflectance mode, using a Model 5000 spectrophotometer from FOSS NIRSystems (Silver Spring, MD, USA) equipped with an immersive optical probe. A reference spectrum for air was obtained before the reaction. The wavelength range scanned was 1100–2498 nm, the spectral resolution was 2 nm and the optical path length was 1 mm. Raw absorbance spectra were exported to NSAS file format by using the software Vision 2.51, then transformed to Matlab file format with The Unscrambler 10.3 (Camo Analytics, Norway) and finally processed for viewing, exploration and multivariate modeling with the software Solo (Eigenvector Research Inc., Wenatchee, WA, USA).

First- and second-derivative spectra were obtained by using the Savitzky–Golay algorithm with a 15-point moving window and a second-order polynomial. Mean-centered data (X-matrix for spectra) were subjected to principal component analysis (PCA). Partial least-squares (PLS) calibration models for mean-centered data (X-matrix for spectra and Y-matrix for reaction yield) were constructed by cross-validation, using the leave-one-out method. Samples for inclusion in the calibration and prediction sets were selected by using the Kennard–Stone method [41]. The quality of the PLS models and their predictive ability were assessed in terms of the root mean square error of calibration (RMSEC) and prediction (RMSEP), defined as:.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i^{pred} - Y_i^{ref}\right)^2}{n}} \tag{3}$$

where n is the number of samples used, Y^{ref} the reaction yield provided by the reference method and Y^{pred} that estimated by the NIR model.

The number of PLS factors required to define each model was chosen from the minimum of a plot of RMSEP vs number of factors. The predictive ability was also assessed by statistical evaluation of the least-squares regression line between the reference GC reaction yield and the NIR-predicted value.

The aim of quantitative multivariate modeling is reducing prediction errors by using the simplest possible model (i.e., that with the fewest factors). PLS models were evaluated over wide or narrow spectral intervals selected from different calculations, namely: regression coefficients, X-loadings, XY correlation vector, variable importance in projection and selectivity ratio. These tools allow simple numerical assessment of the usefulness of each X-variable in a regression model. The higher were the results obtained with these calculations over the significant threshold, the greater was the usefulness of the variables for regression and prediction. Selection of the spectral interval was combined with different spectral processing methods, first in the absorbance mode, and then as their first and second derivatives, and also upon standard normal variate SNV-based scaling. The combination of SNV followed by derivatization was also tested.

3. Results and discussion

3.1. Fatty acid composition of the waste cooking oil

Local public waste management companies typically collect and mix large amounts of waste cooking oil (WCO). As a result, WCO composition can vary widely depending on consumption patterns and regional or local cooking traditions. In Spain, most oil based cooking is done with olive or sunflower oil, which leave vast amounts of oleaginous waste [42].

Because it was obtained from a public waste management company not implementing traceability, the oil used here was of unknown origin. Analysis of its fatty acid composition was performed as described in Section 2.6 to compare it with published profile of olive and sunflower oil and results exhibited that the fatty acid profile for the WCO was similar to that for olive oil (Table 1) [43]. The acidity of WCO, which is another relevant factor having influence on the performance of the biocatalyst during transesterification, has already been analyzed in a previous publication exhibiting a low acidity value (0.77%) suggesting that oil was not extensively cooked [14].

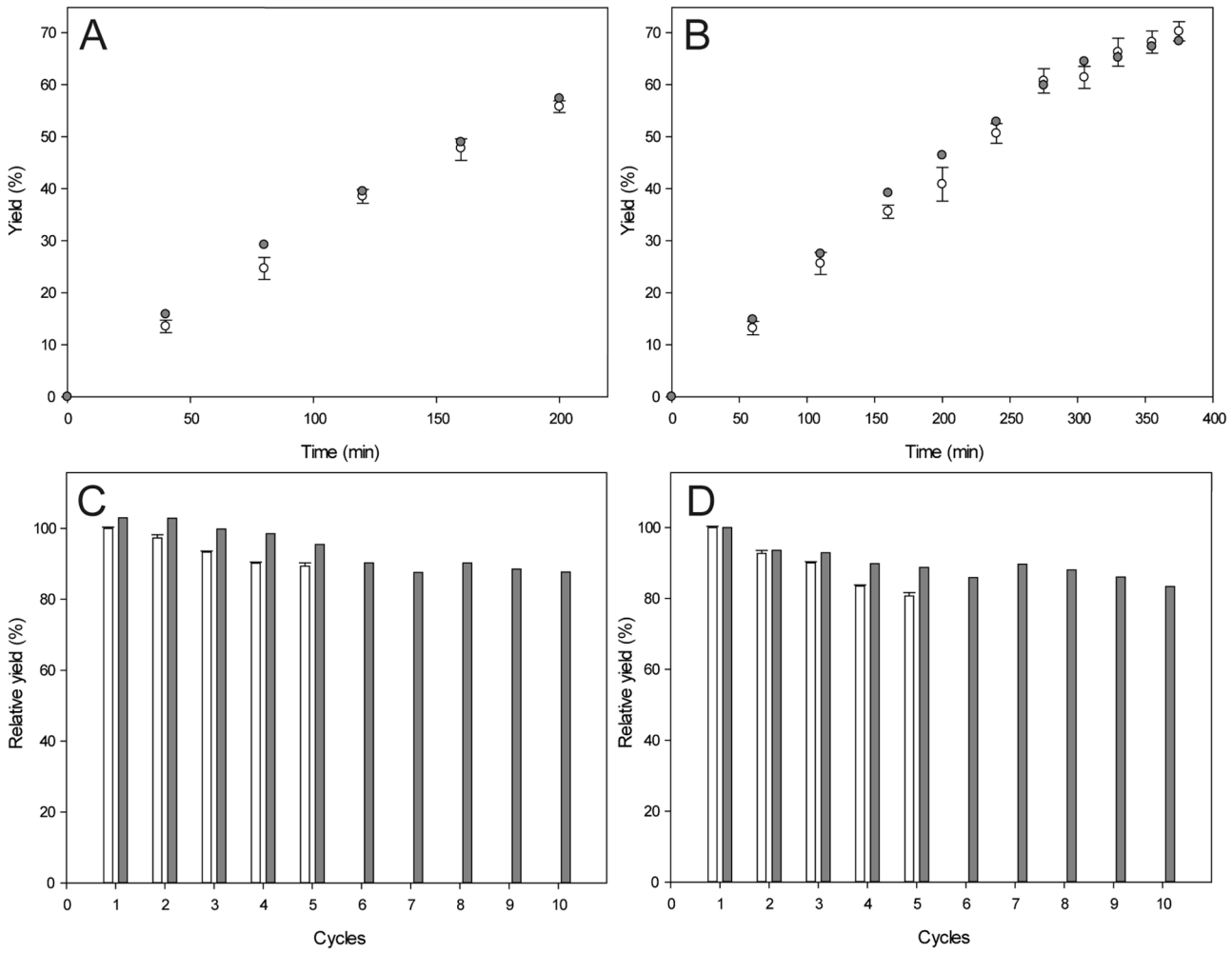


Fig. 1. Yield profile for WCO transesterification with 5 pulses of ethanol (A) and 10 of methanol (B) in the presence of EO-rROL. Relative yield of consecutive transesterification cycles with 5 pulses of ethanol (C) and 10 of methanol (D). White bars and points correspond to the reaction in 10 mL vials and grey bars and points to that in the laboratory minireactor. Relative yields were calculated against that of the first cycle (100%).

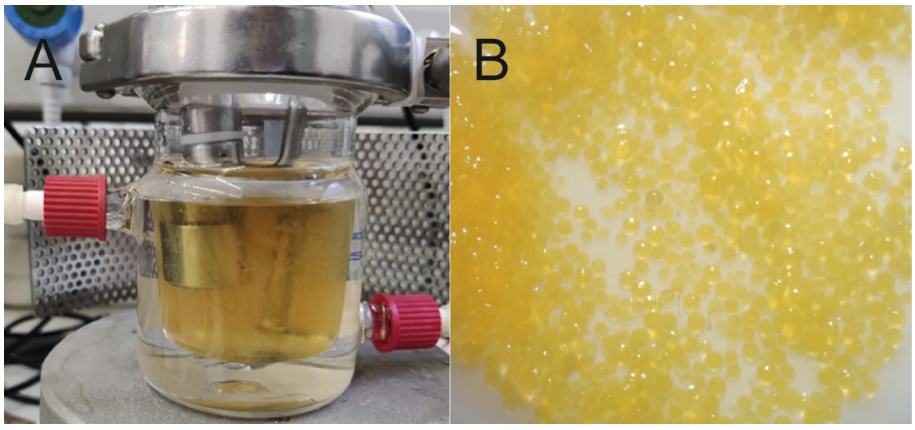


Fig. 2. (A) Customized reactor for use of the NIR probe. (B) The EO-rROL biocatalyst after 10 consecutive reaction cycles as seen under a binocular loupe.

3.2. Transesterification reaction: scale-up, productivity and half-life

The transesterification reactions initially conducted in 10 mL vials under orbital stirring were successfully scaled up to a 50 mL laboratory reactor under mechanical stirring (mimicking industrial large-scale conditions), which was specially designed for use with a NIR probe. Already optimized conditions, regarding the biocatalyst and the alcohol addition strategy for biodiesel production from WCO were employed [14], and as can be seen from Fig. 1A and B, the yield profiles for the first reaction batch were almost identical with ethanol and methanol in both the vial and the reactor —because of the 1,3-regioespecificity of ROL, the theoretical maximum yield in the absence of acyl-migration phenomena is 66%, resulting in the formation of the corresponding alkyl esters and 2-monoacylglycerol [44]. The results therefore suggest that

Table 2
EO-rROL productivity and half-life in biodiesel production with ethanol and methanol as calculated by fitting to a first-order exponential decay (Eq. (1)) and two-component first-order exponential decay (Eq. (2)) deactivation model.

Acyl acceptor	Model equation	R^2	Half-life (h) *	Productivity (μmol min^{-1})
Ethanol	1	0.87	170 [51]	327
	2	0.96	146.7 [44]	
Methanol	1	0.87	225 [36]	219
	2	0.95	337.5 [54]	

* The numbers in square brackets represent the number of cycles for those of reaction hours.

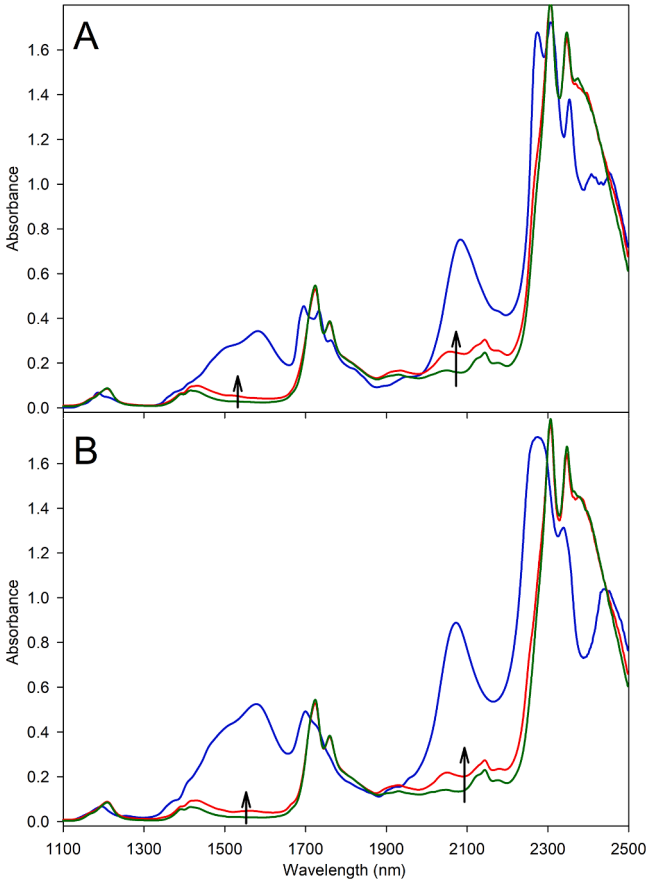


Fig. 3. NIR absorbance spectra recorded at the start (green) and end (red) of the reaction as compared with those for each pure alcohol (blue): ethanol (A) and methanol (B). The arrows represent evolution of the reactions.

mass transfer and reaction performance in the vials and the reactor were essentially identical under the employed conditions, and also that the presence of the NIR probe did not detract from homogeneity in the reaction medium. Thus, the 50 mL reactor with a customized lid and tank, and a second stirrer (Fig. 2A), reproducibly echoed the results obtained with 10 mL vials. In fact, both systems led to similar relative yields after 5 consecutive cycles (Fig. 1C and 1D).

Biodiesel productivity with each alcohol was calculated as the combined figure for all 10 cycles performed in the laboratory-scale reactor. As expected, FAEE productivity (ethanol as acyl-acceptor) was 1.5 times higher than FAME productivity (methanol as acyl-acceptor) as a result of the longer reaction times required for transesterification with the latter alcohol (Table 2), even though, the final transesterification yields with methanol were slightly higher than those obtained with ethanol (Fig. 1A and B). In fact, splitting the amount of alcohol used into 10 pulses is unproductive with alcohols such as ethanol —it has scarce

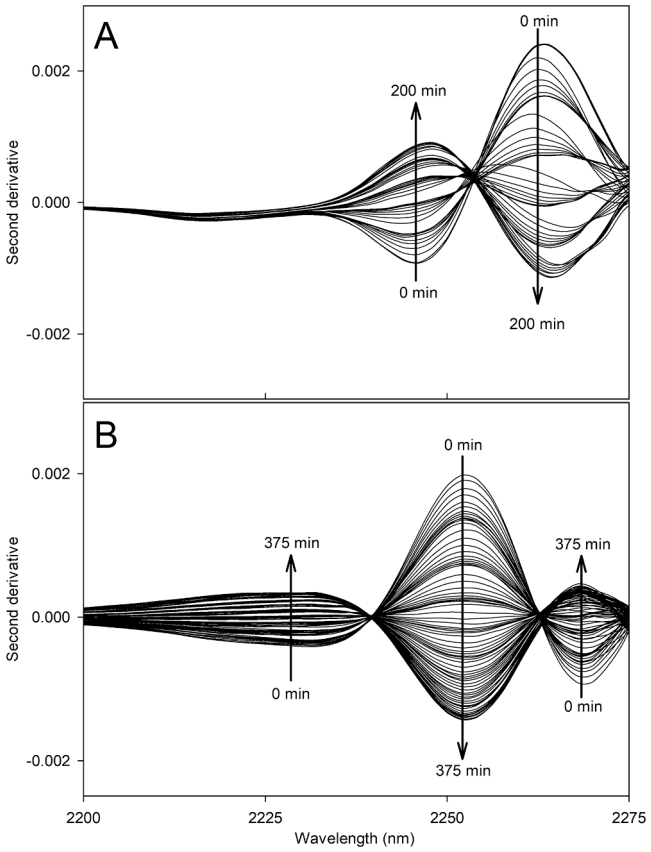


Fig. 4. Second-derivative NIR spectra obtained in the first reaction batch with (A) ethanol and (B) methanol as acyl-acceptor. The arrows represent reaction time.

Table 3
Dimensions of the *X*-dataset and *Y*-dataset.

	Reaction with ethanol	Reaction with methanol
Reaction time (min)	200	375
***X*-matrix** (1100–2350 nm)	322 rows and 626 columns	684 rows and 626 columns
***Y*-matrix** (Reaction yield %)	123 rows and 1 column	243 rows and 1 column

negative impact on the operational stability of the biocatalyst— but can be useful with others such as methanol with severe influence on enzyme stability [17,45]. Thus, the biocatalyst half-life in hours with methanol as acyl-acceptor was about 1.5–2.5 times higher than it was with ethanol as a result of the amount of methanol used being split into more pulses than that of ethanol (Table 2). However, when reaction cycles corresponding to these hours were compared, half-lives ($t_{1/2}$ values into square brackets in Table 2) tended to be similar with both alcohols as a consequence of the reaction times differing between the two.

ROL has exhibited widely variable operational stability in biodiesel production from various substrates. For instance, ROL covalently immobilized onto Relizyme^{TM} OD403 was used to transesterify olive pomace oil with methanol and the relative yield found to decrease to 60% after 26.7 h reaction (7 batches) [46]. In another study with soybean oil, the transesterification yield with glutaraldehyde cross-linked whole-cell biocatalysts decreased from 84 to 65% after 2520 h reaction (35 batches) [47]. By contrast, the ROL catalyzed transesterification of WCO has been the subject of little study, so comparison with the previous results is challenging [15]. Therefore, although a number of half-life values for ROL based biocatalysts used in biodiesel transesterification reactions have been reported with several substrates, comparison with WCO would be more accurate since this substrate may

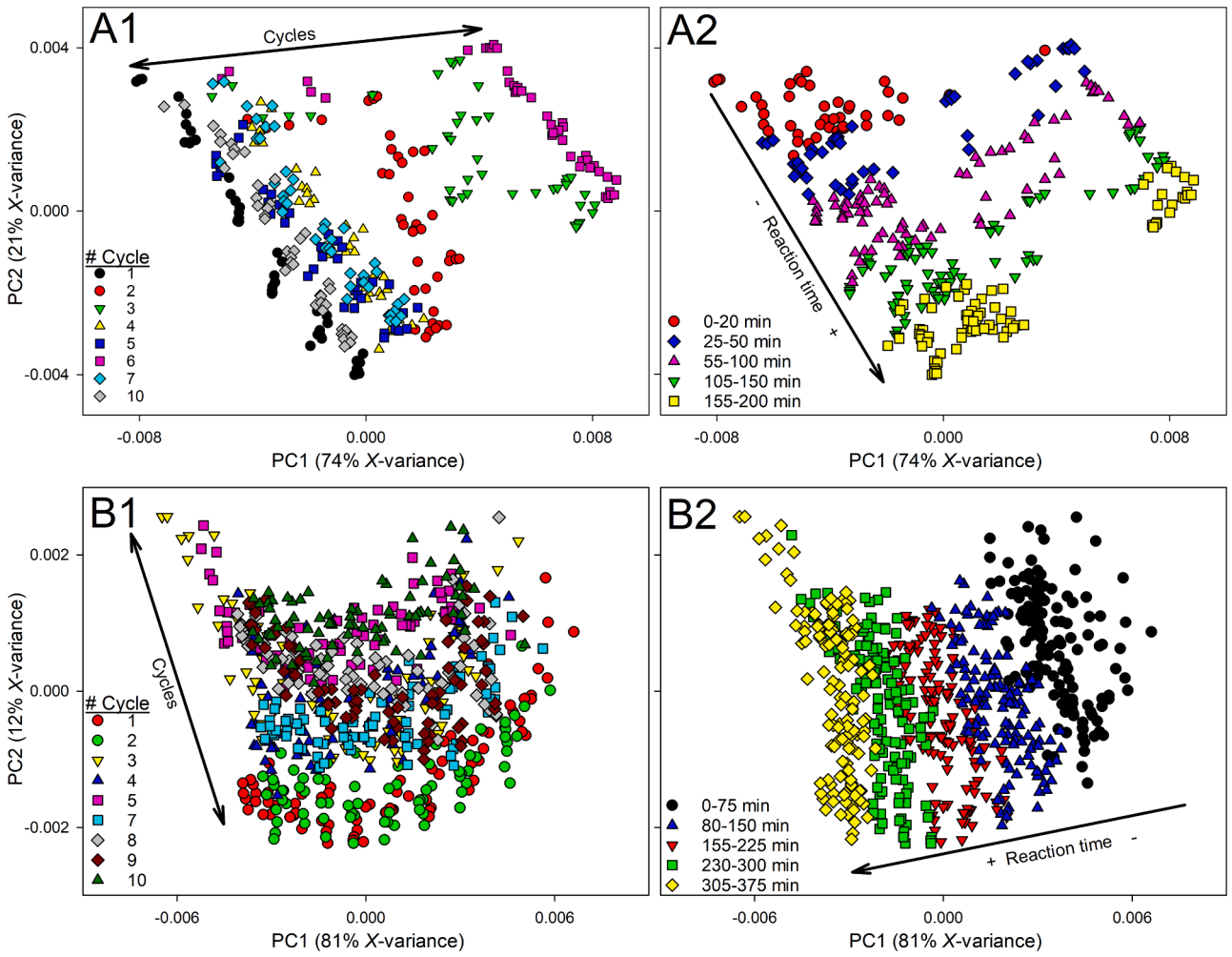


Fig. 5. PCA scores plot constructed from second-derivative spectra over the range 1100–2350 nm obtained with ethanol (A1 and A2) and methanol as acyl-acceptor (B1 and B2). The 1 and 2 subplots stand for different symbol related to cycles and reaction time, respectively.

deactivate the enzyme more markedly as the likely result of its containing detrimental components such as phenolics or varying widely in acidity [10,48,49].

The biocatalysts were imaged with a binocular loupe after the whole sequence of reaction with each acyl-acceptors in order to check whether they had retained their structure without fracturing by effect of mechanical stress under agitation (Fig. 2B). No substantial breaks were detected after the ten cycles, which confirms the strength of the support and its suitability for industrial use, as indicated by the supplier (Purolite®).

3.3. Near infrared (NIR) spectra: A useful tool for the biodiesel industry

Once the biocatalyst suitability for biodiesel production under conditions mimicking industrial work was confirmed, the development of a system enabling inline monitoring and compliance with CPV and PAT principles was sought. In this sense, NIR spectroscopy was employed by inserting a probe in a customized reactor in combination with two different calibration models.

3.3.1. Spectral data

Absorbance spectra acquired across the NIR spectral range (1100–2498 nm) were visually inspected to identify gross outliers and noisy spectral regions. The range from 2352 to 2498 nm exhibited little variability and considerable noise typical of fiber optic probes, so only that from 1100 to 2350 nm was used. Fig. 3 shows the first and last spectra obtained in a reaction run with ethanol and methanol as compared with those for the pure alcohols. Ethanol and methanol (Fig. 3, blue plots) exhibited the following major spectral bands: 1425–1475 nm (OH first overtone), 1650–1750 nm (CH first overtone), 2000–2100 nm (OH combinations) and 2200–2450 nm (CH + CH combinations and CC + CH combinations). Expectedly, methanol spectra presented less bands than that of ethanol at 1650–1750 nm and 2200–2450 nm due to the absence of CH_2; the bands corresponding with CH first overtone, CH + CH and CH + CC combinations are only related with the CH_3 group from methanol. The spectra at the start of the reaction (Fig. 3, green plots) exhibited the following spectral bands: high absorbance at 1650–1750 nm (CH first overtone), very low absorbance at 1875–1950 nm (CO second overtone), low absorbance at 2000–2100 nm (OH combinations) and very high absorbance at 2200–2450 nm (CH + CH and CC + CH combinations). The spectra at the end of the reaction (Fig. 3, red plots) showed higher absorbance than that of the starting point almost throughout the whole studied spectral range (1100–2400 nm). The final yield after the two batches showed in Fig. 3 was 61.0% with ethanol and 65.9% with methanol.

The typically extensive overlap of bands and fairly low resolution of the NIR technique required increasing resolution by using a spectral derivative treatment. The spectral ranges 1400–1700 nm and 1900–2300 nm were adequate to visualize the spectral changes during the reaction because of the low absorbance of the WCO and high absorbance of the alcohols. Fig. 4 shows the second derivative spectra of the first cycle with ethanol (Fig. 4A) and with methanol (Fig. 4B), in the

Table 4
Number of spectra used to construct the calibration and prediction sets, figures of merit of the calibration models and predictive ability.

	Reaction with ethanol	Reaction with methanol
Calibration		
Number of spectra	24 (4 cycles)	27 (3 cycles)
PLS factors	2	2
Yield range (%)	0.0–61.0	6.2–68.9
Cumulative *Y*-variance explained		
Factor 1	66.9	91.4
Factor 2	98.1	99.6
NIR vs GC regression		
Slope	0.98 ± 0.06	0.99 ± 0.03
Intercept	0.6 ± 2.1	0.2 ± 1.2
Correlation coefficient (*r*)	0.990	0.998
RMSEC (%)	2.4	1.3
Prediction		
Number of spectra (total)	298 (8 cycles)	657 (9 cycles)
Number of spectra (with GC reaction yield value)	99	203
Number of spectra (without GC reaction yield value)	199	454
Yield range (%)	0.0–59.4	6.3–68.8
NIR vs GC regression		
Slope	0.98 ± 0.03	0.99 ± 0.01
Intercept	1.0 ± 1.1	–0.3 ± 0.6
Correlation coefficient (*r*)	0.989	0.996
RMSEP (%)	2.1	2.0

wavelength range between 2200 and 2275 nm. As can be seen, the bands were highly ordered in terms of yield and reaction time. This result is also observed throughout the entire wavelength range of 1100–2350 nm.

Table 3 summarizes the dimensions of the *X*- and *Y*- datasets. The number of rows coincided with that of measured spectra and the number of columns with that of variables (wavelengths and reaction yield). The reaction with methanol took almost twice as long as that with ethanol (375 min vs 200 min). Since the spectrum acquisition frequency was identical with both alcohols (1 spectrum every 5 min), the available *X*-matrix for the reaction with methanol contained many more rows than that for ethanol (684 vs 322). The sampling frequency for GC analysis was approximately 1 sample every 15 min. As a result, the number of rows in the *Y*-matrix for which a yield value was available was roughly one-third of all (viz., 123 out of 322 for ethanol and 243 out of 684 for methanol).

Data from second-derivative spectra spanning the range 1100–2350 nm were initially explored through PCA method. One PCA for each alcohol was calculated by using the whole *X*-matrix. The scores plot of the second principal component against the first (PC2 vs PC1) explained approximately 95% of the variance in *X* (Fig. 5). Fig. 5A were calculated from the reactions with ethanol, and 5B with methanol. There were two main sources of spectral variability, namely: reaction time and reaction cycle. For the reaction with ethanol (5A1 and 5A2), the reaction cycle is mostly explained by the PC1, accounting for 81% of the *X*-variance, and the reaction time by the PC2 (21%). For the reaction with methanol (5B1 and 5B2), the reaction time is mostly explained by the PC1 (74%) and the reaction cycle by the PC2 (12%). Neither standard normal variate (SNV) scaling nor first-derivative treatment, whether individually or in combination, provided better results for the reaction with ethanol —at least not on a par with the scores plots for methanol. The spectral variability of the reaction cycle depended on the normal laboratory conditions of sample and instrument.

3.3.2. Calibration

The Kennard–Stone method selects a subset of calibration samples which provide a very uniformly distributed network of selected points over the dataset and includes samples on its boundary. Table 4 shows the number of spectra used for calibration and prediction. The calibration set for the ethanol reaction included 24 samples from reaction cycles 1, 2, 3 and 6, which spanned an FAEE yield range of 0.0–61.0%, while the set for methanol comprised 27 samples from reaction cycles 2, 3 and 10, with an FAME yield of 6.2–68.9%. No spectrum for a near-zero yield was available for methanol owing to instrumental issues arising at the beginning of the reaction. All other spectra were included in the prediction set.

Reaction yields were quantified by using PLS calibration models constructed from independent calibration and prediction sets of NIR spectra. Models spanning a narrow spectral range failed to reduce calibration or prediction errors relative to the whole spectral range. Also, narrow ranges failed to reduce the number of factors required by each model. Second-derivative spectra proved the best choice in any case.

Table 4 summarizes the calibration and prediction results obtained with the PLS model for each alcohol. Both were constructed with two factors and explained a cumulative *Y*-variance higher than 98%. The only difference between the two was the *Y*-variance captured by the first factor, which was 66.9% with ethanol and 91.4% with methanol. This was a result of the distribution of PCA scores and the two sources of variability observed (reaction time and reaction cycle; Fig. 5). The *Y*-variable (yield) was explained mainly by PC1 in the reaction with methanol, and by both PC1 and PC2 in the reaction with ethanol. The model using PC1 alone was more robust for methanol than it was for ethanol. With PC1 and PC2 jointly, however, the two models captured almost the same *Y*-variance (98.1% for ethanol and 99.6% for methanol). The upper range of biodiesel yield for the reaction with methanol (68.9%) was higher by effect of the increased experimental reaction yield. Calibration errors (RMSEC), expressed in the same units as yield, were 2.4% for ethanol and 1.3% for methanol and the correlation coefficients were higher than 0.99.

The prediction results of Table 4 were obtained from pure independent sets which were not employed during calibration. Regression lines for NIR yield versus GC yield plots were evaluated through different tests and analysis. In this sense, the correlation coefficients were high (viz., 0.989 and 0.996 for ethanol and methanol, respectively). The residuals were randomly distributed around 0 and their normal distribution was demonstrated with the Anderson-Darling, Ryan-Joiner and Kolmogorov-Smirnov tests (95% of significance). The Mandel test for linearity was calculated for the two regressions and the F-tests (95% of significance) concluded that quadratic models do not performe better than linear ones. The absence of significant differences between the two techniques was demonstrated through the t-tests (95% of significance) of the slope

Table 5
RMSEP (%) value for each reaction cycle.

RMSEP	Cycle										Overall
	1	2	3	4	5	6	7	8	9	10	
Reaction with ethanol	2.4	2.0	2.7	2.1	1.9	2.1	1.6	N/A	N/A	1.6	2.1
Reaction with methanol	2.0	1.0	1.9	2.1	2.1	N/A	2.5	1.9	2.0	2.3	2.0

N/A: not available.

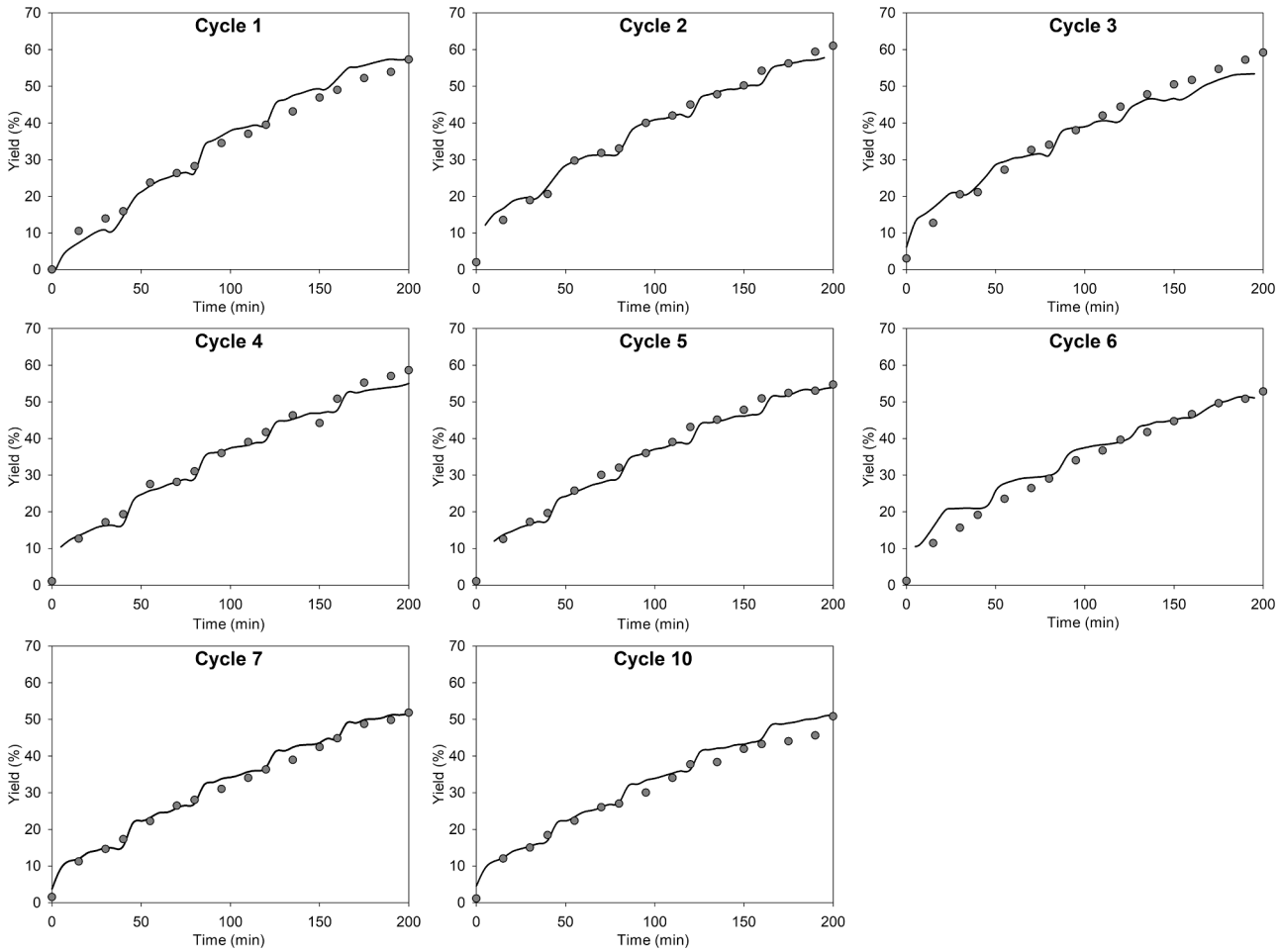


Fig. 6. NIR prediction of reaction yields (solid line) with ethanol as acyl-acceptor. Dots represent reference GC yield values.

and intercept of the NIR versus GC regressions. In fact, the slopes and intercepts were not significantly different from 1 and 0, respectively, which highlights the accuracy of the NIR method. Finally, the satisfactory predictive ability was also assessed with the low RMSEP prediction errors (2.1% for ethanol and 2.0% for methanol). As can be seen in Table 5, RMSEP for the individual reaction cycles ranged from 1.0 to 2.7%. Although RMSEC for the methanol model was lower than for the ethanol model (1.3% vs 2.4%), there were virtually no differences in predictive ability between the two on an individual batch basis.

3.3.3. Validation

Figures 6 and 7 show the NIR profiles of yield vs time used to validate the results of the reaction cycles with the two alcohols. The figures represent the NIR predictions from the whole set of prediction spectra (with or without a matching reaction yield obtained by GC), namely: 298 spectra for ethanol and 657 for methanol. These spectra were distributed among the 8 reaction cycles examined for ethanol and the 9 for methanol. As can be seen, there was high correlation between the NIR and GC results. Uncertainty in the prediction of each spectrum was assessed via the estimated error, which is a measure of goodness of prediction and calibration error, and was used in reaction yield units. This error measure is software specific (Solo) and it uses the equation 9 of a previously reported work [50]. The 298 NIR predictions of the reaction with ethanol were subject to an estimated average error of 2.94% (min = 2.84, max = 3.13), while the 657 NIR predictions of the reaction with methanol had an estimated average error of 1.51% (min = 1.47, max = 1.65). The magnitude of the error was not constant throughout the reaction, however; rather, it was smallest in the middle of each batch, and peaked at the beginning and end. This result is consistent with the typical distribution of interpolation errors observed with least-squares regression. The PCA scores plot (Fig. 5) affected the larger estimated error for the reaction with ethanol compared with the reaction with methanol. The *Y*-variable (yield) in the reaction with ethanol was explained by both PC1 and PC2. This led to a PLS model for the reaction of ethanol with a *Y*-explained variance of the first factor of 66.9% (Table 4). The yield in the reaction with methanol was explained mostly by PC1 and the *Y*-explained variance of the PLS model was 91.4%. The PLS for both reactions used two factors, achieving 98.1% and 99.6% of the *Y*-explained variance. The previous results testify to the robustness of NIR spectroscopy for monitoring biodiesel production by enzymatic transesterification with ethanol or methanol as acyl-acceptor.

4. Conclusions

Waste cooking oil (WCO), which based on its fatty acid composition was probably of olive oil origin, was used to obtain third-generation biodiesel with immobilized *Rhizopus oryzae* lipase and to successfully scale up the reaction from orbitally stirred 10 mL vials to a mechanically stirred 50 mL laboratory reactor obtaining similar yields and operational stability. EO-rROL exhibited a high operational stability in terms of half-life ($t_{1/2}$), over 35 reaction batches with both ethanol and methanol. However, methanol-based reaction showed decreased productivity as a result of the increased reaction times (219 vs 327 μmol min^{-1}). Besides, transesterification reaction with both alcohols was monitored by NIR spectroscopy in the 50 mL reactor. The GC and NIR results were highly correlated, with a prediction error (RMSEP) of 2.0% for methanol and

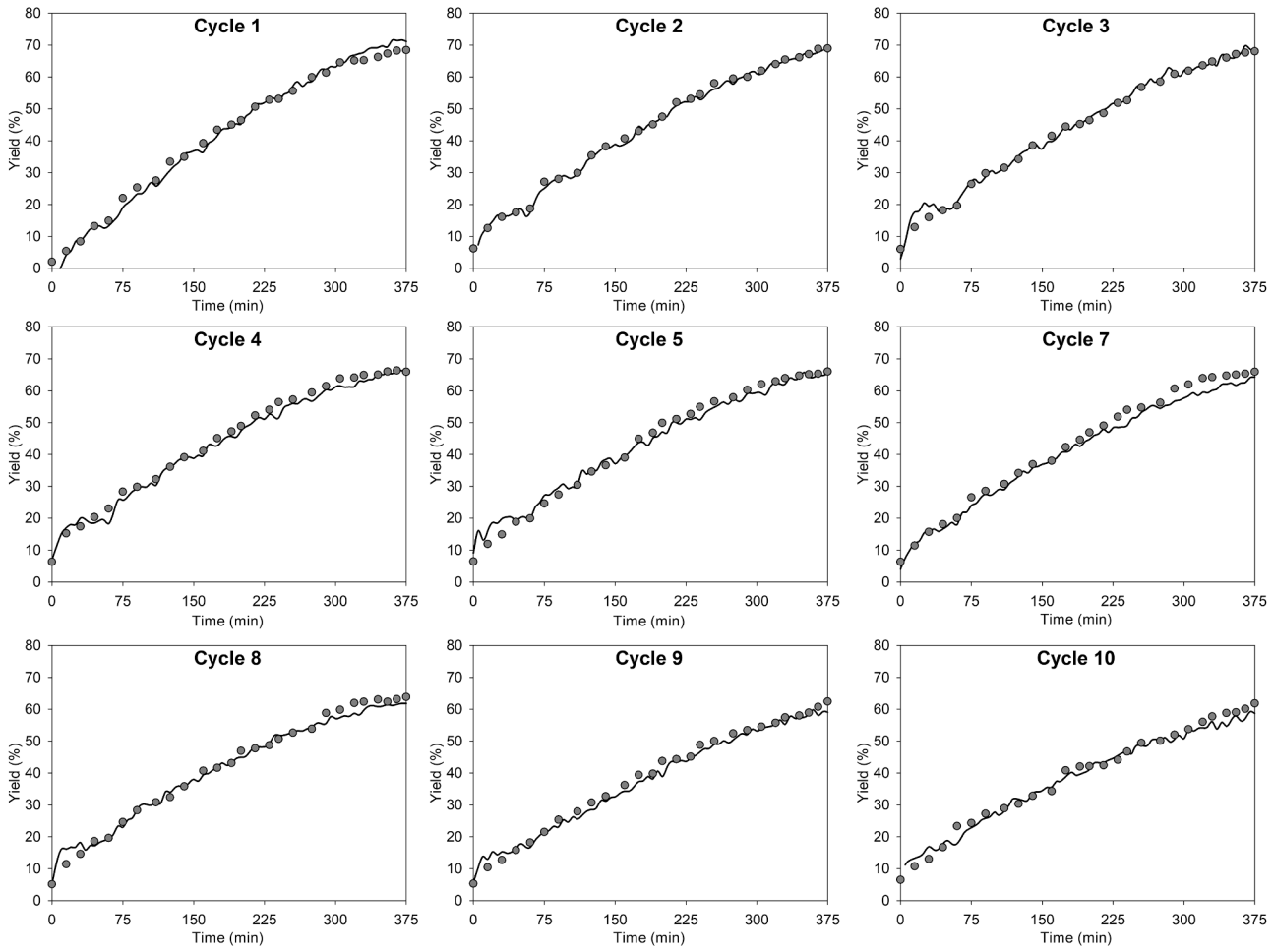


Fig. 7. NIR prediction of reaction yields (solid line) with methanol as acyl-acceptor. Dots represent GC yield values.

2.1% for ethanol. Based on these results, using immobilized enzymes and adding the alcohol stepwise, which might have increased background noise in spectral measurements, was no issue here. Consequently, NIR spectroscopy stands as a robust tool for monitoring industrial biodiesel production in compliance with CPV and PAT principles. Inline monitoring of the reaction can help identify the most suitable time for addition of alcohol pulses and renewal of the biocatalyst for optimal performance.

CRediT authorship contribution statement

Josu López-Fernández: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Visualization. **Desirèe Moya:** Methodology, Investigation. **María Dolors Benaiges:** Conceptualization, Methodology, Validation, Writing – review & editing, Supervision, Project administration. **Francisco Valero:** Conceptualization, Methodology, Validation, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Manel Alcalà:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by the Spanish Ministry of Science and Innovation (Project PID2019-104666GB-100). JLF was additionally awarded a Basque Government Scholarship for the training of predoctoral researchers (PRE_2017_1_0110).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fuel.2022.123794.

References

[1] Connolly D, Lund H, Mathiesen BV. Smart Energy Europe: the technical and economic impact of one potential 100% renewable energy scenario for the European Union. Renew Sustain Energy Rev 2016;60:1634–53. https://doi.org/10.1016/j.rser.2016.02.025.
[2] Kumar B, Verma P. Biomass-based biorefineries: an important architype towards a circular economy. Fuel 2021;288:119622. https://doi.org/10.1016/j.fuel.2020.119622.
[3] Hoang AT, Nižetić S, Ölçer AI. 2,5-Dimethylfuran (DMF) as a promising biofuel for the spark ignition engine application: a comparative analysis and review. Fuel 2021;285:119140. https://doi.org/10.1016/J.FUEL.2020.119140.
[4] Fadhil AB, Nayyef AW, Sedeeq SH. Valorization of mixed radish seed oil and *Prunus armeniaca* L. oil as a promising feedstock for biodiesel production: evaluation and analysis of biodiesels. Asia-Pacific J Chem Eng 2020;15(1). https://doi.org/10.1002/APJ.2390.
[5] Mercy Nisha Pauline J, Sivaramakrishnan R, Pugazhendhi A, Anbarasan T, Achary A. Transesterification kinetics of waste cooking oil and its diesel engine performance. Fuel 2021;285:119108.

[6] Unglert M, Bockey D, Bofinger C, Buchholz B, Fisch G, Luther R, et al. Action areas and the need for research in biofuels. Fuel 2020;268:117227. https://doi.org/10.1016/j.fuel.2020.117227.

[7] Ogunkunle O, Ahmed NA. A review of global current scenario of biodiesel adoption and combustion in vehicular diesel engines. Energy Rep 2019;5:1560–79. https://doi.org/10.1016/j.egyr.2019.10.028.

[8] Günay ME, Türker L, Tapan NA. Significant parameters and technological advancements in biodiesel production systems. Fuel 2019;250:27–41. https://doi.org/10.1016/j.fuel.2019.03.147.

[9] Singh D, Sharma D, Soni SL, Sharma S, Kumar Sharma P, Jhalani A. A review on feedstocks, production processes, and yield for different generations of biodiesel. Fuel 2020;262:116553–68. https://doi.org/10.1016/j.fuel.2019.116553.

[10] Bonet-Ragel K, Canet A, Benaiges MD, Valero F. Synthesis of biodiesel from high FFA *alperujo* oil catalysed by immobilised lipase. Fuel 2015;161:12–7. https://doi.org/10.1016/j.fuel.2015.08.032.

[11] Biernat K, Matuszewska A, Samson-Bręk I, Owczuk M. Biological methods in biodiesel production and their environmental impact. Appl Sci 2021;11:10946. https://doi.org/10.3390/APP112210946.

[12] Harding KG, Dennis JS, von Blottnitz H, Harrison STL. A life-cycle comparison between inorganic and biological catalysis for the production of biodiesel. J Clean Prod 2008;16:1368–78. https://doi.org/10.1016/J.JCLEPRO.2007.07.003.

[13] Cavalcante FTT, Neto FS, Rafael de Aguiar Falcão I, Erick da Silva Souza J, de Moura Junior LS, da Silva Sousa P, et al. Opportunities for improving biodiesel production via lipase catalysis. Fuel 2021;288:119577.

[14] López-Fernández J, Dolors Benaiges M, Valero F. Second- and third-generation biodiesel production with immobilised recombinant *Rhizopus oryzae* lipase: influence of the support, substrate acidity and bioprocess scale-up. Bioresour Technol 2021;334:125233. https://doi.org/10.1016/j.biortech.2021.125233.

[15] López-Fernández J, Benaiges MD, Valero F. *Rhizopus oryzae* lipase, a promising industrial enzyme: biochemical characteristics, production and biocatalytic applications. Catalysts 2020;10:1277. https://doi.org/10.3390/catal10111277.

[16] Rodrigues A, Bordado JC, Dos Santos RG. Upgrading the glycerol from biodiesel production as a source of energy carriers and chemicals - A technological review for three chemical pathways. Energies 2017;10:1817–53. https://doi.org/10.3390/en10111817.

[17] Bonet-Ragel K, Canet A, Benaiges MD, Valero F. Effect of acyl-acceptor stepwise addition strategy using *alperujo* oil as a substrate in enzymatic biodiesel synthesis. J Chem Technol Biotechnol 2018;93:541–7. https://doi.org/10.1002/jctb.5399.

[18] Dunn RO. Effects of monoacylglycerols on the cold flow properties of biodiesel. JAOCS, J Am Oil Chem Soc 2012;89:1509–20. https://doi.org/10.1007/s11746-012-2045-7.

[19] Feltes MMC, de Oliveira D, Block JM, Ninow JL. The production, benefits, and applications of monoacylglycerols and diacylglycerols of nutritional interest. Food Bioprocess Technol 2013;6:17–35. https://doi.org/10.1007/s11947-012-0836-3.

[20] Pinzi S, Alonso F, García Olmo J, Dorado MP. Near infrared reflectance spectroscopy and multivariate analysis to monitor reaction products during biodiesel production. Fuel 2012;92:354–9. https://doi.org/10.1016/j.fuel.2011.07.006.

[21] Tian M, Fu J, Wang Z, Miao C, Lv P, He D, et al. Enhanced activity and stability of *Rhizomucor miehei* lipase by mutating N-linked glycosylation site and its application in biodiesel production. Fuel 2021;304:121514. https://doi.org/10.1016/J.FUEL.2021.121514.

[22] Syed MB. Analysis of biodiesel by high performance liquid chromatography using refractive index detector. MethodsX 2017;4:256–9. https://doi.org/10.1016/J.MEX.2017.07.002.

[23] Morgenstern M, Cline J, Meyer S, Cataldo S. Determination of the kinetics of biodiesel production using proton nuclear magnetic resonance spectroscopy (^{1}H NMR). Energy Fuels 2006;20:1350–3. https://doi.org/10.1021/ef0503764.

[24] de Lima FW, Corgozinho CNC, Tauler R, Sena MM. Monitoring biodiesel and its intermediates in transesterification reactions with multivariate curve resolution alternating least squares calibration models. Fuel 2021;283:119275. https://doi.org/10.1016/j.fuel.2020.119275.

[25] Rosas JG, Blanco M, González JM, Alcalà M. Real-time determination of critical quality attributes using near-infrared spectroscopy: a contribution for Process Analytical Technology (PAT). Talanta 2012;97:163–70. https://doi.org/10.1016/j.talanta.2012.04.012.

[26] Gorsky I. Process validation stage 3: continued process verification. Princ Parenter Solut Valid A Pract Lifecycle Approach 2020:209–32. https://doi.org/10.1016/B978-0-12-809412-9.00008-3.

[27] De Lima SM, Silva BFA, Pontes DV, Pereira CF, Stragevitch L, Pimentel MF. In-line monitoring of the transesterification reactions for biodiesel production using NIR spectroscopy. Fuel 2014;115:46–53. https://doi.org/10.1016/j.fuel.2013.06.057.

[28] Blanco M, Alcalá M, González JM, Torras E. A process analytical technology approach based on near infrared spectroscopy: tablet hardness, content uniformity, and dissolution test measurements of intact tablets. J Pharm Sci 2006;95:2137–44. https://doi.org/10.1002/jps.20653.

[29] Palou A, Miró A, Blanco M, Larraz R, Gómez JF, Martínez T, et al. Calibration sets selection strategy for the construction of robust PLS models for prediction of biodiesel/diesel blends physico-chemical properties using NIR spectroscopy. Spectrochim Acta - Part A Mol Biomol Spectrosc 2017;180:119–26.

[30] Rocha W, Sheen D. Determination of physicochemical properties of petroleum derivatives and biodiesel using GC/MS and chemometric methods with uncertainty estimation. Fuel 2019;243:413–22. https://doi.org/10.1016/j.fuel.2018.12.126.

[31] Richard R, Dubreuil B, Thiebaud-Roux S, Prat L. On-line monitoring of the transesterification reaction carried out in microreactors using near infrared spectroscopy. Fuel 2013;104:318–25. https://doi.org/10.1016/j.fuel.2012.07.054.

[32] Sales R, da Silva NC, da Silva JP, França HH, Pimentel MF, Stragevitch L. Handheld near-infrared spectrometer for on-line monitoring of biodiesel production in a continuous process. Fuel 2019;254:115680. https://doi.org/10.1016/J.FUEL.2019.115680.

[33] Arnau C, Ramon R, Casas C, Valero F. Optimization of the heterologous production of a *Rhizopus oryzae* lipase in *Pichia pastoris* system using mixed substrates on controlled fed-batch bioprocess. Enzyme Microb Technol 2010;46:494–500. https://doi.org/10.1016/j.enzmictec.2010.01.005.

[34] Guillén M, Benaiges MD, Valero F. Biosynthesis of ethyl butyrate by immobilized recombinant *Rhizopus oryzae* lipase expressed in *Pichia pastoris*. Biochem Eng J 2012;65:1–9. https://doi.org/10.1016/j.bej.2012.03.009.

[35] Resina D, Serrano A, Valero F, Ferrer P. Expression of a *Rhizopus oryzae* lipase in *Pichia pastoris* under control of the nitrogen source-regulated formaldehyde dehydrogenase promoter. J Biotechnol 2004;109:103–13. https://doi.org/10.1016/j.jbiotec.2003.10.029.

[36] Rodrigues J, Canet A, Rivera I, Osório NM, Sandoval G, Valero F, et al. Biodiesel production from crude Jatropha oil catalyzed by non-commercial immobilized heterologous *Rhizopus oryzae* and *Carica papaya* lipases. Bioresour Technol 2016;213:88–95.

[37] Aymard C, Belarbi A. Kinetics of thermal deactivation of enzymes: a simple three parameters phenomenological model can describe the decay of enzyme activity, irrespectively of the mechanism. Enzyme Microb Technol 2000;27:612–8. https://doi.org/10.1016/S0141-0229(00)00258-1.

[38] López-Fernández J, Benaiges MD, Valero F. Constitutive expression in *Komagataella phaffii* of mature *Rhizopus oryzae* lipase jointly with its truncated prosequence improves production and the biocatalyst operational stability. Catal 2021;11:1192. https://doi.org/10.3390/CATAL11101192.

[39] Canet A, Dolors Benaiges M, Valero F. Biodiesel synthesis in a solvent-free system by recombinant *Rhizopus oryzae* lipase. Study of the catalytic reaction progress. JAOCS, J Am Oil Chem Soc 2014;91:1499–506. https://doi.org/10.1007/s11746-014-2498-y.

[40] Vicente G, Martínez M, Aracil J. Integrated biodiesel production: a comparison of different homogeneous catalysts systems. Bioresour Technol 2004;92:297–305. https://doi.org/10.1016/j.biortech.2003.08.014.

[41] Ng W, Minasny B, Malone B, Filippi P. In search of an optimum sampling algorithm for prediction of soil properties from infrared spectra. PeerJ 2018;2018.. https://doi.org/10.7717/peerj.5722.

[42] Consumo de aceites en España|Anierac https://anierac.org/consumo-en-espana/?lang=en (accessed September 17, 2021).

[43] Orsavova J, Misurcova L, Vavra Ambrozova J, Vicha R, Mlcek J. Fatty acids composition of vegetable oils and its contribution to dietary energy intake and dependence of cardiovascular mortality on dietary intake of fatty acids. Int J Mol Sci 2015;16:12871–90. https://doi.org/10.3390/ijms160612871.

[44] Canet A, Benaiges MD, Valero F, Adlercreutz P. Exploring substrate specificities of a recombinant *Rhizopus oryzae* lipase in biodiesel synthesis. N Biotechnol 2017;39:59–67. https://doi.org/10.1016/j.nbt.2017.07.003.

[45] Lotti M, Pleiss J, Valero F, Ferrer P. Effects of methanol on lipases: molecular, kinetic and process issues in the production of biodiesel. Biotechnol J 2015;10:22–30. https://doi.org/10.1002/biot.201400158.

[46] Bonet-Ragel K, López-Pou L, Tutusaus G, Benaiges MD, Valero F. Rice husk ash as a potential carrier for the immobilization of lipases applied in the enzymatic production of biodiesel. Biocatal Biotransformation 2018;36:151–8. https://doi.org/10.1080/10242422.2017.1308498.

[47] Sun T, Du W, Liu D. Comparative study on stability of whole cells during biodiesel production in solvent-free system. Process Biochem 2011;46:661–4. https://doi.org/10.1016/j.procbio.2010.11.006.

[48] Lin YH, Luo JJ, John Hwang SC, Liau PR, Lu WJ, Lee HT. The influence of free fatty acid intermediate on biodiesel production from soybean oil by whole cell biocatalyst. Biomass Bioenergy 2011;35:2217–23. https://doi.org/10.1016/j.biombioe.2011.02.039.

[49] Tan Y, Chang SKC, Zhang Y. Comparison of α-amylase, α-glucosidase and lipase inhibitory activity of the phenolic substances in two black legumes of different genera. Food Chem 2017;214:259–68. https://doi.org/10.1016/j.foodchem.2016.06.100.

[50] Faber NM, Bro R. Standard error of prediction for multiway PLS: 1. Background and a simulation study. Chemom Intell Lab Syst 2002;61:133–49. https://doi.org/10.1016/S0169-7439(01)00204-0.