

Article



Quality Monitoring of Biodiesel and Diesel/Biodiesel Blends: A Comparison between Benchtop FT-NIR versus a Portable Miniaturized NIR Spectroscopic Analysis

Luísa L. Monteiro ¹, Paulo Zoio ^{1,2}, Bernardo B. Carvalho ³, Luís P. Fonseca ^{1,*,†} and Cecília R. C. Calado ^{2,*,†}

- ¹ Institute for Bioengineering and Biosciences (iBB), The Associate Laboratory Institute for Health and Bioeconomy–i4HB, Instituto Superior Técnico (IST), Universidade de Lisboa (UL), Av. Rovisco Pais, 1049-001 Lisboa, Portugal; luis4lopesmonteiro@gmail.com (L.L.M.); pjzoio@gmail.com (P.Z.)
- ² CIMOSM—Centro de Investigação em Modelação e Optimização de Sistemas Multifuncionais, ISEL—Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa, R. Conselheiro Emídio Navarro 1, 1959-007 Lisboa, Portugal
- ³ Instituto de Plasmas e Fusão Nuclear, Instituto Superior Técnico (IST), Universidade de Lisboa (UL), Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
- * Correspondence: luis.fonseca@tecnico.ulisboa.pt (L.P.F.); cecilia.calado@isel.pt (C.R.C.C.)
- + Both authors contributed equally to this work.

Abstract: A methodology such as near-infrared (NIR) spectroscopy, which enables in situ and in real-time analysis, is crucial to perform quality control of biodiesel, since it is blended into diesel fuel and the presence of contaminants can hinder its performance. This work aimed to compare the performance of a benchtop Fourier Transform (FT) NIR spectrometer with a prototype of a portable, miniaturized near-infrared spectrometer (miniNIR) to detect and quantify contaminants in biodiesel and biodiesel in diesel. In general, good models based on principal component analysis-linear discriminant analysis (PCA-LDA) of FT-NIR spectra were obtained, predicting with high accuracies biodiesel contaminants and biodiesel in diesel (between 75% to 95%), as well as good partial least square (PLS) regression models to predict contaminants concentration in biodiesel and biodiesel blends, with high coefficients of determination (between 0.83 and 0.99) and low prediction errors. The miniNIR prototype's PCA-LDA models enabled the prediction of target contaminants with good accuracies (between 66% and 86%), and a PLS model enabled the prediction of biodiesel concentration in diesel with a reasonable coefficient of determination (0.68), pointing to the device's potential for preliminary analysis of biodiesel which, associated with its potential low cost and portability, could increase biodiesel quality control.

Keywords: biodiesel; portable NIR spectrometer; contaminants; adulteration; diesel/biodiesel blends

1. Introduction

Every passing year registers significant growth in energy consumption, with 80% of the global energy consumption coming from fossil fuels, aggravating pollution levels and thus contributing to the environmental and energy crises [1]. Recently, energy policies have focused on diversifying primary energy sources and efficient use of resources to reduce greenhouse gases' emissions, without compromising the economic aspect of their commercialization or the quality of citizens' lives [2]. Such is the case of biofuels, fuels produced from biomass and that, in global balance, have lower exhaust emissions than fossil fuels [3]. Biodiesel is a biofuel made up of fatty acid alkyl esters resulting from the transesterification of triglycerides found in oils with alcohols in the presence of a catalyst, with glycerol as a by-product [4]. Its self-lubricant [1], devoid of sulfur and aromatic compounds (making emissions associated with engine wear very low) and the presence of oxygen in the ester molecules make for a combustion that emits significantly fewer particulates, hydrocarbons, CO, and CO₂ than fossil diesel [5]. It offers a similar



Citation: Monteiro, L.L.; Zoio, P.; Carvalho, B.B.; Fonseca, L.P.; Calado, C.R.C. Quality Monitoring of Biodiesel and Diesel/Biodiesel Blends: A Comparison between Benchtop FT-NIR versus a Portable Miniaturized NIR Spectroscopic Analysis. *Processes* **2023**, *11*, 1071. https://doi.org/10.3390/pr11041071

Academic Editors: Pedro Fernandes, Carla C. C. R. de Carvalho and Clayton Jeffryes

Received: 27 January 2023 Revised: 21 March 2023 Accepted: 30 March 2023 Published: 3 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). performance to fossil diesel, with which it shares similarities in terms of cetane number, viscosity, and calorific value [1]. Since it has a higher flash point than fossil diesel, it is also safer to handle, transport, and store [6,7].

Given their similarities, biodiesel can also be blended with fossil diesel. Diesel/biodiesel blends are designated "B", followed by a number indicating the percentage of volume biodiesel mixed in the diesel. When blended with diesel, biodiesel increases the lubricity and the cetane number of the overall fuel, reducing ignition delay [8]. Currently, B7 is the maximum blend permitted by the Fuel Quality Directive for sale across the EU, and the legislation does not require a minimum biodiesel percentage [5]. In the USA, B20 is a commonly used blend, balanced in terms of cost, emissions, and cold flow properties.

Like all fossil fuels, biodiesel is subjected to quality controls to ensure it is safe to use in engines, whether by itself or in diesel/biodiesel blends. Crude biodiesel may have some residual methanol and glycerol after the reaction is concluded, which is removed through washing procedures so that the biodiesel can comply with the concentration limit for methanol and glycerol, for example, imposed by the European Standard EN 14214, both of which is 0.2%(w/w) [9]. Diesel/biodiesel blends can sometimes display a high raw fatty oil content, either due to incomplete transesterification of the biodiesel or to adulteration with cheaper supplies such as vegetable oil (VO) or used cooking oil (UCO) to profit from subsidies involved in biodiesel production [1,2]. These VOs are viscous and less volatile than biodiesel, causing carbon deposits, injection blocking, and incomplete combustion of the fuel [10], which have negative consequences on engine start-up control, acceleration, and emission of gases [11]. Therefore, the rising demand for biodiesel has induced a rise in demand for analytical procedures that quickly and effectively detect anomalies during synthesis, contaminants, and blend adulteration.

While adequate, the use of chromatographic techniques for biodiesel quality control has some drawbacks, as they involve sample destruction, generate large amounts of toxic waste, require off-line sample preparation, and are often time-consuming [2]. In later years, the industry has strived for the implementation of quick, non-destructive, in- and online monitoring methods that do not require sample preprocessing [4]. For example, when applied to in situ monitoring of transesterification reactions, these control methodologies could detect any deviation of the process, allowing its mitigation, without compromising the outcome of the reaction [12]. Such is the case with near-infrared (NIR) spectroscopy, which reflects a combination of tones and overtones of fundamental vibrations of functional groups (e.g., NH, OH, and CH groups) [6]. For an in situ and in real-time analysis, an optical fiber associated with a transflectance probe is typically used, which combines transmittance and reflectance operation modes [13]. In the case of spectrometers, the Fourier Transform (FT) NIR spectrometer is usually preferred to dispersive equipment, due to its greater wavelength accuracy, spectral quality, and reproducibility [14], covering regions between 12,820 and 4000 cm^{-1} (780–2500 nm). NIR spectroscopy has been successfully implemented in the biodiesel industry in reaction monitoring [4,12], quality assessment (methyl ester content and properties) [15–18], contaminant detection (water, methanol, and glycerol) [15,19,20], VO detection [21], and to determine biodiesel content in biodiesel/diesel blends [14,22–27], as described in Table 1.

Table 1. Predictability of spectra partial least square (PLS) regression models of biodiesel contaminants, biodiesel content in diesel blends, biodiesel parameters of the biodiesel synthesis, and ester content. By default, except when stated otherwise, all analyses were made with a dispersive NIR benchtop spectrometer using a transflectance probe coupled to a fiber optic cable enabling in situ analysis.

Aim of Analysis	Type of Analysis	R ²	RMSEP	Ref	
VO content—Soybean oil		0.986	3.40% (<i>w</i> / <i>w</i>)		
VO content—Palm oil		0.996	2.50% (<i>w</i> / <i>w</i>)	[19]	
VO content—Rapeseed oil		0.977	5.40% (<i>w</i> / <i>w</i>)		
NO sectors	FT	0.992	0.30% (v/v)	[22]	
vO content	Miniaturized	0.973	0.46% (v/v)	[23]	
Contaminant content—Water		0.903	111 mg/kg	[19]	
Contaminant content—Water		0.986	87 mg/kg	[20]	
Contaminant content—Methanol		0.996	73 mg/kg	[20]	
Contaminant content—Methanol		0.993	61 mg/kg	[28]	
	Miniaturized	0.998	1.80 wt%	[22]	
	FT	0.994	0.26% (v/v)	[23]	
	Portable and micro	0.988	0.32% (v/v)	[23]	
Biodiasal content in diasal /biodiasal	Off-line transmission	0.9922	0.662% (v/v)	[24]	
blends	Off-line reflection	0.999	0.83% (v/v)	[26]	
	FT Off-line transmission	0.994	0.18% (v/v)	[27]	
		0.999	0.24%	[29]	
	FT 0.96 0.4%		0.4%	[30]	
	Miniaturized	0.88	1.01%	[00]	
Parameter—Density Parameter—S content Parameter—Distillation point T50% Parameter—Distillation point T85%	FT Online transmission (a flow cell)	0.990 0.930 0.880 0.940	0.56 kg/m ³ 0.02% (w/w) 2.12 °C 3.37 °C	[16]	
Parameter—Oxidative stability		0.910	1.28 h	[17]	
Parameter—Iodine value		0.991	3.1 g I ₂ /100 g	[10]	
Parameter—Acid number		0.856	0.22 mg KOH/g		
Reaction monitoring—Ester content		0.989 0.996	2.1% 2.0%	[12]	
Reaction monitoring—Ethyl ester content	FT	0.985	4.10%	[31]	
Reaction monitoring—Glycerides conversion to methyl esters		-	0.74% of conversion	[32]	
Reaction monitoring—Alkyl ester content	Miniaturized	0.79	2.15 wt%	[33]	
Reaction monitoring—Methyl ester content	FT	0.98	2.25%	[34]	
Methyl ester content		0.895	0.9%	[15]	

Recently, portable and some even miniaturized NIR (miniNIR) spectrometers, working mostly in at-line formats, i.e., rapid off-line, have been developed for the quality control of fuels [35]. Their ergonomic design (small size that allows its implementation in portable, pocket-size devices [36]), easy operation, and compliance with the principles of green chemistry [22] significantly reduce time and cost of analysis, which ultimately will increase biofuel monitoring at all stakeholders, including distributors and final consumers, which can increase trust and consequently biodiesel utilization. However, to achieve this, miniaturized equipment must have acceptable monitoring quality characteristics in relation to

the non-miniaturized equipment, such as the more sophisticated ones based on FT-NIR spectroscopy. Some examples of applications of miniaturized NIR spectrometers in recent years are in the detection of counterfeit drugs, the assessment of wood properties, and the examination of contaminated soil for remediation application [37]. These devices have become especially prevalent in the agri-food sector, in the determination of quality parameters of final food products, detection of adulteration of meat and dairy, prediction of nutritional values (protein, lipids, and moisture), discrimination between sources of the food products, and much more [38]. Their application for fuel analysis is not as widespread, but portable and miniaturized spectrometers have been used to quantify ethanol in gaso-line [39], biodiesel in diesel/biodiesel blends [22,23,30], VO content in diesel [23], determine gasoline properties [30], and transesterification reaction monitoring [31–34] (Table 1).

The present work aims to evaluate the simultaneous in situ analysis of all the following variables: vegetable oil, used cooking oil, methanol, and glycerol in biodiesel and of biodiesel in diesel, based on a benchtop FT-NIR spectrometer with a transflectance probe. These analyses will be compared with the ones obtained with a prototype of a portable miniaturized NIR spectrometer, based on spectral regions focusing on part of the vibration combination and the first overtone regions, usually not explored in this type of equipment, such as the ones presented in Table 1. The higher absorptivity and less broad bands associated with these regions in relation to second and third overtone could theoretically lead to better predicting models. Therefore, it is also aimed to compare the models' output obtained with the benchtop FT-NIR spectrometer with the prototype of a miniaturized NIR spectrometer.

2. Materials and Methods

2.1. Sample Preparation for the Detection and Quantification of Contaminants in Biodiesel and of Biodiesel in Diesel/Biodiesel Blends

Biodiesel blends with one contaminant each (the BX blends) and a blend of diesel with biodiesel were prepared: five sets consisted of biodiesel with 1.25, 2.5, 5.0, 10.0, and 50.0% (m/m) of rapeseed oil (RSO), two different types of UCO, methanol, and glycerol; two sets consisted of biodiesel with 0.08, 0.16, 0.313, and 0.625% (m/m) of methanol and glycerol; the eighth consisted of diesel 2.5, 5.0, 10, 12.5, and 15.0% (v/v) of biodiesel (DPB), which is in a similar range to the maximum limit for diesel/biodiesel blends stipulated by the law (Table 2).

		Contaminant Concentration/% (w/w)					
Blends	Description	1	2	3	4	5	
BC	Biodiesel with RSO	1.25	2.5	5	10	50	
BU	Biodiesel with UCO (Type 1)	1.25	2.5	5	10	50	
BUA	Biodiesel with UCO (Type 2)	1.25	2.5	5	10	50	
BM	Biodiesel with methanol	1.25	2.5	5	10	50	
BMM	Biodiesel with methanol	0.08	0.16	0.313	0.625	-	
BG	Biodiesel with glycerol	1.25	2.5	5	10	50	
BGA	Biodiesel with glycerol	0.08	0.16	0.313	0.625	-	
		Biodiesel Concentration/% (v/v)					
	_	1	2	3	4	5	
DPB	Diesel with biodiesel	2.5	5	10	12.5	15	

Table 2. Sets of blends prepared and their respective samples' concentrations.

2.2. Infrared Spectroscopy

The benchtop FT-NIR spectroscopic analysis was conducted as described by Sales et al. [13]. Spectra were acquired using a NIR transflection fiber optic probe IN-271P (Bruker Optics, Ettlingen, Germany), with a path length of 2 mm, coupled to a Vertex-70 spectrometer (Bruker Optics, Ettlingen, Germany). A reference atmospheric air spectrum was acquired before the probe insertion in the samples. NIR spectra were collected the

9000–4000 cm⁻¹ (1111–2500 nm) range, consisting of 32 co-added scans with 8 cm⁻¹ resolution. The scanner velocity was set to 20 kHz and the aperture setting defined was 6 mm.

A prototype of a portable miniaturized NIR spectrometer, called nirU, operating in the spectral region between 1300–2100 nm (7692–4762 cm⁻¹), was used (Figure 1). This in-shop assembled instrument, powered (5 V) and controlled via USB port of a computer, originally employs a matrix of six independent LED working in the following wavenumbers: 1300 nm (7692 cm⁻¹), 1400 nm (7143 cm⁻¹), 1600 nm (6250 cm⁻¹), 1700 nm (5882 cm⁻¹), 1900 nm (5263 cm⁻¹), and 2100 nm (4762 cm⁻¹). The nirU resolution is 133 nm at 1000 nm. The LED matrix is synchronized with a built-in preamplifier detector (InGaAs) with internal cooling (both provided by LED Microsensor NT, represented by MicroTech BG Ltd., formerly based in Saint-Petersburg, Russia). The sample solution (16 μ L) was pipetted into the demountable cuvette (optical length, 0.5 mm), placed on the fixed holder between the LED array and the detector. Spectra bands were obtained by averaging 1000 scans with an integration time of 2.4 ms resulting in a total measurement time of 2.4 s per sample.



Figure 1. Schematic representation of the nirU, a prototype of a portable miniaturized NIR spectrometer.

2.3. Spectra Preprocessing and Processing Methods

Several spectra preprocessing methods, such as atmospheric compensation, baseline correction, mean centering, area normalization, maximum normalization, multiple scattering correction (MSC) and 2nd order Savitsky-Golay derivative based on 2nd degree polynomial with a 15-point window were used. Spectral principal component analysis (PCA) was applied to discover data patterns, to individual BX blends, to mixtures of the BX blends, and to the DPB blend. A PCA-linear discriminant analysis (PCA-LDA) was also conducted.

Diverse PLS regression models were developed for each of the BX blends to predict contaminant concentration, according to Table 2. The models were built after outlier removal, identified based on the hoteling T² analysis conducted on PCA at a 95% confidence interval. Spectral data were randomly partitioned into the calibration (Cal) and validation (Val) sets, containing 70% and 30% of spectral data, respectively. The models' performance was evaluated by the root mean square error of prediction (RMSEP), the coefficient of determination of validation (prediction R-squared), and the number of latent variables (LVs). Spectra atmospheric compensation and baseline correction were conducted with OPUS[®] software, version 6.5 (Bruker, Germany), while remaining spectra preprocessing and processing methods were conducted with Unscrambler [®] X 10.4 software (CAMO software AD, Oslo, Norway).

3. Results and Discussion

3.1. FT-NIR Spectroscopy of Biodiesel Contaminants

The in situ FT-NIR spectroscopic analysis was conducted with a benchtop spectrometer using a transflectance probe associated to a fiber optic cable. Figure 2 depicts the FT-NIR spectra of the BX blends (detailed in Table 2) based on baseline correction (Figure 2a) and 2nd order derivative (Figure 2b). The bands in the combination region from 4000 to 5300 cm⁻¹ and in the first overtone region from 5000 to 6700 cm⁻¹ belong to vibrations mainly from CH_X bonds. PCA based on baseline correction (data not shown) and 2nd order derivative were conducted (Figure 2c), using the whole spectra (data not shown) and the 4000 to 6800 cm⁻¹ sub-region, which presents less noise and showed the best data patterns according to target contaminant.



Figure 2. FT-NIR spectra of the BX blends preprocessed with (**a**) baseline correction; (**b**) 2nd order derivative, and (**c**) the corresponding PCA of 2nd order derivative, conducted between 4000 and 6800 cm⁻¹, of pure biodiesel (B) and biodiesel with varying concentrations of rapeseed oil (BC), UCO (BU and BUA), methanol (BM and BMM), and glycerol (BG and BGA).

Since PCA is an unsupervised algorithm that finds directions of maximum variance regardless of labels, supervised PCA-LDA were conducted to determine the presence of each target contaminant, among a pool of spectra representing pure biodiesel and biodiesel with a high diversity of contaminant types.

This analysis was conducted between 4000 and 6800 cm⁻¹, based on 2nd order derivative (SV2), area normalization (AN), and 2nd derivative coupled with area normalization (SV2 + AN) (Table 3). The 2nd order derivative was chosen to resolve superimposed bands, consequently highlighting the characteristics of chemical analyte signals [40], while area normalization was used to correct differences in the global intensity of the signals, giving each signal the same weight [41] and reducing the concentration effect.

Preprocessing Method	#Components	BC + BU + BUA + BM + BMM + BG + BGA	BC + BU + BUA	BC	BU + BUA	BM + BMM	BG + BGA
	2	38	83	78 70	79 70	78	57
SV2	3	40	83	79	79	84	76
	4	65	78	71	79	//	64 72
	5	68	80	74	79	89	73
	2	26	84	80	72	89	65
4 N T	3	68	95	80	79	91	66
AN	4	75	79	84	73	81	61
	5	75	80	84	75	91	78
SV2 + AN	2	56	48	32	43	49	62
	3	50	78	79	78	66	61
	4	67	81	70	78	64	66
	5	68	78	74	76	89	73

Table 3. Accuracy (%) of PCA-LDA analysis based on FT-NIR spectra preprocessed by second derivative (SV2), area normalization (AN), and second derivative and area normalization (SV2 + AN), to predict if biodiesel is contaminated, if it is contaminated with VOs (BC + BU + BUA), RSO (BC), UCO (BU + BUA), methanol (BM + BMM), or glycerol (BG + BGA).

In general, the PCA-LDA models accurately predicted the presence of contaminants among a library of spectra of biodiesel with different contaminants, with accuracy values ranging between 71% and 95%. The models built to detect contaminants in biodiesel displayed the lowest accuracies, with the highest one being 75%, most likely due to the diversity of contaminants included. The models developed to predict what type of contaminant is present, following contaminant detection, showed higher accuracy. It was possible to build PCA-LDA models with the following accuracies: VOs at 95%, RSO at 84%, UCO at 79%, methanol at 91%, and glycerol at 78% (Table 3). The models based on normalized spectral data often showed the highest accuracy values, likely because any spectral differences caused by contaminant concentration are eliminated, thus allowing the analysis to focus on differences between molecular structure.

The PCA of non-normalized, 2nd order derivative spectra of individual contaminants (Figure A1) shows data patterns formed according to the target contaminant's concentration, with PC2 often separating different clusters. To further explore this, diverse partial least square regression (PLS) models were built, for contaminant concentrations ranging from 0 to 10% (w/w), for the blends BC, BU, BUA, BM, and BG, and from 0 to 0.625% (w/w), for the blends BMM and BGA, based on the whole spectra (data not shown) and between 4000 and 6800 cm⁻¹ (Table 4), and based on the following spectra preprocessing: unprocessed (None), mean centering (MC), baseline correction (Base corr), and 2nd order derivative (SV2).

Overall, the developed PLS models showed very good predictive ability, even for contaminants present in low concentrations (BMM and BGA), as based on a low number of LVs, presented for both the calibration and validation datasets high R-squared values and low RMSE values.

The performance of the models developed to quantify the amount of RSO in biodiesel (BC blend) is comparable to the performance of the model developed by Oliveira et al. [42] (Table 1) to quantify VO in biodiesel in a 0–5% (w/w) range, with an RMSEP value of 0.238% (w/w), a prediction R-squared of 0.97, based on 10 LVs. The models obtained for the BMM blend are comparable to the ones obtained by Felizardo et al. [28] (Table 1), which were built using the 9000–4500 cm⁻¹ region from spectra of biodiesel samples contaminated with methanol in a 0.002–0.286% (w/w) range and obtained RMSEP values lower than 0.02% (w/w) and prediction R-squared values of over 0.97.

Although several preprocessing methods were used, the models derived from the unprocessed spectral data show results just as good as the ones from the models derived from the preprocessed data, which show that NIR spectroscopy is a robust technique for contaminant quantification.

Blend	Preprocessing Method	None	MC	Base Corr	SV2
BC	Latent variables R ² calibration	4 0.99	3 0.98	5 0.98	4 0.99
0–10% (<i>w</i> / <i>w</i>)	R ² external validation RMSEC/% (w/w) RMSEP/% (w/w)	0.99 0.43 0.56	$0.98 \\ 0.40 \\ 0.49$	0.97 0.79 0.71	0.99 0.48 0.68
BU 0–10% (w/w)	Latent variables R^2 calibration R^2 external validation RMSEC/% (w/w) RMSEP/% (w/w)	3 0.97 0.98 0.80 0.75	5 0.99 0.99 0.32 0.33	4 0.98 0.96 0.74 0.84	3 0.98 0.99 0.69 0.70
BUA 0–10% (w/w)	Latent variables R^2 calibration R^2 external validation RMSEC/% (w/w) RMSEP/% (w/w)	3 0.99 0.99 0.58 0.60	3 0.99 0.99 0.36 0.33	5 0.99 0.99 0.48 0.55	3 0.99 0.99 0.59 0.67
BM 0–10% (w/w)	Latent variables R^2 calibration R^2 external validation RMSEC/% (w/w) RMSEP/% (w/w)	3 0.86 0.73 1.70 3.24	7 0.90 0.41 1.03 3.24	5 0.82 0.83 2.21 2.09	2 0.87 0.73 1.64 3.27
BMM 0–0.625% (w/w)	Latent variables R^2 calibration R^2 external validation RMSEC/% (w/w) RMSEP/% (w/w)	4 0.99 0.99 0.03 0.03	3 0.98 0.98 0.03 0.03	5 0.98 0.99 0.04 0.04	4 0.99 0.99 0.03 0.03
BG 0–10% (w/w)	Latent variables R^2 calibration R^2 external validation RMSEC/% (w/w) RMSEP/% (w/w)	6 0.97 0.97 0.84 0.84	7 0.95 0.94 0.74 0.86	7 0.97 0.97 0.82 1.01	5 0.95 0.92 1.20 1.33
BGA 0–0.625% (w/w)	Latent variables R^2 calibration R^2 external validation RMSEC/% (w/w) RMSEP/% (w/w)	7 0.98 0.83 0.05 0.09	7 0.97 0.84 0.04 0.08	7 0.96 0.90 0.07 0.10	6 0.97 0.72 0.06 0.11

Table 4. PLS models based on FT-NIR spectra per BX blend, according to the concentration range of each contaminant and preprocessing method.

3.2. FT-NIR Spectroscopy of Biodiesel in Diesel (DPB)

Figure 3 shows the FT-NIR spectra of the DPB blend based on baseline correction (Figure 3a) and 2nd order derivative (Figure 3b), and PCA based on 2nd order derivative between 4000 and 6800 cm⁻¹ (Figure 3c), with concentrations of biodiesel according to Table 2. The PCA reveals the grouping of replicas of samples according to biodiesel concentration, with PC1 capturing the concentration differences and PC2 distinguishing between the lower and higher biodiesel concentrations.

A PCA-LDA based on spectra normalized to the area and between 4000 and 6800 cm⁻¹ was also conducted, which enabled the prediction of biodiesel in diesel with a 97% accuracy with 5 components (data not shown). Furthermore, the PLS models built to quantify the biodiesel concentration in the DPB blend, based on different preprocessing methods, conducted between 4000 and 6800 cm⁻¹, similar to the region used by Alves et al. [26] in the same type of analysis (between 4400 and 6200 cm⁻¹), showed excellent predictive ability, with calibration and validation R-squared values of 0.99 and low RMSE values (0.10% (v/v)), based on 3 LVs (Table 5). The performance of these models can be compared to ones previously developed by Fernandes et al. [24] and Pimentel et al. [27] (Table 1) to

predict biodiesel content in diesel/biodiesel blends from NIR spectral data preprocessed with second derivative. Similar values of calibration and prediction R-squared and LVs were obtained, and RSMEP values around 0.70% (v/v) [24] or below 0.30% (v/v) [27] were reported.



Figure 3. FT-NIR spectra of pure diesel and the DPB blend preprocessed with (**a**) baseline correction; (**b**) 2nd order derivative, and (**c**) the corresponding PCA of 2nd order derivative, between 4000 and 6800 cm^{-1} , of pure diesel (DP) and diesel with varying concentrations of biodiesel.

Table 5. PLS models based on FT-NIR spectra between 6800 and 4000 cm⁻¹, with diverse preprocessing methods to determine biodiesel content in the DPB blend (0 to 15% (v/v)).

Preprocessing Method	None	MC	Base Corr	SV2
Latent variables	3	3	4	2
R ² calibration	0.99	0.99	0.99	0.99
R ² external validation	0.99	0.99	0.99	0.99
RMSEC / % (v / v)	0.10	0.43	0.73	0.41
RMSEP/% (v/v)	0.10	0.47	0.76	0.44

3.3. nirU of Biodiesel Contaminants

Figure 4 depicts a PCA based on nirU analysis of the BX blends exhibiting a data pattern formed according to the target contaminant. Several preprocessing methods were evaluated, including diverse normalization, and multiple scattering correction. The best data pattern was obtained with baseline correction coupled with maximum normalization. The PCA-LDA models (Table 6) showed a reasonable predictive power of contaminant presence, with accuracy values ranging from 68% to 86%. The best model to predict presence of contaminants (without targeting any specific contaminant) showed a reasonable accuracy (i.e., 78%), especially considering the high diversity of possible contaminants. As previously observed with FT-NIR spectral data (Table 3), the models based on normalized data, in general, showed the highest accuracies.



Figure 4. PCA based on nirU data, preprocessed with baseline correction and maximum normalization, of pure biodiesel (B) and biodiesel with varying concentrations of rapeseed oil (BC), UCO (BU and BUA), methanol (BM and BMM), and glycerol (BG and BGA).

Table 6. Accuracy (%) of PCA-LDA analysis based on nirU data with diverse preprocesseing methods to predict if biodiesel is contaminated, if is contaminated with vegetable oils (BC + BU + BUA), rapeseed oil (BC), UCO (BU + BUA), methanol (BM + BMM), or glycerol (BG + BGA).

Preprocessing Method	#Components	BC + BU + BUA + BM + BMM + BG + BGA	BC + BU + BUA	BC	BU + BUA	BM + BMM	BG + BGA
	2	29	55	48	68	67	68
Linnerssand	3	35	58	50	59	48	67
Unprocessed	4	46	66	72	60	53	66
	5	75	70	84	49	52	76
	2	31	55	49	67	47	67
Decelie a compation	3	48	54	44	58	45	67
baseline correction	4	48	66	72	55	53	73
	5	75	72	86	54	57	76
	2	29	56	48	68	58	58
Area	3	72	61	47	71	58	55
normalization	4	76	66	73	71	53	79
	5	78	66	74	71	57	77
Devilian	2	71	55	48	68	50	51
Baseline correction	3	72	55	45	71	47	51
and maximum normalization	4	68	66	61	71	45	69
	5	69	69	78	74	57	78
	2	47	51	45	40	56	71
Baseline correction	3	66	56	60	40	55	44
and MSC	4	71	49	59	61	68	41

The PLS models based on nirU data (Table A1), with unprocessed data (None), baseline correction (Base corr), or baseline correction and maximum normalization (Base corr + Max norm) performed poorly, presenting low R-squared values and high RMSE. Of exception are the models built from unprocessed data of the BUA and BGA blends, with validation R-squared values of 0.81 and low RMSEP values, and the model built from data of the BG blend preprocessed with baseline correction and maximum normalization, which showed a validation R-squared value of 0.70. These results could result from the LEDs used associated with the low resolution, which, however, contributes to the low-cost of the device. Even

so, the nirU spectrometer could potentially be used to perform preliminary analysis of biodiesel, which could then be validated by FT-NIR spectroscopy.

3.4. nirU Spectroscopy of Biodiesel in Diesel (DPB)

The PCA of nirU data based on the baseline correction of the DPB blends (Figure 5) reveals data clustering according to biodiesel concentration, with PC1 capturing the concentration effect of biodiesel and PC2 distinguishing between lower and higher biodiesel concentrations. Of the evaluated preprocessing methods, baseline correction provided the best separation between sample groups. A PCA-LDA predicted the presence of biodiesel in diesel with an 80% accuracy with 5 components (data not shown).



Figure 5. nirU data PCA based on baseline correction, conducted on the whole spectra, of pure diesel (DP) and diesel with varying concentrations of biodiesel.

A PLS models built to quantify the biodiesel concentration in the DPB blend (Table 7) based on data preprocessed with baseline correction and maximum normalization showed a validation R-squared value of 0.71 and a 5.19% (v/v) RMSEP.

Table 7. PLS models based on nirU data, with diverse preprocessing methods to determine biodiesel content in the DPB blend (0 to 15% (v/v)).

Preprocessing Method	None	Base Corr	Base Corr + Max Norm
Latent variables	4	5	3
R ² calibration	0.70	0.72	0.67
R ² external validation	0.68	0.69	0.71
RMSEC / % (v/v)	5.02	4.91	5.34
RMSEP/% (v/v)	5.41	5.32	5.19

The loadings of PC3, PC4, and PC5 of PCA-LDA of FT-NIR spectra normalized to the area of all the BX blends reveals the bands at 4435, 4840, and 5790 cm⁻¹ (Figure A2) as relevant to predict the presence of contaminants in biodiesel, while the equivalent analysis of the DPB blend (data not shown) also pointed to the band at 4435 cm⁻¹ as relevant to the prediction of biodiesel content in diesel. Despite these bands being near the LEDs at 4760, 5260 and 5880 cm⁻¹, especially considering the LEDs' low resolution, other sets of LEDs could be evaluated in the future. It is worth highlighting that the low cost of the device depends on the LEDs utilization.

4. Conclusions

The development of methodologies to enable in situ in the real-time detection of biodiesel contaminants or to predict biodiesel content in diesel blends, as based on FT-NIR

spectrometry, is crucial to increasing the quality control of the product and the biodiesel synthesis process. This is further strengthened by the development of portable equipment and miniaturized equipment, based on in situ or at-line analysis, by enabling a very economic, rapid and in loco analysis, which will increase the biodiesel monitoring at all stakeholders, as distributors and final consumers. This eventually increases quality control, promotes trust, and consequently promotes biodiesel use.

In the present work, a benchtop FT-NIR spectroscopic analysis, based on a transflectance probe with a fiber optic cable enabled to predict *in-situ* the presence of target contaminants, such as VO, methanol, and glycerol, based on the PCA-LDA models with accuracies between 75% and 95%. It was also possible to develop regression models to predict the contaminant concentration, with R-squared for the independent validation dataset between 0.83 and 0.99, and low RMSEP. Moreover, based on in situ FT-NIR spectroscopy, it was possible to predict the biodiesel concentration in diesel, for the validation dataset, with R-squared of 0.99 and an RMSEP of 0.10%(v/v).

Compared to the benchtop FT-NIR spectrometer, the nirU prototype presented reasonable models that could predict the target contaminant, with slight lower accuracies (between 68% and 86%), while the presence of biodiesel in diesel was predicted with an 80% accuracy. The present nirU prototype setup did not enable the development of reasonable PLS models to predict the target contaminant concentration, except for UCO and glycerol concentrations, and biodiesel concentration in diesel. In the near future, the LEDs' optimization should be conducted to improve the predictive power of the models. Due to potential low costs and dimension of the current device, the equipment would potentially be use for a first screen of contaminants and blends of biodiesel in diesel, and performing a second round of analysis, if necessary, with more sophisticated equipment to predict the contaminant concentration or the biodiesel content in diesel.

Author Contributions: Conceptualization of the nirU device, P.Z. and B.B.C.; conception of blends and blend concentration to analyze: L.L.M., P.Z. and L.P.F.; multivariate data analysis: L.L.M., P.Z. and C.R.C.C.; original draft preparation: L.L.M. and C.R.C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Fundação para a Ciência e a Tecnologia projects DSAIPA/DS/ 0117/2020, UIDB/04565/2020 and UIDP/04565/2020, by the Associate Laboratory Institute for Health and Bioeconomy—i4HB project LA/P/0140/2020, and by the Instituto Politécnico de Lisboa project NeproMD/ISEL/2020. The present work was partially conducted in the Engineering and Health Laboratory, which resulted from a collaboration protocol established between the Universidade Católica Portuguesa and the Instituto Politécnico de Lisboa.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A



Figure A1. PCA based on the 2nd order derivative FT-NIR spectra between 4000 and 6800 cm⁻¹ of pure biodiesel (B) and blends (a) BC; (b) BU; (c) BUA; (d) BM; (e) BMM; (f) BG; (g) BGA.

Table A1. PLS regression models per blend based on nirU datal data according to the concentration range of each contaminant and the preprocessing method.

Blend	Preprocessing Method	None	Base Corr	Base Corr + Max Norm
	Latent variables	4	5	2
	R ² calibration	0.62	0.61	0.58
BC	R ² external validation	0.56	0.57	0.59
	RMSEC/% (w/w)	3.57	3.62	3.73
	RMSEP/% (w/w)	3.67	3.64	3.57

Blend	Preprocessing Method	None	Base Corr	Base Corr + Max Norm
	Latent variables	6	5	5
	R ² calibration	0.56	0.56	0.56
BU	R ² external validation	0.59	0.58	0.58
	RMSEC/% (w/w)	3.27	3.27	3.27
	RMSEP/% (w/w)	3.25	3.27	3.30
	Latent variables	4	4	5
	R ² calibration	0.71	0.69	0.74
BUA	R ² external validation	0.81	0.74	0.67
	RMSEC/% (w/w)	0.11	0.11	2.45
	RMSEP/% (w/w)	0.09	0.11	3.04
	Latent variables	6	6	5
	R ² calibration	0.61	0.62	0.61
BM	R ² external validation	0.57	0.58	0.57
	RMSEC/% (w/w)	3.08	3.04	3.07
	RMSEP/% (w/w)	3.02	2.97	3.02
	Latent variables	2	2	4
	R ² calibration	0.53	0.50	0.54
BMM	R ² external validation	0.36	0.36	0.28
	RMSEC/% (w/w)	$2.45 imes 10^{-3}$	2.51×10^{-3}	0.24
	RMSEP/% (w/w)	$2.30 imes10^{-3}$	$2.30 imes10^{-3}$	0.24
	Latent variables	6	6	4
	R ² calibration	0.69	0.71	0.72
BG	R ² external validation	0.66	0.68	0.70
	RMSEC / % (w/w)	2.59	2.52	2.47
	RMSEP / % (w / w)	3.28	3.19	3.11
	Latent variables	4	4	3
221	R ² calibration	0.80	0.83	0.83
BGA	R ² external validation	0.82	0.79	0.81
	RMSEC/% (w/w)	$1.49 imes 10^{-3}$	$1.41 imes 10^{-3}$	0.14
	RMSEP/% (w/w)	$1.33 imes10^{-3}$	$1.43 imes10^{-3}$	0.13

Table A1. Cont.



Figure A2. Loadings of (a) PC3, (b) PC4, and (c) PC5 of FT-NIR spectra PCA of the BX blends based on area normalization between 4000 and 6800 cm^{-1} .

References

- Bozbas, K. Biodiesel as an alternative motor fuel: Production and policies in the European Union. *Renew. Sustain. Energy Rev.* 2005, 12, 542–552. [CrossRef]
- Ferrero, G.O.; Faba, E.M.S.; Rickert, A.A.; Eimer, G.A. Alternatives to rethink tomorrow: Biodiesel production from residual and non-edible oils using biocatalyst technology. *Renew. Sustain. Energy Rev.* 2019, 150, 128–135. [CrossRef]
- Verma, P.; Sharma, M.; Dwivedi, G. Impact of alcohol on biodiesel production and properties. *Renew. Sustain. Energy Rev.* 2016, 56, 319–333. [CrossRef]
- Gelinski, E.K.; Hamerski, F.; Corazza, M.L.; Santos, A.F. Biodiesel Synthesis Monitoring using Near Infrared Spectroscopy. Open Chem. Eng. J. 2018, 12, 95–110. [CrossRef]
- European Biodiesel Board (EBB), about Biodiesel. European Union. Available online: https://ebb-eu.org/about-biodiesel/ (accessed on 11 August 2022).
- 6. Zhang, W.-B. Review on analysis of biodiesel with infrared spectroscopy. *Renew. Sustain. Energy Rev.* **2012**, *16*, 6048–6058. [CrossRef]
- Christopher, L.P.; Kumar, H.; Zambare, V.P. Enzymatic biodiesel: Challenges and opportunities. *Appl. Energy* 2014, 119, 497–520. [CrossRef]
- Alternative Fuels Data Center, Diesel Vehicles Using Biodiesel. U.S. Department of Energy. Available online: https://afdc.energy. gov/vehicles/diesel.html (accessed on 11 August 2022).
- 9. European Standard EN 14214; CEN—European Committee for Standardization: Brussels, Belgium, 2003.
- Pontes, M.J.C.; Pereira, C.F.; Pimentel, M.F.; Vasconcelos, F.V.C.; Silva, A.G.B. Screening analysis to detect adulteration in diesel/biodiesel blends using near-infrared spectrometry and multivariate classification. *Talanta* 2011, 85, 2159–2165. [CrossRef]
- 11. Corgozinho, C.N.; Pasa, V.M.; Barbeira, P.J. Determination of residual oil in diesel oil by spectrofluorimetric and chemometric analysis. *Talanta* **2008**, *76*, 479–484. [CrossRef]
- López-Fernández, J.; Moya, D.; Benaiges, M.D.; Valero, F.; Alcalà, M. Near Infrared Spectroscopy: A useful technique for inline monitoring of the enzyme catalyzed biosynthesis of third-generation biodiesel from waste cooking oil. *Fuel* 2022, *319*, Article 123794. [CrossRef]
- Sales, K.C.; Rosa, F.; Sampaio, P.N.; Fonseca, L.P.; Lopes, M.B.; Calado, C.R. In Situ Near-Infrared (NIR) Versus High-Throughput Mid-Infrared (MIR) Spectroscopy to Monitor Biopharmaceutical Production. *Appl. Spectrosc.* 2014, 69, 760–772. [CrossRef]
- Mazivila, S.J. Trends of non-destructive analytical methods for identification of biodiesel feedstock in diesel-biodiesel blend according to European Commission Directive 2012/0288/EC and detecting diesel-biodiesel blend adulteration: A brief review. *Talanta* 2018, 180, 239–247. [CrossRef] [PubMed]
- 15. Baptista, P.; Felizardo, P.; Menezes, J.C.; Correia, M.J.N. Multivariate near infrared spectroscopy models for predicting the methyl esters content in biodiesel. *Anal. Chim. Acta* 2008, 607, 153–159. [CrossRef]
- Lira, L.F.B.; Vasconcelos, F.V.C.; Pereira, C.F.; Paim, A.P.S.; Stragevitch, L.; Pimentel, M.F. Prediction of properties of diesel/biodiesel blends by infrared spectroscopy and multivariate calibration. *Fuel* 2010, *89*, 405–409. [CrossRef]
- 17. Canha, N.; Felizardo, P.; Menezes, J.C.; Correia, M.J.N. Multivariate near infrared spectroscopy models for predicting the oxidative stability of biodiesel: Effect of antioxidants addition. *Fuel* **2012**, *97*, 352–357. [CrossRef]
- Canha, N.; Felizardo, P.; Correia, M.J.N. Controlling the Oxidative Stability of Biodiesel Using Oils or Biodiesel Blending or Antioxidants Addition. *Environ. Prog. Sustain. Energy* 2017, 37, 1031–1040. [CrossRef]
- 19. Baptista, P.; Felizardo, P.; Menezes, J.C.; Correia, M.J.N. Monitoring the quality of oils for biodiesel production using multivariate near infrared spectroscopy models. *J. Near Infrared Spectrosc.* **2008**, *16*, 445–454. [CrossRef]
- 20. Felizardo, P.; Baptista, P.; Uva, M.S.; Menezes, J.C.; Correia, M.J.N. Monitoring biodiesel fuel quality by near infrared spectroscopy. *J. Near Infrared Spectrosc.* **2007**, *15*, 97–105. [CrossRef]
- Du, Q.; Zhu, M.; Shi, T.; Luo, X.; Gan, B.; Tang, L.; Chen, Y. Adulteration detection of corn oil, rapeseed oil and sunflower oil in camellia oil by in situ diffuse reflectance near-infrared spectroscopy and chemometrics. *Food Control* 2021, 121, Article 107577. [CrossRef]
- 22. Correia, R.M.; Domingos, E.; Cáo, V.M.; Araujo, B.R.; Sena, S.; Pinheiro, L.U.; Fontes, A.M.; Aquino, L.F.M.; Ferreira, E.C.; Filgueiras, P.R.; et al. Portable near infrared spectroscopy applied to fuel quality control. *Talanta* **2018**, *176*, 26–33. [CrossRef]
- 23. Paiva, E.M.; Rohwedder, J.J.R.; Pasquini, C.; Pimentel, M.F.; Pereira, C.F. Quantification of biodiesel and adulteration with vegetable oils in diesel/biodiesel blends using portable near-infrared spectrometer. *Fuel* **2015**, *160*, 57–63. [CrossRef]
- 24. Fernandes, D.D.S.; Gomes, A.A.; Costa, G.B.; Silva, G.W.B.; Véras, G. Determination of biodiesel content in biodiesel/diesel blends using NIR and visible spectroscopy with variable selection. *Talanta* **2011**, *87*, 30–34. [CrossRef] [PubMed]
- 25. Silva, G.W.B.; Gomes, A.A.; Silva, P.; Costa, G.B.; Fernandes, D.D.S.; Fontes, M.M.; Veras, G. Biodiesel/Diesel Blends Classification
- with Respect to Base Oil Using NIR Spectrometry and Chemometrics Tools. J. Am. Oil Chem. Soc. 2012, 89, 1165–1171. [CrossRef]
 Alves, J.C.L.; Poppi, R.J. Biodiesel content determination in diesel fuel blends using near infrared (NIR) spectroscopy and support vector machines (SVM). Talanta 2012, 104, 155–161. [CrossRef]
- 27. Pimentel, M.F.; Ribeiro, G.M.; Cruz, R.S.; Stragevitch, L.; Filho, J.G.A.P.; Teixeira, L.S. Determination of biodiesel content when blended with mineral diesel fuel using infrared spectroscopy and multivariate calibration. *Microchem. J.* **2006**, *82*, 201–206. [CrossRef]

- 28. Felizardo, P.; Baptista, P.; Menezes, J.C.; Correia, M.J.N. Multivariate near infrared spectroscopy models for predicting methanol and water content in biodiesel. *Anal. Chim. Acta* 2007, 595, 107–113. [CrossRef] [PubMed]
- 29. Alves, J.C.L.; Poppi, R.J. Quantification of conventional and advanced biofuels contents in diesel fuel blends using near-infrared spectroscopy and multivariate calibration. *Fuel* **2016**, *165*, 379–388. [CrossRef]
- Silva, N.C.; Cavalcanti, C.J.; Honorato, F.A.; Amigo, J.M.; Pimentel, M.F. Standardization from a benchtop to a handheld NIR spectrometer using mathematically mixed NIR spectra to determine fuel quality parameters. *Anal. Chim. Acta* 2017, 954, 32–42. [CrossRef]
- 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*, 318–325. [CrossRef]
- 32. Killner, M.H.; Rohwedder, J.J.; Pasquini, C. A PLS regression model using NIR spectroscopy for on-line monitoring of the biodiesel production reaction. *Fuel* **2011**, *90*, 3268–3273. [CrossRef]
- 33. Sales, R.; Silva, N.C.; Silva, J.P.; França, H.H.; Pimentel, M.F.; Stragevitch, L. Handheld near-infrared spectrometer for on-line monitoring of biodiesel production in a continuous process. *Fuel* **2019**, 254, Article 115680. [CrossRef]
- Lima, S.M.; Silva, B.F.A.; Pontes, D.V.; Pereira, C.F.; Stragevitch, L.; Pimentel, M.F. In-line monitoring of the transesterification reactions for biodiesel production using NIR spectroscopy. *Fuel* 2014, 115, 46–53. [CrossRef]
- Alcalà, M.; Blanco, M.; Moyano, D.; Broad, N.W.; O'Brien, N.; Friedrich, D.; Pfeiffer, F.; Siesler, H.W. Qualitative and quantitative pharmaceutical analysis with a novel hand-held miniature near infrared spectrometer. J. Near Infrared Spectrosc. 2014, 21, 445–457. [CrossRef]
- Kranenburg, R.F.; Weesepoel, Y.; Alewijn, M.; Sap, S.; Arisz, P.W.; Esch, A.; Keizers, P.H.; Asten, A.C. The importance of wavelength selection in on-scene identification of drugs of abuse with portable near-infrared spectroscopy. *Forensic Chem.* 2022, 30, Article 100437. [CrossRef]
- Beć, D.K.B.; Grabska, D.J.; Huck, P.C.W. Principles and Applications of Miniaturized Near-Infrared (NIR) Spectrometers. *Chemistry* 2021, 27, 1514–1532. [CrossRef] [PubMed]
- Beć, K.B.; Grabska, J.; Huck, C.W. Miniaturized NIR Spectroscopy in Food Analysis and Quality Control: Promises, Challenges, and Perspectives. *Foods* 2022, 11, 1465. [CrossRef]
- 39. Lutz, O.; Bonn, G.; Rode, B.; Huck, C. Reproducible quantification of ethanol in gasoline via a costumized near-infrared spectrometer. *Anal. Chim. Acta* 2014, 826, 61–68. [CrossRef] [PubMed]
- 40. Zimmermann, B.; Kohler, A. Optimizing Savitzky–Golay Parameters for Improving Spectral Resolution and Quantification in Infrared Spectroscopy. *Appl. Spectrosc.* 2013, 67, 892–902. [CrossRef]
- 41. Zeaiter, M.; Rutledge, D. 3.04—Preprocessing Methods. In *Comprehensive Chemometrics*; Brown, S.D., Tauler, R., Walczak, B., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; Volume 3, pp. 121–231.
- Oliveira, F.C.; Brandão, C.R.; Ramalho, H.F.; Costa, L.A.; Suarez, P.A.; Rubim, J.C. Adulteration of diesel/biodiesel blends by vegetable oil as determined by Fourier transform (FT) near infrared spectrometry and FT-Raman spectroscopy. *Anal. Chim. Acta* 2007, 587, 194–199. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.