Contents lists available at SciVerse ScienceDirect

# Talanta



iournal homepage: www.elsevier.com/locate/talanta

# An analytical method for lubricant quality control by NIR spectroscopy

# David Zamora<sup>a</sup>, Marcelo Blanco<sup>a,∗</sup>, Manel Bautista<sup>b</sup>, Rufí Mulero<sup>b</sup>, Miquel Mir<sup>b</sup>

<sup>a</sup> Department of Chemistry, Faculty of Sciences, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain <sup>b</sup> Development and Quality Assurance Department, CRODA S.A. (Mevisa Site), C-35 Km 72, 08945 Fogars de la Selva, Barcelona, Spain

# a r t i c l e i n f o

Article history: Received 11 October 2011 Received in revised form 20 December 2011 Accepted 23 December 2011 Available online 2 January 2012

Keywords: Lubricants Viscosity NIR spectroscopy POEs Cascading libraries identification PLS models

# A B S T R A C T

The excellent compatibility of polyol esters (POEs) with general fluids coolants such as CO<sub>2</sub> and nonchlorinated coolants (e.g. hydrofluorocarbons, HFCs) in terms of solubility, miscibility and chemical stability has fostered their use as lubricants by the refrigeration industry. The most widely used POEs bases consist of esters of polyalcohols – such as pentaerythritol, dipentaerythritol and neopentylglycol – with mixtures of carboxylic acids of 4–10 carbon atoms. Their thermophysical properties (viscosity mainly) are crucial with a view to producing efficient lubricants for specific applications. Usually, POE formulations, which usually contain several bases, are characterized in terms of global indices.

In this work, we developed a methodology based on NIR spectroscopy for the characterization and analysis of lubricant formulations. The products, including lubricant bases and lubricant formulations, are characterized by reference to two spectral libraries that are used to identify as the starting lubricant bases as well as the lubricant formulations. It has been proposed to build libraries in cascade for the differentiation of lubricant formulations without and with low content in additives. Once the identification of the formulation is applied PLS multivariate models are used to determine the components of a lubricant formulation and its viscosity.

© 2011 Elsevier B.V. All rights reserved.

#### **1. Introduction**

Mechanical devices require the use of lubricants to prevent overheating and wear of solid surfaces in contact. In fact, friction during the motion of two solid surfaces in contact produces heat and wear. The lubricant is a liquid product applied to moving parts in order to lubricate them and prevent their wear and sealing. Lubricant formulations are usually mixtures of various components that are supplied with additives intended to improve specific performancerelated properties such as thermal and/or chemical stability and the response to viscosity and/or surface tension changes [\[1\].](#page-5-0) Although lubricants consisting of mineral oils are widely used for a variety of purposes, they are scarcely biodegradable and, frequently, toxic. Therefore, they can have a strong impact on living beings if they reach the environment [\[2,3\].](#page-5-0)

There has been much progress in the development of new lubricants to meet the requirements of a number of industrial applications (refrigeration, air conditioning, hydraulic fluids) with improved specifications and minimal environmental impact in recent years.

For 50 years, chlorofluorocarbons (CFCs) were used as cooling fluids in refrigeration systems; however, their adverse effect on the ozone layer posed the need to replace them with environmentally compatible alternatives such as  $CO<sub>2</sub>$  or hydrofluorocarbons (HFCs). The new refrigerators require also new, compatible synthetic lubricants with a minimal environmental impact. Such new lubricant formulations are usually mixtures of various chemical bases including polyalpha olefins (PAOs), polyol esters (POEs) and polyalkylene glycols (PAGs). Based on some studies, these lubricant formulations can be deemed "green" as they improve energy efficiency, reduce part wear and are biodegradable [\[1,4,5\].](#page-5-0)

Cooling fluids such as  $CO<sub>2</sub>$  and HFCs are widely used with polyol esters (POEs) on account of their good lubricity and miscibility. In addition, POEs possess a low volatility and an excellent viscosity–temperature relationship in addition to biodegradability. However, they are easily hydrolysed to acids that can be aggressive to the metal surfaces to be lubricated.

Commercially available POE-based lubricants are usually obtained by esterifying one or two polyfunctional alcohols with a mixture of acids in the present of a catalyst (a tin or titanium complex) or a strong acid (e.g. p-toluenesulphonic acid) at  $200-250$  °C. The characteristics of the resulting product depends on the starting polyol (pentaerythritol, dipentaerythritol, neopentylglycol, trimethylolpropane) and the acid or a mixture of acids of variable chain length (usually 4–10 carbon atoms), whether linear or branched, used in the synthesis. The product can be identified by gas chromatography–mass spectrometry [\[6\].](#page-5-0) However, the most common lubricant formulations are complex mixtures, and their characterization and analysis is difficult. By way of example,



<sup>∗</sup> Corresponding author. Tel.: +34 93 581 1367; fax: +34 93 581 1367. E-mail address: Marcel.Blanco@uab.es (M. Blanco).

<sup>0039-9140/\$</sup> – see front matter © 2011 Elsevier B.V. All rights reserved. doi:10.1016/j.talanta.2011.12.067

reacting pentaerythritol, which possesses four hydroxyl groups, with a mixture of two monocarboxylic acids gives up to five esters – position isomers and products of incomplete esterification excluded. Also, the determination of individual components lacks practical interest at the industrial level since the proportion of each component depends on the acid proportions used in the synthetic procedure. Industries use a wide variety of POE for reaction with mixtures of polyols with mixtures of acids in variable proportions. This adds to the complexity of the resulting lubricant formulation. The information potentially provided by the GC–MS technique in this situation is rather complex and of little use since it allows neither the overall performance of the product nor its characteristics to be determined. Moreover, lubricant formulations typically contain additives including antioxidants, anticorrosion agents, copper deactivators or acid catchers, which make analyses even more complicated.

For the previous reasons, lubricant formulations are usually characterized in terms of global indices such as viscosity or acid value. The most important parameter in this context is viscosity as it governs performance and fatigue resistance in bearing and gearing mechanisms, as well as its changes over the range of working temperatures. The ideal lubricant formulation is one whose viscosity is scarcely affected by temperature changes within the operating interval [\[7–10\].](#page-5-0) Viscosity determinations are slow and cumbersome, so it would be interesting to be able to characterize lubricant formulations in a simpler, more expeditious manner (e.g. by determining their bases and composition).

Near infrared spectroscopy (NIRS) has proved especially useful for the analysis of mixtures with no reagent or sample pretreatment [\[11–13\].](#page-5-0) In this work, we assessed its potential for determining the lubricant's bases of a lubricant formulation and its kinematic viscosity at variable temperatures. To this end, we constructed PLS multivariate calibration models to determine the components of the mixture and its viscosity. Also, we constructed spectral libraries to identify the lubricant bases and lubricant formulation compositions.

#### **2. Experimental**

#### 2.1. Lubricant formulations

The lubricant formulation studied contained two POEs (lubricant base) in appropriate proportions plus additives and/or antioxidants including butylhydroxytoluene (BHT), tricresyl phosphate (TCP) and Silicone Fluid. Each lubricant base consisted of the esters of a single polyol (pentaerythritol, dipentaerythritol or neopentylglycol) [\(Fig.](#page-2-0) 1) with a mixture of monocarboxylic acids of 5–10 carbon atoms in variable proportions. A total of 9 lubricant bases were studied. [Table](#page-2-0) 1 shows the polyol and acid mixture used for each lubricant base.

#### 2.2. Samples

The lubricant formulation chosen as example for quantitation was a mixture of two lubricant bases consisting of pentaerythritol esterified with a mixture of pentanoic acid (C5) and heptanoic acid (C7), lubricant base 2 (50.60%), and another obtained by esterifying neopentylglycol with heptanoic acid (C7), lubricant base 8 (48.65%). Laboratory samples were prepared by mixing the two lubricant bases in variable proportions spanning a range of  $\pm 5\%$ (w/w) around the nominal content of each in the lubricant formulation. In addition, the lubricant formulation is supplied with two additives at the end of the production process, namely: TCP (0.7%) and BHT (0.05%).

#### 2.3. Instrumental

Near infrared spectra were acquired in the transmittance mode, using cylindrical vials of 6.5 mm inner diameter, on a FOSS NIRSystems 5000 spectrophotometer equipped with a Vial Heater Module/Pair Fiber s/n 9372. The temperature of the vial heater was set at 75 ◦C. Spectra were recorded at 2 nm intervals over the wavelength range of 1100–2500 nm and the instrument was controlled via the software Vision v. 2.51.

Viscosity measurements were made with a Stabinger SVM 3000 viscometer from Anton-Paar GmbH (Graz, Austria). This is a cylindrical rotational viscometer consisting of a fast rotating outer tube and a slower rotating inner measuring rotor. Dynamic viscosity values are calculated from the rotor speed at which the propelling shear forces acting on the sample equilibrate with the braking effect of the current induced by the rotating magnet inside the viscometer.

#### 2.4. Construction of spectral identification libraries

The high similarity of the NIR spectra for the lubricating bases and lubricant formulation, reflected in their correlation coefficients, required the use of highly discriminating pattern recognition methods. We chose to use the residual variance model to identify the products. This method relies on a principal component analysis (PCA) of the samples defining a class. The ensuing factors were used to reconstruct the spectrum for each sample and calculate its residual. If the residual of the reconstructed spectrum for a new sample was below the preset limit for the model, then the sample was assumed to be identified by the class concerned.

One crucial choice towards ensuring adequate selectivity in a spectral library is the choice of an appropriate threshold; in our case, the threshold was taken to be the maximum residual variance allowing a spectrum to be unequivocally assigned to a given class. All spectra exhibiting a residual variance below the threshold for a class were assumed to belong to such a class. Too high threshold can lead to ambiguity (does not allow the distinction between two or more classes because more than a product of the library have lower values) and not allow the correct identification, while a very low threshold will not recognize the classes of the library.

The great similarity between the spectra of the lubricant formulations with and without additives prevents correct differentiation and to solve this problem have been proposed the application of cascading libraries [\[14\];](#page-5-0) to build it, the sample spectra (with and without additives) which have ambiguities are split in two different classes (sub-library) and the correct identification is done in the same way, but defining a new threshold. This methodology (cascading library identification) as it involves identifying the lubricant formulation against the general library and, if the result is inconclusive, using a sub-library for qualification.

This methodology is suggested as a step prior to quantification to prevent the possible mistake of applying a model of quantification which is not applicable to the formulation to determine; firstly the sample is identified as a particular formulation with or without additives and secondly it is determined with the right model.

The procedure to be followed in constructing the library involves of 5 steps (omitting step 4 for the lubricant bases library):

- 1. Recording the NIR spectra using sets of samples of known identity.
- 2. Choosing the spectra for each class (product), selecting samples from different batches.
- 3. Construction of the library by choosing the pattern recognition method to be applied (residual variance in our case). One must choose the spectral pre-treatment, wavelength range, variance

<span id="page-2-0"></span>



<sup>a</sup> Different ratio in acids employed.

explained and threshold value to be used. The following step involves an internal validation, in order to check for the possible existence of miss-identified, unidentified or ambiguously spectra. Based on validation results, whether some change in the threshold, spectral range, pre-treatment, etc., must be introduced is decided upon iteratively until assure the correct identification of the products.

- 4. Constructing sub-cascading libraries. The samples of those classes that result in an ambiguous identification in the general library are divided into sub-classes that apply a similar construction procedure followed in the construction of the general library.
- 5. External validation. The general library and sub-libraries must be validated by checking that external spectra (validation set) not used in the construction of the library, are correctly, unambiguously identified.

# 2.5. Data processing

Data were processed with the chemometric software Unscrambler v. 9.8 from Camo Process (Trondheim, Norway). Firstand second-derivative spectra were obtained by using the Savitzky–Golay algorithm with 11-point moving window and a second-order polynomial. Partial least-squares (PLS1) models were constructed by cross-validation, using the leave-one-out method. The quality of the results was assessed in terms of the relative standard error of calibration (RSEC) or validation (RSEP), and the root mean square error (RMSE):

RSE 
$$
(\%) = \sqrt{\frac{\sum_{i=1}^{m} (y_i^{\text{pred}} - y_i^{\text{ref}})^2}{\sum_{i=1}^{m} (y_i^{\text{ref}})^2}} \times 100
$$

$$
RMSE = \sqrt{\frac{\sum_{i=1}^{m} (y_i^{pred} - y_i^{ref})^2}{n}}
$$

where  $y^{pred}$  is the predicted value,  $y^{ref}$  the reference value and n is the number of calibration samples for RMSEC or validation samples for RMSEP.

The model selected in each case was that leading to the lowest RSE and RMSE for a test set consisting of samples not used to construct it.

# **3. Results and discussion**

The lubricant bases studied were polyol esters with very similar NIR spectra and no significant differences between their absorption bands [\(Fig.](#page-3-0) 2). The spectra exhibited two strong bands in the region of 1700-1800 nm (combination band of C-H bonds) in addition to two smaller bands at 1200 nm corresponding to the second overtone of C-H bonds and at 1400 nm corresponding to the first



**Fig. 1.** Chemical structure of pentaerythritol, dipentaerythritol and neopentylglycol.

<span id="page-3-0"></span>

**Fig. 2.** Near-infrared (NIR) spectra of the 9 lubricant bases.

overtone for such bonds. One other major band was observed at 2100–2200 nm due to combination band for  $C$ –H and  $C$ =O bonds. The signal saturated above 2200 nm owing to the large path length of the transmission cell used. The high similarity between spectra reflected in their correlation coefficients, all of which exceeded 0.9921 and exposed the difficulty of constructing qualitative models affording the distinction of individual bases and quantitative models allowing their determination in the mixtures.

The 9 starting bases were used to establish 23 lubricant mixtures in terms of the desired final properties. The spectra for the end products were also highly correlated, with a coefficient of at least 0.9913.

The primary aims were to identify and discriminate each lubricant base and preparation from spectral libraries, and to confirm whether the mixtures contained any of the additives typically added to the formulations. In addition, we used PLS models to determine the components of each formulation and the viscosity of the final lubricants at 40 and 100 ◦C.

#### 3.1. Discriminating lubricant bases and preparations

In this study, we have constructed two libraries: The first one has been built for the identification of the lubricant bases. The second library has been constructed for the identification of the final lubricant formulations and also capable of identifying whether a mixture containing additives that are added to the formulation.

The lubricant bases identification library (or first spectral library) was constructed from 168 second-derivative spectra for various production batches of samples of the 9 lubricant bases, using the wavelength region of 1100–2200 nm, a threshold of 0.950 and an explained variance of 96%. These conditions allowed the accurate identification of the lubricant bases included in the library, which was validated both internally to avoid confusion between spectra for the lubricant bases and externally with 38 lubricant bases spectra not included in the library. It has been shown the correctidentification of all lubricant bases and the absence of ambiguities in the identification process.

The identification of lubricant formulations without and with antioxidant additives added, was accomplished by using a second spectral library with the cascading method. This second spectral library, general library, contained 1034 second-derivative spectra spanning the wavelength range of 1100–2200 nm (threshold, 0.952; explained variance, 98%) for 23 different lubricant formulations with and without additives. Some compounds exhibited identification values below the threshold for more than one class and were thus identified ambiguously – more accurate identification is required using sub-libraries. The spectral library used to identify the lubricant formulations was validated both internally to avoid confusion between spectra for the formulations and externally with 186 lubricant formulation spectra not included in the library. All formulations were thus identified and qualified unambiguously – lubricants containing no additives included.

An example that illustrates the ability of cascading libraries differentiation is the resolution of one ambiguity arisen between two lubricant formulations (lubricant formulation 1 and lubricant formulation 2) differing only slightly in their proportions of neopentylglycol and C8 acid (lubricant base 9) on the one hand, and pentaerythritol and C5–C7 acids (lubricant base 2). Both contained 0.05% BHT, but only lubricant formulation 2 included Silicone Fluid (0.012%). The correlation coefficient between the spectra for these lubricant formulations was >0.999. Their sub-library was constructed at a threshold of 0.86 and an explained variance of 99%. [Table](#page-4-0) 2 shows the identification results obtained with the general library and the sub-library. The identification results in the general library show the ambiguity between the two lubricant formulations; it is necessary to employ the sub-library to identify each lubricant formulation correctly. As can be seen, the identification values in the sub-library were both lower than the threshold and correctly identified (positive identification), and the products were not identified correctly when they exceeded it (negative identification).

# 3.2. Construction of the calibration models

Once the identification of the lubricant formulation are been made, the system selects the appropriate calibration model for determining the composition and/or viscosity. Among 23 lubricant formulations, one has been chosen as an example to show the calibrations models built, for the determination of lubricant bases composition, additives composition and viscosity at 40 ◦C and  $100 °C$ 

Lubricant formulation viscosity and composition were assessed in the selected sample consisting of a pentaerythritol base esterified with a mixture of pentanoic (C5) and heptanoic acid (C7), lubricant base 2 (50.6%), and other consisting of neopentylglycol esterified with heptanoic acid (C7), lubricant base 8 (48.65%). The NIR spectra for the two lubricant bases were very similar, with correlation coefficients above 0.997. The high spectral similarity between the two precluded the construction of a calibration model for accurately predicting samples from the industrial process.

A total of 21 samples were prepared by mixing the lubricant bases in proportions spanning a range of  $\pm 5\%$  (w/w) around the nominal value for each. The samples were split into a calibration set that was used to construct the individual model for each variable (composition and viscosity) and a validation set consisting of samples not included in the previous one that were used to assess the predictive ability of the ensuing models. The samples included in the calibration set for viscosity and each lubricant formulation spanned the studied range in the corresponding variable.

The calibration models for each variable were constructed from both absorbance spectra and processed spectra (SNV, first and second derivative with Savitzky–Golay smoothing). Also, models spanned the whole spectral region (1100–2200 nm) and various ranges excluding non-absorbing zones in order to reduce spectral noise. [Table](#page-4-0) 3 shows the best calibration models for lubricant formulation composition and viscosity at 40 and 100 ◦C. All models were very simple; thus, 2 PLS factors sufficed with second-derivative spectra spanning the wavelength ranges of 1100–1850 and 2000–2200 nm. The table also shows the figures of merit of the PLS calibration models, which testify to the predictive ability of the four models. A regression plot of NIR concentration values against reference values was linear, and had a slope and

# <span id="page-4-0"></span>**Table 2**

Identification values of lubricant formulation 1 and lubricant formulation 2 samples identified ambiguous in the general library (threshold 0.952). In the sub-library the threshold is fixed at 0.86. Positive identifications (ID value < 0.86) are in bold face.



<sup>a</sup> Samples 1, 2 and 3 belong to lubricant formulation 1; samples 4 and 5 belong to lubricant formulation 2.

# **Table 3**

Figures of merit for the four partial least squares (PLS) calibration models (two lubricant bases composition and viscosity at 40 and 100℃) corresponding to a lubricant formulation.



Wavelength range of 1100–1850, 2000–2200 nm, with 2 PLS factors and second derivative spectral pretreatment.

### **Table 4**

Figures of merit for the four partial least squares (PLS) calibration models (two lubricant bases and two additives) corresponding to a lubricant formulation.



Wavelength range of 1100–1850, 2000–2200 nm, second derivative as spectral pretreatment.

<span id="page-5-0"></span>intercept not significantly different from 1 and 0, respectively, at the 95% confidence level.

The predicted values for the composition of the lubricant formulation and viscosity of 16 production batches were not significantly different from the specifications, testifying again to the high predictive ability of the proposed models.

The experimental work was supplemented with tests involving the determination of the additives typically added at the end of the lubricant formulation production process. A total of 26 samples were prepared by mixing lubricant's bases spanning a concentration range around  $\pm 5\%$  (w/w) of their nominal value and two additives having a nominal concentration of 0.7 and 0.05%, respectively, that were used at concentrations from 0 to 1.5% and 0 to 0.08%, respectively. As before, the samples were split into a calibration set and a validation set to construct calibration models for the composition of the two lubricant's bases and antioxidants. [Table](#page-4-0) 4 shows the ensuing calibration models. All were simple – albeit more complicated that those for lubricant base composition alone – and included 3 PLS factors for the lubricant's bases and 7 for the additives (second-derivative spectra, wavelength ranges of 1100–1850 and 2000–2200 nm). [Table](#page-4-0) 4 additionally shows the figures of merit of the PLS calibration models, which reflect the good predictive ability of the four PLS models. A regression plot of NIR concentration values against reference values was linear and had a slope and intercept not significantly different from 1 and 0, respectively, at a 95% confidence level.

The predicted values of the composition of lubricant bases and additives of 16 production batches were not significantly different from the specifications, testifying again to the high predictive ability of the proposed models.

#### **4. Conclusions**

As shown in this work, NIR spectroscopy in combination with multivariate calibration models affords analytical control of composition and physical properties. The proposed method allows both the identification of components of lubricant formulations with very similar spectra, the detection of antioxidants present in the formulations and the simultaneous determination of components and additives in a lubricant formulation for which no analytical control method had previously been reported. The construction of cascading libraries allows the identification of the presence of low contents of additives in formulations. The proposed method, which affords the accurate simultaneousdeterminationof all sample components in a short time (1 min), requires no sample pretreatment; also, it uses no reagents and produces no waste.

# **Acknowledgements**

The authors gratefully acknowledge funding of this work by Spain's Ministry of Science and Technology (Project CTQ2009- 08312). David Zamora also wishes to acknowledge additional support in the form of a scholarship from Universitat Autònoma de Barcelona. The authors are also indebted to the firm Croda (Fogars de la Selva, Spain) for kindly supplying the samples.

#### **References**

- [1] G.E. Totten, S.R., Westbroock, R.J. Shah, Fuels and Lubricants Handbook: Technology, Properties, Performance, and Testing, ASTM Manual Series: MNL37WCD, West Conshohocken, PA, USA, 2003.
- [2] F. Haus, J. German, G.A. Junter, Chemosphere 45 (2001) 982–990.
- [3] A. Willing, Chemosphere 43 (2001) 89–98.
- [4] K.N. Marsh, M.E. Kandil, Fluid Phase Equilibr. 199 (2002) 319–334.
- [5] T. Wakabayashi, S. Suda, J. Jpn. Pet. Inst. 51 (2008) 134–142.
- [6] A. Lucia, Q. Luo, Adv. Environ. Res. 6 (2002) 123–134.
- [7] A.S. Pensado, M.J.P. Comuñas, L. Lugo, J. Fernández, Ind. Eng. Chem. Res. 45 (2006) 2394–2404.
- [8] A.S. Pensado, M.J.P. Comuñas, J. Fernández, Ind. Eng. Chem. Res. 45 (2006) 9171–9183.
- [9] O. Fandiño, M.J.P. Comuñas, L. Lugo, E.R. López, J. Fernández, J. Chem. Eng. Data 52 (2007) 1429–1436.
- [10] A.S. Pensado, A.A.H. Pádua, M.J.P. Comuñas, J. Fernández, J. Supercrit. Fluids 44 (2008) 172–185.
- [11] P. Baptista, P. Felizardo, J.C. Menezes, M.J. Neiva Correia, Talanta 77 (2008) 144–151.
- [12] M. Blanco, S. Maspoch, I. Villarroya, X. Peralta, J.M. González, J. Torres, Anal. Chim. Acta 434 (2001) 133–141.
- [13] M. Blanco, D. Zamora, J. Planells, R. Mulero, Anal. Bioanal. Chem. 395 (2009) 839–844.
- [14] M. Blanco, M.A. Romero, Analyst 126 (2001) 2212–2217.