APPLICATION OF FTIR SPECTROMETRY USING MULTIVARIATE ANALYSIS FOR PREDICTION FUEL IN ENGINE OIL

Marie Sejkorová

1Department of Transport Means and Diagnostics, Jan Perner Transport Faculty, University of Pardubice, Studentská 95, 532 10 Pardubice, Czech Republic

Abstract


This work presents the potentiality of partial least squares (PLS) regression associated with Fourier transform infrared spectroscopy (FTIR spectrometry) for detecting penetration of diesel fuel into the mineral engine oil SAE 15W-40 in the concentration range from 0 % to 9.5 % (w/w).

As a best practice has proven FTIR-PLS model, which uses the data file in the spectral range 835–688 cm\(^{-1}\). The quality of the model was evaluated using the root mean square error of calibration (RMSEC) and cross validation (RMSECV). A correlation coefficient \(R = 0.999\) and values of RMSEC, RMSECV were obtained 0.11 % and 0.38 % respectively. After the calibration of the FTIR spectrometer, the contamination engine oil with diesel fuel could be obtained in 1–2 min per sample.

Keywords: engine oil, FTIR spectrometry, multivariate analysis, lubricant analysis, lubricant quality, diesel fuel, partial least squares (PLS) regression.

INTRODUCTION

Reliability and a length of a service life of vehicles or other machines depend not only on operational use of the given device but also on features of the oil that is used for lubrication (Stodola and Mazal, 2016; Raffai et al., 2015; Glos and Sejkorová, 2016).

Lubricants play a key role in extending the working life of industrial machines and transport means. Lubrication oils are during their operation in the lubrication system of an engine exposed to high temperatures (Pošta et al., 2016), pressures and contamination by undesired materials from outside. These factors cause troubles with the operational quality of oils and materials of mechanical components of a lubrication system. The origin of degradation products and the contamination of oil from outside cause the additives to decrease.

Among the most basic observed parameters of the engine oils that have a direct relation with their performance and therefore also with their change interval, we can include viscosity, viscosity index (VI) and a total base number (TBN) (Al-Ghouti et al., 2010).

Based on a change of viscosity, it is possible to determine the engine oil degradation as a result of thermo-oxidation reactions, destruction of viscosity modifiers due to shear forces, contamination by mechanical dirt, water, glycol and fuel (Burg et al., 1997; Knothe and Steidley, 2007). The effect of lubricating oil degradation to the viscosity followed Kumbár and Dostál (2013), Kosiba et al. (2013).

In diesel engines fuel dilution is usually caused by excessive idling, defective injectors, or loose connections. Most engine oil producers define the maximum allowed volume of fuel in the oil as 4–5 % (Glos, 2015).

Veselá et al. (2014) evaluated the effect of biofuels E85 and fuel with 95 octane values types on the quality and purity of engine oil Mobil 10W-40 used in three types of cars. Fluid purity is also one of the important features in the entire process (Máchal et al., 2013). Hönig and Hromádko (2014) analyzed the impact of biofuels on the quality of oil used to
fill diesel engines, and established an optimal oil change interval.

Penetration of fuel into the engine oil also affects another physical characteristic of oil – the flash point. A new oil must have a low resistance and therefore a high flash point level, usually between 230–240 °C. Penetration of fuel into the oil causes that the flash point decreases. The limit value for diesel engines is usually defined on a level of 180–190 °C.

The amount of gasoline in the lubricant can be determined by distillation and gas chromatography (GC) (Mujahid and Dickert, 2012). There are also sensors to observe intrusion of contaminants in the motor oil such as fuel (Capone et al., 2008) which indicates a defect of the engine, usually the fuel injection. There analyzer fuel dilution meter (Sejkorová et al., 2016), which employs a surface acoustic wave vapor microsensor to measure the concentration of fuel in used lubricating oil samples by sampling the head-space in the sample bottle.

In recent years, research has been carried out to apply FTIR spectrometry with chemometrics algorithm for the quality control of lubricant oils. Chemometrics is the science discipline that applies the mathematic-statistical methods to the results of chemical measuring for the purpose of extracting as much relevant information as possible from the chemical data (Sejkorová, 2014). The most applied multivariate methods are classical least squares (CLS), inverse least squares (ILS), principal component regression (PCR) and partial least squares (PLS) regression. According to Sejkorová (2013) and Bassbasi et al. (2013) FTIR spectrometry as methods for lubricating oil analysis it is generally rapid, can be automated and can reduce the need for solvents and toxic reagents associated with wet chemical methods.

Al-Ghouti et al. (2010) dealt with application of chemometry in combination with the FTIR spectrometry to define the viscosity index and the total base number. Caneca et al. (2006) applied the FTIR spectrometry in connection with multiple linear regression (MLR), PCR and PLS in order to predict a viscosity of the motor oil suitable for compression-ignition (diesel) engines. FTIR spectroscopy in combination with a PLS (PLS) used Borin and Poppi (2005) for the quantitative determination of the penetration of contaminants (gasoline, ethylene glycol and water) into the lubricating oil.

Glos (2015) and Glos and Svboda (2015) applied special FTIR analyzer Spectro Oil Q400 for the determination of contamination and depletion of additives in used mineral engines oils.

**MATERIALS AND METHODS**

**Oil samples**

In total, 26 samples of motor oil with added diesel were prepared for the experiment. The set consisted of 10 samples of motor oil Shell Helix HX3 15W-40 contaminated by diesel with added rapeseed oil methyl ester (RME) in concentration of 0–7.62% of mass, and 16 samples of motor oil Rubia 15W-40 contaminated by a pure diesel in concentration of 0–9.5% of mass.

For weighting, the laboratory scales Denver Instrument SI-2002 accurate to ±0.01 g were used.

**Acquisition of IR spectra**

Infrared spectra were recorded by a FTIR spectrometer Nicolet iS10 (Thermo Scientific) equipped with deuterated triglycerine sulfate detectors (DTGS) and a horizontal attenuated total reflectance (ATR) with ZnSe crystal. The device is equipped with an OMNIC software that allows setting parameters of the spectrometer, controls the measurement itself and runs subsequent operations with the spectra. MIR spectra of the motor oil samples were collected over the spectral range 4,000–650 cm⁻¹. Measurement parameters: resolution 4 cm⁻¹, number of spectrum accumulations 64.

Obtained data were processed by TQ Analyst version 8 (Thermo Scientific, Inc., USA), QC Expert version 2.5 (TriloByte, Czech Republic).

**Data analysis**

In the research, the PLS regression algorithm was employed to build quantitative calibration model for concentration diesel fuel in mineral engine oil. Helland (1990), Yeniay and Goktas (2002) describes the principle of the PLS regression.

The evaluation of the calibration performance was assessed by root mean squared error of calibration (RMSEC) and correlation coefficient (R) (Sejkorová, 2014).

The FTIR-PLS calibration model was obtained with optimum number of latent PLS factor; it was selected based on the predicted residual error sum of squares (PRESS) by the cross-validation results. PRESS (Yeniay and Goktas, 2002) is described in Eq. 1

\[
PRESS = \left[ \frac{1}{n} \sum_{i=1}^{n} (C_i - \bar{C_i})^2 \right]^{1/2} \min
\]

where \(\bar{C_i}\) are values predicted by FTIR-PLS model, \(C_i\) are actual values and \(n\) is a number of calibration standards. PRESS diagnostics tested a change of value of root mean square error of cross validation (RMSECV) when a factor is added to the calibration model (Helland, 1990). The performance of the final PLS model is evaluated according to RMSECV, a leave-one-sample-out cross-validation (Juránek et al., 2012) is performed: the spectrum of one sample
of the calibration set is deleted from this set and the PLS model is built with the remaining spectra of the calibration set. The left-out sample is predicted with this model and the procedure is repeated with leaving out each of the samples of the calibration set.

RESULTS AND DISCUSSION

The infra red (IR) spectra of two new engine oils: Shell Helix HX3 15W-40 (red) and Rubia 15W-40 (blue) are shown in Fig. 1. The main peaks observed in Fig. 1 are associated with major functional groups present in this type of engine oils. Infrared peaks at 2,954, 2,920 and 2,848 cm\(^{-1}\) correspond to CH stretching of saturated n-alkyl groups (Hirri et al., 2013).

The band at 1,707 cm\(^{-1}\) can be referred to as polymethacrylate in the oil samples. According (Al-Ghouti and Al-Atoum, 2009) polymethacrylate is being used as a viscosity modifier and a pour-point depressant additive for engine oil and typically exhibits the bands at 1,707 and 1,156 cm\(^{-1}\). According to Kupareva (2012) IR spectra at 1,456 cm\(^{-1}\) show CH deformation of CH\(_2\) and CH\(_3\) groups, which are associated with the carbonates of overbased sulfonates. Vibrations at 1,376 cm\(^{-1}\) and 1,156 cm\(^{-1}\) correspond to CH\(_3\) vibrations of n-alkanes. Spectral band associated with the P-O-C bond of Zinc dialkyl dithiophosphates (ZDDPs) is around 974 cm\(^{-1}\). ZDDPs are organometallic compounds, which are one of the most effective antioxidants. The characteristic frequency of ZDDPs (P = S bond) is also around 660 cm\(^{-1}\).

The PLS regression method was used to construct the calibration models. These methods work with the whole spectrum or its parts; their basis is to compress the spectral information into new variables – PLS factors that represent linear combinations of original absorbances and that have no mutual correlation. Each component is a linear combination of original absorbances; its fundamental attribute is the degree of variability – dispersion.

Previous studies of the author's paper (Sejkorová, 2014) have shown that spectral region and the number of latent PLS factor are key parameters for construction of good calibration models.

The two statistics were used for identifying outliers in the calibration set, it was sample leverage and the studentized residuals (Marinović et al., 2012a). One sample of 26 set samples was identified as an outlier. For model was chosen spectral region where the change in absorbance correlates the most with varying the concentration of fuel in the engine oil. The calibration model for the spectral range 835–688 cm\(^{-1}\), i.e. matrix of range (25×147), was tested. Fig. 2 shows spectrum for new oil and diesel fuel in the studied spectral region. The bands in IR spectrum between 700 and 850 cm\(^{-1}\) can be associated to large number of skelatal vibration of various ring structures (Marinović et al., 2012b).

Borin and Poppi (2005) state that for gasoline determination the region from 650–784 cm\(^{-1}\), attributed to aromatics, was selected.

In determination of the optimal number of factors RMSECV is very important, because it shows the changes in the RMSECV value with an increasing number of factors that are used in model development of each single characteristic. When a PLS model is calibrated, all of the relevant concentration information of diesel fuel and spectral information in the analysis region or regions of the calibration standards is condensed into a set of factors. The parameter RMSECV reached the minimum value 0.38% with use of 8 PLS factors (Fig. 3).

In Fig. 4, we observe the excellent correlation coefficient (R = 0.999) between actual and
predicted values for calibration sets. The value RMSEC = 0.11%.

The calibration, set used to construct FTIR-PLS model, to determine the concentration diesel fuel in engine oil was also tested by cross-validation with sequential leaving one calibration point out. An important result is a correlation coefficient of cross validation (Rcv) (Sejkorová, 2014) that reached the value of Rcv = 0.989 with RMSECV = 0.38%.

Borin and Poppi (2005) reached the almost the same errors when testing external standard lubricating oil with gasoline (RMSEP = 0.34%).

Values of diesel fuel predicted by the FTIR-PLS model were then statistically compared by QC Expert software against results the actual values. The pair t-test (α = 0.05) proved that differences between actual and predicted values were not statistically significant.
CONCLUSION

Fast instrumental methods which provide qualitative as well as quantitative data about degradation processes in the exploited oil fill include FTIR spectrometry. With the development of information technologies and the expansion of software equipment of spectrometers with highly-demanding mathematic-statistical methods, chemometric predictive models are developed which allow for obtaining many parameters of non-chemical (most often physical) character, from a single spectrum by means of multi-component analysis of oils.

In the paper, we present a methodology for construction of the FTIR-PLS regression model for purposes of determination of contamination of the mineral engine oil by a pure diesel and by a diesel with added rapeseed oil methyl ester. The region from $835–688 \text{ cm}^{-1}$, attributed to aromatics, was selected for diesel fuel determination. In the determination, the RMSEC = 0.11 %, $R = 0.999$ and RMSECV = 0.38 %, $R_{cv} = 0.989$.

If the developed methodology for other monitored parameters of degradation and contamination of the engine oil (e.g. viscosity, viscosity index, total base number, total acid number, ethylene glycol and water), it would be possible within 1–2 min of a single spectrum to provide several monitored quantitative parameters per sample.

Common evaluation of engine oil samples by means of classical standardized methods does not lead to results in time that would be short enough; this delays remedial actions in operation. Disadvantages of FTIR-PLS procedures include lower precision as compared to primary methods and considerable experimental demands of creating a quality model.

REFERENCES


