Slavica Marinović, Ante Jukić, Dubravka Doležal, Branka Špehar, Marko Krištović

PREDICTION OF USED LUBRICATING OILS PROPERTIES BY INFRARED SPECTROSCOPY USING MULTIVARIATE ANALYSIS

Abstract

Infrared spectroscopy (IR) in combination with partial least squares regression (PLS) algorithm was applied to develop calibration models for rapid and accurate simultaneous determination of the properties of used lubricating oils: density, viscosity, and pour point. Training sets were composed of hundred used lubricating oil samples. PLS calibration models for each property have been developed by using IR transmission spectra without spectral pretreatment. Cross validation technique was applied for validating models. The PLS models showed very good correlation with corresponding reference methods. Results indicated PLS models based on infrared spectra as practical and reliable analytical methods for routine monitoring of used lubricating oil.

1. Introduction

The basic role of lubricants i.e. lubrication is to reduce friction and hence prevent the wear of material surfaces as a consequence of relative mutual movement. However, it is essential that the lubricant also has other functional properties that will ensure its efficient application. These are: good oxidation and thermal stability, corrosion protection property, compatibility with different materials, low foaming, ability to release air, good detergent-dispersant properties, etc. [1] Therefore, monitoring the lubricating oil condition is essential for preservation of activity and longevity of engines. One of the most important properties of lubricating oil in maintaining the lubricating film between moving parts is its viscosity. Significant deviation (+ or – 20%) from viscosity of the fresh oil may indicate a build up of oxidation products due to overheating, contamination with fuel, water, etc. Measurement of density can be used to confirm is the product in use correct, by comparison of the results against product specification. The pour point, or the temperature at which a lubricant ceases to flow, is important property of lubricating oils at low temperature. The pour point depends of the oil chemical structure. The mentioned properties of lubricating oils are usually determined by the standard methods such as ASTM, ISO and others.

These methods can be time consuming and require large amounts of sample for analysis. An alternative to these procedures is the use of infrared spectroscopy techniques (IR, NIR) combined with multivariate analysis. Measurements of IR data are fast, they can be easily automated and need only a few milliliters of sample. That makes IR spectroscopy a powerful method for studying the degradation products and monitoring the conditions of used lubricating oils.

Vibrational spectroscopy (IR, NIR, Raman) in combination with multivariate analysis and neural networks [2] has been widely applied for monitoring physico-chemical properties of naphtha [3], heavy petroleum products [4], and motor fuels [5-10]. Usually, infrared spectroscopy and the partial least squares algorithm (PLS) are combined to determine properties of diesel fuels like cetan number, cetan index, density, viscosity, aromatics etc. Andrade et al. [11] applied Raman spectroscopy and multivariate calibration to evaluate physico-chemical properties of aviation fuel. Some researches have already been conducted on the analysis of lubricating oils by infrared spectroscopy. Adams et al. [12] used IR spectroscopy in combination with principal components and partial least squares algorithms to determine antioxidant content and total acid number in synthetic aviation lubricating oils. Infrared spectroscopy proved to be a useful alternative to conventional analytical methods for monitoring aging of oil samples from aviation engines. The developed calibration models produced prediction errors comparable to the standard methods. Canaca et al. [13] applied infrared spectroscopy and multivariate calibration to predict viscosity of diesel-lubricating oils. In this case, MLR (multiple linear regression), PCR (principal component regression) and PLS techniques were applied in the calibration. The using attenuated total reflectance (ATR) measurements in the infrared spectral range, RMSEP (root mean square error of prediction) and relative average error were 3.8 cSt and 3.2 %, respectively. These values can be satisfactory for monitoring the condition of lubricants in service. Borin et al. [14] used mid-infrared spectroscopy and interval PLS (iPLS) for the quantification of contaminants (gasoline, ethylene glycol and water) in lubricating oil.

In the study described here, three properties: density, viscosity, and pour point of different used lubricating oils (gas engine and turbine oils) were determined using combination of infrared spectroscopy and partial least squares algorithm in development of calibration, prediction and validation models. It was expected that the obtained results will provide insight into the change of lubricant's performances in very short time, using less than 1 mL of the sample. Recording only one IR spectrum of used lubricating oils should be sufficient to determine these important physical properties.

2. Experimental

2.1. Sampling

A set of 100 samples of used lubricating oils (different types of gas engine and turbine oils in different stages of use) was used for determination of density and viscosity. The 85 samples were used to create calibration models for the pour point.

The samples were collected from gas compressor station in the inland part of Croatia and compressor platform in the northern Adriatic Sea during a period of two years. All samples were stored in tightly sealed glass bottles at max. temperature of 4 °C. Before instrumental analysis, samples were equilibrated at room temperature (22 ± 5 °C).

2.2. Standard methods
The properties of lubricating oils used for calibration models were mainly defined using standard test methods based on the EN / ISO / ASTM procedure.

2.2.1. Density
Density was determined using an oscillating U-tube densimeter (EN ISO 12185). A small portion of the oil sample (0.7 mL) was introduced into a temperature-controlled cell. The oscillation frequency was noted, and the density of the test sample calculated using cell constants. The constants were previously determined by measuring the oscillation frequencies when the cell is filled with calibration fluid of known density [15].

2.2.2. Viscosity
The viscosity of each sample was determined by Stabinger viscometer (ASTM D 7042). The sample was introduced into the measuring cells, which were under known and carefully controlled temperature. The measuring cells consisted of pair of rotating concentric cylinders and an oscillating U-tube. The dynamic viscosity was determined from the equilibrium rotational speed of the inner cylinder under the influence of the shear stress of the test specimen and an eddy current brake in conjunction with adjustment data. The kinematic viscosity was calculated by dividing the dynamic viscosity by the density [16].

2.2.3. Pour Point
The pour point of used lubricating oils samples was determined using standard method for determination of pour point in all petroleum products, ASTM D 97 [17]. After preliminary heating, the sample was cooled at a specified rate and examined at intervals of 3 °C for flow characteristics. The lowest temperature at which movement of the sample has been observed was taken as pour point.

2.3. Instrumentation and software
Infrared transmission spectra were obtained on Nicolet 6700 Fourier transform instrument (Thermo Fisher Scientific Inc, USA). The spectra were recorded using NaCl transmission cell with 0.1 mm path length, 30 scans with a resolution of 4 cm^{-1}, covering the 4000-650 cm^{-1} spectral range, DTGS (deuterated triglycine sulfate) detector and KBr beamsplitter. A liquid cell was chosen rather than using attenuated total reflection cell because it gave a longer optical path length for the sample. The longer path length results in stronger absorbance for the samples and so gives larger peaks in the weakly absorbing regions. It has a disadvantage that some regions of the spectrum become too strongly absorbing to be useful. These saturated spectral regions are those which are expected to be unchanged with the

changes in oil condition and their influence on the prediction model would be insignificant. The infrared spectral information for the calibration model was improved by blanking out areas of high absorption (absorbance values that exceed 1.5 a.u.) (Fig. 1.). The areas blanked out and so not used in the model were from 3010 to 2820 cm^{-1} (high absorbance due to C-H stretching vibrations) and from 1495 to 1350 cm^{-1} (high absorbance due to C-H bending vibrations). Some degradation products of used oil (Fig. 1.) can affect the physical properties such as viscosity, pour point, etc. Oxidation and nitration products were formed after organic compounds were exposed to high temperatures and pressures in the presence of oxygen and nitrogen. Oxidation products like carboxylic acids contributes to the acidity of the oil which causes corrosion and viscosity increase. Infrared spectroscopy determines the level of oxidation by a general response in the carbonyl (C=O) region of between 1800 to 1670 cm^{-1}. A rise of the nitration index of oil can reflect on operating condition, i. e. high loads, low operating temperature etc. Nitration products have characteristic absorbance between 1650 to 1600 cm^{-1}, the region immediately below the one of oxidation products (Fig. 1.). Soot build-up (absorbance intensity at 2000 cm^{-1}, Fig. 1) is a problem in lubrication oil since it causes combustion problems and also increases the viscosity [18].

PLS calibration models were developed using TQ Analyst 8.0 (PLS1) software included in the Omnic Software^{TM} (version 8.1). The 20 factor models were created using "leave-1-out" cross validation. This means that the models were built 100 times with 99 samples, leaving out a different sample every time. The model then predicts the sample left out and an error value is obtained for the model containing 1, 2, and up to 20 factors. The resulting residual error was used to determine the optimal number of factors for models. The only preprocessing parameter performed was the mean centering of the data obtained by averaging the spectrum of all 100 samples and subtracting resulted average from each of the spectra. This allows the PLS model to operate on true spectral variance in the data set.
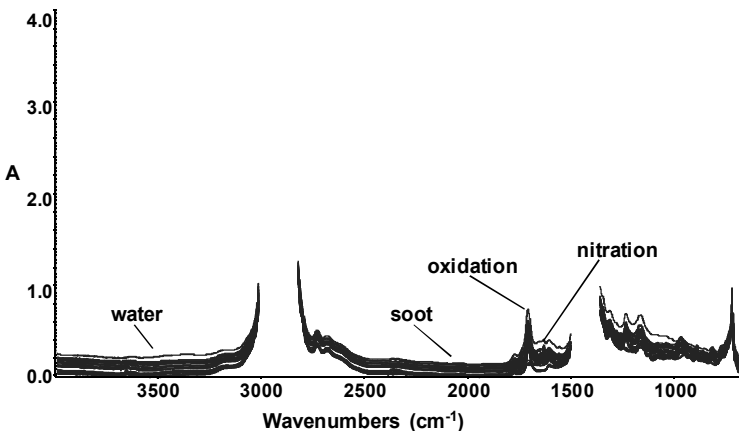

Figure 1: Infrared transmission spectra of used lubricating oil samples

3. Results and discussion

3.1. Identification of outliers

The sample leverage and the studentized residuals were used for identifying outliers in the calibration set. The data points in the studentized residuals vs. leverage plot should be distributed evenly throughout the plot. A data point that is isolated from the others indicates that the corresponding sample is different from other samples in the models. The samples with small studentized residuals and small laverage are typical in the calibration set. Samples with large studentized residual and small leverage are suspect for possible concentration errors (mislabeled simples). Large sample leverages and small studentized residuals may be caused by spectral error or by perfectly good samples that happen to lie at edge of the experimental design. Therefore, these samples should be examined closely before they are removed from the calibration set.

Fig. 2. shows the relationship between studentized residuals and leverage in determination density, viscosity and pour point of used lubricating oil by infrared spectroscopy. The leverage value of one data point at Fig. 2. (in case of all three properties) is significantly higher than the leverage values of the other data points. The difference in the leverage value may be due to a characteristic of the sample that is important to the analysis or it may be caused by a spectral error. However, because the spectral error for this standard was observed, this data point is identified as outlier and excluded from the model. The data points with high studentized residuals (> 3) (Fig. 2) are evidently outliers and because of great concentration differences in comparison to the others and observed concentration errors they had an influence on calibration accuracy. Therefore they were also excluded from the model.

3.2. PLS models

Spectral region and the number of factors are key parameters for construction of good calibration models. In accordance with this observation, only spectral regions containing valid data were used (4000-3097, 1864-1552, and 1320-665 cm^{-1}). The excluded spectral regions contained no spectral data or had absorbance values that exceeded 1.5 a.u. (blank out areas). In determination of the optimal number of factors, root mean square error of cross validation (RMSECV) diagnostic test is very important, because it shows the changes with increasing number of factors that are used in model development of each single characteristic. When PLS model is calibrated, all the relevant concentration and spectral information from analysis regions of the calibration standards are condensed into a set of factors. Each factor represents an independent source of variation in the data. Factors are ranked by the amount of variance they describe. The first factor describes most of the variation in the calibration standards. Each additional factor describes most of the remaining variation. Each time a factor that represents useful information is added to the calibration model, the error is reduced and RMSECV value decreases.
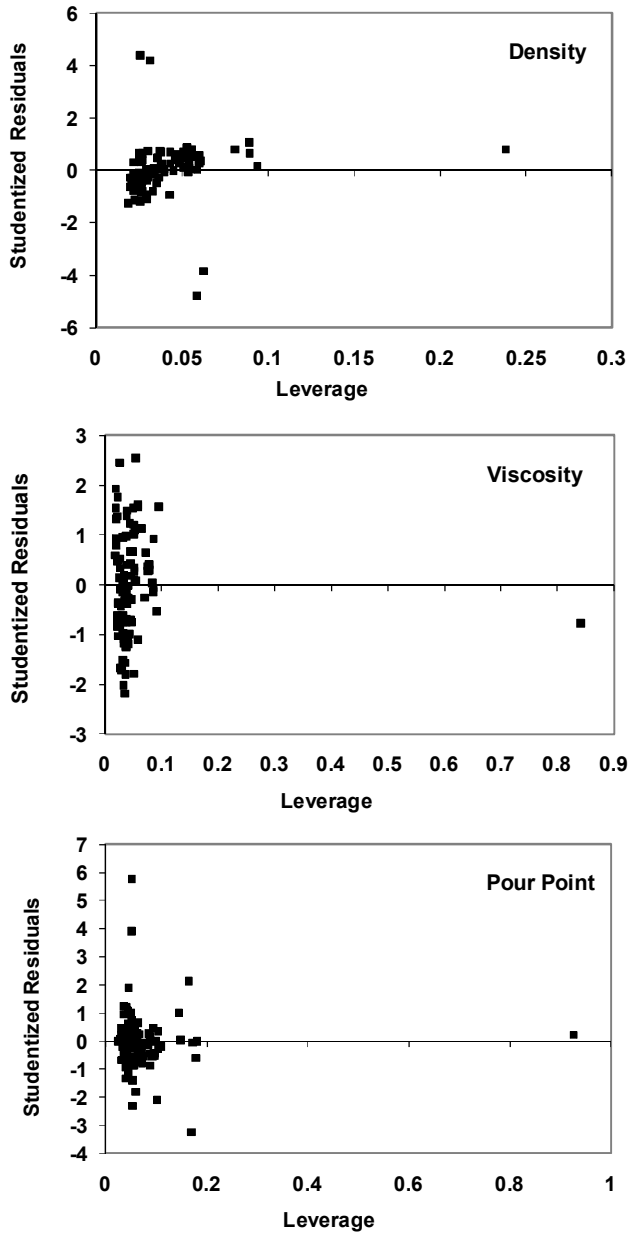
Figure 2: Studentized residual for PLS model as a function of leverage value in identification of outliers by infrared spectroscopy

The optimal number of factors for a model is indicated by minimum of RMSECV vs. number of factors plot. Fig. 3. shows relation between RMSECV and number of PLS factors for determination of density, viscosity and pour point of used lubricating oil. As expected, RMSECV rapidly decreases with the first few factors, while the decrease becomes considerably slower for each additional factor. With an additional increase in the number of factors, RMSECV starts to increase after 12 and 15 factors for density and viscosity, i.e. after 9 factors for pour point determination (Fig. 3). The increase in the RMSECV value indicates that the model is over-fitted i.e. that spectral information unrelated to selected property is included. This increase is particularly evident in the pour point determination. For that reason, the optimal number of factors for determination of used lubricating oils property was 12 and 15 for density and viscosity, respectively, while 9 factors were required for pour point determination. The results of calibration models for each property of used lubricating oils are given in Table 1. The correlation coefficient (R_{cal}^2) and root mean square error of calibration (RMSEC) are used to indicate the goodness of fit between the predicted and actual values. The leave-1-out cross validation technique was applied for validating model. This diagnostic quantifies each calibration standard as if it is a validation standard. Validation data are better indicators of model accuracy than calibration data because calibration standards are used to generate the model while validation standards are used only for method validation. However, the same samples were used to estimate correlation coefficient (R_{val}^2) and cross validation error (RMSECV). The numbers of required factors for the determination of each property of used oil are also listed in Table 1. The results based on correlation coefficients and calibration errors indicate very good correlation between standard methods and the developed FTIR model. Exception are the calibration results for pour point determination. The data given in table also shows that better results were obtained in a calibration data in comparation to the cross-validation data because cross validation data are true indicator of method accuracy. Scatter plots showing correlation between standard (actual) and IR predicted values for each of the used oil properties obtained by PLS model for calibration and validation data set are shown in Fig. 4. It is evident that IR model correlate very well with the standard test methods, while generally, calibration data are slightly closer to the ideal line in comparison to the validation data. Worse correlations were achieved only for pour point determination, which is probably the result of a fewer number of available samples for calibration. Grouping the data in the lower and upper part of the calibration line is the result of using different types of lubricating oils (gas engine and turbine oils) for the model construction (Fig. 4.).

Table 1: PLS calibration parameters

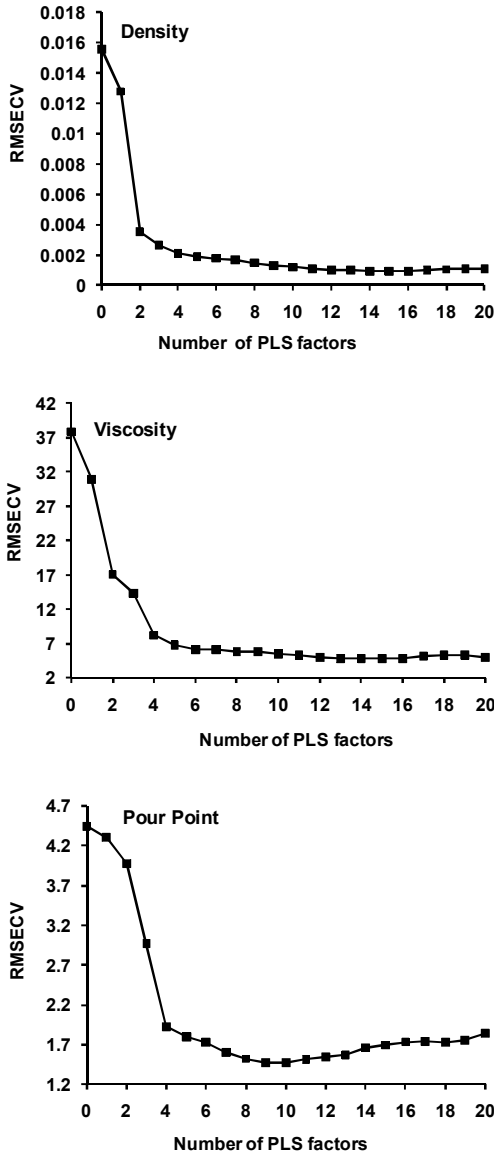| Oil property	Number of factors	R_{cal}^2	RMSEC	R_{val}^2	RMSECV
Density	12	0.9991	$0.673*10^{-3}$	0.9978	0.00105
Viscosity	15	0.9972	2.88	0.9916	5.00
Pour point	9	0.9622	1.31	0.9424	1.61

Figure 3: Root mean square error of cross validation (RMSECV) for PLS model as a function of number of PLS factors
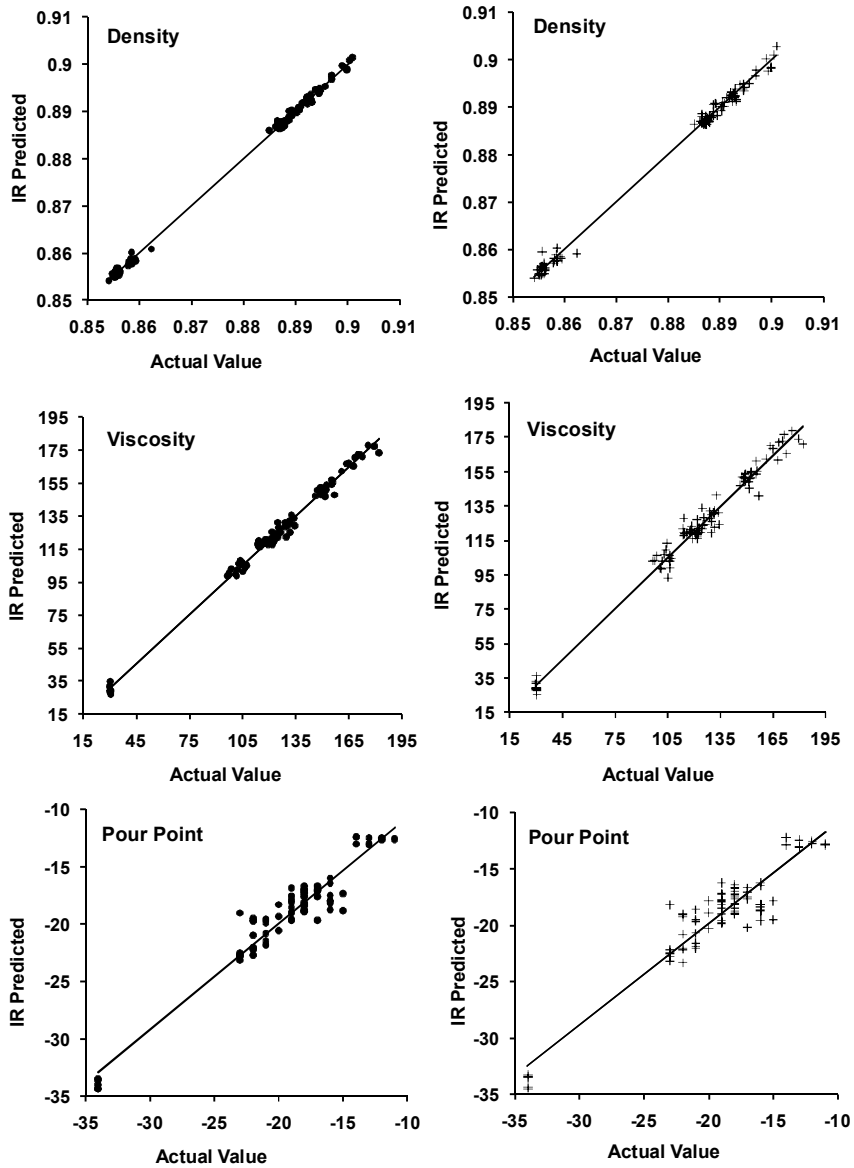
Figure 4: Predicted vs. actual values for each of the used oil properties obtained by PLS model for (•) calibration and (+) validation set

4. Conclusion

In this study, the combination of infrared spectroscopy and partial least squares regression algorithm was applied in determination of density, viscosity and pour point of different lubricating oils. The developed model has relatively good correlations with standard methods. Worse correlations were achieved only for pour point determination, which was probably result of less samples used for the model construction. Obtained models can be used as own analytical method for the rapid monitoring of the used oils since it allows reliable, fast and simultaneous determination of important oil properties by using less than 1 mL of sample.

References

1. Novina B., Tomrlin B., *Goriva i maziva*, **41**, 227-243, 2002.
2. Marinović S., Bolanča T., Ukić Š., Rukavina V., Jukić A*., Chemistry and Technology of Fuels and Oils*, **48**, 1, 67-74, 2012.
3. Ku M. S., Chung H., *Applied Spectroscopy*, **53**, 557-564, 1999.
4. Chung H., Ku M. S., *Applied Spectroscopy*, **54**, 239-245, 2000.
5. Fodor G. E., Kohl K. B., Mason, R. L., *Analytical Chemistry*, **68**, 23-30, 1996.
6. Cooper J. B., Wise K. L., Groves J., Welch W. T., *Analytical Chemistry*, **67**, 4096-4100, 1995.
7. Cooper J. B., Wise K. L., Welch W. T., Bledsoe R. R., Sumner M. B*., Applied Spectroscopy*, **50**, 917-921, 1996.
8. Cooper J. B., Wise K. L., Welch W. T., Sumner M. B., Wilt B. K., Bledsoe R. R., *Applied Spectroscopy*, **51**, 1613-1620, 1997.
9. Gómez-Carracedo M. P., Andrade J. M., Calvińo M., Fernández E., Prada D., Muniategui S., *Fuel*, **82**, 1211-1218, 2002.
10. Marinović S., Jukić A., Krištović M., Špehar B., Rukavina V., *Journal of Analytical Chemistry*, **67**, 939-949, 2012.
11. Andrade J. M., Garrigues S., de la Guardia M., Gómez-Carracedo M., Prada D., *Analytica Chimica Acta*, **482**, 115-128, 2003.
12. Adams M.J., Romeo M.J., Rawson P., *Talanta*, **73**, 629-634, 2007.
13. Caneca A. R., Pimentel M. F., Galvão R. K. H., da Matta C. E., de Carvalho F. R., Raimundo Jr. I. M., Pasquini C., Rohwedder J. J. R., *Talanta*, **70**, 2, 344-352, 2006.
14. Borin A., Poppi R. J., *Vibrational Spectroscopy*, **37**, 27-32, 2005.
15. EN ISO 12185-06 Crude petroleum and petroleum products – Determination of density – Oscillating U-tube method.
16. ASTM D7042-04 Standard Test Method for Dynamic Viscosity and Density of Liquids by Stabinger Viscometer (and the Calculation of Kinematic Viscosity).
17. ASTM D97-09 Standard Test Method for Pour Point of Petroleum Products.
18. Robinson N., *Wearcheck – Technical Bulletin*, **18**, 1-8, 2000.

Authors
Slavica Marinović[1], Ante Jukić[2], Dubravka Doležal[1], Branka Špehar[1], Marko Krištović[1]
[1]INA Petroleum Industry Ltd., Lovinčićeva bb, 10000 Zagreb, Croatia
[2]University of Zagreb, Faculty of Chemical Engineering and Technology, Marulićev trg 19, 10000 Zagreb, Croatia