Wear and Tear: A Data Driven Analysis of the Operating Condition of Lubricant Oils

Roney Malaguti, Nuno Lourenço, Cristovão Silva

Wear and Tear: A Data Driven Analysis of the Operating Condition of Lubricant Oils

Roney Malaguti[1,3][0000-0001-5555-1625], Nuno Lourenço[2][0000-0002-2154-0642], and Cristovão Silva[3][0000-0002-7693-9570]

[1] Stratio Automotive - R&D Department, Rua Pedro Nunes - Quinta da Nora, Ed.D 3030 - 199 Coimbra
[2] University of Coimbra, CISUC, Department of Informatics Engineering
[3] University of Coimbra, CEMMPRE, Department of Mechanical Engineering
roney@stratioautomotive.com, naml@dei.uc.pt, cristovao.silva@dem.uc.pt

Abstract. Intelligent lubricating oil analysis is a procedure in condition-based maintenance (CBM) of diesel vehicle fleets. Together with diagnostic and failure prediction processes, it composes a data-driven vehicle maintenance structure that helps a fleet manager making decisions on a particular breakdown. The monitoring or control of lubricating oils in diesel engines can be carried out in different ways and following different methodologies. However, the list of studies related to automatic lubricant analysis as methods for determining the degradation rate of automotive diesel engines is short. In this paper we present an intelligent data analysis from 5 different vehicles to evaluate whether the variables collected make it possible to determine the operating condition of lubricants. The results presented show that the selected variables have the potential to determine the operating condition, and that they are highly related with the lubricant condition. We also evaluate the inclusion of new variables engineered from raw data for a better determination of the operating condition. One of such variables is the kinematic viscosity which we show to have a relevant role in characterizing the lubricant condition. Moreover, 3 of the 4 variables that explaining 90% of the variance in the original data resulted from our feature engineering.

Keywords: Condition-Based Maintenance (CBM) · Lubricating Oils · Diesel Vehicle · Intelligent Data Analysis.

1 Introduction

Fleets of vehicles are critical elements for the operations of companies acting in the transportation sector. Thus, identifying vehicles breakdown signatures and building models to predict them before they occur is of major importance and has been an object of study for practitioners involved in the maintenance area [8]. Based on maintenance techniques and methodologies used in failure prediction, modern industries are increasingly using lube oil testing as a crucial factor in maintaining the condition of vehicles. In recent years, efforts have been invested in the development and research of diagnostic and failure prediction systems based on lubricant testing, with the aim of providing early warning of machine failure and to extend the operational life of lubricating oil, avoiding unnecessary oil change costs, waste, or environmental pollution [14] [3]. Following this growing body of research in the maintenance field, with a focus on diesel engine lubricants, we developed a series of analysis on data acquired from five vehicles used in passenger transportation of different brands and models to determine performance patterns of these variables, identify anomalies in vehicle behavior and showing that the selected parameters can be used to develop a failure prediction model for vehicle lubrication systems. The paper is subdivided into four sections. Section two is dedicated to a literature survey, in the third section there is an evaluation and modeling of the lubricant data from the vehicles and finally the fourth section contains results and discussion.

2 Literature Review

Condition-Based Maintenance (CBM) is a strategy that monitors the current condition of an asset, helping to decide when a given maintenance operation is required, considering the evolution pattern of certain indicators. Checking these indicators on a machine can include non-intrusive measurements for example: visual inspection, analysis at performance data of the vehicles, and scheduled tests like temperature and vibration analysis [13]. Condition data can be collected at certain intervals or continuously (embedded internal sensors). Some studies present maintenance decision guidelines for improving CBM using detailed analysis of used oil data as a source of information to determine whether the fluid in the system is healthy and suitable for further service, in addition to observing the condition of the equipment through this analysis. Case studies present models for monitoring the condition of diesel engine oil in urban buses, which, by monitoring the evolution of its degradation, aim to implement a predictive maintenance policy or a CBM [8]. Based on lubrication condition monitoring (LCM), the intervals for oil replacement can be increased, which directly implies an increase in availability and a reduction in the costs associated with maintenance [8]. [14] presents a detailed approach to recent research trends in the development of LCM-based approaches applied for maintenance decision support and applications in equipment diagnosis and prognosis. This study reviews and classifies LCM tests and parameters into several categories, which include physicochemical, elemental, contamination, and additive analyses. These studies show that building data-driven models accurately identify the state of wear and tear in a component or system, and make the condition-based maintenance methodology more effective and reliable. Furthermore the study of lubricants can determine the rate of equipment degradation, and the results of lubricant analysis to determine the behavior of wear levels as a function of operating time and operating typology are important tools for developing a predictive maintenance methodology [8]. [6] presents the physicochemical tests in the analysis of oils that are used for the evaluation of their condition and reports that the interpretation of waste oil analysis is very complex because the individual analyses are interdependent. For this reason it is necessary to know the whole oil analysis and make conclusions based on results of individual analyses. According to [3], the analysis of oil aging or degradation have continually evolved to provide a more accurate diagnosis of the remaining life of lubricants or any latent or impending failure in the equipment or processes in which they are being used. To analyze lubricating oils, it is necessary to collect samples during and after operation, and to take into consideration the parameters that will determine the quality of the lubricant and the condition of the equipment [11]. According to [5], the main parameters related to oil degradation are: oxidation, nitration, viscosity, depletion of antioxidant additives, anti-wear, TAN, TBN and RUL. Another important parameter besides those mentioned above is the so-called Oil Stress Factor, which has been used as an estimator of the potential stress suffered by the oil and allows some correlation with its possible degradation, considering the relationship between the specific power per displacement, the oil change period and the oil sump volume [5]. Another important variable is the oil viscosity considering perfect operating conditions which represents the fluidity of the oil at given temperature. In this sense the more viscous oils are thicker, while the less viscous ones are more fluid. [12] explains that the influence of operating conditions on the wear of both the equipment and the lubricants occurs because, due to many short trips and including a cold start, automotive engines often run without reaching their nominal temperature. The relationship between oil temperature and viscosity is known to have a strong effect on engine friction. The Oil stress factor (OSF) is a method for quantifying the stress under which engine oil is placed and allows for the prediction of oil degradation as a function of engine conditions. For this study, 2 different ways will be used to determine OSF to verify the relationship between them and determine the most appropriate for the problem at hand: Oil_z: described in [7] and OSFv3 described in [2] According to [1], in the maintenance of internal combustion engines, different lubricant analysis techniques are used to determine the condition of the oil and the system. A rapid method of analysing the condition of the lubricant is the so-called oil slick test, which allow to estimate the degradation of the lubricant and the presence of contaminants and it was used in this study to provide the prediction model labels. Figure 1 presents an example of an oil slick test. The central zone of the spot is identified, characterized by its dark and uniform intensity, and the intermediate or diffusion zone of the slick that indicates the degree of dispersion of the carbon particles and presence of contaminants. With these parameters defined, information about the dispersion and detergency conditions of a lubricant can be obtained by making an oil stain on filter paper. After applying the oil drop to the test paper, it is possible to see an image with up to 4 circles based on the chromatographic effect. Although the slick test provides results on a scale of 0 to 9 for solid particle contamination and for condition. We convert this scale to a binary problem: 0 indicates that the oil is good for operating conditions, and 1 indicates that the oil is not good for operating conditions.

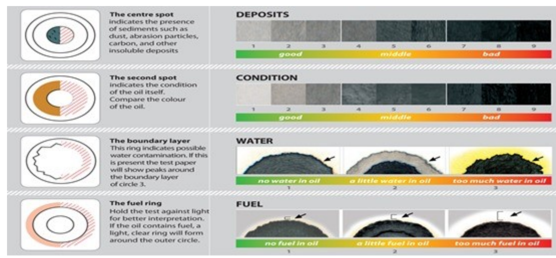


Fig. 1: Example of an oil slick text

3 Data Analysis

The five vehicles considered in this study have a set of sensors which combined with an acquisition hardware, developed by Stratio Automotive [9], were used to collect the required data. The collected variables are: Engine Oil Temperature (°C), Coolant Temperature (°C), Engine Oil Pressure (mbar), Vehicle speed (km/h) and Engine speed (rpm). Missing value analysis helps to address several concerns caused by incomplete data, since they can lead to misleading interpretations. A descriptive analysis was initially performed on the collected variables to check the percentage of missing values and we observed a negligible percentage (inferior to 5%).

A visual analysis of the variables behaviour allowed to determine that the behavior of the variables was similar among the selected cars despite being of different brands, models and years of manufacture (see Fig. 2). From the analysis

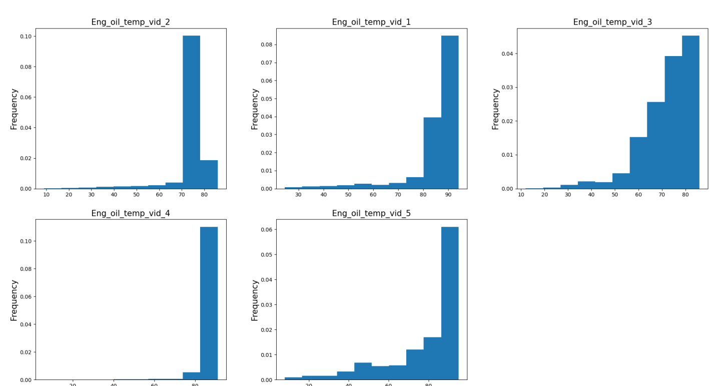
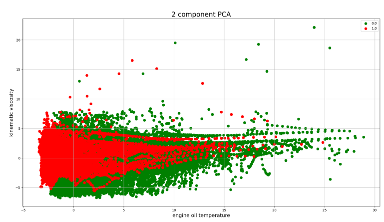


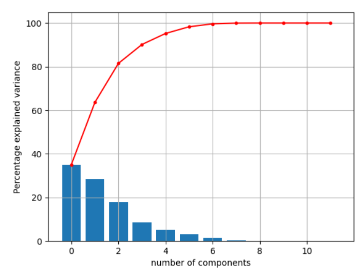
Fig. 2: A histogram grouped into frequency of values

of the histograms it can be seen that the data is asymmetrical. The asymmetry indicates that the data may not be normally distributed, but indicates the range of values at which these parameters work most of the time and confirms the limits and optimal operating values. To Confirm confirm whether there was a strong or weak relationship between the distribution of values over the months and the vehicles, we applied use the Kolmogorov-Smirnov test (K-S test). This test confirmed that there is no plausible difference in the data distribution to creating a failures alerts.

Principal Component Analysis (PCA) is a mathematical procedure for converting a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components[10]. The Fig. 3 show the PCA analysis from the one vehicle data and determine the lubricant's operating condition(0) or not(1) based on the two most important variables identified (Engine oil temperature x kinematic viscosity). From the analysis of



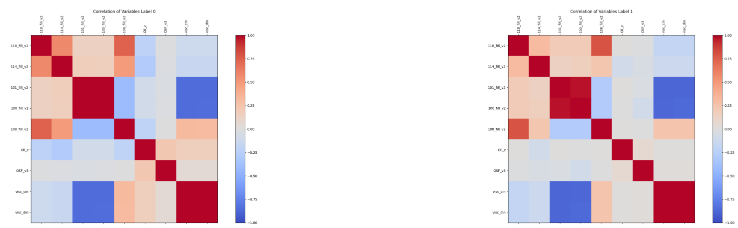
(a) PCA analysis for two principal components



(b) the cumulative explained variance ratio as a function of the number of components

Fig. 3: PCA Analysis

the Fig. 3 (a), we can identifier clear zones of differentiation between labels 0 and 1, it is possible to visualize in the zone between -4 and -2 of the x-axis the non-operational condition of the lubricant. However, attention should be paid to overlapping zones when developing future models. It is worth noting that the PCA analysis was performed for all vehicles and the results were cross-sectional, with no plausible distinction between models and brands that would justify changing the variables. Another important point that is drawn from the PCA analysis is the cumulative variance of the parameters, which determines the loss of information with the simplification of variables. According to Fig. 3 (b), the blue bars show the percent variance explained by each principal component and the red line shows the cumulative sum. From the graph, we can read the percentage of the variance in the data explained as we add the principal components. So the first principal component explains 35% of the variance in the data set, the first 2 principal components explain 63%, and so on. This chart shows that we need only 4 (Engine oil temperature, kinematic viscosity, dynamic viscosity and Oil_z) of the 9 principal components to explain 90% of the variance in the original data. A correlation analysis was performed to better understand the relationships among variables. In statistics, dependence or association is any statistical relationship, causal or otherwise, between two random variables or bivariate data (see Fig. 4).



(a) Correlation for the lubricant operation condition (Label 0)

(b) Correlation for the non-operational condition of the lubricant (Label 1)

Fig. 4: Variables Correlation

In this way we can see, when analyzing the correlation graphs of the variables, several types of relationships, whether they indicate a positive operating condition or a negative operating condition: 1. Oil_z in relation to all the other variables, has correlation values for label 0, something that is almost not identified for Label 1. This information determines that the calculated variable oil_z is an important and decisive variable for determining the oil's non-operating condition, since it is related to all the variables selected for solving the problem. 2. Oil pressure in relation to oil and liquid temperature, decreases its correlation value when it transitions from label 0 to label 1. This information indicates that the control for the relation of the amount of oil necessary for complete lubrication of the engine, according to its operating temperature, is no longer fully functioning and may not be sufficient. 3. Oil pressure versus engine speed increases its correlation value when it transitions from label 0 to label 1. This information determines that the engine needs to rotate faster to try to meet the lubrication demand. 4. The viscosities in relation to the engine speed decrease their correlation values when transiting from label 0 to label 1. This information determines that the film required for optimal engine lubrication at that speed is no longer being achieved. 5. Oil pressure versus viscosities increase their correlation values when transitioning from label 0 to label 1. This information determines the need for greater lubricant pressure to meet lubrication needs, since oil degradation causes it to lose its properties, viscosity being one of the most important oil properties.

We are also interested in looking into which variables may or may not be relevant as input to develop a model(see Fig. 5). Thus, we can extract from this correlation graph that the variables oil temperature and liquid temperature have a high correlation value with the target, which is identified because the operating temperatures are directly linked to the calculations made by the Centralina for the engine's lubrication needs, and the higher the engine's operating temperature, the lower the kinematic viscosity and the worse the oil's lubrication power.

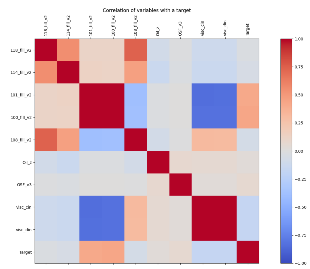


Fig. 5: Correlation between the input variables with the output variable.

Another important point to mention is that due to the high temperature relationship, there is consequently a correlation value of the kinematic and dynamic viscosity with the target, even if to a lesser extent. These viscosity correlation values are due to the lubricant properties, and viscosity control is one of the most important factors in determining the lubricant's operating condition.

4 Conclusion

An important component of engine maintenance is to establish the most cost-effective oil change and overhaul intervals in terms of cost, wear, and failure diagnosis efficiency. From the present study, some conclusions can be drawn and could be subdivided into: **Response time:** it is undeniable that lubricant analysis is of utmost importance to determine the wear of the equipment and present better thresholds for equipment replacement. Real-time data analysis techniques are capable of high-throughput analysis with typical analysis times of less than two minutes compared to laboratory analyses that face both large sample volumes and the constraint of fast turnaround times; **Evaluation parameters:** It is common knowledge in the analyses performed, the influence of temperature on the viscosity of lubricating oils, however as already mentioned in other parts of this object of study, this despite being an important parameter is not the only one that should be analyzed, [4] concludes in his study that to obtain a complete picture of the changes occurring in the oil, it is advisable to interpret the kinematic viscosity values together with the dynamic viscosity, as well as the degree of oxidation and acidity; **Analysis performed:** Through the PCA analysis, it was confirmed that the selected variables can determine the answer to the stated problem, but it is necessary to have at least 4 variables so that the variance of the data is not lost. The kinematic viscosity variable proved to be preponderant for characterizing the lubricant condition or not, and 3 of the 4 variables that explaining 90% of the variance in the original data were new calculated variables, proving that necessary calculate other variables based on the original data to the improvement of the analysis. In addition, the correlation analysis identified relationships that confirm the theory described in this study and consequently pave the way for the development of a model for identifying the operating conditions of lubricants in diesel vehicles.

Acknowledgments

This work is partially funded by national funds through the FCT - Foundation for Science and Technology, I.P., within the scope of the project CISUC - UID/CEC/00326/2020 and by European Social Fund, through the Regional Operational Program Centro 2020 and by national funds through FCT - Fundação para a Ciência e a Tecnologia, under the project UIDB/00285/2020.

References

1. Delgado Ortiz, J., Saldivia Saldivia, F., Fygueroa Salgado, S.: Sistema para la determinación de la degradación del lubricante basado en el tratamiento digital de manchas de aceite de motores diesel. Revista UIS Ingenierías **13**(1), 55–61 (2014)
2. Lee, P.M., Stark, M.S., Wilkinson, J.J., Priest, M., Smith, J.R., Taylor, R.I., Chung, S.: The degradation of lubricants in gasoline engines: Development of a test procedure to evaluate engine oil degradation and its consequences for rheology. Tribology and Interface Engineering Series **48**, 593–602 (2005)
3. Lopez, P., Mabe, J., Miró, G., Etxeberria, L.: Low cost photonic sensor for in-line oil quality monitoring: Methodological development process towards uncertainty mitigation. Sensors (Switzerland) **18**(7) (2018)
4. Loska, A.: The kinetics of changes in kinematic viscosity of engine oils under similar operating conditions. Eksploatacja i Niezawodnosc **19**(2), 268–278 (2017)
5. Macián-Martínez, V., Tormos-Martínez, B., Gómez-Estrada, Y.A., Bermúdez-Tamarit, V.: Revisión del proceso de la degradación en los aceites lubricantes en motores de gas natural comprimido y diesel. Dyna (Spain) **88**(1), 49–58 (2013)
6. Perić, S., Nedić, B., Grkić, A.: Applicative monitoring of vehicles engine oil. Tribology in Industry **36**(3), 308–315 (2014)
7. Prasad, M.V., Lakshminarayanan, P.A.: Estimation of Oil Drain Life of Engine Oils in New Generation Diesel Engines in Tropical Conditions. SAE International Journal of Fuels and Lubricants **5**(2) (2012). https://doi.org/10.4271/2011-01-2405
8. Raposo, H., Farinha, J.T., Fonseca, I., Ferreira, L.A.: Condition monitoring with prediction based on diesel engine oil analysis: A case study for urban buses. Actuators **8**(1), 1–15 (2019)
9. S.A., S.: Homepage - stratio. `https://stratioautomotive.com/`, (Accessed on 03/19/2021)
10. Shahariar, G., Bodisco, T., Chu Van, T., Surawski, N., Sajjad, M., Suara, K.A., Ristovski, Z., Brown, R.: Optimisation of driving-parameters and emissions of a diesel-vehicle using principal component analysis (pca). In: Proceedings of the 22nd Australasian Fluid Mechanics Conference AFMC2020. The University of Queensland (2020)
11. Sharma, B.C., Gandhi, O.P.: Performance evaluation and analysis of lubricating oil using parameter profile approach. Industrial Lubrication and Tribology **60**(3), 131–137 (2008)
12. Tauzia, X., Maiboom, A., Karaky, H., Chesse, P.: Experimental analysis of the influence of coolant and oil temperature on combustion and emissions in an automotive diesel engine. International Journal of Engine Research **20**(2), 247–260 (2019)
13. Vališ, D., Žák, L., Pokora, O.: Failure prediction of diesel engine based on occurrence of selected wear particles in oil. Engineering Failure Analysis **56**, 501–511 (2015)

14. Wakiru, J.M., Pintelon, L., Muchiri, P.N., Chemweno, P.K.: A review on lubricant condition monitoring information analysis for maintenance decision support. Mechanical Systems and Signal Processing **118**, 108–132 (2019)