A Machine-Learning Approach for the Assessment of Quantitative Changes in the Tractor Diesel-Engine Oil During Exploitation

Radočaj, Dorijan; Plaščak, Ivan; Jurišić, Mladen

Source / Izvornik: **Poljoprivreda, 2024, 30, 108 - 114**

Journal article, Published version
Rad u časopisu, Objavljena verzija rada (izdavačev PDF)

https://doi.org/10.18047/poljo.30.1.14

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:151:030649

A Machine-Learning Approach for the Assessment of Quantitative Changes in the Tractor Diesel-Engine Oil During Exploitation

Pristup strojnoga učenja za procjenu kvantitativnih promjena ulja u dizelskome motoru traktora tijekom eksploatacije

Radočaj, D., Plaščak, I., Jurišić, M.

Poljoprivreda / Agriculture

ISSN: 1848-8080 (Online)
ISSN: 1330-7142 (Print)

https://doi.org/10.18047/poljo.30.1.14

Fakultet agrobiotehničkih znanosti Osijek, Poljoprivredni institut Osijek

Faculty of Agrobiotechnical Sciences Osijek, Agricultural Institute Osijek

ISSN 1330–7142
UDK=631.372:004.453
https://doi.org/10.18047/poljo.30.1.14

A MACHINE-LEARNING APPROACH FOR THE ASSESSMENT OF QUANTITATIVE CHANGES IN THE TRACTOR DIESEL-ENGINE OIL DURING EXPLOITATION

Radočaj, D., Plaščak, I., Jurišić, M.

Original scientific paper
Izvorni znanstveni članak

SUMMARY

To evaluate the potential of a machine-learning approach in the assessment of quantitative changes in the tractor diesel-engine oil during exploitation, this study aspired to propose a machine-learning regression method to reduce the frequency of expensive and time–consuming engine oil sampling. The input engine-oil datasets with fresh engine oil (Samples A) and with the engine oil subsequent to 250 working hours (Samples B) were sampled for twelve elements in a two–year exploitation study at the Belje company. The field data collection was performed having deployed six heavy, four–wheel-drive FENDT 930 Vario agricultural tractors, each monitored for 1,500 working hours, during which an engine oil was sampled every 250 working hours. The evaluated machine-learning prediction methods, based on a tenfold cross–validation, achieved a moderately high prediction accuracy, with a slightly higher coefficient of determination (R^2), in the range of 0.51–0.73, for the Samples B, than those in the range of 0.49–0.64 for the Samples A. These results strongly suggest that none of the machine learning methods constantly achieved high prediction accuracy and that the selection of optimal machine-learning models should be mandatory, having also confirmed a high potential of machine-learning methods in the detection of quantitative changes in tractor diesel-engine oil during exploitation.

Keywords: agricultural tractor, pollution, variable importance, optimization, exploitation

INTRODUCTION

Because the combustion process in diesel engines is a source of pollution, it is extremely important to study the combustion of diesel fuel, especially those parts of the combustion related to the processes that produce the pollutants themselves (Kozina et al., 2020). There are several reasons why the chemical processes of soot formation have been an area of active research within the combustion research community for a long time (Carbone et al., 2023). Soot emissions indicate poor combustion and reduced efficiency in the operation of internal combustion engines, and soot deposition can have extremely detrimental effects on the maintenance and service life of such equipment. However, what has really increased the number of studies in the world related to the formation of soot in the last few decades are the harmful effects of soot on the health and the harmful effects of soot on the environment (Carbone et al., 2023). To prevent the negative effects on the environment and to improve the utilization potential, agricultural tractors should be subject to regular maintenance (Dhananjayan et al., 2020). Nonetheless, since this often implies an expensive and time–consuming process, there is a potential in using novel machine-learning methods to optimize this process and streamline the monitoring of tractor use.

As engine-oil quality is dynamic and affected by multiple elements such as wear debris, contaminants, and operating conditions, a thorough and flexible methodology is required for a rapid evaluation. With the ability to identify the intricate patterns in large datasets, machine-learning algorithms provide a reliable way to model and predict how engine-oil properties change

Dorijan Radočaj, Ph. D. (dradocaj@fazos.hr), Assoc. Prof. Ivan Plaščak, Prof. Dr. Mladen Jurišić – Josip Juraj Strossmayer University of Osijek, Faculty of Agrobiotechnical Sciences Osijek, Vladimira Preloga 1, 31000 Osijek, Croatia

relative to a tractor exploitation load (Song et al., 2023). These algorithms make it possible to detect the changes that could be the signs of contamination, engine wear, or oil-quality degradation. This predictive capability makes it easier to identify the problems early, enabling preventive maintenance to avert the engine damage and to reduce the environmental risks associated with deteriorating oil spills. Moreover, it is possible to customize the machine-learning models to account for a variety of operating conditions, which can help improve the accuracy and adaptability of assessment concerning the quantitative changes in tractor diesel-engine oil (Beni et al., 2023). Incorporating the machine-learning techniques into the engine-oil condition monitoring has a potential to significantly improve the environmental sustainability of the agricultural industry while reducing the negative impact of the used oil on the ecosystems and promoting an effective resource management.

This study's objective was to propose a machine-learning regression method to determine the quantitative changes in engine oil during exploitation in a diesel engine of an individual tractor by considering the interaction of combustion quality and tractor diesel-engine oil parameters. This approach aspired to reduce the frequency of expensive and time–consuming engine-oil sampling by evaluating the interaction of combustion quality and tractor diesel-engine oil parameters.

MATERIAL AND METHODS

A two–year field exploitation study at the Belje company was performed having deployed six heavy, four–wheel-drive *FENDT 930 Vario* agricultural tractors. The tractors were equipped with MAN's 6871 cm^3 liquid–cooled, four–stroke, six–cylinder diesel engines (DO836LE510 type) having a direct injection and rated power of 210 kW at 2250 revolutions per minute. During the field study, the tractors were exposed to conventional exploitation loads for this tractor class on five Belje farming centers. They were primarily used in basic tillage (winter plowing), autumnal and vernal soil sowing preparation, sowing of hibernal and vernal crops, and transport operations. Each of the six studied tractors was monitored for 1,500 working hours (i.e., for 9,000 working hours in total), during which engine oil was sampled every 250 working hours. Oil analysis was performed according to the ASTM D5185 method (Bol'shakov et al., 2021), which quantified the sample concentrations of iron (Fe), chromium (Cr), copper (Cu), aluminum (Al), nickel (Ni), titanium (Ti), molybdenum (Mo), magnesium (Mg), zinc (Zn), barium (Ba), boron (B), and silicon (Si). These samples contained the metallic particles formed by wear, particles of oil impurities, and the particles that indicate a degree of oil-additive degradation. The fresh engine oil sampled prior to its incremental replacement was labelled Samples A, while the sampled oil subsequent to 250 working hours was labelled Samples B. Additionally, a combustion quality measurement was performed while determining the blackening coefficient and wear particle concentration (WPC). The machine-learning regression method proposed in the study evaluated the metals from the Samples A and B as dependent variables, predicted according to three covariates: tractor working hours, WPC, and the blackening coefficient.

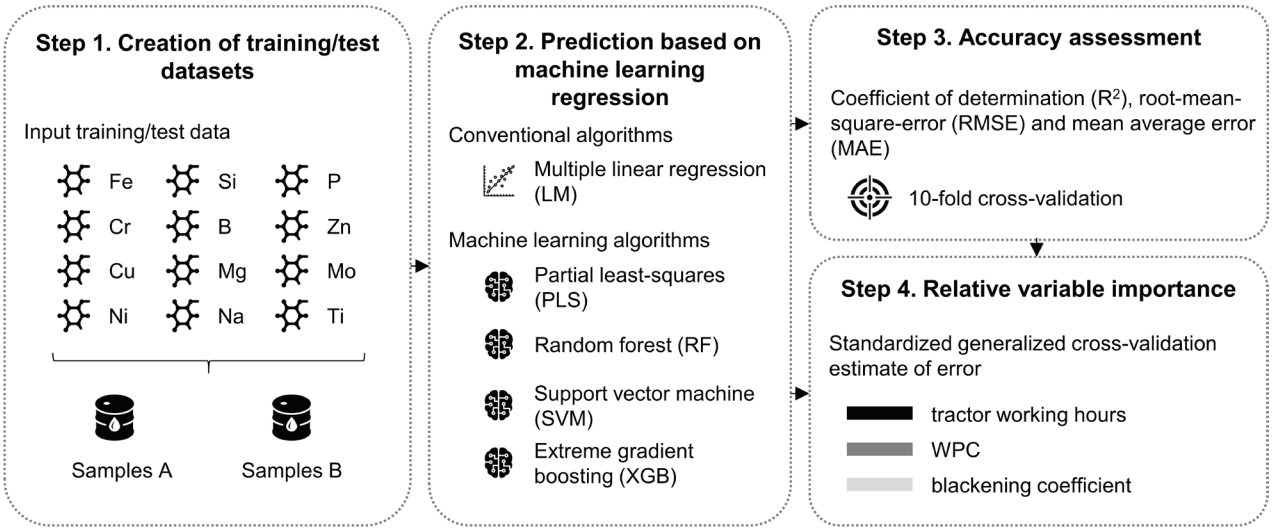

Figure 1. The workflow of the proposed prediction approach, based on a machine-learning regression.

Slika 1. Tijek rada predloženoga pristupa predikcije korištenjem regresije temeljene na strojnome učenju.

Based on the machine-learning regression, the workflow of the proposed prediction approach consists of four fundamental steps (Fig. 1): 1) the creation of training/test datasets from the engine-oil samples classified into Samples A and Samples B datasets; 2) the evaluation of one conventional and four machine-learning regression algorithms for the prediction of input elements; 3) accuracy assessment based on cross–validation; and 4) a relative variable importance calculation according to the three selected prediction covariates .

Five machine-learning regression algorithms were evaluated for the prediction, including the multiple linear regression (LM), partial least–squares (PLS), random forest (RF), support vector machine (SVM), and the extreme gradient boosting (XGB). While the LM is a conventional regression approach, the PLS, RF, SVM, and the XGB machine-learning algorithms achieved a high accuracy and robust prediction in previous studies (Radočaj et al., 2023; Varga et al., 2023), representing the state–of–the–art solutions for machine-learning regression. A library "caret" in R v4.2.2 was used for the machine-learning regression, accuracy assessment, and variable importance calculation.

The LM applied a linear function to represent a connection between the input variables and the output parameters. The approach minimized the sum of squared residuals between the anticipated and the actual values to estimate the coefficients (slopes) of linear equation. This was accomplished while using the ordinary least squares, in which the objective is to identify the coefficients that minimize the sum of squared discrepancies between the anticipated and the actual values. The PLS regression solved the LM drawbacks in prediction with multicollinearity and large–scale datasets, adopting a combination of feature extraction and linear regression. It pinpointed the latent variables, which account for the most variance in both the input and the output variables, while the original variables were combined linearly to create these components. Applying covariance with the input variables, the PLS calculated the direction (component) that has a maximum covariance with the output variable, decreasing the dimensionality of the data and minimizing the collinearity problems while projecting the input variables onto these elements. The RF, a decision-tree–based, ensemble learning algorithm, built several decision trees and combined their predictions. At each split, a part of the input characteristics was retained by each decision tree, which was trained on a data subset and randomly chosen by virtue of replacement—that is, by virtue of bootstrap sampling (Cutler et al., 2012). Decision trees divided the data during training depending on the input criteria and feature subsets as long as the specific termination requirements were satisfied, including the maximum depth and the minimum number of samples per tree. The final regression result was calculated during prediction by averaging each individual tree prediction. Additionally, the RF could handle the high–dimensional datasets, detect feature importance, and handle the missing values and outliers effectively by aggregating the predictions from multiple trees, which also reduced overfitting and provided a reliable estimation of the target variable (Cutler et al., 2012). The SVM regression calculated the hyperplane that optimized a distance between the support vectors and the regression line in this converted feature space, signifying that the data points with the greatest distance between the regression line and the samples were separated by the hyperplane (Awad & Khanna, 2015). The linear kernel function was used with the SVM for the conversion of the input variables into a higher–dimensional feature space. The algorithm's capacity to recognize the complicated patterns depended on the kernel function selected, which also affected how flexible the decision boundary was. This allowed the SVM to solve a quadratic optimization problem, which can be time–demanding to compute for big datasets. The XGB built up the weak models incrementally while applying a boosting approach to produce a strong ensemble model by repeatedly merging the predictions of the weak models (Ayyadevara, 2018). Having fitted an initial decision-tree model to the input data, it improved further models by computing the residuals of the earlier models on an iterative basis. By changing the weights of the weak models, the XGB employed a gradient–based optimization to minimize a mean squared error loss function, eventually lowering the total error. To avoid overfitting and improve generalization, the XGB was combined with an L2 regularization approach. The relative variable importance of the most accurate machine-learning algorithm per element was calculated using the "caret" package, representing a standardized, generalized error cross–validation estimate.

The accuracy assessment of the five machine-learning regression algorithms was performed using a tenfold cross-validation in the "caret" library. The coefficient of determination (R^2), root–mean–square–error (RMSE), and the mean average error (MAE) were selected to quantify the prediction accuracy. The combination of these metrics provided a comprehensive accuracy evaluation of machine-learning regression algorithms, in which the R^2 quantified the proportion of variability explained by the model, while the RMSE and MAE quantified the average prediction error (Radočaj et al., 2023). The R^2 was calculated using the following formula (1):

$$R^2 = 1 - \frac{SSR}{SST}, \quad (1)$$

where the *SSR* were residuals, or the sum of the squared differences between the predicted values and the mean of input variable, whereas the *SST* was a total sum of the squared differences between the actual values and the mean of the input variable. The RMSE was calculated as the square root of the mean of the squared differences between the predicted values and the actual values according to the formula below (2):

$$RMSE = \sqrt{\frac{\sum_1^n (y_{predicted} - y_{input})^2}{n}}, \quad (2)$$

where n was the number of samples, $y_{predicted}$ were the predicted values, and y_{input} were the input values. The RMSE was sensitive to large errors due to the squaring operation, and it was measured in the same units as the input variable. Unlike the RMSE, the MAE did not square the errors, having resulted in a metric that was less sensitive to large errors. The MAE measured the average absolute difference between the predicted values and the actual values according to the following formula (3):

$$MAE = \frac{\sum_1^n \left| y_{predicted} - y_{input} \right|^2}{n}. \quad (3)$$

RESULTS AND DISCUSSION

Figure 1 represents the boxplot of the input datasets containing the concentrations of elements from the Samples A and B. The used engine oil from Samples B consistently exhibited higher concentrations of Fe, Cr, Cu, Ni, Zn, Mo, and Ti subsequent to 250 working hours while also having a larger range or variability in these concentrations. On the opposite, the fresh engine oil from the Samples A displayed higher concentrations of Si and B when compared to the Samples B. Mg, Ba, and P had the similar median values across the sampling datasets, while Mg and P resulted in a relatively large value range in the Samples B. Generally, these values strongly indicate that the quantitative changes in engine-oil parameters were represented in the Samples A and B, which is an observation similar to that of Dörr and colleagues (2019).

The accuracy assessment for the Samples A indicated a preference toward a decision-tree–based RF for Si, Ba, P, and Mo (Table 1), but its comparatively lower R^2 values for certain elements may indicate the limitations in capturing subtle variations in concentration. The challenges encountered in predicting the Si concentrations underscore the inherent complexity of these elements and may necessitate a specialized feature engineering or additional covariates for an improved model performance. The XGB was, most frequently, the optimal model for the prediction of elements in the Samples B (Table 2). Of particular note is its performance in the prediction of concentrations of Cr, Cu, and Ni, where it achieved the highest R^2 values, indicating the robust predictive capabilities of these elements. The RF also excelled in accurately predicting B, Mg, and Zn, demonstrating precision in capturing their concentrations with the lower RMSE and MAE values when compared to the other models. Despite a generally high prediction accuracy, challenges arose while predicting the concentrations of elements such as Ba, P, and Ti, where all models had the comparatively lower R^2 values, highlighting the areas for potential improvement, or for alternative modeling approaches.

In general, the evaluated machine-learning methods achieved a moderately high accuracy, with a slightly higher R^2 in the range of 0.51–0.73 for the Samples B and a slightly lower range of 0.49–0.64 for the Samples A. These results clearly suggest that none of the machine-learning methods constantly achieved a high prediction accuracy and that the selection of optimal prediction results based on the evaluation of several models should be mandatory, in conformity with Sharma and colleagues (2021). Moreover, this study's results confirm the assumptions that the machine-learning regression performance strongly depends on the properties of the input data, as well as on the quantity of input samples and the types of input covariates (Cersonsky et al., 2021).

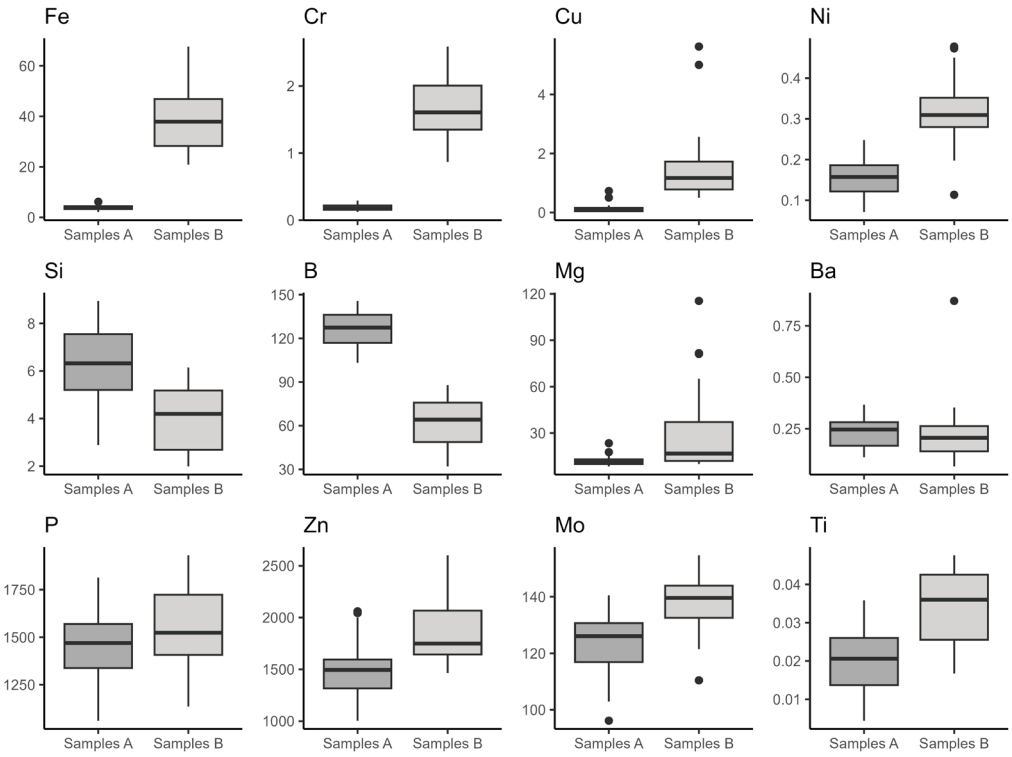

Figure 2. The boxplots of the input engine-oil datasets from the Samples A (engine oil after 250 working hours) and B (fresh engine oil), with the concentrations expressed in ppm.

Grafikon 2. Dijagrami ulaznih skupova podataka o motornome ulju iz uzoraka A (motorno ulje nakon 250 radnih sati) i B (svježe motorno ulje), s koncentracijama izraženima u ppm.

Table 1. Machine-learning prediction accuracy of engine-oil elements for the Samples A (fresh engine oil).

Tablica 1. Točnost predikcije korištenjem strojnoga učenja za elemente u motornome ulju za uzorke A (svježe motorno ulje).

Metrics / *Metrike*	Models / *Modeli*	Elements / *Elementi*											
		Fe	Cr	Cu	Ni	Si	B	Mg	Ba	P	Zn	Mo	Ti
R^2	LM	0.46	0.56	0.30	0.50	0.38	0.37	0.55	0.40	0.47	0.57	0.44	0.48
	PLS	0.52	0.28	0.56	0.47	0.47	0.46	0.67	0.39	0.47	0.42	0.38	0.48
	RF	0.57	0.38	0.24	0.44	0.49	0.58	0.33	0.64	0.61	0.59	0.52	0.51
	SVM	0.26	0.42	0.44	0.44	0.45	0.59	0.44	0.42	0.37	0.62	0.40	0.59
	XGB	0.59	0.57	0.43	0.56	0.47	0.55	0.61	0.49	0.37	0.47	0.47	0.55
RMSE (ppm)	LM	0.92	0.04	0.12	0.04	1.82	10.4	2.74	0.07	193.4	259.2	10.3	0.01
	PLS	0.89	0.04	0.10	0.04	1.82	11.4	2.44	0.07	177.3	275.5	11.1	0.01
	RF	0.85	0.04	0.12	0.04	1.87	10.8	2.86	0.06	223.7	298.4	11.3	0.01
	SVM	0.88	0.05	0.11	0.04	1.90	10.6	2.65	0.08	193.1	234.7	10.6	0.01
	XGB	1.10	0.05	0.20	0.05	2.29	12.1	3.33	0.07	265.5	306.8	14.2	0.01
MAE (ppm)	LM	0.76	0.04	0.09	0.03	1.55	8.87	2.26	0.06	161.5	213.7	8.87	0.01
	PLS	0.78	0.04	0.08	0.04	1.61	10.3	2.10	0.06	152.8	241.7	9.65	0.01
	RF	0.72	0.04	0.10	0.03	1.66	8.82	2.36	0.06	192.9	243.4	9.52	0.01
	SVM	0.74	0.05	0.08	0.04	1.66	9.22	2.26	0.07	159.6	198.6	9.73	0.01
	XGB	0.93	0.04	0.14	0.04	2.07	10.0	2.50	0.07	228.7	260.9	12.2	0.01

The most accurate R^2 values per an analyzed element are bolded.

Table 2. Machine-learning prediction accuracy of engine-oil elements for the Samples B (engine oil after 250 working hours).

Tablica 2. Točnost predikcije korištenjem strojnoga učenja za elemente u motornome ulju za uzorke B (motorno ulje nakon 250 radnih sati).

Metrics / *Metrike*	Models / *Modeli*	Elements / *Elementi*											
		Fe	Cr	Cu	Ni	Si	B	Mg	Ba	P	Zn	Mo	Ti
R^2	LM	0.57	0.33	0.31	0.51	0.62	0.57	0.58	0.48	0.49	0.31	0.47	0.41
	PLS	0.65	0.35	0.54	0.58	0.53	0.53	0.58	0.45	0.41	0.30	0.68	0.49
	RF	0.50	0.56	0.46	0.47	0.43	0.70	0.72	0.35	0.46	0.57	0.34	0.31
	SVM	0.70	0.51	0.37	0.60	0.58	0.39	0.63	0.53	0.49	0.40	0.57	0.49
	XGB	0.69	0.62	0.59	0.61	0.46	0.37	0.50	0.48	0.51	0.51	0.37	0.56
RMSE (ppm)	LM	10.2	0.45	0.95	0.08	1.06	14.8	22.1	0.10	208.0	313.7	9.66	0.01
	PLS	10.5	0.44	0.90	0.07	1.41	15.8	21.7	0.10	212.1	318.9	10.6	0.01
	RF	10.9	0.52	1.08	0.09	1.21	15.7	21.2	0.11	240.8	301.3	9.54	0.01
	SVM	9.68	0.45	0.84	0.07	1.06	14.8	20.0	0.10	202.1	322.2	9.55	0.01
	XGB	11.2	0.64	1.61	0.11	1.52	19.9	32.2	0.10	270.9	472.6	13.1	0.01
MAE (ppm)	LM	8.75	0.41	0.72	0.06	0.96	12.7	17.0	0.08	179.4	276.1	8.21	0.01
	PLS	9.19	0.39	0.71	0.06	1.20	14.0	17.6	0.08	183.2	282.6	8.72	0.01
	RF	8.72	0.47	0.81	0.07	1.08	13.1	16.0	0.09	213.2	265.3	8.00	0.01
	SVM	7.68	0.39	0.64	0.06	0.98	12.1	12.9	0.08	169.7	264.0	8.00	0.01
	XGB	8.66	0.58	1.23	0.09	1.28	17.0	21.2	0.08	224.0	395.3	11.5	0.01

The most accurate R^2 values per an analyzed element are bolded.

The tractor working hours were the most impactful covariate based on the relative variable importances for both the Samples A and the Samples B, having demonstrated a dominant importance for Fe, Ni, and Mo for both datasets (Fig. 3). This underscores a temporal influence on these elements, implying that the extended operational durations significantly impact their concentrations. The WPC also strongly impacted the prediction of Si, Ba, P, and Zn, while the blackening coefficients had a generally minor importance except for the Cr prediction, where they were a dominant predictor in comparison to the tractor working hours and the WPC. To aid the present solutions for the reduction of gaseous emissions in the internal combustion engines (Kozina et al., 2020), the approach proposed in this study supports the inclusion of additional covariates, with a high potential of further improvement of prediction accuracy.

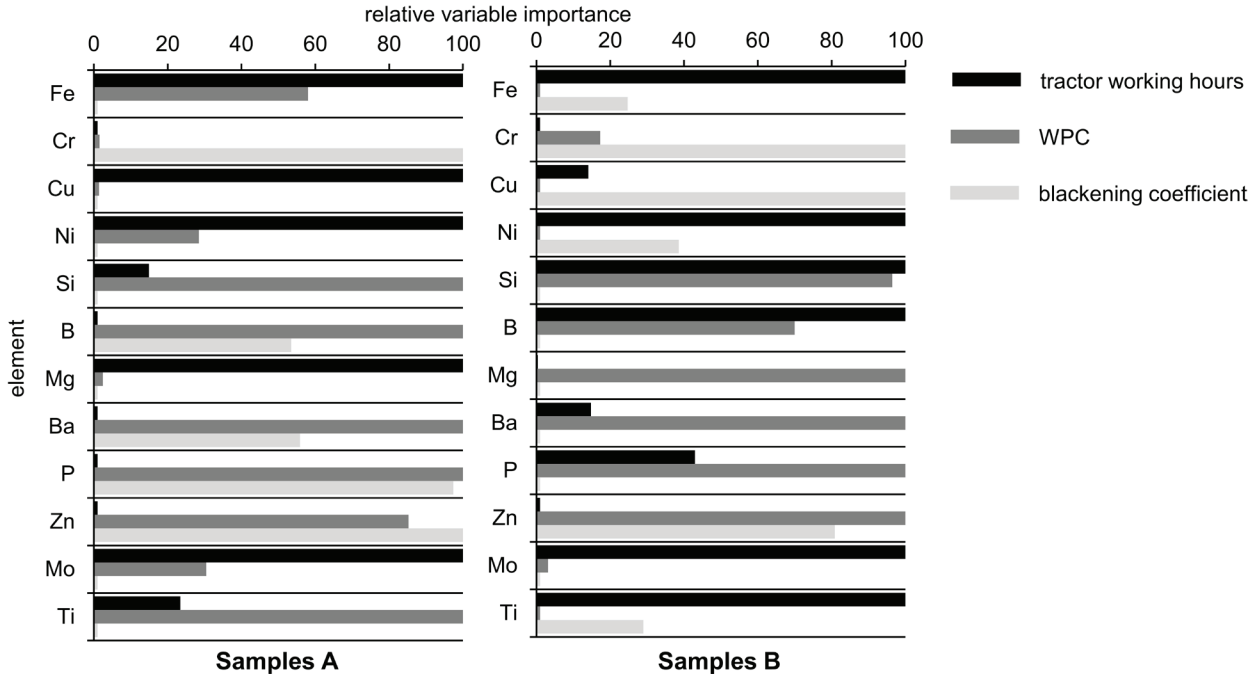

Figure 3. Relative variable importance for the most accurate machine-learning prediction model per element.
Grafikon 3. Relativna važnost varijabli za najtočniju predikciju modela strojnoga učenja po elementu.

CONCLUSION

This study's results clearly suggest that none of the machine-learning method constantly achieved a high prediction accuracy, and that the selection of optimal machine-learning models should be mandatory, having also confirmed a high potential of machine-learning methods in the detection of quantitative changes in tractor diesel-engine oil during exploitation. According to the study results, this approach established a solid foundation to reduce the frequency of expensive and time–consuming engine-oil sampling by evaluating an interaction between combustion quality and tractor diesel-engine oil parameters. The study's results confirm the assumptions that the performance of machine-learning regression strongly depends on the input-data properties, as well as on the quantity of input samples and the types of input covariates. Therefore, further studies of heterogeneous engine-oil datasets on multiple study locations are necessary to prove the robustness of the proposed machine-learning approach. Additionally, the input covariates consisting of the tractor working hours, WPC, and blackening coefficient enabled a moderately high prediction accuracy, with a potential of implementing additional covariates which could provide an optimization and streamline the supervision of agricultural tractor exploitation load. These findings can be applied as a screening process in the management of agricultural machinery to avoid the frequency of performing more expensive and time–consuming laboratory analyses. Likewise, the proposed machine-learning framework should support various activities within a regular machinery maintenance, promoting the efficiency of the process and reducing a possibility of pollution during their exploitation.

REFERENCES

1. Awad, M., & Khanna, R. (2015). Support Vector Machines for Classification. In M. Awad & R. Khanna (Eds.), *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers* (pp. 39–66). Apress. https://doi.org/10.1007/978–1–4302–5990–9_3
2. Ayyadevara, V. K. (2018). Gradient Boosting Machine. In V. K. Ayyadevara (Ed.), *Pro Machine Learning Algorithms: A Hands–On Approach to Implementing Algorithms in Python and R* (pp. 117–134). Apress. https://doi.org/10.1007/978–1–4842–3564–5_6
3. Beni, M. S., Gholami Parashkoohi, M., Beheshti, B., Ghahderijani, M., & Bakhoda, H. (2023). Application of machine learning to predict of energy use efficiency and damage assessment of almond and walnut production. *Environmental and Sustainability Indicators*, *20*, 100298. https://doi.org/10.1016/j.indic.2023.100298
4. Bol'shakov, A. A., Pandey, S. J., Mao, X., & Liu, C. (2021). Analysis of liquid petroleum using a laser–induced breakdown spectroscopy instrument. *Spectrochimica Acta Part B: Atomic Spectroscopy*, *179*, 106094. https://doi.org/10.1016/j.sab.2021.106094
5. Carbone, F., Gleason, K., & Gomez, A. (2023). Chapter 2 – Soot research: Relevance and priorities by mid–century. In K. Brezinsky (Ed.), *Combustion Chemistry and the Carbon Neutral Future* (pp. 27–61). Elsevier. https://doi.org/10.1016/B978–0–323–99213–8.00007–2
6. Cersonsky, R. K., Helfrecht, B. A., Engel, E. A., Kliavinek, S., & Ceriotti, M. (2021). Improving sample and feature selection with principal covariates regression. *Machine Learning: Science and Technology*, *2*(3), 035038. https://doi.org/10.1088/2632–2153/abfe7c

7. Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random Forests. In C. Zhang & Y. Ma (Eds.), *Ensemble Machine Learning: Methods and Applications* (pp. 157–175). Springer. https://doi.org/10.1007/978–1–4419–9326–7_5
8. Dhananjayan, V., Jayakumar, S., & Ravichandran, B. (2020). Conventional Methods of Pesticide Application in Agricultural Field and Fate of the Pesticides in the Environment and Human Health. In R. K. R., S. Thomas, T. Volova, & J. K. (Eds.), *Controlled Release of Pesticides for Sustainable Agriculture* (pp. 1–39). Springer International Publishing. https://doi.org/10.1007/978–3–030–23396–9_1
9. Dörr, N., Brenner, J., Ristić, A., Ronai, B., Besser, C., Pejaković, V., & Frauscher, M. (2019). Correlation Between Engine Oil Degradation, Tribochemistry, and Tribological Behavior with Focus on ZDDP Deterioration. *Tribology Letters*, *67*(2), 62. https://doi.org/10.1007/s11249–019–1176–5
10. Kozina, A., Radica, G., & Nižetić, S. (2020). Analysis of methods towards reduction of harmful pollutants from diesel engines. *Journal of Cleaner Production*, *262*, 121105. https://doi.org/10.1016/j.jclepro.2020.121105
11. Radočaj, D., Tuno, N., Mulahusić, A., & Jurišić, M. (2023). Evaluation of Ensemble Machine Learning for Geospatial Prediction of Soil Iron in Croatia. *Poljoprivreda*, *29*(2), 53–61. https://doi.org/10.18047/poljo.29.2.7
12. Sharma, A., Jain, A., Gupta, P., & Chowdary, V. (2021). Machine Learning Applications for Precision Agriculture: A Comprehensive Review. *IEEE Access*, *9*, 4843–4873. https://doi.org/10.1109/ACCESS.2020.3048415
13. Song, H., Li, G., Li, X., Xiong, X., Qin, Q., & Mitrouchev, P. (2023). Developing a data–driven hydraulic excavator fuel consumption prediction system based on deep learning. *Advanced Engineering Informatics*, *57*, 102063. https://doi.org/10.1016/j.aei.2023.102063
14. Varga, I., Radočaj, D., Jurišić, M., Markulj Kulundžić, A., & Antunović, M. (2023). Prediction of sugar beet yield and quality parameters with varying nitrogen fertilization using ensemble decision trees and artificial neural networks. *Computers and Electronics in Agriculture*, *212*, 108076. https://doi.org/10.1016/j.compag.2023.108076

PRISTUP STROJNOGA UČENJA ZA PROCJENU KVANTITATIVNIH PROMJENA ULJA U DIZELSKOME MOTORU TRAKTORA TIJEKOM EKSPLOATACIJE

SAŽETAK

Kako bi bio procijenjen potencijal pristupa strojnoga učenja u evaluaciji kvantitativnih promjena u ulju dizelskoga motora traktora tijekom eksploatacije, ova studija je imala za cilj predložiti regresijsku metodu strojnoga učenja kako bi se smanjila učestalost skupoga i dugotrajnog uzorkovanja motornoga ulja. Ulazni skupovi podataka motornoga ulja sa svježim motornim uljem (Uzorci A) i s motornim uljem nakon 250 radnih sati (Uzorci B) uzorkovani su za dvanaest elemenata u dvogodišnjoj eksploatacijskoj studiji u Belju d. d. Prikupljanje podataka na terenu obavljeno je na šest teških poljoprivrednih traktora* FENDT 930 Vario *s pogonom na sva četiri kotača, od kojih je svaki praćen 1500 radnih sati, pri čemu je svakih 250 radnih sati uzorkovano motorno ulje. Procijenjene metode predviđanja strojnoga učenja temeljene na desetorostrukoj unakrsnoj validaciji postigle su umjereno visoku točnost predviđanja, s višim koeficijentom determinacije (R^2) u rasponu 0,51–0,73 za uzorke B negoli u rasponu 0,49–0,64 za uzorke A. Ovi rezultati snažno sugeriraju da niti jedna metoda strojnoga učenja nije dosljedno postigla visoku točnost predviđanja i da bi odabir optimalnih modela strojnoga učenja trebao biti obvezan, a istovremeno potvrđuju visok potencijal metoda strojnoga učenja u otkrivanju kvantitativnih promjena u ulju dizelskoga motora traktora tijekom eksploatacije.

Ključne riječi: poljoprivredni traktor, onečišćenje, važnost varijabli, optimizacija, eksploatacija

(Received on January 17, 2024; accepted on February 15, 2024 – *Primljeno 17. siječnja 2024.; prihvaćeno 15. veljače 2024.)*