**The ANFIS Anticipation of Diesel Locomotive Engine on the Improvement of Lubricating Oil Analysis**

Bakhodir Gayratov1, a), Otabek Khamidov1, b), Shohijakhon Kudratov1, c) and Shakhboz Samatov1

1*Tashkent State Transport University, 1 Temiryulchilar St., Tashkent 100167, Uzbekistan*

a) Corresponding author: bgayratoff@gmail.com  
b) otabek.rustamovich@yandex.ru   
c) kudratov.shohijaxon@bk.ru

**Abstract.** The enhancement of the reliability of systems of rail transport and its maintenance quality are some of the most basic issues in the exploitation of contemporary locomotives. More specifically, locomotive engine potential malfunctions are detected early and such aspect is significant because it helps the locomotive engines operate efficiently and safely. This work suggests a new method of engine status assessment with regard to the detailed study of the changes in the physicochemical characteristics of lubricant oil and it is based on the Adaptive Neuro-Fuzzy Inference System (ANFIS). The primary input parameters are kinematic viscosity, total acid number (TAN), wear metal concentration and water contamination level. The fuzzy inference rules coupled with the learning features of neural networks in the proposed model makes it useful in predicting malfunction severity more accurately. The ANFIS-based model is very flexible, capable of modeling nonlinear relationships, and can provide early warning in real-time as compared to the common statistical or threshold-based approaches. This method can be used as a successful solution to the creation of the condition-based and predictive maintenance strategies of locomotive diesel engines. Subsequently, the efficiency of operations can be boosted greatly.

**Keywords:** diesel locomotive engines, lubricating oil, predictive diagnostics, fault detection, condition monitoring, fuzzy logic, engine diagnostics, intelligent systems, railway transport, Adaptive Neuro-Fuzzy Inference Systems (ANFIS).

# Introduction

The Adaptive Neuro-Fuzzy Inference System (ANFIS) was adopted in order to propose a predictive diagnostics model of the diesel locomotive engine in this study accompanied by the lubricant oil analysis. Combining the technologies of fuzzy logic with neural networks, the model has been able to identify nonlinear associations between important parameters of oil like viscosity factor, acid, metal contamination and water content [1]. It showed good diagnostic accuracy (91.2) and a mean square error of 0.083 with a high AUC value of 0.95. This shows that it is a reliable method of detecting and classification of engine faults. It makes use of non-invasive data on oil that allows monitoring the engine in real-time with continuous reports on the state of the engine thus dramatically decreasing the dependency on manual inspection [2]. The ANFIS-based technique has more flexibility, the ability to learn and resistance to diverse working conditions than the traditional threshold-based diagnostics. Such characteristics qualify it as an effective prospective support management tool in the railway fleets. The future upgrade can involve the involvement of more sensor data and real-time operating parameters and come up with an interactive interface to the technical staff. These enhancements will further improve or enhance the diagnostic accuracy and enable it to become accessible to the railway maintenance systems in large number [3]. To sum up, the ANFIS model suggested has a good potential to facilitate the condition-based maintenance policy regarding engine of locomotives, to minimize unscheduled stoppages and to increase the machine lifetime.

# Literature review

The rising focus on larger reliability, safe run, and economic efficiency of locomotive engines has increased the need of predictive diagnostic approaches. Out of many methods available to monitor engine conditions, lubricant analysis has been found to be a credible and non-destructive technique which tells in real time situation about condition of the engine inside. Alteration of the physicochemical characteristics of used oil - reduced viscosity, the achievement of metal wear particles (e.g., Fe, Cu, Pb), the level of oxidation, the content of soot carbon - the direct indicator of the wear of internal components, mechanical loads, and overload on the engine device. Monitoring the conditions of oil has been recently gaining a lot of popularity because of the fact that it allows detecting the malfunctions that happen at the earliest time without disassembling or other intrusive operations. The analytical systems based on the Python programming language and tools such as Pandas and Scikit-learn have also been successfully used to process the data of oil and predict mechanical issues (Sharma et al., 2020). On the same note, Zhang et al. (2019) have proven that tendencies of oil viscosity and temperature analyzed with the help of machine learning would be able to detect the precursors of malfunctions. As the Internet of Things (IoT) technologies go more in use, the diagnostic capabilities of oil analysis are even greater. As an example, Wang et al. (2018) activated real-time surveillance equipment based on IoT-based sensors and intelligent algorithms. Through this system, ongoing remote evaluation of lubricant state and calculative maintenance can be done, particularly when it concerns geographically diffuse locomotive fleets. The recent literature also refers to the use of hybrid and deep learning when diagnosing oil. Li et al. (2021) integrated long short-term memory (LSTM) neural networks and the mature technology of data preprocessing to enhance the precision of the remaining useful life (RUL) predictions of the lubricants. Moreover, authors like Zhang et al. (2022) and Wang et al. (2022) applied models to both the spectroscopic and the sensor-based sets of oil to classify faults and diagnose the condition of engine components based on Random Forest, Support Vector Machines (SVM), and Convolutional Neural Networks (CNNs ). These results justify using the state of the art in artificial intelligence (AI) in oil-based predictive maintenance[4]. The majority of the present-day research, however, is based mostly on such fields, as automotive, wind energy, or industrial engines, with little reference to the peculiarities of the railway locomotive. Efficiency demands specific diagnostic systems to overcome the adverse operating conditions of the engines of the railway - one is constant mechanical stress and another is variable thermal loads. Further, most of the literature done in the past follows a univariate analysis ignoring the intricate interactions between various oil parameters. This shows that very little research has been conducted in the way of practical, multi-feature, AI-driven diagnostic models specialized to railway systems. It is notable that some of the international railway operators - SNCF (France), Amtrak (USA) and Via Rail (Canada) have embarked into the use of smart engine monitoring structures. IBM Watson IoT and Siemens Expert-on-Alert are some of the available solutions that use sensor networks to gather and break down operational data to help predict mechanical failures, as well as minimize downtime that is not predictable. However, such services are prone to using generic forms of monitoring as opposed to oil-specific diagnostic procedures. To manage this gap, the research paper suggests a predictive diagnostic system of diesel locomotive engines using Adaptive Neuro-Fuzzy Inference Systems (ANFIS) in combination with the analysis of lubricants. The combination of professional expertise and data modelling will allow enhancing early fault detection, optimizing maintenance solutions and consequently the service life of engines under railway operating conditions[5].

# Materials and Methods

The set of key input parameters in the proposed diagnostic model are drawn using the analysis of the samples of used lubricating oils in the mainline locomotive diesel engines. These signs are highly sensitive pointers to wear conditions, pollution and degradation of an engine. Other regulatory standards regulate their assessment and classification of such oils, being the most notable GOST 17479.1-85 (GOST 17479.1-2015 rev.). This standard introduces a complete system of classification of motor oils depending on their kinematic viscosity and thresholds of performance under different conditions of thermal and mechanical loading [6]. Moreover, the methodological base of GOST 20759-90 is constructed with the definition of the technical diagnostics and forecast of the further service life of diesel engine, especially with the rules of spectrochemical analysis of used lubricating oils. Such a method of analysis allows identifying the abnormal wear processes and locating the possible malfunctions in the early stages with the help of quantitative evaluation of the concentration of wear-associated metals (e.g., Fe, Cu, Cr, Al, Pb) and contaminants. These standards integrated guarantee the consistency as well as reproducibility of the diagnostic outcomes that is the foundation of predictive maintenance approaches in locomotive fleet management systems [7]. In order to pursue such an approach, a set of physicochemical parameters of used oils is applied as input of a diagnostic model. These are attributes like viscosity, acidity, flash point and level of mechanical and chemical impurities which play a major role in determining the state of oil and consequently wear in the engine. The list of these input parameters applied in the diagnostics of diesel engines in the oil analysis is provided in Table 1.

The comparative evaluation of major physicochemical parameters of used samples of lubricating oil in relation to the regulatory values is shown in figure 1. This visualization illustrates anomalies in the parameters like Viscosity, flash point, Acidity, and Mechanical impurities and this serves as a good example of how the diagnostic indicators are applied in the model [8].

**FIGURE 1.** Physicochemical properties and minimum limits of lubricant

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **TABLE 1.** Input variables used in oil analysis for engine diagnostics   |  |  | | --- | --- | | **Variable Description** | | |  | Kinematic viscosity at 100 °C, mm²/s  Flash point, °C  Content of mechanical impurities, %  Water content (%)  Acidity (mg KOH/g of oil)  Content of water-soluble acids and alkalis  Density at 20 °C, kg/m³ | |

The value of the output variable is the level of the engine malfunction which is specified as a categorical attribute [9] or a continuous fault score y=[0,3].

**ANFIS Architecture**

The ANFIS architecture was built with a seven layer structure to form the predictive model:

**Layer 1** - **Fuzzification**: The input value xi is given membership functions that are fuzzy, e.g., Low, Medium, High (e.g. the Gaussian) [10].

**Layer 2 – Rule Activation**: The firing strength of any fuzzy rule Rk is taken as:

|  |  |
| --- | --- |
|  | () |

**Layer 3 – Normalization**: Each rule’s firing strenght is normalised.:

|  |  |
| --- | --- |
|  | () |

**Layer 4 – Output of the rule**: Output of the rule: Each fuzzy rule output a linear feedback of the input variables. This will enable the so that the intricate relations could be estimated using a piecewise linear formula, in the context of the ANFIS architecture:

|  |  |
| --- | --- |
|  | () |

**Layer 5 – Aggregation**: The last product is the mean of all rule outputs weighted together:

|  |  |
| --- | --- |
|  | () |

**Training Procedure**

Fuzzification tables - fuzzy memberships parame- ters and linear coefficients are adaptive using a hybrid learning method [9]:

LS Least Squares Estimation (LSE) for linear output coefficts. The Regression Of Optimum Gradient; Descent (GD) of fuzzy subscription parameters.

The training set is a historical data of the oil samples of locomotive engines in operation along with the diagnosis of faults [10].

# Results and Discussion

ANFIS toolbox in MATLAB was utilized in the development of the model prototype and trained with a data set of important physicochemical lubricating oil properties. The cross-validation technique (10-fold) was employed to evaluate its predictive quality with the main idea of dividing the dataset into ten subsets after which each subset would be used once to perform the validation and the rest of them to perform training.

The ANFIS-based diagnostic model performance was measured with the help of the standard performance measures, such as Mean Squared Error (MSE), accuracy, precision and recall, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). These indicators provide a general evaluation on the performance of the model to analyze the conditions of the engines using oil as indicator. The quantitative values of these metrics of evaluation are summarized in Table 2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **TABLE 2.** Significant Evaluation Indicators of the ANFIS Based Diagnostic Model | | | | | |
| **Metric** |  |  |  |  | **Value** |
| Mean Squared Error (MSE) 0.083  Accuracy (classified) 91.2%  Precision (avg.) 0.89  Recall (avg.) 0.92  AUC (ROC curve) 0.95 | | | | | |

According to results the accuracy of detecting and classifying the levels of engine malfunction on the basis of the oil data is high. The Adaptive Neuro-Fuzzy Inference System (ANFIS encoded into this proposed diagnostic model) is the representative of a hybrid intelligent method. It involves the expert knowledge fuzzy logic and the data-driven neural learning. This synergetic combination can be used to develop an auto adapting, real time, condition monitoring system which can be used to evaluate the degradation of the lubricating oil in diesel engines. However, input parameters used to build the model include key parameters of Physicochemical techniques so that the diagnostics can be realized without engine stoppage and in a non-invasive manner [11]. Fuzzy rule is the reasonability shown by the model stated in the form: When viscosity is high and iron (Fe) concentration is also high and total acid number (TAN) is also high, then the degree of malfunction is 2.8 equivalent to an almost critical engine health degree. In comparison with the conventional boundary-oriented or exclusively statistical approaches to diagnostic procedures, the suggested model has a number of salient pluses:

Extreme flexibility in responding to changes in the engine load as well as the environment without the need to manipulate manually;

Ability to formulate nonlinear relationship complex relations that are usually ignored in linear models when dealing with oil degradation indicators;

A good generalization skill, despite which it is possible to predict faults in the early stage using past trends and correlations among a number of variables [12]. Regardless of these positive attributes, there are numerous opportunities that can be used to improve the diagnostic capability and functionality of the model: Improvement of input parameters and diagnostic accuracy through fusion of the input data of several sources of sensors among them being the engine load, coolant temperature, fuel intake and vibration implications; Increasing the number of the training data set with the introduction of high accuracy field records obtained when operating a fleet of locomotives under differing operational modes;

Creation and provision of a user-friendly HMI, which will allow understanding the diagnostic outcomes easily and assist specialists serving at maintenance depots to make accurate decisions in time. All in all, such advances make the given suggested diagnostic system much more reliable, scalable and universal when put in the real-life scenario of railway maintenance. The combination of fuzzy logic inference and adaptive neural network learning introduces the model as an attractive way of using predictive maintenance policies and increasing the time of service of a particular engine [13, 14, 15].

The accuracy of the developed model in terms of classification in 4 different categories of faults is represented in figure 2. Whereas the findings have shown high effectiveness in attaining healthy and critical conditions, there is a little lower degree of accuracy found in the case of minor and moderate faults. This has determined the spheres in which the model can be improved more in future [11, 12, 13, 14].

|  |
| --- |
|  |

**FIGURE 2.** The rate of fault diagnosis accuracy per engine fault level.

# Conclusion

Within the frame of the current study, the idea of the predictive diagnosing of the diesel locomotive engines with the help of the Adaptive Neuro-Fuzzy Inference System (ANFIS) was proposed along with the analysis of the lubricating oil. The model has managed to capture nonlinear relationship among the major oil parameters like viscosity, acidity, metal contamination, and water content by integrating fuzzy logic together with neural networks. The model produced a considerable diagnostic accuracy (91.2%) and low mean square error (0.083) and high value of AUC (0.95). This serves to verify that it is quite capable of the detection and classification of engine malfunctions. It can constantly monitor real-time engine condition because of its utilization of non-destructive oil data and greatly minimizes the need to rely on manual checking. In comparison with traditional threshold-based diagnostics procedure, the ANFIS-based approach offers great adaptability, self-learning feature, and resiliency in regards to a wide range of operating conditions. It is appropriate to be used as predictive maintenance tool of railway fleets because of its features. Extended functionalities can incorporate the use of other types of sensor information and real time parameters of the operation and building an intuitive interface to the technical staff. These developments will also enhance the accuracy of the diagnosis procedure and make it easy to adopt in the railway system maintenance. To summarize, the suggested ANFIS model has been proven to be highly effective in aid of condition-based maintenance of locomotive engines, minimized unscheduled downtimes, and increased service life.

# References

1. C. V. Zaharia, R. Niculescu, V. Iorga, C. Ducu, A. Clenci, and B. Aron, in *CONAT 2016 International Congress of Automotive and Transport Engineering*, edited by A. Chiru and N. Ispas (Springer, 2017), pp. 319–327.
2. S. Galbadrakh, P. Jamyan Osor, and O. Natsag, Mong. J. Agric. Sci. **16**, (38) (2021), <https://doi.org/10.5564/mjas.v16i38.3131>.
3. M.-R. Pourramezan and A. Rohani, SAE Int. (2004), <https://doi.org/10.4271/04-18-01-0005>.
4. MDPI Lubricants Editorial Team, Lubricants **13**, 169 (2021).
5. H. Wang, J. Liu, and X. Zhang, Mech. Syst. Signal Process. 140, 106653 (2020).
6. H. Raposo, J. T. Farinha, I. Fonseca, and D. Galar, Tribol. Int. **135**, 65–74 (2019).
7. O. Khamidov, N. Zayniddinov, S. Kudratov, and B. Erkinov, E3S Web Conf. **538**, 01027 (2024), <https://doi.org/10.1051/e3sconf/202453801027>.
8. O. Khamidov, S. Kudratov, O. Khamidov, B. Turdimurodov, and M. Norbotayeva, AIP Conf. Proc. **3045**, 050030 (2022).
9. A. Yusufov, O. Khamidov, N. Zayniddinov, and S. Abdurasulov, E3S Web Conf. **401**, 03041 (2023).
10. S. Jamilov, A. Yusufov, S. Samatov, and S. Kudratov, ICTEA Int. Conf. Therm. Eng. **1**, 1 (2024).
11. O. Khamidov, A. Yusufov, S. Kudratov, and A. Yusupov, E3S Web Conf. **365**, 05004 (2023).
12. S. Abdurasulov, N. Zayniddinov, A. Yusufov, and S. Jamilov, E3S Web Conf. **401**, 04022 (2023).
13. S. Jamilov, O. Ergashev, M. Abduvaxobov, S. Azimov, and S. Abdurasulov, E3S Web Conf. **401**, 03030 (2023), <https://doi.org/10.1051/e3sconf/202340103030>.
14. S. Fayzibaev, Z. Muxammedova, and I. Raxmiddinov, E3S Web Conf. **458**, 01015 (2023), <https://doi.org/10.1051/e3sconf/202345801015>.
15. B. Gayratov and D. Insapov, AIP Conf. Proc. **3286**, 060007 (2025), https://doi.org/10.1063/5.0279777.