Vol. 45, No. 3 (2023) 487-502, DOI: 10.24874/ti.1491.06.23.08

Tribology in Industry

www.tribology.rs

RESEARCH

Prediction of Gearbox Oil Degradation Based on Online Sensor Data and Machine Learning Algorithms

Kunal Kumar Gupta[a] and S.M. Muzakkir[b,*]

[a]*Nagarjuna College of Engineering and Technology, Bangalore, India,*
[b]*Department of Mechanical Engineering, Jamia Millia Islamia, New Delhi, India.*

Keywords:

Lubricating ageing
Additive depletion
Wear
LR
DTR
KNNR
RFR
ANN

* Corresponding author:

S. M. Muzakkir
E-mail: smmuzakkir@jmi.ac.in

Received: 6 June 2023
Revised: 16 July 2023
Accepted: 22 August 2023

A B S T R A C T

In most of the gearboxes, mixed lubrication conditions prevail, and to avoid the wear of gear surfaces, oil additives like extreme pressure, anti-wear, anti-rust and antioxidant additives are used. The lubricant additives form a lubricant film on gear surfaces, minimize metal-to-metal contact and protect the surfaces. But in this process, the lubricant additives are consumed, and oil quality deteriorates causing degradation of wear-prevention lubricant functionality. The degradation of lubricant with time, even without its usage is alarming and it has been reported in the present manuscript. To observe the consequence of degraded gear-oil on gear surface, an experimental setup has been developed. The results of experiments, conducted on commercially available two-gear oils have been detailed. Three cases of single stage spur-gear pair were considered: (1) Loaded with 40 Nm torque value and operated at 1200 rpm for 198 hours duration. (2) Loaded with 50 Nm torque value and operated at 500 rpm and. (3) Accelerated conditions generated by adding 0.0025 %v of mild (36% concentration) Hydrochloric acid in the lubricant in addition to accelerated conditions specified in case 2. For the cases 2 & 3, the setup was run for 90 minutes duration. The dataset of this study includes five parameters namely time, humidity, temperature, oil quality and generated Fe debris. Machine learning techniques have been used to reduce the actual number of experiments by applying LR, DTR, KNNR, RFR, ANN and SVM in predicting the Oil degradation rate.

1. INTRODUCTION

Ideally, gears work under EHL (Elasto-hydrodynamic lubrication) condition, whereby elastic deformation of surface asperities shown in Figure 1-a and piezo-viscous effect of lubricant at local contact take place. In reality, mixed lubrication conditions prevail and increases the chances of gear surfaces getting worn out. To avoid such wear phenomenon, extreme pressure additives are used. Gear lubricants are categorized into three classes by AGMA 9005-F16: [1] compounded oils, anti-scuffing, and rust and oxidation inhibited oils. To impart qualities like anti-wear and anti-scuff properties under high pressure circumstances, lubricants are combined with additive packages. The extreme pressure (EP) additive molecules, under the

action of extreme pressure and relatively high temperature [2], react with gear surface to form protective sacrificial layer [3]. The sacrificial layer, known as tribo-film, protects the gear surfaces from direct asperities contact as shown in Figure 1-b.

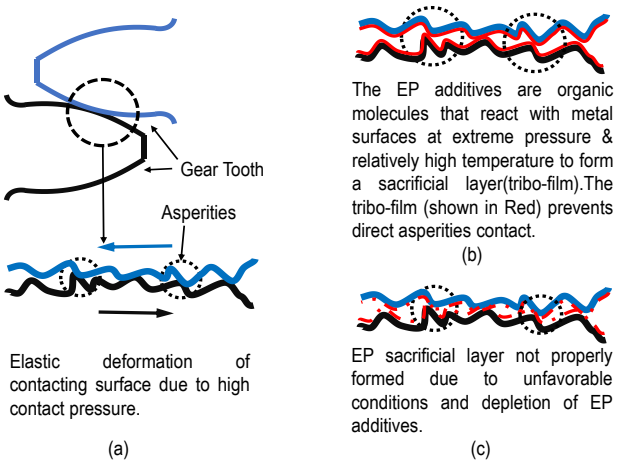


Fig. 1. (a) Surface interaction of gear surfaces under EHL lubrication (b) Formation of protective tribo-films to avoid direct metal to metal contact (c) Metal to metal contact under depleted additives leading to surface wear.

With the usage of additives (in manner of sacrificial layer formation and neutralizing the formed acid), additives deplete continuously. In addition, the presence of air, moisture, and particles in the lubricant cause entrapment of additives molecules, thereby causing accelerated fall in additive level. Under such conditions, the formed tribo-film would not be sufficient to protect the wear of gear flank as shown in Figure 1-c. The wear of the journal is affected by the geometrical features [4], addition of multi-walled carbon nano-tubes and zinc nano-particles in lubricant minimizes the wear [5-7], the viscosity of lubricant that affects the wear can be actively controlled by using smart materials [8], optimization methods that converges fast are utilized [9-11], the reduction in viscosity of lubricant due to shear thinning must be incorporated [12], and a suitable design methodology must be used [13-16]. For the purpose of monitoring the status of the gearbox, several techniques based on vibration, acoustic, wear debris, and lubricant analysis have been developed. [17-21]. On the basis of machine learning techniques and the internet of things (IoT), several approaches for diagnostics and evaluation of the gearbox's remaining usable life were developed. [22-27]. In the present paper, Machine learning calculations using the lubricant degradation dataset have been used to accelerate the prediction of gear conditions. The particle size affects the shear stress in the lubricant [28,29], and in certain conditions the sliding surfaces must be separated from each other to reduce wear [30-33]. The root cause of failure must be identified [34].

Gas chromatography can be used to examine the lubricating oil's oxidation stability and determine its composition, including the amount of heptadecanes, tetracontanes, and octadecanes [35]. The pressure differential scanning calorimetry (PDSC), thermogravimetric analysis (TGA), micro-oxidation rotating pressure vessel test (RPVOT), and IP 48 methods of measurement were all examined by Jain et al. [36] to determine how different antioxidants affected the thermo-oxidative stability of lubricant oils and found that PDSC provide better and fast results as compared to the other tools. The amine and phenolic based antioxidant were good under different temperature ranges. Wolak et al. [37] studied the effect of time on the degradation of the different oil. They prepared the statistical model on the basic of TAN value change with the operating hours. Rivas et al. [38] employed numerous models, such as random forest, support vector machine, linear model, etc. to develop a model to predict the TAN value of used oil. They measured the TAN value and the FTIR (Fourier transform spectroscopy) to detect the depletion of various components in the engine oil.

Avoiding any faults is crucial for ensuring the gearbox's safe and reliable operation. The machine learning and IoT based algorithms have been utilised to predict the occurrence of fault and useful life (RUL). Popular method using machine learning starts with selection of features by algorithms like NCA- Neighbourhood component analysis, principal component analysis etc. and followed by prediction of faults severity by algorithm like SVR support vector regression, decision tree, SBL-sparse Bayesian learning, HT- hypothesis testing, CNN- conventional neural network, LSTM- long and short term memory network, SSAE- stacked sparse autoencoder model, ANN artificial neural networks, deep learning, and Dempster-Shafer etc., and optimization like PSO- particle swarm optimisation, multi-objective optimization etc. Inturi et al. [39] used Hurst exponent to analyse non-stationary and aperiodic data based on self-similarities and scale-invariance properties of the data to detect the fault in multi-stage gearbox operating under non-stationary operating conditions.

Although, many studies have been carried out using sophisticated machine learning algorithms to predict residual life of gearbox, but impact of oil degradation on gear wear is missing. The oil analysis and wear debris analysis have been treated independently. In the present research the effect of oil degradation on wear debris have been studied. An experimental setup has been developed. The online oil sensor suite and metallic wear debris sensor have been placed to monitor progressive wear of the gear. The results of experiments, conducted on commercially available two-gear oils have been detailed. Three cases have been considered: (1) Single stage spur-gear pair loaded with 40 Nm torque value and operated at 1200 rpm for 198 hours duration. (2) Accelerated wear conditions generated by increase torque and reducing speed. (3) Accelerated conditions generated by adding 0.0025 %v of mild (36% concentration) Hydrochloric acid in the lubricant in addition to load and speed conditions specified in the case 2. For the cases 2 & 3, the setup was run for 90 minutes duration, at 50 Nm and 500 rpm. The dataset of these studies includes five parameters namely time, humidity, temperature, oil quality and generated Fe debris. The actual number of experiments has been decreased through the application of machine learning techniques like LR, DTR, KNNR, RFR, ANN and SVM in predicting the Fe and Oil quality parameters.

2. EXPERIMENTAL SETUP

Figure 2 shows the photograph of the experimental setup, consisting of single-stage spur gearbox driven by a DC electric motor (30 kW). The speed of the motor can be varied from 0 to 3000 rpm by the help of a speed regulator. The load on the gearbox is applied by eddy current dynamometer (capacity 0-75 Nm).

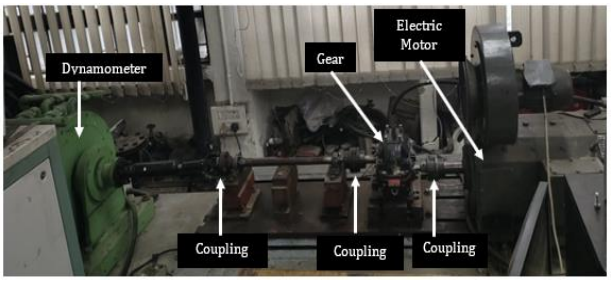


Fig. 2. Experimental setup.

The module of the pinion and the gear is 2. There are 27 and 53 number of teeth on pinion and gear respectively. The pitch diameter of pinion and gear are 54 and 106 mm respectively. The base diameter of pinion and gear are 50.7 and 99.6 mm respectively. The center distance is 80 mm, face width is 33 mm, pressure angle is 20°, and contact ratio is 1.697. The material is EN24 steel (with HRC of 20, Poisson's ratio of 0.28 and Youngs Modulus of 207 GPa).

The lubrication was provided by API GL4 SAE 90 during the test. The lubricant consists of extreme pressure additives, anti-rust and corrosion additive package. The TAN value of the lubricant was measured. The measured TAN value (0.064) was found within the prescribed limits defined by AGMA for the R&O lubricants, this value indicates the acidic nature of the lubricant. To confirm this situation a second commercially available lubricant (API GL4 EP90) was purchased and processed for the measurement of the TAN value. TAN value was found to be 0.0595 which was well within the limits, but it ensures oil towards acidic side. To confirm the effect of aging of the oil the second packet of the oil 2 was purchased from the market. The packet 2 of the oil 2 was nearly 10 months older than packet 1. The TAN value was measured, found to be 0.069. The oil specifications are listed in Table 1.

Table 1. TAN values for the fresh and HCL mixed lubricant.

Grade of the oil used	Date of manufacturing	Fresh oil TAN value (mg KOH/g)	(Oil + 0.0025 %v HCL) TAN value (mg KOH/g)	Additive package (according to OEM)	Market price of the one liter of lubricant (INR)
API GL-4 SAE 90 (Oil 1)	April, 2019	0.064	0.0865	Extreme pressure, anti-rust and anti-corrosion additives	220
API GL-4 EP90 (Oil 2)	March 2022 (Packet 1)	0.0595	0.1135	Extreme pressure additives for moderate service condition and anti-rust and corrosion protection	230
	June 2021 (Packet 2)	0.069	0.121		

A metallic wear debris sensor manufactured by Kittiwake (AS-19144-KW) has been used for measuring the total wear mass (WM) (µg/hr), particle size in different size bands: ferrous (ranges from 40µm to 400µm) and non-ferrous (ranges from 135µm to 450µm) particles. The sensor works on the principle of Biot-Savarts's law and relates change in the magnetic field with the size of wear debris (smallest diameter of the wear debris). The sensor is low sampling frequency 10 samples per 60 second. The framework and the front panel of the sensor software are shown in Fig. 3. The panel shows the alarm value for different parameters: the particle per minute, wear mass, ferrous and non-ferrous particle cumulative count. The software saves the data in excel files for postprocessing operation.

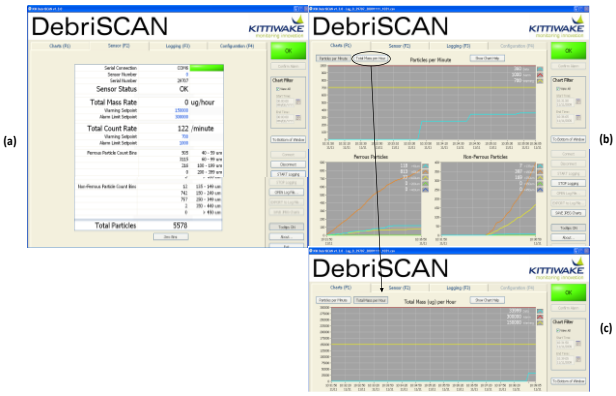


Fig. 3. Front panel of the metallic wear debris sensor data acquisition software. (a) Details of login, data count, and warning or alarm value of the parameters, (b) showing the panels for the particle per minute, ferrous particle, non-ferrous particles and (c) wear mass (µg/hour).

An oil sensor suite, ANALEXrs, has been used for measuring the total ferrous wear debris content (Fe ppm), oil quality (OQ) (1-100), and oil temperature (OT) (-20 °C to 120 °C). The suite consists of different sensor units installed to monitor performance parameters. The sensor uses Tan Delta dielectric sensing and thick capacitance film and integrates those values using algorithm to calculate the oil quality, temperature of lubricant and Fe ppm. The sensor is low sampling sensor with one sample per 60 second. The framework and the front panel of the sensor suite are shown in Fig. 4. There are two panels: One for ferrous content and other for the oil quality, oil temperature and relative humidity. The data are saved in excel files.

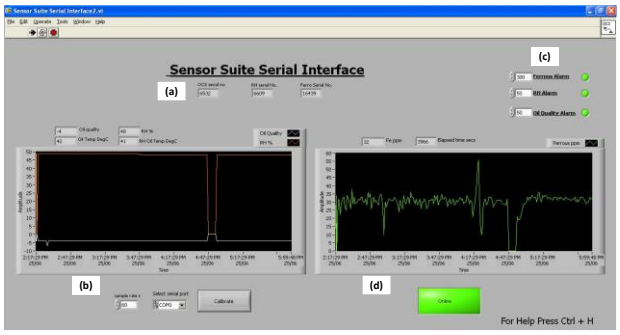


Fig. 4. Image showing the front panel of the oil sensor suite program. (a) serial number of the active sensors, (b) oil quality, oil temperature and oil humidity sensor reading panel, (c) warning reading of the sensors, and (d) ferrous sensor reading.

A 5 MP ultra-mini-camera is used as an inspection camera to capture the periodic image of the gear teeth. The image of gear teeth was captured after every 50 hours. The camera was positioned inside with the help of a flexible tube so that the image can be captured inside and closer to the gear tooth at any angle.

Anton paar rheometer MCR 102 was used for measuring the viscosity of the used lubricant oil. The 20 mm diameter cone-plate was used to measure the oil viscosity at a gap of 1mm and shear rate of 100 s-1 was used. The rheometer has a maximum torque limit of 200 n-Nm and temperature range of -5 to 200 °C. The accuracy limit of the instrument is 0.45 n-Nm torque.

REMI CPR- 24 plus centrifuge was used for particle separation. The maximum rotation rpm is 7000 and the weight at the maximum rpm was 4.26 kg. The centrifuge help in the deposition of the metallic particles on the side walls of the tubes. The oil was drained and the particles were collected and washed with ethanol to remove the oil layer.

3. METHODOLOGY

The acquired data from the online sensor was mix of signal and noise. It is required to de-noise the data for extracting the useful information. To simplify the data processing, three cases of dataset: case 1, case 2, and case 3 have been considered. five parameters (time (minutes), torque (nm), speed (rpm), oil temp, rh%) are considered as input based on which oil quality and FE ppm is estimated. the details of the dataset are given in the table 2.

Table 2. Information and parameters present in the 3 cases of data.

Case	No of values of time (minutes)	No of values load (Nm)	No of values of oil temp	No of values of RH%	No of values of output (Oil Quality)	No of values of output (Fe PPM)
1	11876	11876	11876	11876	11876	11876
2	90	90	90	90	90	90
3	90	90	90	90	90	90

Six machine learning algorithms, namely, Linear Regression (LR), Decision Tree (DTR), KNN Regression (KNNR), Random Forest Regression (RFR), Artificial Neural Network (ANN), Support Vector Regression (SVR) were separately applied to estimate the values of the Oil Quality and Fe ppm. The evaluation metric like MAE, MSE, and RMSE are used to find the best suitable algorithm for the predictions.

Oil quality is estimated by following the below mentioned pre-processing steps.

i) The entire datasets are divided into training and testing set in the ratio of 80:20 (Table 3).

Table 3. Training and testing set of 3 different datasets.

Dataset	Training Set	Testing Set
1	9501	2375
2	72	18
3	72	18

ii) After dividing the training set, the training and testing input parameters are separately standardized by subtracting the mean and scaling each feature to unit variance considering the standard normal distribution.

iii) Then the features were fed to into the different machine learning algorithms to predict the Oil Quality and FE PPM.

Table 4. Hyper parameter of the different machine learning models.

Model	Hyper parameters
LR	-
DTR	CART Algorithm
KNNR	K =5
RFR	bootstrap=False, max_features=1, n_estimators=50
ANN	• 2 Hidden layer with Relu activation function • Each Hidden Layer contains 6 hidden neurons • Batch Size =32 • Epoch = 100 • Adam Optimiser
SVR	C=0.8, epsilon=0.2, kernel='poly'

iv) In the next step, different machine leaning regression algorithms were applied to predict the Oil Quality and FE PPM. The hyper parameters of the different models are presented below in the Table 4.

4. RESULT AND DISCUSSIONS

4.1 Case 1

The experiment was conducted at constant rotational speed of 1200 rpm and torque 40 Nm for a test duration of 198 hours. The commercially available lubricant oil (Oil-1) was used. The offline and online study was conducted. In online study, two sensors, Oil sensor suite and metallic wear debris sensor were used to acquire the data.

For the online oil analysis, the data acquired from the sensor was processed using the Machine learning algorithms as described earlier. The trend of increase in oil quality increases with operating time was observed, which confirms the deterioration of the oil quality with time. No significant increase/decrease in temperature and relative humidity was noticed.

Table 5. Fe particle size.

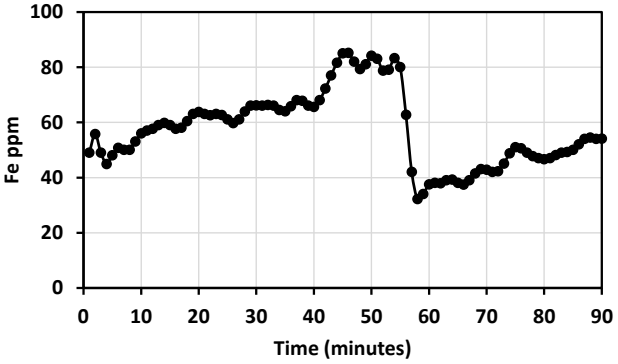
Particle size	Particle Concentration	Test Duration (hours)
Greater than 40 µm	51660	6
	7135	50
	6435	100
	12573	150
	12134	198
Greater than 60 µm	13000	6
	3794	50
	3546	100
	20684	150
	51537	198
Greater than 100 µm	178	6
	148	50
	241	100
	1247	150
	4095	198

For the online wear debris analysis, the data related to the ferrous debris concentration, and wear mass per hour (µg/ hour) were recorded, as given in Table 5. Increase in the Ferrous debris concentration and wear mass (µg/hour) was observed while more fluctuations were observed towards the end of the experiment. The smaller particle number (>40 µm and <60 µm) was higher in the initial running, which also indicate running in process in gear-lubrication does not occur as particle size lesser than 40 micro-meter were negligible.

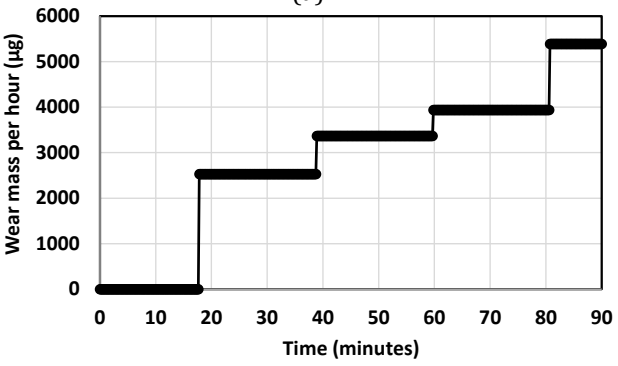
The gear tooth's failure mode observed on the gear surface was burr, scuffing, initial pitting, and spall.

4.2 Case 2 and 3

The tests were conducted at constant rotational speed of 500 rpm and 50 Nm torque for a duration of 90 minutes. The commercially available lubricant oil API GL4 EP90 was used. In case 2 no accelerating degradation agent was used and in the case 3 the oil is mixed with 0.0025%v of HCL. The online data was acquired using oil sensor suite and metallic wear debris sensor. The oil samples of 30ml were collected after every 10 minutes. The collected oil samples were processed for the viscosity measurement.



(a)

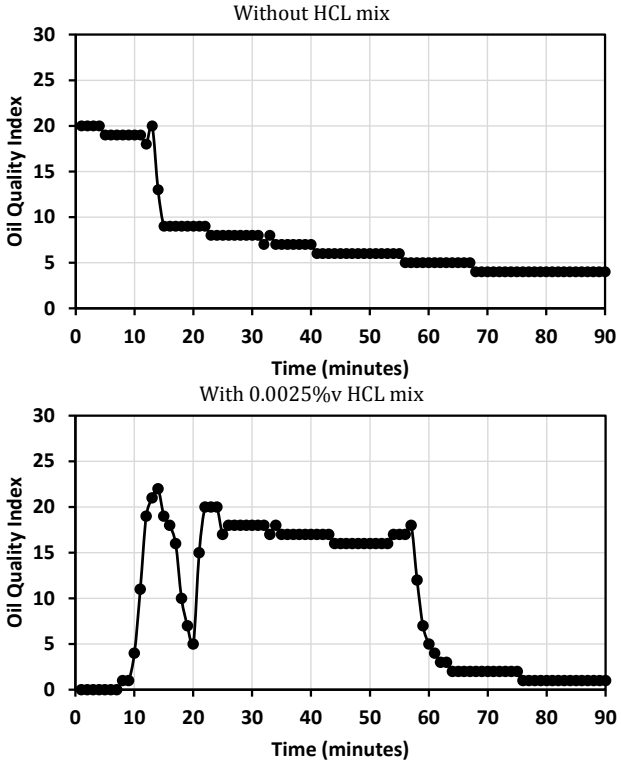


(b)

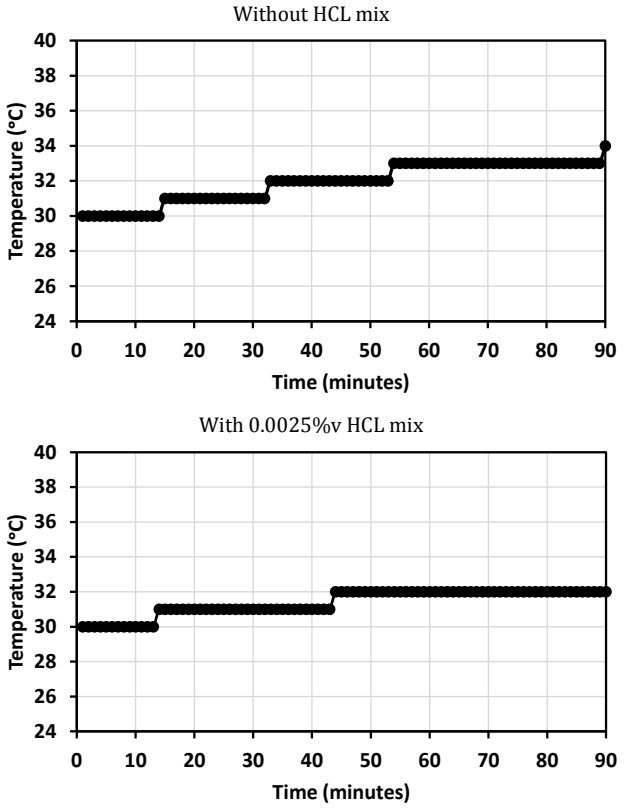
Fig. 5. Graph showing the Fe ppm and wear mass per hour variation in of the case study 1 for 90 minutes.

In all the three cases, it was observed that the system shows the progressive wear and no-run in stage was observed. In figure 5, the ferrous concentration (Fe ppm) and wear mass per hour have been plotted and the plot shows trend of increasing both the parameters with time. To confirm this the accelerated tests were planned and conducted.

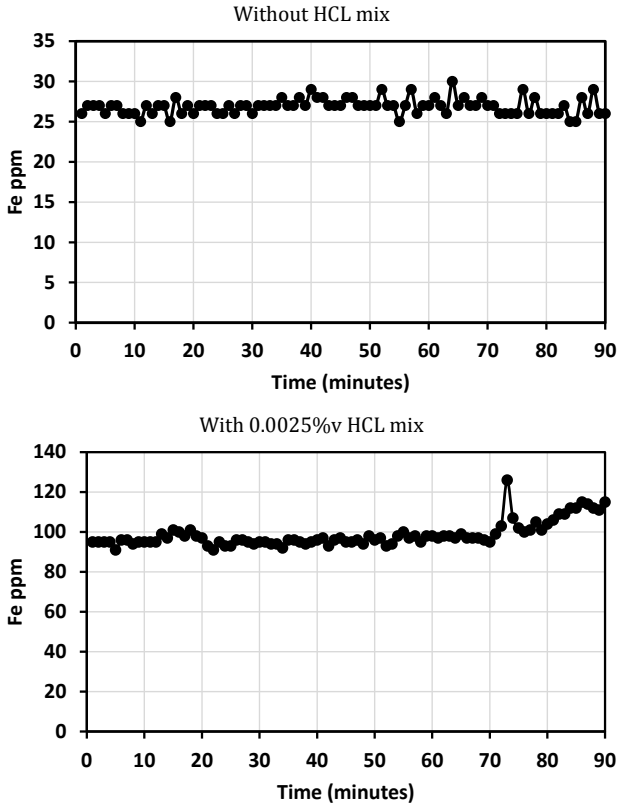
For online oil analysis, it is observed from figure 6 (a), that the oil quality index value is initially large and shows a decreasing behavior for both the tests (without degradation agent and with degradation agent). The oil quality index is the function of water ingress, TAN value and change in viscosity. As from the figure 6 (c), it can be observed that the relative humidity shows the increasing trend for lubricant oil the without HCL mix and saturated value of 100% for the with 0.0025%v HCL mix test. The saturated value indicate that further addition of the water will be free and form the emulsion in lubricant. In this situation two phenomena "viscosity decrease due to increased RH and viscosity increase due to oil oxidation happen simultaneously. To confirm the quality of the oil, the offline oil tests were conducted by measuring the pH and the viscosity of the lubricant oil. The results obtained from the offline measurement of the pH and viscosity is provided in the manuscript. The results show the deterioration of the oil quality as the kinematic viscosity decreases with the time for both the tests. From figure 6 (b), the increased trend of temperature was also observed.



(a)

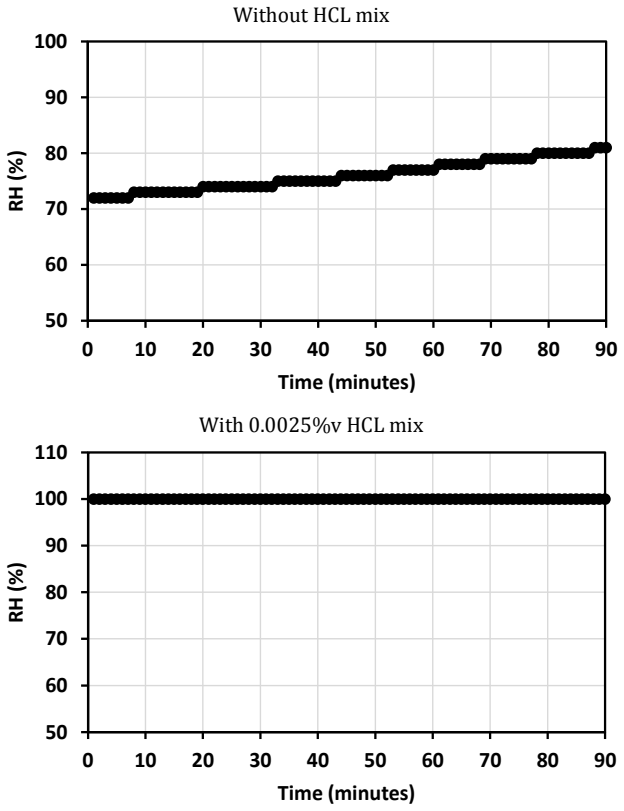


(b)

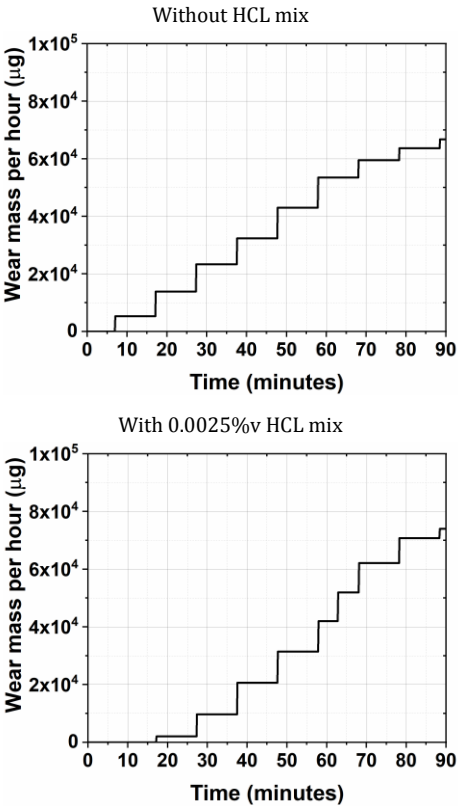


(a)

Fig. 6. Oil quality index, oil temperature, and relative humidity for normal lubrication and 0.0025% HCL mixed lubricant tests.



(c)

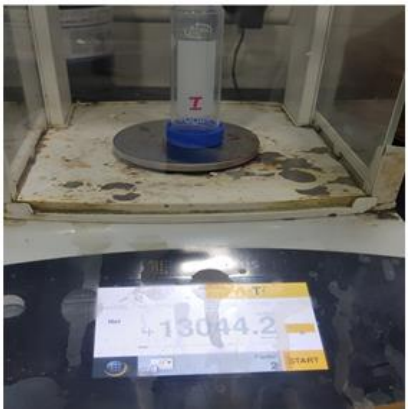


(b)

Fig. 7. Ferrous particles concentration and wear mass per hour for test-1 (without HCL mix) and test-2 (0.0025% HCL mixed).

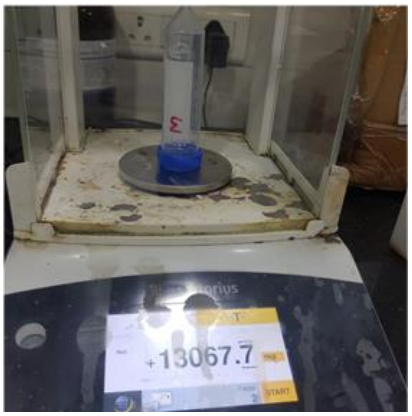
For online wear debris analysis, it is observed from Figure 7(a), that an increasing trend of Ferrous particle concentration (Fe ppm) for both the cases can be observed. The number of Fe ppm increased nearly four time the initial value from case 2 (without HCL mixed) to case 3 (with 0.0025%v HCL mixed), which confirms the accelerated wear in the system. From the Fig. 7(b), it was observed that the wear mass per hour shows an increase from 66769µg to 73969µg. To confirm the increase in wear mass the wear particles were extracted from the oil with the help of mechanical centrifuge at 7000 rpm and processed for the weighing. The measured values were found 138400 µg and 322400 µg for the without HCL and with 0.0025%v HCL (refer figure 8), which is 2.32 times. The wear debris sensor's measurement of the wear particle mass was below the level of the system's actual deterioration. It appears that detection rate of Fe sensors becomes a limitation in accelerated wear test (with 0.0025% v HCL).

Case 2
Weight of falcon tube without wear particles



Case 3
Weight of falcon tube with wear particles

Case 3
Weight of falcon tube without wear particles



Case 3
Weight of falcon tube with wear particles

Fig. 8. Graphs showing the wear mass of the extracted wear particles from the oil for case 2 and case 3.

It can be observed from figure 9, that the overall particles increased in oil with depletion agent test as compared to without accelerating agent mix. The particle size >100 µm was the dominating contributor in the wear of the gear in both the tests (case 2 and case 3).

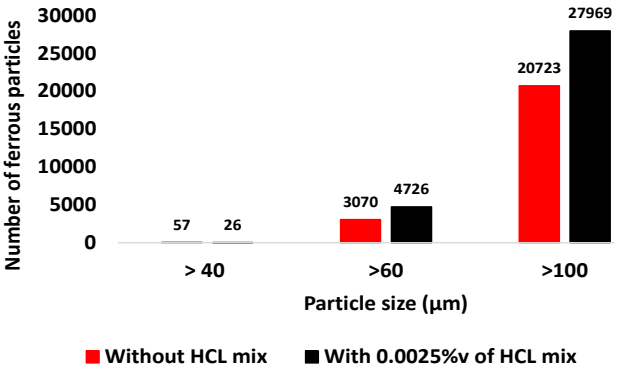


Fig. 9. Ferrous particle size distribution.

For offline analysis (pH and viscosity measurements), the periodic sampling (every 10 minutes) of the lubricating oil was done to check the quality of the lubricant for both cases 2&3. The pH and the viscosity of the oil was measured to check the quality of the oil. The decreasing trend of the pH value was observed in both the cases. The measured value for the case 2 (without HCl mix) decreased from 5.856 to 4.979 and for case 3 (with 0.0025%v HCL mix) decreased from 1.550 to -2.280 (the existence of the negative pH is confirmed from the article in ref.35).

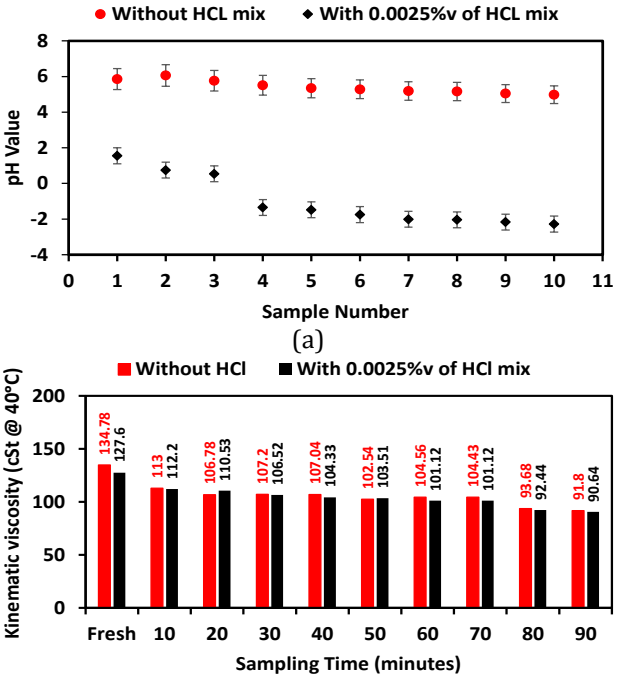


Fig. 10. pH value and kinematic viscosity (with and without degradation agent).

The decreasing trend of the pH value can be observed from fig. 10(a), the sample 1 indicate the sample of the oil before start of the test and sample 10 indicate the sample after 90 minutes of operation. From fig. 10(b), it can be inferred that the kinematic viscosity decreases for both the cases 2&3. Both pH and kinematic viscosity shows the decreasing trend.

4.3 Prediction of degradation parameters using Machine learning techniques

At first, the LR is applied to predict the Oil Quality from the dataset 1, and then it is applied on dataset 2 and dataset 3. Like other algorithm, the LR is used to predict the Oil Quality from the input parameters Time (minutes), Load (Nm), Speed (RPM), Oil Temp, and RH%.

The evaluating models are examined and reported the best model based on MAE, MSE, and RMSE, equations 1 to 3:

$$\boldsymbol{MAE} = \frac{\sum_{1=1}^{n} |\boldsymbol{yi} - \boldsymbol{xi}|}{\boldsymbol{n}} \quad (1)$$

$$\boldsymbol{MAE} = \frac{\sum_{1=1}^{n} |\boldsymbol{yi} - \boldsymbol{xi}|}{\boldsymbol{n}} \quad (2)$$

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{1=1}^{n} |\boldsymbol{xi} - \boldsymbol{yi}|^2|}{\boldsymbol{n}}} \quad (3)$$

In equations 1 to 3, y_i is the predicted value and x_i is the actual value of the parameter. The results of the estimated oil quality using the different algorithm are given in Table 6.

Table 6 reports the minimum MAE of the Oil Quality prediction in case of ANN as compared to the other models. As the dataset 1 contains a good number of samples, so almost all the models perform well in dataset 1. The RFR, KNNR, and DTR also produce same results with a minimum MAE of 2.16, 2.2, and 2.15, respectively. By comparing the values of MAE, MSE, and RMSE, the ANN model outperforms the other model in the case of dataset 1. The ANN model is implemented by considering the 2 hidden layers with Relu activation Function in each layer. With a batch size of 32, the input information is compiled for 100 epochs with early stopping call back to avoid the over fitting issue of the model. An Adam optimizer is used to optimize the model. The performance of the ANN model is further evaluated by calculating

the training and validation loss of the model for 100 epochs, Figure.11. The Figure 11 indicates the more errors in the model in the initial epochs but after that it is gradually decreasing with the epoch and learning process. The scatter plot association between the actual and anticipated values of oil quality is used to further assess the model's veracity.

Table 6. Results of the different algorithm on the 3 different datasets for oil quality prediction.

Algorithm	Predicted output	Dataset	MAE	MSE	RMSE
LR	Oil Quality	Dataset 1- CASE 1	3.09	19.78	4.44
DTR			2.15	13.19	3.63
KNNR			2.2	15.04	3.87
RFR			2.16	13.19	3.63
ANN			0.87	1.50	1.22
SVR			2.70	16.47	4.05
LR		Dataset 2- CASE 2	2.03	6.50	2.55
DTR			0.41	0.52	0.72
KNNR			0.42	0.52	0.72
RFR			0.41	0.52	0.72
ANN			7.2	66.53	8.15
SVR			2.95	28.71	5.35
LR		Dataset 3- CASE 3	7.60	64.82	8.05
DTR			5.45	53.81	7.33
KNNR			5.43	54.05	7.35
RFR			5.47	53.81	7.33
ANN			10.55	171.35	13.09
SVR			7.03	68.98	8.30

The visualization difference between the predicted and actual test data is presented in Figure 12.

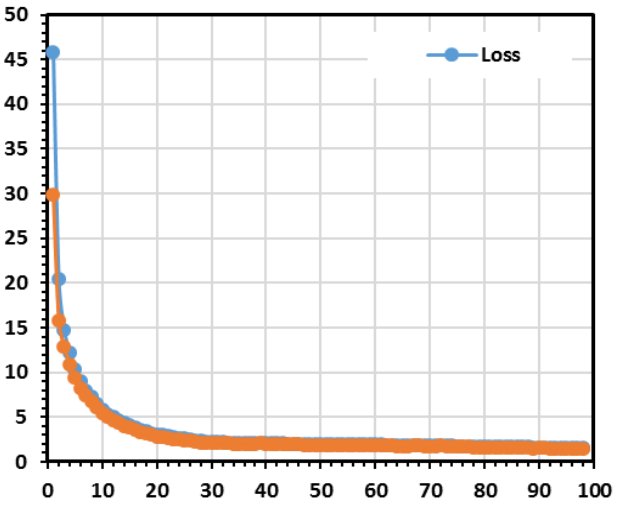


Fig. 11. Training and validation loss of the ANN model.

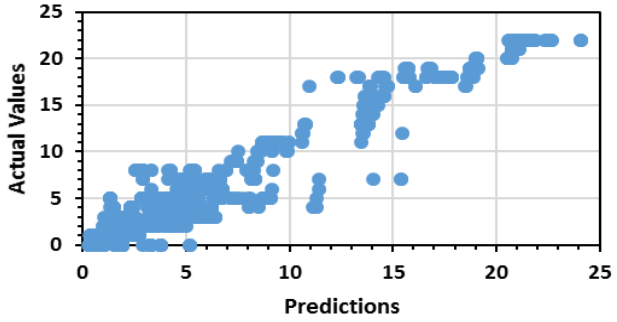


Fig. 12. Difference between the predicted and actual test data.

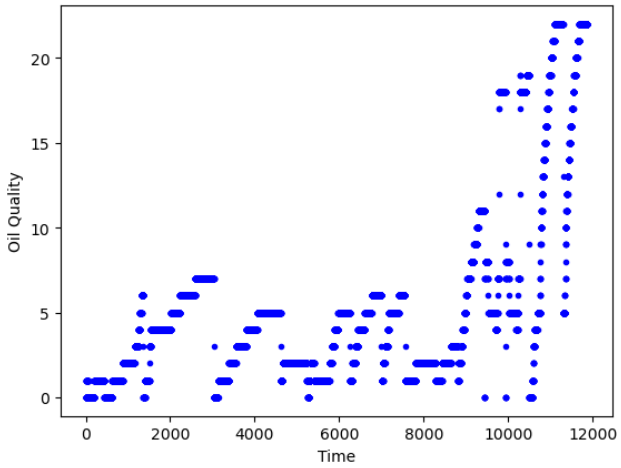


Fig. 13 Oil quality for dataset 1.

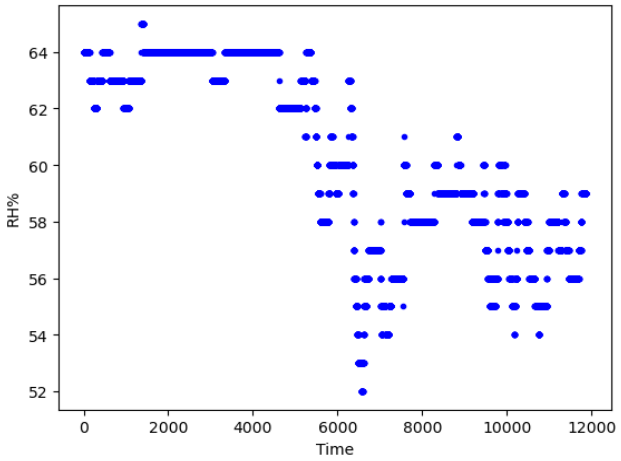


Fig. 14. RH % for dataset 1.

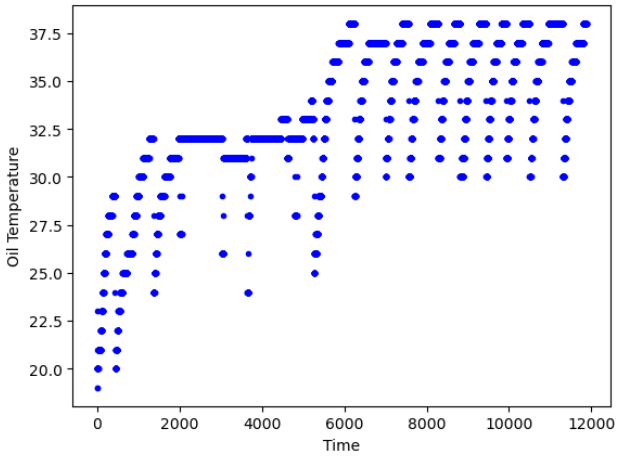


Fig. 15. Oil temperature variation dataset 1.

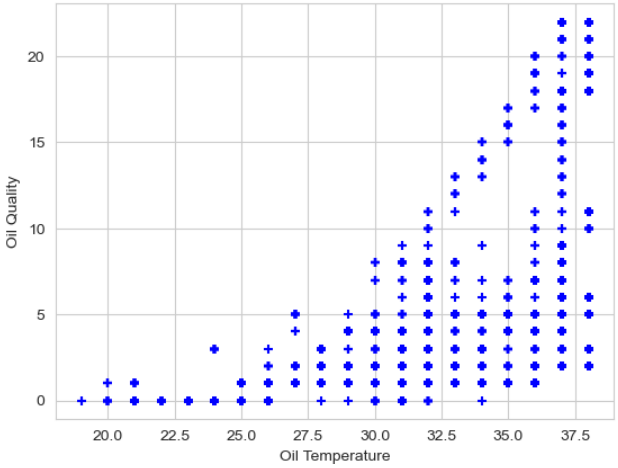


Fig. 16. Dependence of oil quality on oil temperature, dataset 1.

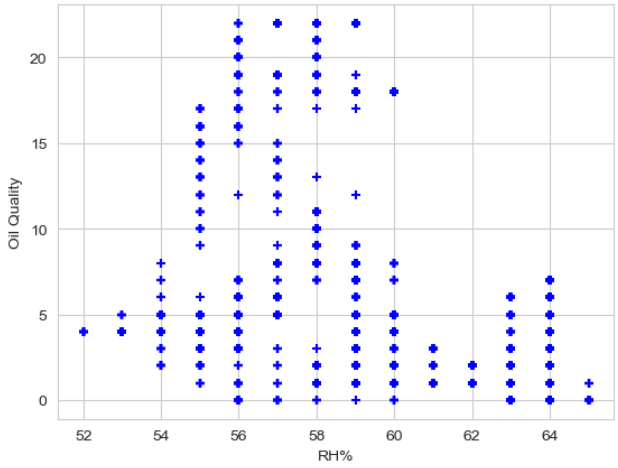


Fig. 17. Dependence of oil quality on humidity, dataset 1.

Based on the distribution of the datasets1 as visualized in the figures 13 to 19, different machine learning models like LR, DT, KNN, RF, ANN, NB etc. is utilized to make prediction of Oil Quality over certain period of time. In the conducted experiment it is observed that models like DT, KNN, RF and ANN is performing very well in terms of accuracy and prediction whereas performance of all other algorithms are not so well for the provided dataset as given in Table 7. Based on that conclusion is made that ML algorithm like DT, KNN and RF perform very well with the distributed dataset.

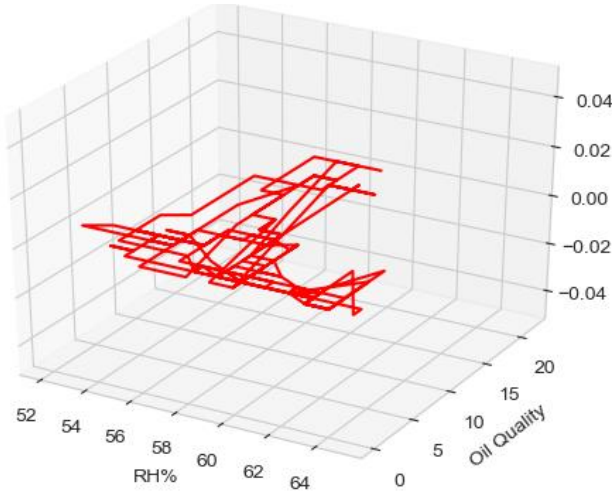


Fig. 18. 3D data distribution of oil quality and humidity, dataset 1.

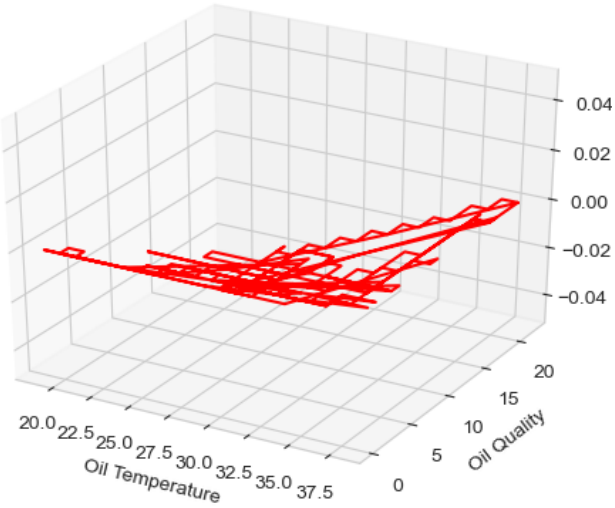


Fig. 19. 3D data distribution of oil temperature and oil quality, dataset 1.

Table 7. Accuracy of oil quality prediction.

Algorithm	Predicted Output	Dataset	Accuracy
LR	Oil Quality	Dataset 1-CASE 1	25.39%
DTR			100%
KNNR			97.03%
RFR			99.03%
ANN			64%
Naive bayes			26.12%

Based on the conclusion in the first experiment only DT, KNN and RF algorithm is utilized for the Dataset2 and Dataset3 for the prediction and it is observed that conclusion is still valid as all three algorithm is performing very well with the dataset2 and dataset3.

Overall conclusion from the experiment is that algorithm like DT, KNN and RF work well with widely distributed dataset.

Like case 1, the DTR, KNNR, and RFR reported almost same MAE. The DTR and RFR report 0.41 MAE whereas the KNNR reported the 0.42. The other three algorithms (LR, ANN, and SVR) showed 2.03, 7.2, and 2.95, respectively. To find the best suitable model for Oil Quality in the case of dataset 2, the accuracy of the DTR, KNNR, and RFR are also calculated in this study, given in Table 8, and it is reported that the DTR showed the best results with an accuracy of 98.41%.

Table 8. Accuracy of DTR, KNNR, and RFR for the prediction of oil quality in dataset 2.

Algorithm	Dataset	Accuracy
DTR	Dataset 2	0.9841
KNNR		0.9840
RFR		0.9840

The DTR, KNNR, and RFR also reported almost same results in the case of case 3, i.e., dataset 3. The MAE of the DTR, KNNR, and RFR is 5.45, 5.43, and 5.47, respectively. The MSE of the DTR, KNNR, and RFR is 53.81, 54.05, and 53.81, respectively.

Like Oil Quality prediction, the different machine learning algorithms, by maintaining the same hyper parameter, are reported to predict the Oil FE PPM. Same evaluating parameters as defined in Table 4 are also reported to analysis and choose the best machine learning model for FE PPM, as given in Table 9.

Table 9. Results of the different algorithm on the 3 different datasets for Fe PPM prediction.

Algorithm	Predicted Output	Dataset	MAE	MSE	RMSE
LR	Oil Fe PPM	Dataset 1- CASE 1	8.64	122.02	11.04
DTR			6.35	80.70	8.98
KNNR			6.92	95.27	9.76
RFR			6.35	80.70	8.98
ANN			7.87	106.60	10.32
SVR			9.19	154.82	12.44
LR		Dataset 2- CASE 2	0.70	0.85	0.92
DTR			0.75	0.90	0.95
KNNR			0.72	0.89	0.94
RFR			0.75	0.90	0.95
ANN			1.12	1.8	1.34
SVR			0.82	1.10	1.05
LR		Dataset 3- CASE 3	3.42	21.08	4.59
DTR			3.23	19.93	4.46
KNNR			2.97	21.69	4.65
RFR			3.21	19.93	4.46
ANN			3.57	28.11	5.30
SVR			3.05	28.53	5.34

Table 9, indicates that the DTR and RFR reported the best result of oil FE PPM prediction with a MAE of 6.35 respectively the case of dataset 1. The models reported the same MSE (80.70) for the FE PPM prediction. In the case 2, the LR reported the best results for prediction with a MAE of 0.70, MSE of 0.72 but the KNNR, DTR, and SVR also reported a very good results with an MAE of 0.75,0.72, and 0.75. The outperformed model LR showed 0.85 MSE with a RMSE value of 0.92. In the case 3, i.e., the KNNR model reported 2.95 MAE with a MSE of 21.69 and RMSE of 4.65. Along with the KNNR, the other model reported the almost the same kind of results for the prediction of oil FE PPM. From the above analysis it is found that DTR, LR, and KNNR works efficiently in the case 2 and case 3 as these datasets have limited number of samples.

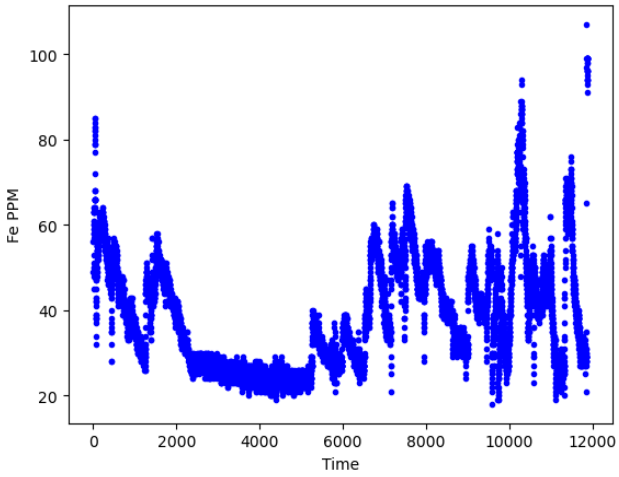


Fig. 20. Variation of Fe ppm, dataset 1.

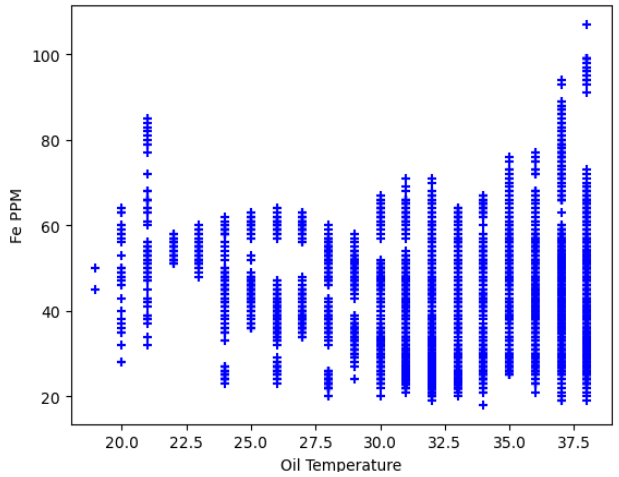


Fig. 21. Variation of Fe ppm with oil temperature, dataset 1.

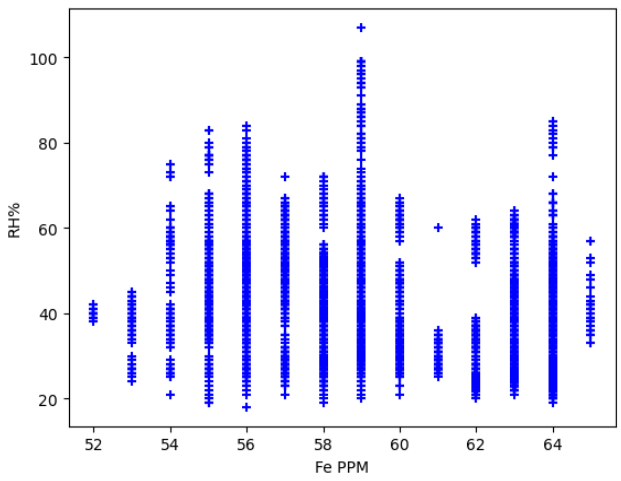


Fig. 22. Variation of Fe ppm with humidity, dataset 1.

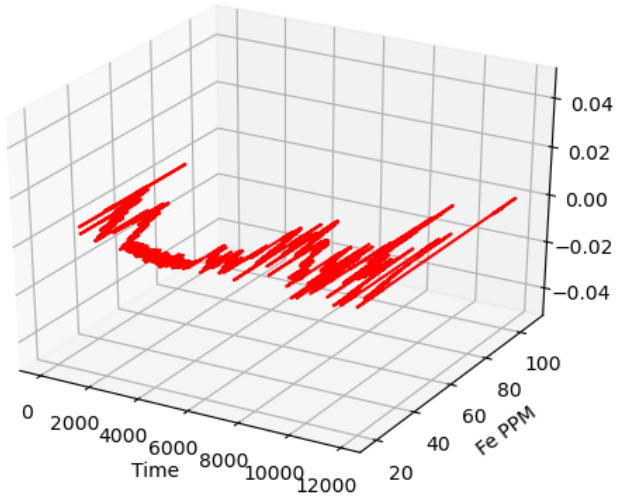


Fig. 23 3D data distribution of Fe ppm, dataset 1.

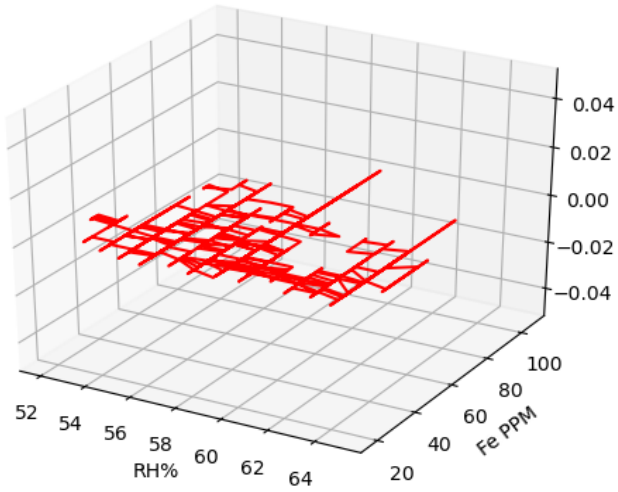


Fig. 24 3D data distribution of Fe ppm and humidity, dataset 1.

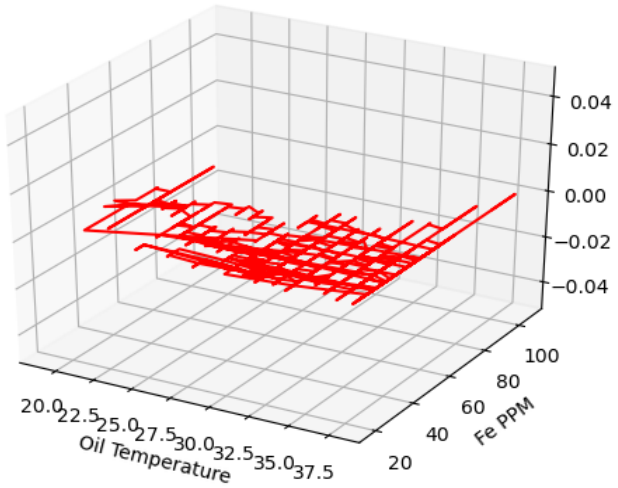


Fig. 25 3D data distribution of Fe ppm and oil temperature, dataset 1.

Based on the distribution of the datasets1 as visualized in the figures 20 to 25, different machine learning models like LR, DT, KNN, RF, ANN, NB etc. is utilized to make prediction of Fe ppm over certain period of time. In the conducted experiment it is observed that models like DT, KNN, RF and ANN is performing very well in terms of accuracy and prediction whereas performance of all other

algorithms are not so well for the provided dataset as given in Table 10. Based on that conclusion is made that ML algorithm like DT, KNN and RF perform very well with the distributed dataset.

Table 10. Accuracy of DTR, KNNR, and RFR for the prediction of Fe ppm in dataset 1.

Algorithm	Predicted Output	Dataset	Accuracy
LR	Oil Quality	Dataset 1-CASE 1	6.39%
DTR			100%
KNNR			98.7%
RFR			96.24%
ANN			56%
Naive bayes			79.6%

Many a times the data available is insufficient, therefore under these conditions DTR method is used, which even though is computationally intensive, but, is able to give a better prediction. The LR method is extremely poor and is thus not suitable to use.

5. CONCLUSION

In the present work, the effect of oil degradation was studied. The oil degradation plays an important role in the failure of gears. The following conclusions are drawn from the above work:

- In all three cases studied in the present study, a progressive increase in the concentration of the ferrous debris was observed. The running-in stage was missing.
- Two different oils and three batches of oils were used in the experiments. In each pH value was lesser than 7.0 and one case fresh oil pH was 5.762 (acidic oil).
- It was observed that the lubricant oil becomes more acidic as the ageing of the oil take place.
- The accelerated tests conducted in cases 2 & 3 confirmed the progressive wear of the gear.
- The wear debris monitoring sensor provides reasonably accurate data, but in accelerated test it provides only the trend. For more reliable results, the wear debris need to be extracted from the lubricating oil using the centrifuging action.
- The pH value monitoring will be the better indicator of the oil deterioration.
- The ML algorithms like DT, KNN, RF and ANN performs very well in terms of accuracy and prediction whereas performance of all other algorithms are not so well for the type of dataset obtained from experiments.
- The ML algorithms like DT, KNN and RF perform very well with the dataset which is a distributed dataset.
- Whenever the collection of large amounts of data is not possible under these conditions the use of LR method is not recommended as it does not give correct prediction.
- Under scarce data conditions, DT method is used, which even though is computationally intensive, but, is able to give a better prediction.

REFERENCES

[1] American National Standard, Industrial Gear Lubrication, ANSI/AGMA 9005- F16, American Gear Manufacturers Association, Alexandria, Virginia, 2016.

[2] N.N. Gosvami, J.A. Bares, F. Mangolini, A.R. Konicek, D.G. Yablon, R.W. Carpick, *Mechanisms of antiwear tribofilm growth revealed in situ by single-asperity sliding contacts*, Science, vol. 348, iss. 6230, pp. 102–106, 2015, doi: 10.1126/science.1258788

[3] I. Minami, *Molecular science of lubricant additives*, Applied Sciences, vol. 7, iss. 5, 2017, doi: 10.3390/app7050445

[4] S.M. Muzakkir, K.P. Lijesh, H. Hirani, G.D. Thakre, *Effect of cylindricity on the tribological performance of the heavily loaded slow speed journal bearing*, Journal of Engineering Tribology, vol. 229, iss. 2, pp. 178–195, 2014, doi: 10.1177/1350650114548053

[5] K.P. Lijesh, S.M. Muzakkir, H. Hirani, *Experimental tribological performance evaluation of nano lubricant using multi-walled carbon nano-tubes (MWCNT)*, International Journal of Applied Engineering Research, vol. 10, no. 6, pp. 14543-14551, 2015.

[6] S.M. Muzakkir, H. Hirani, G.D. Thakre, *Lubricant for Heavily Loaded Slow-Speed Journal Bearing*, Tribology Transactions, vol. 56, iss. 6, pp. 1060–1068, 2013, doi: 10.1080/10402004.2013.823530

[7] H. Hirani, K. Athre, S. Biswas, *Rapid and globally convergent method for dynamically loaded journal bearing design*, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, vol. 212, iss. 3, pp. 207–213, 1998, doi: 10.1243/1350650981542010

[8] C. Sarkar, H. Hirani, *Development of a magnetorheological brake with a slotted disc*, Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, vol. 229, iss. 14, pp. 1907–1924, 2015, doi: 10.1177/0954407015574204

[9] H. Hirani, *Multiobjective optimization of journal bearing using mass conserving and genetic algorithms*, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, vol. 219, iss. 3, pp. 235–248, 2005, doi: 10.1243/135065005X9844

[10] S.S. Goilkar, H. Hirani, *Design and development of a test setup for online wear monitoring of mechanical face seals using a torque sensor*, Tribology Transactions, vol. 52, iss. 1, pp. 47–58, 2009, doi: 10.1080/10402000802163017

[11] H. Hirani, K. Athre, S. Biswas, *A Hybrid Solution Scheme for Performance Evaluation of Crankshaft Bearings*, Journal of Tribology, vol. 122, iss. 4, pp. 733–740, 2000, doi: 10.1115/1.1286271

[12] H. Hirani, K. Athre, S. Biswas, *Lubricant shear thinning analysis of engine journal bearings*, Tribology Transactions, vol. 44, iss. 1, pp. 125–131, 2001, doi: 10.1080/10402000108982435

[13] H. Hirani, K. Athre, S. Biswas, *Comprehensive design methodology for an engine journal bearing*, Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, vol. 214, iss. 4, pp. 401–412, 2000, doi: 10.1243/1350650001543287

[14] S.S. Goilkar, H. Hirani, *Parametric study on balance ratio of mechanical face seal in steam environment*, Tribology International, vol. 43, iss. 5–6, pp. 1180–1185, 2010, doi: 10.1016/j.triboint.2009.12.019

[15] H. Hirani, K. Athre, S. Biswas, *Dynamic analysis of engine bearings*, International Journal of Rotating Machinery, vol. 5, iss. 4, pp. 283–293, 1999, doi: 10.1155/S1023621X99000251

[16] H. Hirani, K. Athre, S. Biswas, *A Simplified Mass Conserving Algorithm for Journal Bearing under Large Dynamic Loads*, International Journal of Rotating Machinery, vol. 7, iss. 1, pp. 41–51, 2001, doi: 10.1155/S1023621X01000045

[17] S. Ebersbach, Z. Peng, N.J. Kessissoglou, *The investigation of the condition and faults of a spur gearbox using vibration and wear debris analysis techniques*, Wear, vol. 260, iss. 1–2, pp. 16–24, 2006, doi: 10.1016/j.wear.2004.12.028

[18] C.K. Tan, P. Irving, D. Mba, *A comparative experimental study on the diagnostic and prognostic capabilities of acoustics emission, vibration and spectrometric oil analysis for spur gears*, Mechanical Systems and Signal Processing, vol. 21, iss. 1, pp. 208–233, 2007, doi: 10.1016/j.ymssp.2005.09.015

[19] T.H. Loutas, D. Roulias, E. Pauly, V. Kostopoulos, *The combined use of vibration, acoustic emission and oil debris on-line monitoring towards a more effective condition monitoring of rotating machinery*, Mechanical Systems and Signal Processing, vol. 25, iss. 4, pp. 1339–1352, 2011, doi: 10.1016/j.ymssp.2010.11.007

[20] P.J. Dempsey, *Integrating oil debris and vibration measurements for intelligent machine health monitoring,* Creative Media Partners, LLC, 2013.

[21] P. Kumar, H. Hirani, A. K. Agrawal, *Online condition monitoring of misaligned meshing gears using wear debris and oil quality sensors*, Ind. Lubr. Tribol., vol. 70, no. 4, pp. 645–655, 2018, doi: 10.1108/ILT-05-2016-0106

[22] H.S. Dhiman, D. Deb, J. Carroll, V. Muresan, M.L. Unguresan, *Wind turbine gearbox condition monitoring based on class of support vector regression models and residual analysis*, Sensors, vol. 20, iss. 23, pp. 1–17, 2020, doi: 10.3390/s20236742

[23] X.J. Zeng, M. Yang, Y.F. Bo, *Gearbox oil temperature anomaly detection for wind turbine based on sparse Bayesian probability estimation*, International Journal of Electrical Power & Energy Systems, vol. 123, 2020, doi: 10.1016/j.ijepes.2020.106233

[24] K.L. de Calle, S. Ferreiro, C. Roldán-Paraponiaris, A. Ulazia, *A context-aware oil debris-based health indicator for wind turbine gearbox condition monitoring*, Energies, vol. 12, iss. 17, pp. 1–19, 2019, doi: 10.3390/en12173373

[25] H. Liu, C. Yu, C. Yu, *A new hybrid model based on secondary decomposition, reinforcement learning and SRU network for wind turbine gearbox oil temperature forecasting*, Measurement, vol. 178, 2021, doi: 10.1016/j.measurement.2021.109347

[26] J. Fu, J. Chu, P. Guo, Z. Chen, *Condition Monitoring of Wind Turbine Gearbox Bearing Based on Deep Learning Model*, IEEE Access, vol. 7, pp. 57078–57087, 2019, doi: 10.1109/ACCESS.2019.2912621

[27] S.R. Saufi, Z.A. Bin Ahmad, M.S. Leong, M.H. Lim, *Gearbox Fault Diagnosis Using a Deep Learning Model with Limited Data Sample*, IEEE Transactions on Industrial Informatics, vol. 16, iss. 10, pp. 6263–6271, 2020, doi: 10.1109/TII.2020.2967822

[28] C. Sarkar, H. Hirani, *Effect of Particle Size on Shear Stress of Magnetorheological Fluids*, Smart Science, vol. 3, iss. 2, pp. 65–73, 2015, doi: 10.1080/23080477.2015.11665638

[29] H. Hirani, C.S. Manjunatha, *Performance evaluation of a magnetorheological fluid variable valve*, Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, vol. 221, iss. 1, pp. 83–93, 2007, doi: 10.1243/09544070JAUTO408

[30] K.P. Lijesh, H. Hirani, *Magnetic bearing using rotation magnetized direction configuration*, Journal of Tribology, vol. 137, iss. 4, 2015, doi: 10.1115/1.4030344

[31] K.P. Lijesh, S.M. Muzakkir, H. Hirani, *Failure mode and effect analysis of passive magnetic bearing*, Engineering Failure Analysis, vol. 62, pp. 1–20, 2016, doi: 10.1016/j.engfailanal.2015.11.033

[32] K.P. Lijesh, H. Hirani, *Design and development of Halbach electromagnet for active magnetic bearing*, Progress In Electromagnetics Research C, vol. 56, pp. 173–181, 2015, doi: 10.2528/PIERC15011411

[33] K.P. Lijesh, H. Hirani, *Modeling and development of RMD configuration magnetic bearing*, Tribology in Industry, vol. 37, no. 2, pp. 225–235, 2015

[34] H. Hirani, *Root cause failure analysis of outer ring fracture of four-row cylindrical roller bearing*, Tribology Transactions, vol. 52, iss. 2, pp. 180–190, 2009, doi: 10.1080/10402000802180151

[35] A.Y. El-Naggar, R.A. El-Adly, T.A. Altalhi, A. Alhadhrami, F. Modather, M.A. Ebiad, A. Salem, *Oxidation stability of lubricating base oils*, Petroleum Science and Technology, vol. 36, no. 3, pp. 179–185, 2018, doi: 10.1080/10916466.2017.1403450

[36] M.R. Jain, R. Sawant, R.D.A. Paulmer, D. Ganguli, G. Vasudev, *Evaluation of thermo-oxidative characteristics of gear oils by different techniques: Effect of antioxidant chemistry*, Thermochimica Acta, vol. 435, iss. 2, pp. 172–175, 2005, doi: 10.1016/j.tca.2005.05.016

[37] A. Wolak, *Changes in Lubricant Properties of Used Synthetic Oils Based on the Total Acid Number*, Measurement and Control, vol. 51, iss. 3–4, pp. 65–72, 2018, doi: 10.1177/0020294018770916

[38] B.L. De Rivas, J.L. Vivancos, J. Ordieres-Meré, S.F. Capuz-Rizo, *Determination of the total acid number (TAN) of used mineral oils in aviation engines by FTIR using regression models*, Chemometrics and Intelligent Laboratory Systems, vol. 160, pp. 32–39, 2017, doi: 10.1016/j.chemolab.2016.10.015

[39] V. Inturi, S.V. Balaji, P. Gyanam, B.P. Pragada, R. S. Geetha, V. Pakrashi, *An integrated condition monitoring scheme for health state identification of a multi-stage gearbox through Hurst exponent estimates*, Structural Health Monitoring, vol. 22, iss. 1, pp. 730–745, 2022, doi: 10.1177/14759217221092828