IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 10, Issue 9, September 2020, ||Series -I|| PP 56-63

Predictive Maintenance of Mining Haul Truck Engines Using Oil Sampling and Telemetry Data

Bharadwaj Madhu[1], Guilherme Santos[2], Mubaroq Ilham[3], Kuncoro Teguh Dwi[3], Fahlefi Ali[4], Bansal Mayur[1], Edson Antonio[4], Malcolm Mcrae[4], Sandro Cerilli[4]

1 Infosys Ltd, No. 350, Hebbal Electronics City, Mysore 570 027, India.
2 IBM Advanced Analytics, Sao Paulo, Brazil.
3 Plant Site DP. 15, Sorowako, South Sulawesi, 92984, Indonesia.
4 18 Rink Street, Copper Cliff, P0M1N0, Sudbury, Canada.
Corresponding Author: Madhu Bharadwaj
Received 15 September 2020; Accepted 30 September 2020

Abstract: Predictive Maintenance approach can play a very important role in reducing the operating costs of mining haul truck engines by extending their Mean Time Between Failure (MTBF). Early prediction of failures also reduces the risk of production impact and enables maintenance team to take corrective actions ahead of time. Machine learning techniques can be used to develop a model using oil sample analysis data, telemetry data and historical replacement data to predict failure window of truck engines. This model considers the threshold of trespassing variables and combines them to form a multivariable analysis. Life prediction is achieved by using a tumbling window approach, where incremental models are created by slicing the component lifetime into number of equal intervals based on the available failure data and using classification algorithm for each window to predict failure. The remaining life is calculated by subtracting the current running hours from the predicted lifetime to failure. Out of 7 replacements that happened during testing, the model could accurately predict 4 out of 5 planned (end of life) replacements and 1 out of 2 premature (breakdown) failure successfully.
Keywords: Predictive Maintenance, Haul Truck Engine, Machine Learning, remaining life prediction, Oil Sample Analysis, Telemetry data.

I. INTRODUCTION

Ore movement in open pit mines is done primarily using heavy equipment like haul trucks. These haul trucks have huge power engine systems and unexpected break downs can hamper the shift production targets. Hence their maintenance is extremely important to reliably meet the production targets and maintain a safe working environment. The operating costs of these trucks are approximately USD 550 per hour per truck, so extending their Mean Time Between Failure (MTBF) can help to bring down the operating costs and maximize the profitability. In addition to this, the repair and maintenance costs are also huge [1] and any unplanned maintenance event may lead to increasing this cost by up to 30%. Thus, having a rigorous condition monitoring system in place is required to take necessary preventative measures before the breakdown happens.

One important health indicator used by maintenance department is to take periodic oil samples from engine and perform lab analysis to check the condition of the oil for various metal content and contamination to gauge the amount of wear in the engine. This process can help to replace oil and extend the life to some extent. Oil analysis process typically takes few minutes for creating a report which provide details about the wear material content, but major time is lost in the shipment of samples and report back and forth. Thus, it is common that, by the time the maintenance supervisor receives a report indicating a problem, the equipment might have already failed. This situation necessitates the proactive condition monitoring rather than reactive monitoring based on oil analysis report. The same conclusion was arrived by Heng [2] in his thesis where he reported that more than half of the maintenance initiatives are ineffective. In this regard, machine-learning algorithm based predictive models can enable forecasting the probable engines that could fail, well in advance. It is achieved by creating a baseline assessment, which when correlated with the experience of maintenance team helps to take appropriate action before failure happens. The successful implementation of this ML based approach can also help in extending lifetime of engines by doing proactive maintenance.

With the advent of Internet of Things (IoT), many companies have also started to capture real-time machine telemetry data which when combined with oil sample analysis data can be used to train the model for better prediction of failure and improve the accuracy of model. This has resulted in an increase of automating the costly and time-consuming traditional methodologies in classifying the health of machines. Carstens et al. [3] has suggested prognostic model approach based on the data from the mining industry. A similar approach for attempted by OlcayAltıntas et al. [4] where they used artificial neural network for maintenance of locomotive engine using engine lube oil. Maria Grazia et al., [5] has attempted to machine learning based techniques to show degradation effects on the engine performance with the aim of predicting the health of aero engine.

Similar approach was used to develop a predictive model using engine oil samples and telemetry data along with historical failure data to predict failure window of haul truck engines.

II. MODEL DEVELOPMENT AND VALIDATION

A. Objectives and Inputs

The main objectives of the current work include:

a. Increase the engine life from 23K to 26K hours
b. Predict premature failure to reduce the downtime and maintenance cost
c. Prioritize component replacement to maximize the useful life and uptime simultaneously
d. Provide early alert for part procurement – minimize inventory of parts

There are 3 types of inputs used for the Machine Learning (ML) algorithm, viz., Scheduled Oil Sample (SOS) Analysis report, Telemetry data collected by onboard vehicle sensors during operation and log of scheduled & unplanned replacements (breakdown). Model uses a combination of SOS data (Oil Analysis) and Telemetry data-based hybrid ML models to predict damage (Failure) even before actual damage occurs in the part. This algorithm is designed to let the owner know probable problem area and alerts them almost 1000 hours before the failure.

Scheduled Oil Sample (SOS) Analysis report

During the operation of engine, the wear debris are continuously produced which can be found by sampling the engine oil. Wear element chemistry forms a valuable source of information about wear mechanism and source of origin. According to the existing established practice, the engine oil is sampled every 250 hours to observe oil conditions through a Spectral Analyzer (SA) device. The SA device uses spectral analysis method to determine the amount of different chemicals in the engine oil. The common oil analysis parameters for engines that are used in the current model as input for machine to learn are given in Table 1

Table 1: Oil Analysis Parameters [6]

Category	Parameters	Chemicals
Wear	Elemental concentration	Al, Fe, Cu, Pb,Sn, Cr, Ni
Fluid contamination	Water	H_2O
	Fuel Soot	ST
	Coolant	Na, K, Si
Oil Degeneration	Viscosity at 100C	V100
	Oxidation, Nitration, Sulfation	OXI, SUL,NIT
Additives		Ca, Zn, Mg, Mo, B

Wear element monitoring provides unequivocal signals of changes in machine health. Data from SA generally shows a steady increase in wear metal concentration over time, giving a little or no warning for undesirable changes in wear rates. Added to this complexity, the concentration of wear particles will be affected by an oil change. A schematic diagram of wear particle concentration with an oil change is shown Figure 1

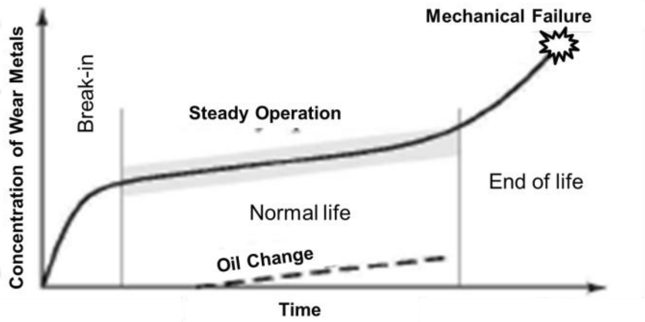

Figure 1:*Schematic illustration of changes in wear particle concentration [7]*

In reality, the above classical way of accumulation of wear partial never happens. Exploratory data analysis has indicated that there is a non-linear interaction between various chemical elements happening continuously. This complexity demands usage of ML algorithm to predict machine health.

Telemetry data

The other data used in the model is telemetry. Telemetry is an automated communication by which measurements are made and transmitted remotely to data collection points. This data has a wealth of information on abuse of engine, slow degradation, instantaneous efficiency etc. usage of this information in conjunction with SOS data provides better accuracy.

The telemetry data obtained from onboard vehicle sensors get stored in a Historian database called OSI PI server®. Table 2 provides different sensors related to the engine state at running time. These time series data are converted into alarms which are in binary (1's and 0's) format indicating whether the alarm is being flagged or not respectively. The definition for different alarms is provided in Table 2. Normally the data is contaminated with sensor noise and hence an alarm is considered as a legitimate alarm if the sensor value is beyond threshold value for more than 3 sec.

Table 2: Telemetric Data and its Definition

Alarm Name	**Definition**
Engine Over Revolution	Value = 1; When Engine Speed >= 2200 rpm
Engine Over-Heat	Value = 1; When Engine Coolant Temperature >= 102 ^{o}C
Engine Low Coolant Temperature	Value = 1; When Engine Coolant Temperature <= 75 ^{o}C
High Engine Boost Pressure	Value = 1; When Boost Pressure >= 345 kPa
Low Engine Boost Pressure	Value = 1; When Boost Pressure <= 96 kPa
Low Engine Oil Pressure	Value = 1; When Engine Oil Pressure <= 172 kPa
High Crankcase Pressure	Value = 1; When Crankcase Pressure >= 34 kPa
Inlet Air Restriction	Value = 1; When Air Filter Restriction > 7 kPa
High Left Exhaust Temperature	Value = 1; When Left Exhaust Temperature >= 750^{o}C
High Right Exhaust Temperature	Value = 1; When Right Exhaust Temperature >= 750 ^{o}C

In addition to sensor noise sometimes data is missing due to various reasons. This can be addressed by data wrangling techniques like forward filling of missing data.

Replacements log

One of the most important input to the ML model is a log of planned & unplanned replacements (breakdown). This data contains details like truck number, running hours at failure or replacement. The ML model uses this data to learn and calibrate itself for future prediction. Hence, accuracy of data is paramount. It is also important to have a dataset that has failure data points spread across different failure modes throughout the lifetime of engine. As more failure events are fed to the algorithm, better equipped the model becomes to predict future failures.

B. Data Preparation and Model Development

Data from all the 3 sources is combined to prepare the data for training the model. Each historical replacement is considered as a primary key to aggregate the oil sample and telemetry data. All the oil sample records for the data ranges when the engine was in operation till the replacement date are tagged by the key for that replacement. Count of telemetry alarms between 2 sequential oil sample data range is taken and added to this dataset. The telemetry alarms counts are then converted to alarm rate by dividing the count by difference in running hours between 2 oil samples.

Once the data is prepared in this format, additional statistical features are added for each of the existing features. In this process the first oil sample is discarded as it does not have any prior historical information. For the second oil sample onwards, historical data from previous oil samples are added. The features added in this process are basic statistics information like sum, mean, standard deviation and variation. In this way the model considers other features in addition to thresholds to form a multivariable analysis which leads to a better AI model.

Since failures can happen at different stages of component usage and root cause of each failure is different and hence a single classification model may not be very effective to predict the failures and remaining useful life. A single classification model may not be able to differentiate different failure modes and may only be able to indicate failure probability at the very best. To address this issue a tumbling window approach is used where incremental models are created by slicing the component lifetime into N number of equal intervals based on the available failure data.

The model is trained with the XGBClassifier classification algorithm with a default threshold value of 0.5 which indicates that any result where probability of Premature failure is >=50% will indicate the transition from Normal Failure to Premature Failure class. The window at which this threshold value is crossed is considered as the total life of engine.

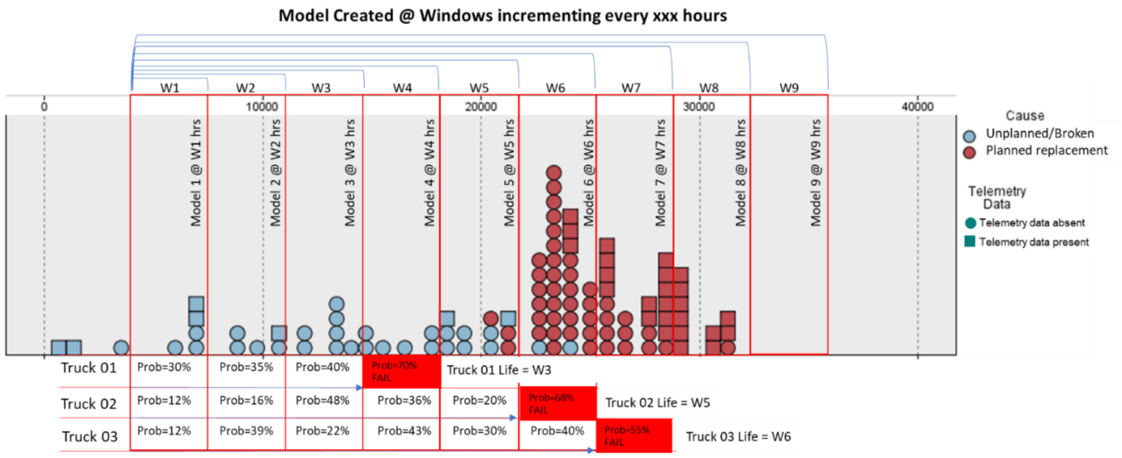

Figure 2:*Labelling & Logic for Tumbling Windows*

For example, if the maximum life of an engine is 28000 hours and the lifespan is divided into 500 hours intervals, then there will be 56 models created in increments of 500 hours. i.e., 500, 1000, 1500 hours till 28000 hours. The model is evaluated in each of the window to check for the window where the threshold value is breached. Say for a Truck 01 engine, this breach happens at window "W4", then the life of that engine is stated as window "W4" hours. Similarly for Truck 02 engine it will be W6 and so on. This logic is illustrated in Figure 2.

C. Model Validation

The oil data used for training the model was from 02-Jan-2014 to 16-Oct-2019, when this tool was rolled out for validation. The telemetry systems were deployed later, hence the telemetry data was available only from 11-Nov-2018 to 16-Oct-2019. The replacement data that has been incorporated was from 27-Dec-2013. Hence it was decided whenever a new replacement data is added, the model will be retrained to increase the accuracy by improving the overlap of all 3 datasets viz, SOS Analysis report, Telemetry data and Replacements log. The active window considered for the model prediction is considered from 5,000 hours to 23,000 hours. Any predictions before 5,000 hours is neglected as the warranty of the machines would take care of any failures. The time 23,000 hours is considered as the end of life for the engines and are normally replaced based on availability of replacement parts and labor. Since one of the key goals is to extend the current engine life from 23,000 to 26,000 hours, the performance of machines in this time period is studied separately. Failure of any engine before 10,000 hours is considered as premature failure. Thus, the performance of the model is studied separately for engines before and after 10,000 hours. Figure 3 shows the distribution of failure over time in use. Here it can be deduced that there are 150 failures or strategic/planned (end of life) replacements data points, out of which 32 failure points had telemetry data for at least one year before failure.

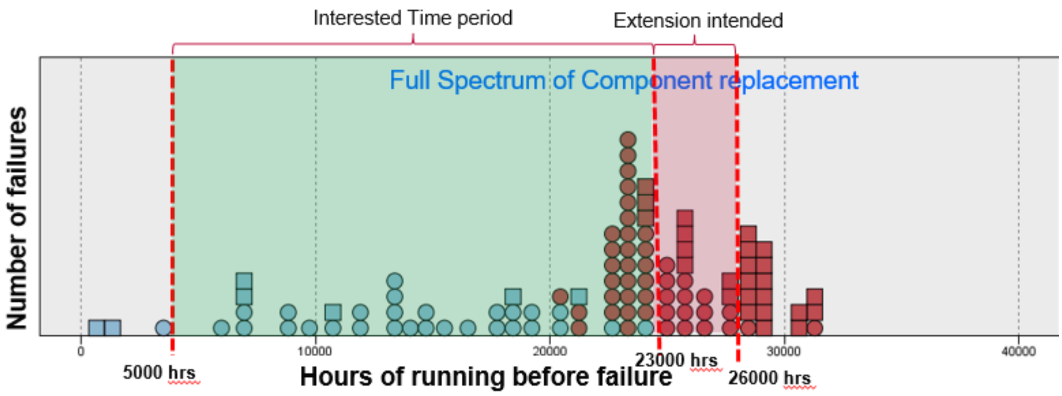

Figure 3:*Distribution of Failure*

The cross validation is performed with 5 number of folds. The error is back calculated as the difference between predicted lives (before 1,000 hours of failure) to actual life at the failure for each truck. The results are as given in

Table 3.

Table 3: Average Error in Prediction

	Avg error (Total)	Avg error (<10K)	Avg error (>10K)	Avg error(>23K)
Full model with telemetry data	3663	10470	2054	733

Based on the

Table 3. there is a huge error for premature failure (life < 10,000 hours), but majority of the fleet in the current mine is above this life. However, since the average error for engines with life greater than 23,000 hours is less than the desired 1000 hours alert window, it was concluded to deploy the tool as a pilot for monitoring the trucks for next 3 to 4 months and validate the model with actual replacements and failure. The condition monitoring team will continue monitoring the fleet in the standard procedure as usual in this time period and this prediction tool will be tested against their decision.

III. PILOT RESULTS

The model was run every week and the results were studied. The data after 3 months was compiled for checking the performance of the model. The performance is categorized into two categories, viz., Planned Replacement and Un-planned Replacement. Totally there were 7 replacements out of which 5 replacements were planned (end of life) and 2 were unplanned (breakdown). The current model could accurately predict 4 planned (end of life) replacements out of 5 and 1 unplanned replacement (breakdown) out of 2. The details of this analysis are explained in the following section.

A. Model Performance Details for Planned Replacement (end of life)

The engine that has reached end of life and needs a replacement is categorized as Planned Replacement.

The graphs in Figure 4 show details about the model prediction for the engines that were analyzed in this category. The '+' marker indicates the Predicted lifetime trend across the running history of the engine. This data is obtained based on the model runs at different time periods through the running life of engine as the oil sampling data was collected and processed. The running hours of the engine is indicated by the blue line '-.- 'marker which starts from 0 and goes up to the end life when the engine was replaced. The failure tolerance zone is indicated by '---' marker where top line is the hours when engine was replaced and bottom is 1000 hours less. Thus, a prediction that lies within this failure zone is considered to be a good prediction and the error indicated by marker '*' will lie as low as possible (< 500 hours) is considered to be good.

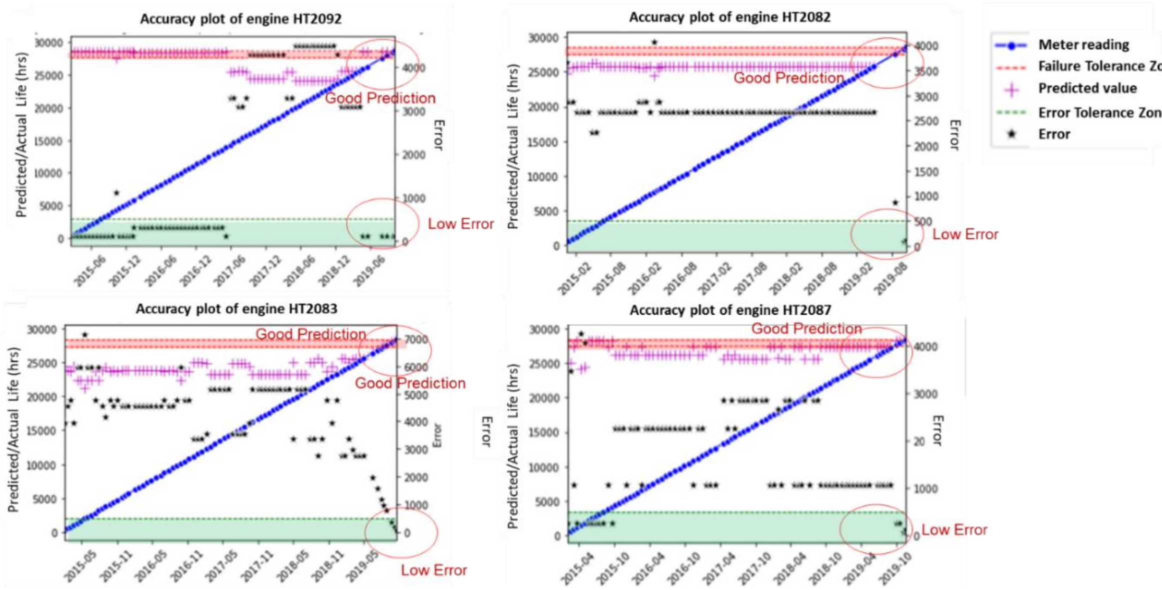

Figure 4: *Engines with Good Prediction for Planned Replacement (end of life)*

The predictions of haul trucks HT2092, HT2082, HT2083 & HT2087 were in line with that of condition monitoring team. Hence the model is working fine as the end of life predicted by the model is within an error of 500 hours.

However, for the truck HT2048 the life predicted was 22,600 hours. The condition monitoring team decided to do an engine replacement at a meter reading of 18,023 hours.

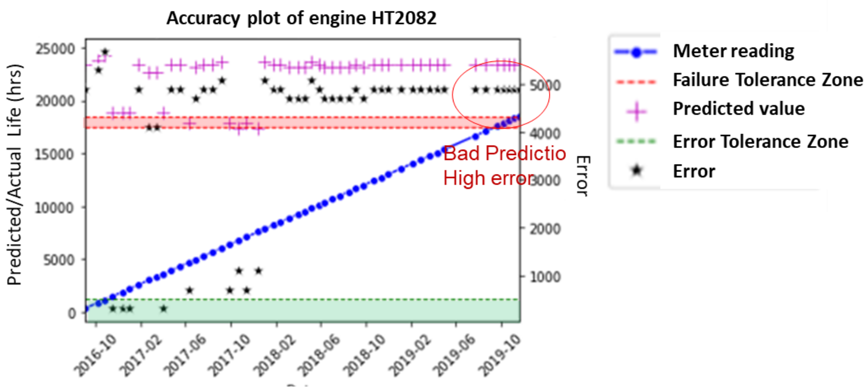

Figure 5:*Engine for which prediction was not accurate*

This decision by the condition monitoring team was taken due to the presence of high quantity of aluminum and iron in filter. However, upon checking the oil data no such indications were found for both chemicals. This provided a good input to improve the model by considering the results of filter checks as an additional feature in the model.

B. Model Performance Details for Un-planned Replacement (breakdown)

The unplanned replacements are the replacements that happen due to failure when the machine is in operation and hence no maintenance can be planned for the same. There were 2 such un-planned breakdowns in 3 months out of which one was predicted accurately and for other one prediction was close but outside of the failure tolerance zone.

In the first instance, actual failure for the truck HT2035 occurred at 24,245 hours and the reason was jammed Filter Base. This haul truck was reported to be critical and had been flagged by the prediction tool on 13-Nov-2019, 13 days before failure. This truck was also supposed to go for planned replacement in December, based on the recommendations from condition monitoring team.

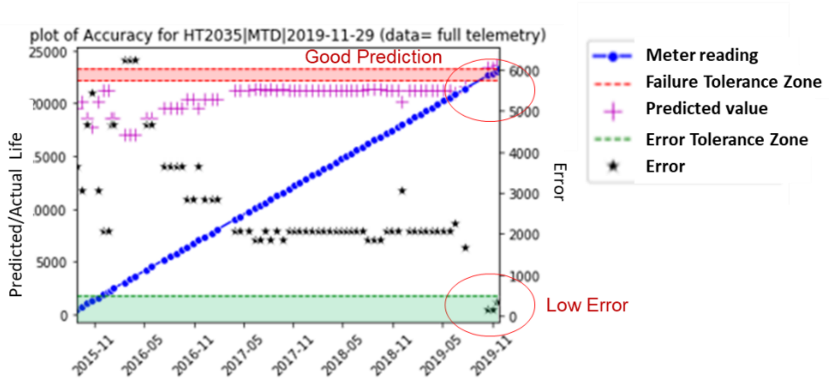

Figure 6:*Accurate Prediction of Unplanned Failure*

In another instance, the actual failure for the truck HT2032 occurred at 10,463 hours. The model had predicted the end of life to be around 11,800 hours, an error of 1,500 hours. The reason for failure was attributed to oil filter base leak that resulted in seized engine. Here although the prediction was outside of failure tolerance zone but still very close to the actual failure.

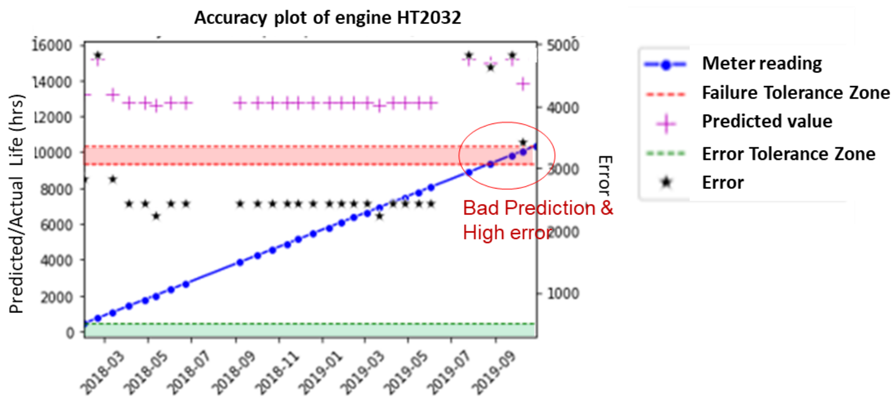

Figure 7: *Failure that was not Accurately Predicted*

IV. CONCLUSION

Based on the validation of the model it is concluded that the predictive model is good to be deployed for production. The error in the premature period of machine life is due to the paucity of information. During early stages, the machine fails mainly due to design problems which is not provided as an input to the model. The current model considers only Oil analysis that captures wear phenomena and alarm data, which is more of machine abuse than design characteristics. However, the model has low error in end of life period and the model showed good agreement with control group.

ACKNOWLEDGEMENT

This solution is part of Vale´s AI Center portfolio of replicable AI products, which are being scaled to the global operations in a strategic effort to advance and spread culture of decision driven by data. This article is related to the application of product Power Train Failure Prediction to Indonesia operations, which were supported by Infosys and IBM efforts and professionals.

REFERENCES

[1]. In-House Oil Analysis Lab Saves Gold Mine Millions in Downtime, Repair and Replacement Costs.https://www.spectrosci.com/resource-center/lubrication-analysis/literature/case-studies/gold-mine-saves-millions-with-on-site-oil-analysis-laboratory/

[2]. Heng, A.S.Y. Intelligent prognostics of machinery health using suspended condition monitoring data. PhD thesis, Queensland University of Technology, Australia. 2009

[3]. W.A. Carstens & P.J. Vlok. "Statistical analysis of caterpillar 793d haul truck engine data and through-life diagnostic information using the proportional hazards model", South African Journal of Industrial Engineering August 2013 Vol 24(2), pp 59-68

[4]. Olcay Altıntas, Murat Aksoy, Emin Ünal, Ogˇuzhan Akgöl, Muharrem Karaaslan, "Artificial neural

network approach for locomotive maintenance by monitoring dielectric properties of engine lubricant", Measurement, Volume 145, October 2019, Pages 678-686

[5]. Maria Grazia, De Giorgi Stefano CampilongoAntonio Ficarella, "A diagnostics tool for aero-engines health monitoring using machine learning technique", Energy Procedia, Vol 148, Aug 2018, Pages 860-867

[6]. Zhao, Y. (2017). Oil Analysis Handbook for Predictive Equipment Maintenance. Spectro Scientific

[7]. Gwidon Stachowiak, A. W. (2004). Experimental Methods in Tribology (Vol. Tribology series). Elsevier.

Bharadwaj Madhu, et. al. "Predictive Maintenance of Mining Haul Truck Engines Using Oil Sampling and Telemetry Data." *IOSR Journal of Engineering (IOSRJEN)*, 10(9), 2020, pp. 56-63.