Materials Research Proceedings **27** (2023) 308-314 https://doi.org/10.21741/9781644902455-40

Machine learning: approaches to predicting reliability and developing maintenance strategies

Subash Singh[1, a *], B. Vien[1], D. Welshby[2], W.K. Chiu[1]

[1]Department of Mechanical and Aerospace Engineering, Monash University, Clayton, Victoria 3800

[2]Institute of Railway Technology, Faculty of Engineering, Monash University, Clayton, Victoria 3800

[a]subash.singh@monash.edu

Keywords: Track Geometry, Track Quality Index (TQI), Track Condition Monitoring Vehicle (TCMV), Side Wear (SW), Side Wear Rate (SWR), Area Wear (AW), Area Wear Rate (AWR), Rail Operations, Terrain, Tracking Bias, Flange Difference (FD) Measurement Data, Machine Learning, Convolutional Theory, Gross Metric Tonnes (GMT), Head-hardened Rail (HH), Through-Hardened Rail (THH), Absolute (abs), Maximum Likelihood Estimator (MLE), Ordinary Least Squares (OLE), Original Equipment Manufacturer (OEM)

Abstract. Current approaches to maintenance of rolling stock bogies are focused on compliance to wear limits as stipulated by OEM specifications. OEM recommendations are critical to providing an industry wide approach to safety and compliance. These are not operation specific and are often not the most cost-effective solutions. A system approach to reliability is an established approach that is applied in less complex systems where the relationships between components are well defined with historical data and predictable conditions. Extending this approach to more complex multi-variate systems where many relationships are not intuitively obvious or mathematically defined presents a challenge. Machine learning techniques have been applied to address such problems with examples in image recognition, tool wear prediction using multiple sensory inputs and estimating railway bogie wear using vibration inputs. [8,9,10] The aim of the study is to extend and adapt machine-learning techniques to the area of developing maintenance strategies for optimal business benefit with a specific focus on railway bogie maintenance. This study aims to present an insight into the variables, which includes bogie tracking condition affecting track side wear rate. A finding is that an in-depth study of each independent variable's individual impact is a necessary step to efficient modelling. These include track geometry, operating and bogie component wear variables. Track side wear, curve radius, superelevation and track top variance have been found to be significant predictors of track side wear rate. These impact predictions are not consistent between the different rail tracks and are not exhaustive. Specifically, the impact of bogie performance requires inclusion. Combining these variables mathematically using statistical inference and convolutional theory with maximum likelihood estimators would establish a predictor for side wear rate for the specific operation. The paper finally discusses the relationship of area wear rate to side wear rate and the influences of grinding frequency and rail material type.

Introduction

Logical reasoning is used for significant maintenance /renewal programmes in the heavy haul rail industry. This approach generally lacks analytical accuracy on determining the optimal outcome regarding business value generated. The challenge in creating the optimal strategy is due to the inherent complexity of relationships between components, assets and business value metrics with

Materials Research Proceedings **27** (2023) 308-314 https://doi.org/10.21741/9781644902455-40

multiple feedback loops and differences in the timing of impacts. A closed form mathematical solution will in general not be able to account for all variables and conditions. Such a solution needs to be uniquely derived for specific operations due to the different environments in which railways operate. These challenges lend itself to computational modelling with real-time learning to account for changing environmental conditions. The research question is if such modelling can be developed by embedding the techniques of deep machine learning to achieve the business value outcome.

A body of research exists on linking maintenance strategies with maximising business value such as improving asset up time and reducing total maintenance cost spend. [1,2,3,4] Topics of research that combine the subject of maintenance strategies with machine learning include fault detection using condition monitoring data [5], sensor inputs to condition monitor tool wear [6] and fault diagnosis in bearings [7]. In these examples, machine learning is utilised to achieve reliable fault detection through condition monitoring. The goal is improved business shareholder value by productivity improvement and/or cost reduction. More specifically, in railway maintenance applications, machine learning has been researched for bogie maintenance [8], track defect detection [9] and for predicting wheel and rail interface wear [10].

This paper discusses the key variables to be analysed and the approaches in developing a learning model to link the impact of physical railway rolling stock, operating and track variables to the track wear rate. These variables are a core aspect of the larger framework to be defined which links the cost of maintenance activity on specific components to the net business value add, usually measured in terms of net revenue. The learning to be applied represents the gathering of condition data that describes the wear patterns of track and rolling stock under different operating and geometrical constraints. The specific focus will be on the impact of bogie condition on track maintenance.

Bogie fleet operational performance is represented by the mean and variance of the fleet flange difference. The current mean+3 standard deviations of the transformed data will be added as a constant term in the regression. This is because the track variables are currently measured across the track network at a single time-stamp, and given that the fleet operates randomly on all sections of the track network. The extension of the modelling to include data at different time stamps is discussed.

Modelling Approaches and Assumptions

Track maintenance and consequential downtime is primarily influenced by grinding or re-railing activity. Re-rail occurs once the track has been ground to minimum acceptable vertical height. Grinding is carried out to restore track profile after wear (both top and gauge) and remove surface defects (rolling contact fatigue). The wear rate due to the rolling stock may be influenced by GMT, asymmetrical loading due to wheel wear and changes in lateral loading due to operating speed, curve radius, superelevation and TCI. Ballast condition and tamping frequency may also play a role in wear rates as an influence on the dynamic behaviour of the rolling stock.

The modelling approach is initially based on single time-stamp data with limited available data at different time-stamps. For single time-stamps, the data is 2-dimensional and explains the relationhips of the variables to SWR spatially across the network. The time-dimension is excluded and more specifically implies that the variable that explains bogie performance across the rolling stock fleet is necessarily represented as a constant. This is a consequence of the entire train fleet operating randomly over all parts of the network. The constant chosen will be represented by the wheel fleet flange difference mean + n*standard deviations calculated for wheel trip range <100 to more closely represent bogie condition and not fleet-wide wheel condition. The value of integer "n" can be in the range of 1 to 3.

Materials Research Proceedings **27** (2023) 308-314 https://doi.org/10.21741/9781644902455-40

The impact of bogie condition is expected to directly influence side wear (gauge) rate (SWR) due to its impact on bias tracking and consequential flange wear difference (FD) across a wheelset. The use of data at multiple time-stamps will enable individual sections of the track to be analysed for the relationship of bogie FD to track SWR where track variables Curve radius, Superelevation and Gradient become constants.

The approach will be to formulate multiple regressions on individual track sections with significant SWR across time stamps. Combining the regression across time-stamps, where SWR is significant, with that at a single time-stamp across the network will result in a set of simultaneous equations to be solved to relate SWR to bogie FD in combination with the other variables of significance. An upgrade to the recording instruments and database storage is in progress which will enable the capturing of more wear rate data across multiple time-stamps.

The use of data across multiple time-stamps for individual track sections may indicate that the analysis at a single time stamp is redundant. However, the data where SWR has significant movement is limited to only certain track sections. The limited time-stamps currently available and the extended periods of time required to collect additional data implies that the number of data points available for regression at each track section is limited by a factor of 100 or more compared to single time-stamp data across the network. The relation of bogie condition to side wear rate on each track section will then be an input to the maintenance plans required for each section respectively. Summing across the rail network will provide the overall downtime impact. A complication is that Area wear which is the sum of side and top wear significantly influences grinding frequency and hence downtime impact.

The selection of regression technique is dependent on the behaviour of SWR with each of the variables. With multiple variables, the solution calls for multiple regression. Linear multiple regression requires each variable to have a linear relationship with SWR. This is not true from the data behaviour discussed further on. An adaptation for applying linear regression is for non-linear relationships to be transformed into linear forms. This can be complex but not impossible where there is a mixture of linear and non-linear relationships to SWR. In applying non-linear regression, the initial step is to choose representative function types for each variable. This can be derived from a study of the data scatter plots. Representative nonlinear functions can be polynomial, power terms, logarithmic, combinations thereof [16] or specialised forms of these described in [17] such as Richards' curve, Gompertz Growth Curve and the Michaelis-Menten Model. Combining these into an additive equation is a further assumption that will require validation. Where a variable is likely dependent on another variable, adding in a multiplication term may compensate [16].

SWR may behave differently for different ranges of each variable. Unique regressions may then be applied for different combinations of variable ranges. The selection of a non-linear function can be simplified where the data range for each variable is split piecewise where less complex polynomial approximations can be developed. For n variables where each variable is split into r_1, r_2, r_3,....,r_n distinct ranges respectively , the total regression combinations required to be analysed is $r_1r_2r_3....r_n$. The total number of regressions necessary can be reduced where a variable has an insignificant relationship to SWR in any of the ranges. In such cases, the variable is removed or the range is limited to only significant ranges of the variable. Each regression equation then predicts the SWR for specific track sections. In combination, the regressions will explain the SWR across the network.

A general multiple non-linear regression equation is as below [17] where Y is the dependent variable, $\underline{X} = \{X_1, X_2, X_3,\ldots,X_p\}$is a p-dimensional vector of independent variables, $\boldsymbol{\beta} = (\beta_0,\beta_1,\beta_2,\beta_3,\ldots,\beta_k)$ is a k-dimensional vector of parameters and ε is the random error term or residual.

Materials Research Proceedings **27** (2023) 308-314 https://doi.org/10.21741/9781644902455-40

$$Y = f(\underline{X},\boldsymbol{\beta})+ \varepsilon \quad (1)$$

The assumptions necessary for non-linear regression are the same as for linear regression [14], [15] and [16]:

1. No correlation between independent variables.
2. Linearity and independence of residuals or no autocorrelation between residuals.
3. Residuals should be normally distributed with zero mean
4. The residuals should be distributed with equal variance (Homoscedasticity)

In non-linear regression, assumptions (3) and (4) are not necessarily satisfied but meet the normality criteria where the sample size is large enough because of asymptotic theory [17]

Methods used in non-linear modelling include Newton's, Gauss-Newton or the Levenberg-Marquardt Methods.[17] A gradient descent algorithm together with a function to optimise will be explored.[17] This is either the ordinary least squares formula (OLS) to minimise or the likelihood estimator to maximise (MLE). Both estimators yield the same result for normally distributed residuals [14].

Data cleansing and analysis

The variable data was initially taken in a single time stamp across the track network. The track network comprises an east track (carrying empty trains to the Mine) and west track (carrying loaded trains back to Port).

For each of the variables analysed, initial regression plots showed significant difference in relationships with SWR for each of the tracks.

Table 1: Summary of individual correlations of independent variables to SWR split between East and West tracks

	East track				West track			
Variable	Pearson	Spearman ρ	Linear/ non linear	R^2	Pearson	Spearman ρ	Linear/ non linear	R^2
Side wear	0.62	0.10	linear	38	0.36	0.13	linear	13
Curve radius	-0.24	-0.24	non-linear	8.9	-0.34	-0.34	linear	11.4
Superelevation abs	0.11	0.10	linear	1.7	0.14	0.25	non-linear	8
Gradient abs	0.01	-0.04		0	0.03	0.06		0
TCI outlier removed	0.16	0.17	non-linear	3.3				0
Line outlier removed	0.11	0.05	not sig	0.5				0
Top outlier removed	0.22	0.20	non-linear	10.9				0
Cross level	-0.11	-0.01		0	0.01	0.02		0
Abs cross level	0.08	0.22		1.2				0.4
number of braking events	0.09	0.15	non-linear	1.2	-0.01	0.12	non-linear	0.86
average speed	-0.10	-0.15	linear	1	-0.08	-0.03	non-linear	1.4
average brake pressure integrated with time	0.11	0.19	non-linear	1.3	0.04	0.09	non-linear	1.24
Total				64.7				36.3

To improve the regression accuracy across the variable range, scatter plot boundaries were defined for changes in the relationship. Estimating these change over points visually lead to the development of piecewise regression estimation in a single variable for each case.

This approach benefited with a more detailed investigation into the relationship which often hightlighted new signficant variables or showed dependence between variables. It also helped to understand and distinguish correlation with and without causation. An example was the study of the variable SW on SWR for the east track. At R squared = 38%, SW potentially represents the most signficant relationship to SWR.

Data range for SW was split into ranges from 0 to 1.37, 1.37 to 4mm and above 4mm. 0 to 1.37 represents 80% of the data points, whilst the range above 4mm represents less data but indicates a change in relationship that shows much higher correlation at R squared = 53%. Individual scatter plots for these regions are shown in Fig 1 and 2.

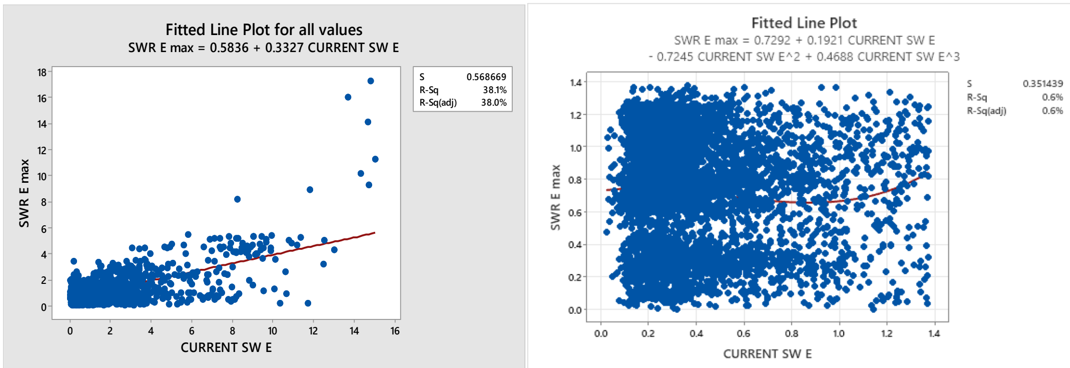

Fig 1 (a) East Track Population *Fig 1 (b) East Track SW < 1.37*

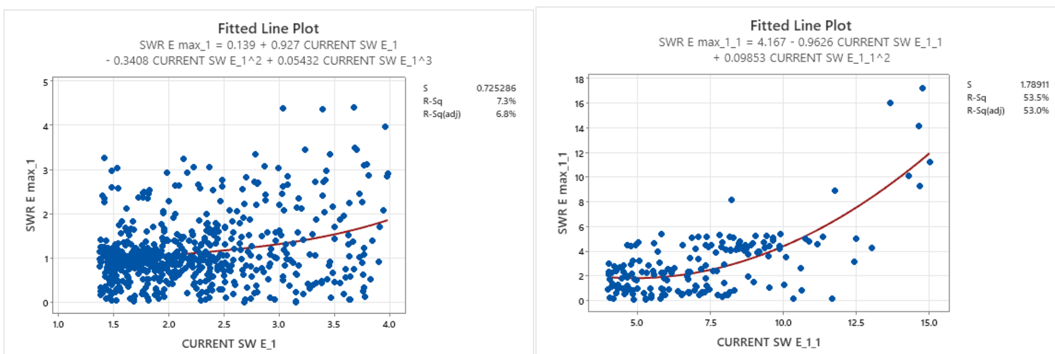

Fig 2 (a) East Track 1.37 < SW <4 *Fig 2 (b) East Track SW > 4*

For range SW < 1.37, two distinct populations are evident and are defined with SWR > 0.6 and SWR<0.4 respectively and were compared across the following variables: Track grinding frequency within last year, curve radius, location, vertical wear, TQI and rail material type. Test used was 2-sample T for each variable except for a proportion test for rail material type.

There was significant difference between the 2 populations with respect to grinding frequency, vertical wear, TQI, curve radius and rail material type. These variables did not show sigificant correlations to SWR with R squared summing to a maximum of 14% . Aspects to investigate include track gauge variation and operating variables of speed and braking.

For SW >4, two populations were also defined with SWR >2.5 and SWR <1.5. Comparing these populations using the same techniques showed significant difference with respect to grinding frequency and rail material type. Seperating the data for THH rail material type showed significant polynomial R^2 correlations of SWR to AWR (32%), AW (22%), Grinding frequency (6.6%), Curve Radius (5%) and Cant (3.4%) except at extreme SWR values where there were 6 data points with large residuals. The outliers had excessive Cant angle for the average speeds recorded through each section.

A strong R^2 correlation between Grinding frequency and AWR (60%) is also evident except at 5 extreme AWR values with large residuals. These points are the same outliers as above. Other variables of significance have since been identified from the previous analyis include grinding frequency, area wear, area wear rate and rail material type. It also shows that the data may be further seperated with respect to rail material type to improve the regression correlations. The same analysis will also need to consider operational and remaining track geometry variables for any significant dependence and determine any further seperation necessary.

Conclusion

For the railway system analysed, the individual relationships were defined intuitively and then tested using regression analysis with the available data. This has led to the classification of significance amongst the variables.

Splitting the analysis between the east and west tracks with different side wear rate behaviour was a significant step in the data analysis. The split resulted in a wide range of correlation significance across the variables. Variables of note include side wear, curve radius, top and superelevation and to a lesser extent braking and speed as depicted in Table 5.

Further splitting of the data where there are distinct sub-populations in the scatter plots that are influenced by other sub-variables helped improve the regression accuracy and exposed any multi-collinearity between variables. An example was the impact of grinding frequency and rail material type (HH and THH) on the dependence of SWR on SW for the east track.

The intent is that each variable is split into a maximum of three distinct ranges to limit the number of regression combinations that would need to be calculated but simultaneously enabling the application of simpler polynomial functions to the multiple non-linear regression method. There is a risk that the approach is data pre-processing biased, reducing the flexibility of the machine-learning model to process raw data. Alternatively, these steps can be automated or a possible trade off developed with fewer but more complex regressions.

Multiple non-linear regression using Python Scipy curve fit optimisation method with a least squares optimisation function will be explored [18]. With remaining variables being analysed initially at a single time-stamp due to limited instances of different time-stamp data for track wear, the bogie fleet will be represented by the mean and standard deviation of the fleet-wide FD relevant to the current time-stamp. Multiple time-stamp data will subsequently introduce bogie fleet performance as a variable in the analysis.

The data analysed is not exhaustive and more variables may likely be included once the initial multiple regression is completed. These include mileage since last wheel turn and wheel hollowing.

References

[1] K. Doganay and M Bohlin, “Maintenance plan optimisation for a train fleet,” WIT Transactions on The Built Environment, vol 114, 2010 IT Press. https://doi.org/10.2495/CR100331

[2] CP Ward, RM Goodall, R Dixon, G Charles, “Condition Monitoring of Rail Vehicle Bogies,” UKACC International Conference on Control, Coventry, UK, pp 1178-1183.

[3] R Santos, PF Teixeira, AP Antunes, “Planning and scheduling efficient heavy rail track maintenance through a Decision Rules Model,” Research in Transportation Economics 54(2015) 20-32, 2015 Elsevier Ltd. https://doi.org/10.1016/j.retrec.2015.10.022

[4] M Gopalakrishnan, A Skoogh, “Machine criticality assessment for productivity improvement, Smart maintenance decision support,” International Journal of Productivity and Performance Management Vol68 No. 5, 2019 pp 858-878, Emerald Publishing Limited. https://doi.org/10.1108/IJPPM-03-2018-0091

[6] F Aghazadeh, A Tahan, M Thomas,”Tool condition monitoring using spectral subtraction and convolutional neural networks in milling process”,The International Journal of Advanced Manufacturing Technology (2018) 98:3217-3227. https://doi.org/10.1007/s00170-018-2420-0

Materials Research Proceedings **27** (2023) 308-314 https://doi.org/10.21741/9781644902455-40

[7] H Liu, L Li, J Ma, “Rolling Bearing Fault Diagnosis Based on STFT-Deep Learning and Sound Signals,” Hindawi Publishing Corporation, Shock and Vibration, Volume 2016, Article ID 6127479, 12 pages. https://doi.org/10.1155/2016/6127479

[8] Y Wu, W Jin, J Ren, Z Sun, “Fault Diagnosis of High-Speed Train Bogie Based on Synchrony Group Convolutions,” Hindawi, Shock and Vibration, Volume 2019, Article ID 7230194, 13 pages. https://doi.org/10.1155/2019/7230194

[9] C Yang, Y Sun, C Ladubec, Y Liu,”Developing Machine Learning-Based Models for Railway Inspection,” Appl. Sci, 2021,11,13. https://doi.org/10.3390/app11010013

[10] A Shebani, S Iwnicki, “Prediction of wheel and rail wear under different contact conditions using artificial neural networks,” Wear 406-407(2018), 173-184, Elsevier Ltd. https://doi.org/10.1016/j.wear.2018.01.007

[11] A Meyer, “WheelView Freight Measurement Parameters,”Beena Vision Systems Inc.,pp 2,5, 2016, www.beenvision.com

[12] A. Troiano, “Track Geometry Measuring System, Operating Manuals On Board Monitoring Systems,” Mermec, P118013A11Y VV 501 Rev 1.0 (2015), pp 16-20

[13] A. Troiano,“Rail Profile Measuring System, Operating Manual On Board Monitoring Systems, „Mermec, P118013A11Z VV 501 Annex C Rev 1.0 (2015), pp 16-21

[14] X. Yan, X. Gang Su,“Linear Regression Analysis, Theory and Computing,“World Scientific Publishing Co. Pte. Ltd.(2009). https://doi.org/10.1142/6986

[15] R. Christensen,“Advanced Linear Modelling“,Springer Science+Business Media, New York, (2001). https://doi.org/10.1007/978-1-4757-3847-6

[16] C Flatt, RL Jacobs, „Principle Assumptions of Regression Analysis: Testing, Techniques, and Statistical Reporting of Imperfect Data Sets, Advances in Developing Human Resources, Vol 21(4) 2019 pp 484-502, Sage. https://doi.org/10.1177/1523422319869915

[17] D.S. Young,“Handbook of Regression Methods“,CRC Press, New York, (2017). https://doi.org/10.1201/9781315154701

[18] Kumar, B, Python SciPy Curve Fit - Detailed Guide, Python Guides, Copyright 2022, https://pythonguides.com/python-scipy-curve-fit