
White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/247328/

Version: Published Version

Article:
Cartwright, S.J., Folorunso, M., Rothwell, B.C. et al. (2026) Feature importance analysis of low adhesion risk model to find key parameters influencing friction. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit. ISSN: 0954-4097

https://doi.org/10.1177/09544097261498963

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.





Check for updates

Original Research Article

Institution of MECHANICAL ENGINEERS
Journal of RAIL AND RAPID TRANSIT

Proc IMechE Part F:
J Rail and Rapid Transit
2026, Vol. 0(0) 1–10
© IMechE 2026



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09544097261498963
journals.sagepub.com/home/pif
Sage

Feature importance analysis of low adhesion risk model to find key parameters influencing friction

Samuel J. Cartwright[1], Morinoye Folorunso[2], Benjamin C. Rothwell[1], Grazziela Figueredo[3], Carol Eastwick[1], Humberto Medina[1], Roger Lewis[2] and Joseph L. Lanigan[2]

Abstract

Low wheel-rail adhesion presents a major safety and economic challenge for the rail industry by causing station overruns and delays. The impact of this (with estimates of costs up to £355 m annually) presents a research opportunity with the potential for significant safety improvements and cost savings. Poor adhesion can be mitigated somewhat using weather forecasting, although this is limited to regional-level resolution and results in reduced averaged speeds. Recent studies have investigated rail cleaning to remove railhead contamination prior to wheel contact, which increases friction and braking performance. Determining when to activate systems for removing contamination on the rail surface is an area of active research with systems combining sensor and camera data with machine learning models recently demonstrating accurate prediction of Coefficient of Friction (CoF) values. While these systems are effective in initial testing, they display a common problem for safety-critical ML systems in a lack of model interpretability due to the "black box" nature of the neural network models used. This study seeks to mitigate poor ML interpretability through the use of Feature Importance (FI) algorithms, to better understand the causes of low wheel-rail friction prediction in ML models. A dataset correlating rail friction measurements against sensor data and rail images is used to train an ensemble of ML models, which are then used alongside a testing dataset to evaluate an array of FI algorithms (SHapley Additive exPlanations, Local Interpretable Model-agnostic Explanations and Permutation Importance). A feature importance fusion algorithm is then applied to assist in interpreting these results. Across a range of feature importance techniques, it is consistently found that atmospheric data such as air temperature and humidity is the strongest predictor of low friction conditions. Tests of similar friction prediction models using only atmospheric data with no photographic images show only a small loss in accuracy. Edge detection is found to be the most useful feature extracted from images, while single-pixel features such as brightness and colour are found to be the least important. This provides insight into best practices for feature selection when designing ML models for the rail industry in future.

Keywords

railway safety, railways, technology, vehicle track interaction, wheel/rail interaction

Received: 15 January 2026; accepted: 8 September 2026

Introduction

One of the most significant challenges at the wheel-rail interface has long been recognised as low adhesion, particularly during the autumn season when wet leaf layers can form. Low friction resulting from railhead contamination leads to critical safety risks and operational delays, primarily due to braking problems that can result in signals passed at danger (SPADs) or station overruns. Additionally, poor traction caused by low adhesion can lead to substantial delays.[1] Low adhesion can be caused by leaf layers or by the "wet-rail" phenomena where oxide on the rail mixes with small amounts of water to form a lubricating paste. Overall, the cost of managing low adhesion in Great Britain during autumn is estimated at £355 million annually for performance impacts, safety costs, and cleaning, management, and prevention.[1]

[1]Mechanical and Aerospace Systems Research Group, Faculty of Engineering, University of Nottingham, Nottingham, UK
[2]The University of Sheffield, Sheffield, UK
[3]School of Medicine, University of Nottingham, Nottingham, UK

Corresponding author:
Roger Lewis, The University of Sheffield, School of Mechanical, Aerospace and Civil Engineering, Mappin Street, Sheffield S1 3JD, UK.
Email: roger.lewis@sheffield.ac.uk

Despite the high cost and safety risks, the railway industry currently lacks real-time friction data to assist in better planning and decision-making. The primary tool used for adhesion forecasting is the Met Office Adhesion Index,[2] which relies solely on weather data such as wind speed, leaf fall, and moisture or ice presence. While useful, this index does not account for the actual condition of the railhead, which is a critical factor in adhesion levels. These limitations underscore the need for more advanced tools with higher accuracy and broader applicability.

Moreover, there is still a lack of comprehensive understanding of the specific parameters that affect railhead friction. While some data is now available, for example the work carried out by Folorunso et al.[3] in analysing rail friction and environmental data from UK rail tracks, this has not been fully leveraged to develop reliable models for real-time friction estimation. Recent advancements in data collection technology, however, now offer an opportunity for a more in-depth analysis of these parameters and their influence on railhead conditions which was explored by Folorunso et al. in designing a machine learning model using field data to estimate railhead friction.[4]

In this work, a dataset consisting of tribometer measurements, railhead photographs and atmospheric sensor data is used to train an ensemble of ML models to predict the wheel-rail friction coefficient. The performance of these models is analysed against a separate testing dataset. These ML models are then used to estimate feature importance coefficients in the training dataset using the SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations) and PI (Permutation Importance) algorithms. The results from these algorithms are normalised and combined using Feature Importance Fusion techniques to reduce the output variance, and the resulting feature importance coefficients are studied to gain insight into optimal data gathering techniques for future friction prediction systems.

Background

Wheel-rail adhesion

Previous work has identified that the presence of third body particles such as rust play a significant role in wheel-rail friction under specific conditions.[5] Experimental investigation has determined that the presence of dry rust particles alone does not significantly modify friction conditions, however with the addition of oil or water friction can vary significantly.[5] Experiments comparing the performance of different rail tribometer designs emphasise the importance of understanding the role of third bodies in obtaining both accurate measurements and models of rail surface friction.[6] Under contamination by oil, surface roughness is also found to play a role in the resulting friction.[7,8] The effect of both surface roughness and temperature becomes much larger when using water as a contaminant.[8,9]

Mathematical modelling of wheel-rail contact has been studied for some time.[10–12] While simplified models of rolling contact between elastic bodies are not computationally complex,[10] extending models to capture real world conditions such as the rail surface texture and presence of both solid and liquid contaminants (and their behaviour under the high pressure contact) is more difficult.[9,13]

The difficulty of analytical modelling of wheel-rail contact under real-world conditions makes the presence of reliable measurements of rail friction valuable[6,14] for the development of numerical rules. The development of practical and compact tribometer tools has enabled a greater volume of data to be gathered.[14] Despite this, the requirement of physical contact with the rail surface for tribometers to operate limits the usefulness of friction measurement in day-to-day operation. Recent work has involved the application of ML systems to leverage datasets of rail friction measurements in order to quickly and remotely predict friction conditions.[4]

Feature importance in engineering

Investigations into the use of ML in safety-critical applications frequently emphasize the importance of interpretable ML models.[15] This is broadly described by the title of "Verifiable AI", defined in 16 as "the goal of designing AI-based systems that have strong, ideally provable assurances of correctness with respect to mathematically specified requirements". One approach to this is the use of post-prediction explanation in the form of Feature Importance (FI) algorithms.[17] Many different algorithms have been proposed and while many of these has been shown to individually effective, arguments have been made that retroactive explanation techniques are fundamentally flawed.[18] An alternative to dedicated verifiable ML architectures is the use of ensemble-based feature importance approaches.[15,19] Applications of this are numerous including obtaining key influencing variables across a variety of fields.[20–23]

Variation in Feature Importance estimates arise from multiple sources including differences in the ML models, usage of different FI algorithms and dependence on the sample of data used to obtain the estimates. Ensemble FI therefore seeks to reduce this variability by combining a multitude of different ML and FI algorithms.[23] The FI algorithms applied in this study are Permutation Importance,[24] SHAP[25] and LIME.[26] These are implemented by the BioFEFI library.[27]

Data overview

Railhead images were acquired using GoPro Hero cameras operating in automatic exposure mode. The camera was positioned at a fixed distance from the railhead, resulting in a consistent field of view focused primarily on the rail surface. Image collection was conducted during daylight or adequately illuminated conditions, reducing

Table 1. Quantity of data collected from each site.

Location	Railway	Image quantity
Butterley, derbyshire	Midland railway	53
Shottle, derbyshire	Ecclesbourne valley railway	16
Swanwick, derbyshire	Midland railway	30
Deepcar, sheffield	Non-heritage	18
Wharncliffe woods, sheffield	Non-heritage	6
Matlock, derbyshire	Peak rail	24
Beeley woods, sheffield	Non-heritage	19
Monk bretton, barnsley	Non-heritage	34
Long marston, stratford-upon-avon	Non-heritage	25
Darley dale, derbyshire	Peak rail	28
Idridgehay, derbyshire	Ecclesbourne valley railway	94
Wirksworth, derbyshire	Ecclesbourne valley railway	94

the occurrence of severely underexposed or overexposed images.[4]

A railhead dataset of over 400 measurement points was recorded between 2019 and 2020 during different seasons to give a variety of in-field railhead conditions. The majority of these were for "wet-rail" related railhead conditions rather than leaf layers (Table 1). The data was collected from several rail tracks in the UK which included some Heritage railways sites. These are shown in Figure 1.

A variety of data was recorded from the different sites visited to give a proper understanding of low adhesion mechanisms. Friction coefficients on the rail surface were measured using a pendulum tribometer. The air, railhead and dew point temperatures, as well as the relative humidity, were measured using an infrared thermometer. The contamination layer thickness was measured using an eddy current paint thickness gauge, and images of the rail surface and forward-facing view were captured using GoPro Hero cameras. Further detail regarding the



(a) GoPro image of the rail surface.

(b) Front view of the local environment.

(c) Measurements obtained using the pendulum tribometer.

Figure 1. Images of the data acquisition.

methodology of the data collection is outlined in the work carried out by Folorunso et al.[3] Figure 1 illustrates some of the images collected and friction measurements in progress.

Data analysis approach

Image feature extraction

Prior to feature extraction, an automated railhead localisation procedure was applied to each image. An initial cropped region containing the railhead was identified, after which a Gaussian filter (σ = 10) and vertical Sobel edge detection were used to determine the railhead position. The Gaussian-filtered image was used solely for localisation purposes, and a consistent railhead region of interest (ROI) was subsequently extracted from the original image. All image features were computed from this cropped ROI rather than from the filtered image.

Before the images could be used as inputs to the Machine Learning (ML) model, features were extracted to reduce the dimensionality of the input.[4] A variety of features were selected targeting different image characteristics including colour, brightness and surface texture, which were proposed on the basis of being likely to correlate with the CoF.[4] The feature extraction methodology was reused to ensure that this work's analysis remains relevant to earlier studies.[4,28]

The extracted image features included mean brightness, mean colour, black-pixel content, orange-pixel content, and vertical edge intensity. Mean brightness and black-pixel content were used to quantify surface darkness, which can be associated with moisture, contamination, or surface films. Orange-pixel content was extracted using HSV colour-space analysis to identify rust-coloured deposits and oxidised material. Mean colour was also calculated in HSV colour space to characterise general railhead surface colour properties, while vertical edge intensity, calculated using a Sobel operator, was included as a measure of surface texture and structural variation. These feature extraction schemes are described further in the following sections.

In order to maintain consistency with the methodology of previous work,[4,28] no explicit exposure normalisation, histogram equalisation, white-balance correction, colour calibration, or denoising procedures were applied prior to feature extraction. The fixed camera geometry, close proximity of the camera to the railhead, and use of HSV-based colour features helped reduce sensitivity to variations in illumination conditions. Since the local environmental conditions play an important role in the formation of rail surface contamination, information which correlates to ambient conditions such as brightness and colour balance remained uncontrolled to allow the ML system to learn from them.

Brightness. The first feature was a measure of the image's brightness. Changes in image brightness are an important descriptor of the rail's local environment and are associated with sunlight exposure and shade, which plays an important role in the onset and evolution of rail surface contamination.[28] Image brightness was calculated by flattening the original RGB image and calculating the mean intensity of each subpixel.

Darkness. Similar to the brightness feature, darkness quantifies the number of sufficiently dark regions within each image, as in sufficient quantity these can be an indicator of contamination on the rail surface.

Each image was first converted to a Hue-Saturation-Value (HSV) format. The saturation and value channels were normalised and a threshold value of 0.3 was applied. A pixel-wise multiplication of the saturation and value channels was carried out, and the feature value for each image was taken as the mean of the resulting array.

Orange pixel detection. The presence of orange pixels was used as an indicator of rail contamination due to the formation of rust. To detect these, the image was first converted to HSV format. The set of pixels are selected such that $0 \leq H < 1/12$. For these pixels, the element-wise product of the hue, saturation and value was calculated and averaged.

Edge detection. Edge detection was applied to measure the texture of the rail surface. In normal operation the rail surface should be smooth metal, however contamination such as rust formation creates a spotty texture. The magnitude of surface texturing on the rail was calculated by approximating the image gradient using a Sobel filter[29] to detect vertical edges. The kernel used was

$$\begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} \quad (1)$$

After the filter was applied to the image to obtain gradients in the horizontal direction, the mean of the absolute values was calculated to extract a quantity which qualitatively measured the visual smoothness of the rail surface.

Colour detection. For the final extracted feature, the image was converted into the HSV format. For each pixel, the product of the hue and saturation was calculated. The mean of this value was calculated for the whole array to obtain a measure of how much colour the image contains.

Construction of dataset

A total of 300 data points were used for the training dataset, where each data point corresponds to one set of features extracted from one image as well as additional measurements such as the air temperature, humidity and the rail temperature. Each image is also associated with a pendulum tribometer reading. The Pendulum

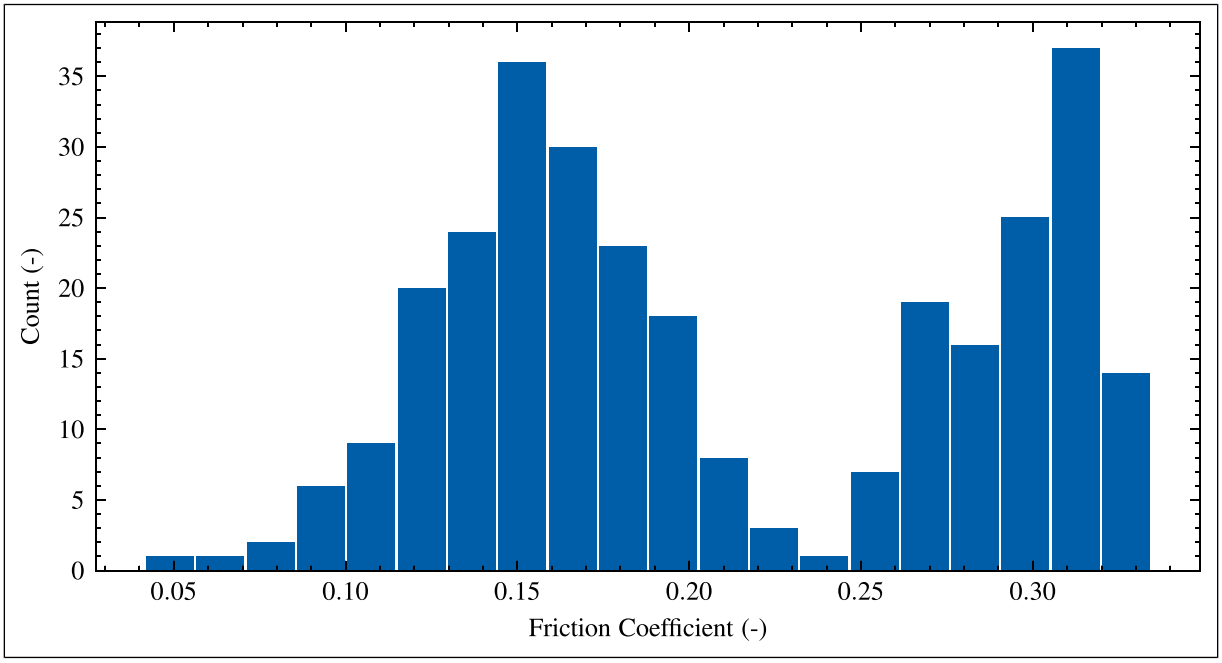


Figure 2. Distribution of the Coefficient of Friction in the training dataset.

Test Value (PTV) is converted into a Coefficient of Friction (CoF) using equation (2)[3]:

$$\mu = \frac{PTV}{195.45} \tag{2}$$

The distribution of the CoF labels within the training dataset is shown in Figure 2. A bimodal distribution was found - this was likely a result of the fact that data gathered on different days was likely to be either under normal (acceptable friction levels) or low friction conditions.

Analysis of feature colinearity was carried out and is illustrated in Figure 3(a). The process for this was adapted from Fang et al.[20] It can be seen that the features obtained from the sensors (air temperature, humidity, dew point and railhead temperature) frequently display high Pearson correlation coefficients. While this is to be expected as temperature, humidity and dewpoint are physically related, mitigation of feature collinearity is desirable as it can lead to incorrect feature importance estimates.[22] Reduction of collinearity was achieved by repeating analysis with a variety of different feature selection sets (visualised in Figure 3), as well as by combining the high collinearity features in two ways as described below. The resulting datasets are summarised in Table 2.

Method 1: The normalised air temperature, humidity, dew point and railhead temperature features were combined into a single feature by summing them. All features in the dataset are then standardised prior to model training using z-score standardisation.

Method 2: In accordance with earlier work regarding the effects of condensation on the rail,[3] a new variable Condensation Index is created and defined as:

$$\begin{aligned}&\text{Condensation Index}\\&= \text{Relative humidity} * \text{Dewpoint} * (T_{air} - T_{rail})\end{aligned} \tag{3}$$

These variables are chosen such that each of them increasing (humidity, dew point and the difference in temperature between the air and rail) results in an increased rate of condensation. All features in the dataset are standardised using the z-score. The effect on the variable collinearity is illustrated in Figure 4.

ML model performance

Before training commenced, a random sample of 20 % of the dataset was selected to be withheld from training and used as a testing dataset. Each feature within the training dataset was preprocessed using z-score standardisation. The training dataset was used to train an ensemble of models using a variety of ML architectures. By using a series of different models and different FI algorithms, a more reliable feature importance estimate can be achieved by reducing reliance on any one method. A linear regression model was used as a baseline for performance comparison against the following ML techniques:

- Random Forest (RF) regression.
- Support Vector Regression (SVR).
- Extreme Gradient Boosting (XGBoost).

Model hyperparameter optimisation was conducted and the selected hyperparameters are shown in Table 3.

The performance of each model in predicting the friction coefficient using the images and sensor data is summarised in Table 4. It can be seen that the dataset using

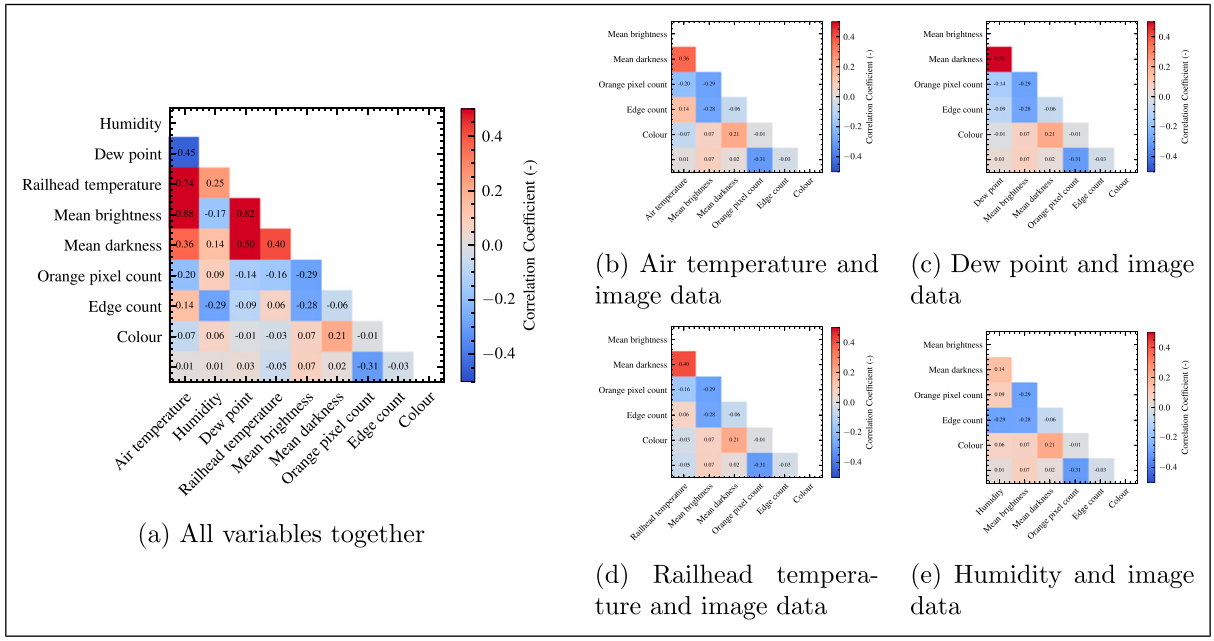


Figure 3. Plots of Pearson correlation coefficient for the different feature-reduced datasets.

the newly defined Condensation Index consistently outperforms the simpler method of feature reduction, across a variety of ML techniques. Optimal performance was obtained using the XGBoost model, with an RMS friction coefficient error of 0.047.

The feature importance fusion technique described in 23 was applied here. Three feature importance algorithms are applied to each ML model; SHapley Additive exPlanations (SHAP),[25] Local Interpretable Model-agnostic Explanations (LIME)[26] and Permutation Importance (PI).[24] The matrix of FI data obtained from this is visualised in Table 5.

The set of FI values corresponding to multiple ML models and FI algorithms was then processed using the feature importance fusion technique from.[27] The sets of FI rankings are combined in the following ways:

- Mean: The final calculated importance of each feature was equal to the mean of the values for that feature from each FI algorithm.
- Median: The importance of each feature was equal to the median of the FI algorithm outputs for that feature.

The result of this was an array of FI values where each feature has one value for each of the fusion methods outlined above. The final step of analysis involves rescaling each feature's values to aid their interpretation.

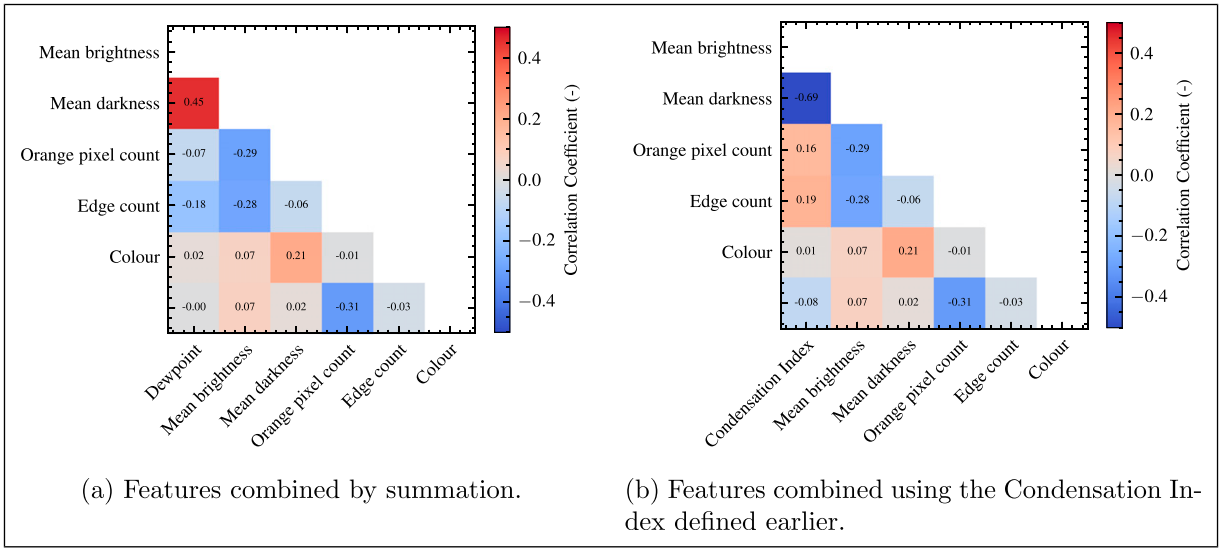
Discussion

The results of the FI fusion analysis are shown for the feature summed dataset and condensation index dataset in Figure 5 respectively. It can be seen that for both datasets, the relative importance of the features constructed from atmospheric and sensor data was ranked as the most important feature using both FI fusion techniques.

Specifically of the data extracted from the image features, the edge detecting feature was ranked as the most important. This is consistent with the idea that low adhesion is a consequence of the formation of rust spots and other surface contamination that alters the nominally smooth surface texture. The other image-extracted features (darkness, brightness, orange pixel detection and colour) are found to have low importance compared to the edge detection and sensor data-based features.

Table 2. List of features in the tested datasets.

Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7
Air temperature, railhead temperature, dewpoint, humidity	Air temperature only	Railhead temperature only	Dewpoint only	Humidity only	Sum of sensor features	Condensation index
		Brightness				
		Darkness				
		Orange pixel fraction				
		Edge detection sum				
		Coloured fraction				



(a) Features combined by summation.

(b) Features combined using the Condensation Index defined earlier.

Figure 4. Correlation matrices for the preprocessed datasets. (a) Features combined by summation. (b) Features combined using the Condensation Index defined earlier.

Combined, this indicates that the measured data (air temperature, railhead temperature, dewpoint and humidity) are strong predictors of the presence of low adhesion. This can be further tested by limiting the training dataset to these four features. From earlier testing, the best performing ML architecture was the XGBoost model with an RMS error of 0.047. Using the same hyperparameters documented in the "ML Model Performance" section but limiting the training features exclusively to the sensor data, this model was found to display a friction coefficient RMS error of 0.055. In contrast, using only the image data to train the same model, the performance degrades further to 0.064.

That atmospheric terms alone are able to predict wheel-rail friction nearly as well as the best case XGBoost model using the full feature set (and better than the feature-reduced model using the feature sum) provides additional evidence that low adhesion is influenced primarily by atmospheric effects which result in condensation.

Table 3. Hyperparameters for the three ML models.

RF	
Number of trees	300
Minimum split samples	2
Minimum leaf samples	1
Max depth	None
SVR	
Kernel	Linear
Degree	2
C	100
XGBoost	
Number of estimators	100
Max depth	None
Learning rate	0.1
Subsample	0.25

The feature importance fusion data, and poorer performance of the XGBoost model using image features only, suggests that the image features extracted using single-pixel colour and brightness data (darkness, brightness, colour and orange pixel detection) are weak predictors of low-adhesion conditions. Conversely, the best performing image feature (edge detection) being the only one to capture multi-pixel elements in the image indicates that more complex image analysis may be required to improve performance or obtain performance parity using only image data. Features that depend on simple extracted characteristics like colour and brightness can be biased by external changes in conditions such as sunlight intensity, angle and cloud cover whereas these have a much smaller effect on texture-based features such as edge detection.

Further work could therefore improve the contribution of images to the low-adhesion model by investigating other multi-pixel feature extraction techniques, such as using a Fourier transform to extract information on the length scale of the texturing visible on the rail. Alternatively, the features extracted from

Table 4. RMS error for each ML model type and dataset. The best performing model for each dataset is indicated in bold.

	Friction coefficient RMS error (−)			
Dataset	Linear	RF	SVR	XGBoost
All variables	0.061	0.048	0.065	**0.047**
Air temperature	0.067	**0.058**	0.073	0.059
Dewpoint	0.075	**0.058**	0.073	0.060
Humidity	0.063	**0.053**	0.065	**0.053**
Railhead temperature	0.072	**0.061**	0.075	**0.061**
Sum of variables	0.073	**0.060**	0.072	0.061
Condensation index	0.072	**0.058**	0.071	**0.058**

Table 5. List of combinations of ML models and FI algorithms.

		FI algorithm		
		SHAP	LIME	PI
ML model	RF	✓	✓	✓
	SVR	✓	✓	✓
	XGBoost	✓	✓	✓

the images can themselves be made a trainable feature of an ML model using a Convolutional Neural Network (CNN) architecture. This allows the model to learn the important features of the image without being limited by the human-written feature extraction algorithms. A possible limitation of this is the increased requirements for training data in accordance with the higher parameter count of the expanded ML model.

A possible limitation of this work is that the data gathered is specific to a small area of the Midlands of the UK. This therefore limits the scope of the model performance and conclusions gained from it as the local climate plays a key role in the formation of contamination by rust or leaf debris. Future work could therefore seek to gather data from a more diverse range of climates. A similar methodology could then be applied to test whether the measured dependence on atmospheric data generalises geographically, in conditions where the key drivers of railhead contamination may differ. Future work could also benefit from the use of uncertainty-tolerant ML architectures to account for the errors introduced at the various stages of this work including experimental data gathering

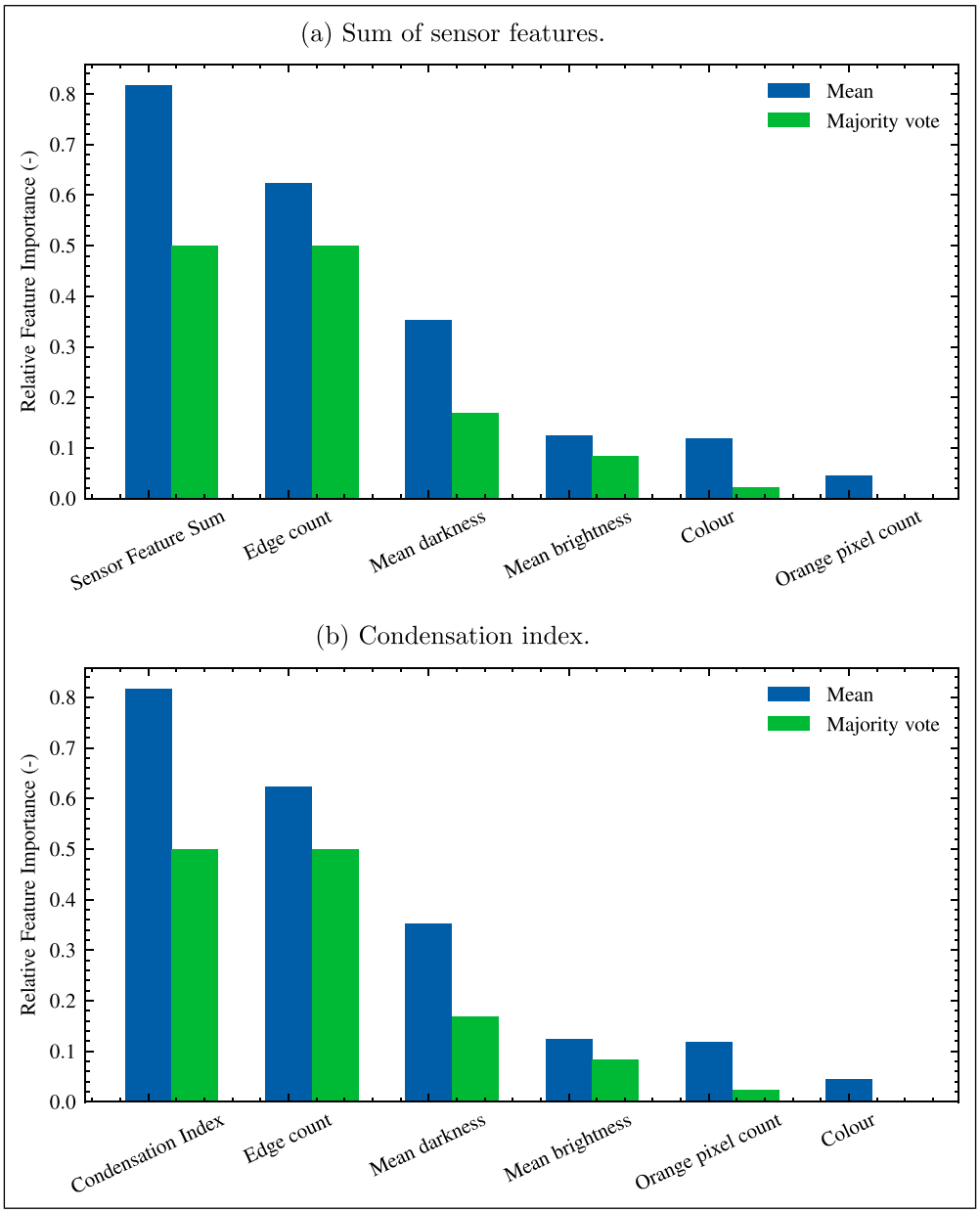


Figure 5. Relative feature importance for feature-reduced datasets: (a) Sum of sensor features; (b) Condensation Index.

and the variance associated with relatively small training datasets.

Conclusions

The management of wheel-rail adhesion presents a significant problem both economically and in terms of safety. Techniques such as weather forecasting can mitigate the impact of low-adhesion conditions somewhat, but the lack of high-resolution data limits the applicability to regional-scale forecasting. Mobile vehicle-mounted systems have been developed to allow real-time ML-based prediction of low wheel-rail friction with positive results however the well documented problem of black box ML models makes interpreting the model performance difficult. In this study, a dataset of sensor data and railhead images was used to develop an ensemble of ML models designed to predict the CoF at the wheel-rail interface. A variety of FI algorithms were then applied to these to produce an FI matrix, which was reduced into a more human-readable form using an ensemble feature importance fusion algorithm.

Application of the FI algorithms to the railhead dataset indicated that atmospheric properties such as air temperature, rail temperature, humidity and dew point were consistently ranked as the most important predictor of low adhesion conditions. This was found to occur across two separate dataset formulations. The relatively low contribution of the image features to the overall model performance was further reinforced by additional testing which showed only a small degradation in model performance when these features were removed from the training dataset entirely.

Of the features extracted from the image data, the edge detection algorithm was found to consistently outperform the other colour- and brightness-based features. A key recommendation for future work in this area is therefore to design feature extraction algorithms to detect larger multi-pixel image features which are more resistant to variation in colour and brightness as a result of local lighting conditions.

ORCID iDs

Samuel J. Cartwright https://orcid.org/0000-0002-3274-7649
Benjamin C. Rothwell https://orcid.org/0000-0003-2503-7232
Carol Eastwick https://orcid.org/0000-0001-5773-6439
Roger Lewis https://orcid.org/0000-0002-4300-0540
Joseph L. Lanigan https://orcid.org/0000-0001-9860-9559

Funding

The authors would like to gratefully acknowledge the EPSRC for the support under the Prosperity Partnership Grant Cornerstone: Mechanical Engineering Science to Enable Aero Propulsion Futures, Grant Ref: EP/R004951/1, and "Friction: The Tribology Enigma", EPSRC Programme Grant EP/R001766/1.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Seasonal Challenges Steering Group. *Managing low adhesion*. 7 ed.. Rail Delivery Group, 2024.
2. El Rashidy R, Chapman V, Lawcett L, et al. *Feasibility of integrating operational data with adhesion forecasts*. RSSB, 2020. Report COF-FCA-03.
3. Folorunso MO, Lewis R and Lanigan JL. Effects of temperature and humidity on railhead friction levels. *Proc Inst Mech Eng F J Rail Rapid Transit* 2023; 237(8): 1009–1024. https://doi.org/10.1177/09544097221148236
4. Folorunso MO, Watson M, Martin A, et al. A machine learning approach for real-time wheel-rail interface friction estimation. *J Tribol* 2023; 145(9): 091102. https://doi.org/10.1115/1.4062373
5. Beagley TM, McEwen IJ and Pritchard C. Wheel/rail adhesion — the influence of railhead debris. *Wear*. 1975; 33(1): 141–152. https://doi.org/10.1016/0043-1648(75)90230-6
6. Harrison HD, McCanney T and Cotter J. Recent developments in coefficient of friction measurements at the rail/wheel interface. *Wear*. 2002; 253(1–2): 114–123. https://doi.org/10.1016/S0043-1648(02)00090-X
7. Zhu Y. *Adhesion in the wheel-rail contact under contaminated conditions*. KTH Royal Institute of Technology, 2011.
8. Zhu Y, Olofsson U and Persson K. Investigation of factors influencing wheel–rail adhesion using a mini-traction machine. *Wear*. 2012; 292-293: 218–231. https://doi.org/10.1016/j.wear.2012.05.006
9. Zhu Y, Olofsson U and Söderberg A. Adhesion modeling in the wheel–rail contact under dry and lubricated conditions using measured 3d surfaces. *Tribol Int* 2013; 61: 1966953119. https://doi.org/10.1016/j.triboint.2012.11.022
10. Kalker JJ. A fast algorithm for the simplified theory of rolling contact. *Veh Syst Dyn* 1982; 11: 2021325992. https://doi.org/10.1080/00423118208968684
11. Kalker JJ. *Three-dimensional elastic bodies in rolling contact*. Springer, 1990.
12. Kalker JJ. Wheel-rail rolling contact theory. *Wear*. 1991; 144(1–2): 243–261. https://doi.org/10.1016/0043-1648(91)90018-P
13. Tomberger C, Dietmaier P, Sextro Wet al. Friction in wheel–rail contact: a model comprising interfacial fluids, surface roughness and temperature. *Wear*. 2011; 271(1–2): 2–12. https://doi.org/10.1016/j.wear.2010.10.025
14. Lewis R, Lewis SR, Zhu Y, et al. The modification of a slip resistance meter for measurement of railhead adhesion. *Proceedings of the Institution of Mechanical Engineers Part F Journal of Rail and Rapid Transit* 2013; 227: 196–200.
15. Rengasamy D, Rothwell BC and Figueredo GP. Towards a more reliable interpretation of machine learning outputs for safety-critical systems using feature importance fusion.

Applied Sciences 2021; 11(24): 11854. https://doi.org/10.3390/app112411854
16. Seshia SA, Sadigh D and Sastry SS. Toward verified artificial intelligence. *Commun ACM* 2022; 65(7): 46–55. https://doi.org/10.1145/3503914
17. Chakraborty S, Tomsett R, Raghavendra R, et al. Interpretability of deep learning models: a survey of results. 2017 SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI, San Francisco, CA, 4 August 2017, pp. 1–6.
18. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 2019; 1(5): 206–215. https://doi.org/10.1038/s42256-019-0048-x
19. Rengasamy D, Jimiama MM, Jimiama Mafeni Met al. Feature importance in machine learning models: a fuzzy information fusion approach. *Neurocomputing*. 2022; 511: 163–174. https://doi.org/10.1016/j.neucom.2022.09.053
20. Fang G, Liu W and Wang L. A machine learning approach to select features important to stroke prognosis. *Comput Biol Chem* 2020; 88: 107316. https://doi.org/10.1016/j.compbiolchem.2020.107316
21. Sanchez S, Rengasamy D, Hyde CJet al. Machine learning to determine the main factors affecting creep rates in laser powder bed fusion. *J Intell Manuf*. 2021; 32: 2353–2373. https://doi.org/10.1007/s10845-021-01785-0
22. Cammarota C and Pinto A. Variable selection and importance in presence of high collinearity: an application to the prediction of lean body mass from multi-frequency bioelectrical impedance. *J Appl Stat* 2021; 48(9): 1644–1658. https://doi.org/10.1080/02664763.2020.1763930
23. Kumar A, Mase JM, Rengasamy Det al. EFI: A toolbox for feature importance fusion and interpretation in python. *International Conference on Machine Learning, Optimization, and Data Science*. Springer, 2022, pp. 249–264.
24. Strobl C, Boulesteix A-L, Zeileis A, et al. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinf* 2007; 8(1): 25. https://doi.org/10.1186/1471-2105-8-25
25. Lundberg SM and Lee S-I. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst*. 2017; 30: 4768–4777.
26. Ribeiro MT, Singh S and Guestrin C. "why should i trust you?" explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, San Francisco, CA, 13-17 August 2016, pp. 1135–1144.
27. Lea D, Aguilar E, Sivakumar K, et al. *Biomaterials-for-medical-devices-ai/biofefi*. Zenodo, 2025.
28. Folorunso MO, Watson M, Tomlinson K, et al. Real-time railhead friction estimation using machine learning: development of an on-board-train data capture system. *J Tribol* 2026; 148(8): 081114. https://doi.org/10.1115/1.4071160
29. Sobel I. An isotropic 3x3 image gradient operator. *Presentation at Stanford A.I. Project*. 1968; 1968.