Brake System Fault Detection via Hot Wheel & Brake Shoe Algorithm

Dean G. Kezurer

Adriano Polpo[1], Debora Correa[2], Melinda Hodkiewicz[3]
1. Mathematics and Statistics
2. Science and Software Engineering
3. Mechanical Engineering
The University of Western Australia

John McArthur, Bradley McCarthy
CEED Client: Rio Tinto

Abstract

This paper evaluates a novel approach to detect faults in wagon brake systems by using an auto-encoder framework on hot wheel & brake shoe data. Auto-encoders are a deep learning architecture that can be used to identify anomalous data sequences, which, in the current study, are characterised by hot-wheel events. The data used within this investigation was sourced daily from 7,670 unique wagons over the course of 2023, containing only 22 unique hot-wheel events. Simulated data is used to test the approach, and early results in real data suggests that model performance is very dependent on hyperparameter optimisation.

1. Introduction

The integrity of brake systems are essential in locomotive operations. Ore trains travel along specified routes, occasionally needing to brake to avoid high speeds or dangerous terrain. Due to a combination of these braking systems and environmental conditions, many wagons experience abnormally elevated wheel temperatures described as “hot wheel events”. These events arise from improper maintenance of wheel brake pads or incorrect driver strategy, and are typified by wheel temperatures of 200 degrees Celsius or higher. “Hot wheel events” are detrimental to the business as they oftentimes lead to mainline delays, which in turn will slow or even halt operations.

1.1 About the Data

Due to the adoption of technological on-track equipment, Rio Tinto Iron Ore (RTIO) possesses masses of real-time data describing their asset fleet. Leveraging the currently available data, it is possible to create a snapshot overview of wagon behaviour over the span of 2023. The current data contains 1,824,768 unique, daily timestamped observations across 7,670 distinct wagons, with key descriptive features including wheel temperature recordings and brake shoe thickness measurements. A hot wheel event boolean indicator has also been integrated into the dataset, with only 22 uniquely recorded events for the full dataset. The features of the resulting data are described as follows:

- Wheel Temperature: The maximum temperature recorded for a wagon on a day
- Speed: The maximum recorded speed into a detector for a wagon on a day
- Wagon Wheel-Set Difference: Difference of mean temperature of left sided wheels and right sided wheels
- Target: The response variable, which was taken as 1 if an event occurred on that day for that wagon.

2. Literature Review

Time series anomaly detection (TSAD) is a contemporary process for identifying anomalous sequences within time-indexed data. TSAD approaches are predicated on the assumption that anomalous data has a unique underlying data representation compared to typical data, and as such these methods can be used to identify time windows that contain abnormal data. Typically, TSAD methods utilise the continuous nature of time-indexed data to develop profiles of expected trends, and then use deviation metrics to identify sections of the input data that differ from the norm.

A recent literature survey examined the application of 71 (with 10 algorithms not producing results) TSAD algorithms on 976 real-world and synthetic time-series datasets. The study discovered that amongst the contemporary literature, there is a significant preference for the application of unsupervised algorithms, with 39 out of 61 applied algorithms being unsupervised and a further 19 being semi-supervised (Schmidl, 2022). Schimidl et al. further posits that reconstruction based approaches can be highly effective in identifying anomalous patterns while being less sensitive to the magnitude of deviations, with the Donut approach scoring an average AUC-ROC score of 0.9891 across all datasets. Despite the performance of these techniques, the study shows that most anomaly detection algorithms were sensitive to their parameter settings and required on average seven specifications per framework.

Encoder-Decoder neural network architectures are unsupervised deep learning frameworks used commonly in anomaly detection. Contemporary literature describes a number of techniques to learn data patterns, with the most simple of the cohort being an Auto-Encoder. Auto-encoder modelling techniques involve the data being read into a model, decomposed into lower dimensional space and then reconstructed based on the learnt feature weights. Auto-encoders are an effective deep learning framework within the current investigative space, as they can accurately model imbalanced anomaly data. Since this technique does not require labelled data, anomaly detection performance is measured by the reconstruction loss, a measure that describes the distances of reconstructed feature values from their true values. As such, an anomaly detection model can be characterised by training an auto-encoder on the 'normal' representation of the data, which consequently results in anomalous test data achieving abnormally large reconstruction loss. By applying this approach to real-world data, a 2021 study demonstrated that an unsupervised autoencoder method achieved a F1 score exceeding 0.95 on two distinct sample datasets (Jennings, 2021). These results underscore the effectiveness of the method, though it is important to recognize that the performance is contingent on the quality and characteristics of the available data.

2.1 Project Objectives

The research questions for this project are:

- How effective is a deep learning autoencoder framework in detecting anomalous sequences within the current data?
- How does the data affect modelling processes, given that it is inherently imbalanced and is time-series in nature?

3. Methodology

3.1 Data Cleaning and Preprocessing

3.1.1 Explanatory Features

Based on scoping discussions with the client, a subset of important data features were identified as essential in describing typical and atypical wagon behaviour. The subset of vital data features includes wagon wheel temperature and brake shoe thickness recordings

Wheel Temperature data is stored as a time-indexed recording within a software called PhoenixCMS. Due to data constraints, an extraction pipeline was designed to process downloaded XML files containing individual wheel temperature recordings per location and generate 'train' objects containing important features. Around two years of measurements were successfully extracted from the available data spanning 2022 and 2023, with a significant amount of missing-not-at-random data within 2022 (see Fig 1). Based on this insight, and the surplus of available wagons, an assumption was made that the subsequent model will be solely trained on data from 2023.

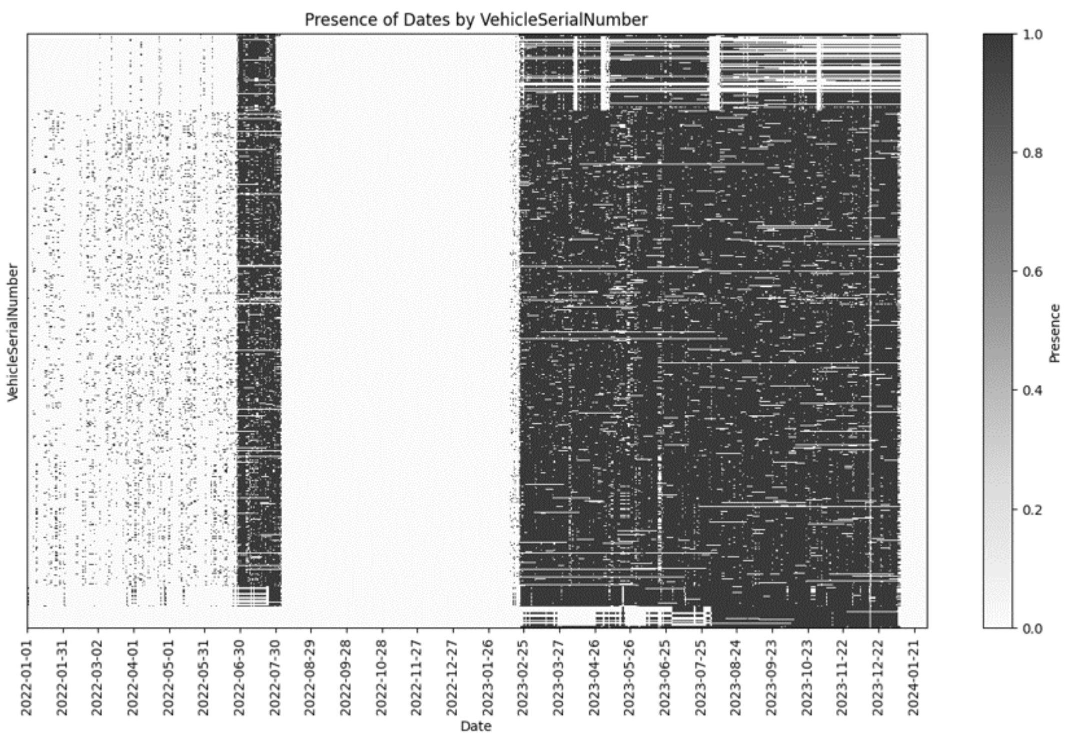

Figure 1 Available wheel temperature data.

Each object was initially extracted at the granularity of an individual wheel however, these recordings were subsequently grouped per wagon on a composite key of site location, time stamp, wagon identification number as well as train identification number. With the temperature data collated and optimised, the extraction pipeline was subsequently enriched by adding an associated trainsheet identification number.

Brake shoe thickness measurements are taken twice per train trip, once as it is leaving the port and again as it returns, and record the size of the thickness at the top as well as the bottom of the brake pad of each wheel. This data was subsequently grouped on trainsheet identification number and wagon identification number, taking the minimum recorded thickness per wheel on a wagon. A minimum thickness was decided as the grouping condition as it better resembled the characteristics of atypical event-related data. The resulting data was then left joined onto the aforementioned temperature data via the composite key of: trainsheet identification number, train identification and wagon identification number.

3.1.1 Response Feature

Anomalous data within this project is signposted by the occurrence of a hot wheel event and stored within RTIO's data warehouse. Once filtered down to 2023, the events table contains 106 uniquely saved hot wheel events. The available information was thereafter extracted and grouped on a per day, per wagon basis to align with the current data schema. Due to the significant filtering a completed data frame only contained 22 unique events across 7,670 unique wagons within 2023.

3.1.2 Model Data Preparation

Data was then discretized at the granularity of a single day as it allowed for regular time stepping intervals whilst maintaining data quantity. Due to the latent missing-at-random data, a basic imputation method of backwards fill, followed by a forward fill was used to ensure regular time steps. This method was chosen over a mean / median fill as it maintained simplicity while ensuring that the data adhered to the existing structure. Sliding windows were subsequently used to define data sequences for anomaly detection. The sliding window size was set at three days prior to a recording and were implemented for all explanatory variables.

3.2 Experiments and Results

The base model used in this investigation was an autoencoder, as its flexibility is preferred for anomaly detection with the current data. An autoencoder framework will ingest the available data, decompose it into a smaller, pre-specified feature space and then rebuild the data based on the learned latent weight representations. By using a Mean Squared Error loss function, it is subsequently possible to calculate a reconstruction error in which a distance measurement is taken between the reconstructed and original features. It is therefore hypothesised that training an autoencoder on typical wagon functionality will result in larger reconstruction errors for anomalous data.

A simulated dataset based on the distributions of the current data proved that an autoencoder is effective in separating anomalous from typical data sequences. By using a basic two-layer autoencoder outlined in Table 1, the deep learning framework output a Gaussian distribution of reconstruction errors centred at 0.644 for normal data and a bimodal distribution of reconstructions errors centred at 4.499 and 8.501 for anomalous data sequences (see Fig 2a).

Hyperparameters	Simulated Model	Current Model
Number of Unique Layers	2	2
Perceptron Sequence	[60, 30, 15]	[60, 30, 15]
Loss Function	MSE	MSE
Epochs	50	50
Learning Rate	1e-3	1e-5
L2 Regularisation	None	1e-5

Table 1 Comparison of hyperparameters between simulated and current models

Implementation of the above framework was effective in simulated conditions, however due to the distribution of the available data, hyperparameter optimisation was required for an applied base model. Model training was also set to 50 epochs per model, as majority of tests plateaued at this stage and a typical 50 epoch model took, on average, 30 minutes to train. By leveraging the capabilities of a 10-fold cross validation grid search, it was found that hyperparameter optimisation was vital in achieving model convergence. This process exposed the optimised hyperparameters described in Table 1, which resulted in the base model being able to begin to separate normal and abnormal data sequences (as seen in Fig 2b). Furthermore, the mean reconstruction error of typical data sequences was 0.59339 in the current model, whereas atypical data sequences was 1.34725.

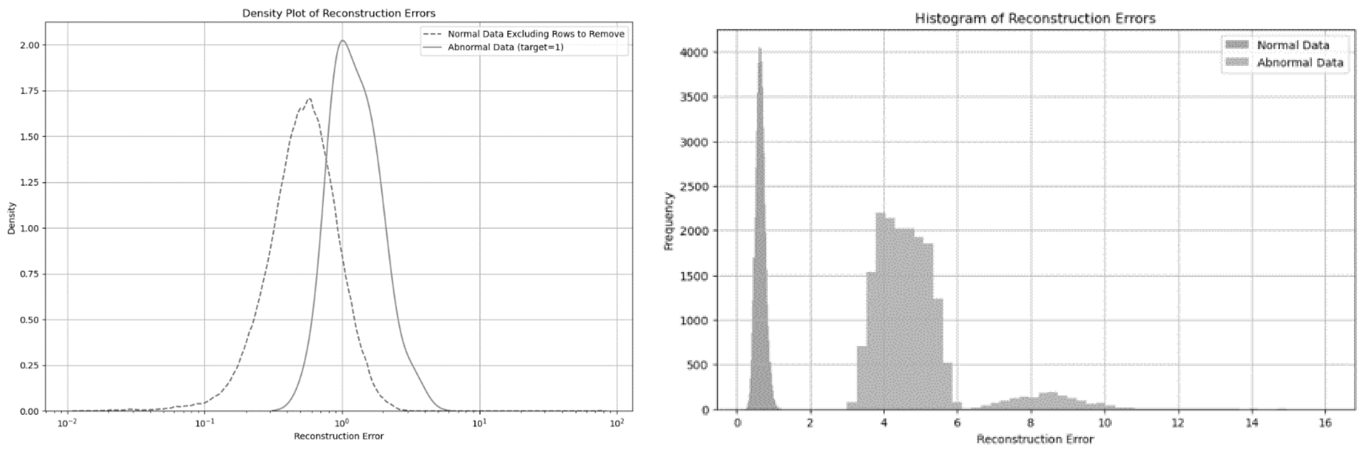

Figure 2 Comparison of simulated and real data results.

Plotting the reconstruction error for a wagon over time highlights the base model's ability to identify abnormal data sequences. The current base model configuration is successful in producing time series plots in which a hot wheel event appears to be at the end of an increasing reconstruction error sequence (as evident in Fig 3a). Despite this, there are also a few wagons that experience hot wheel events which are not visually located at a reconstruction error peak, suggesting that hyperparameter and model architecture optimisation is needed (see Fig 3b).

4. Conclusions and Future Work

The current results highlight the efficacy of an autoencoder framework on the available data. Despite having used simple imputation techniques and assumptions within the data pre-processing pipelines, a basic autoencoder model was able to successfully learn the latent representation of typical wagon data. Although the model is promising it does need significant fine tuning, as there is still overlap between the normal and abnormal wagon data. Future work

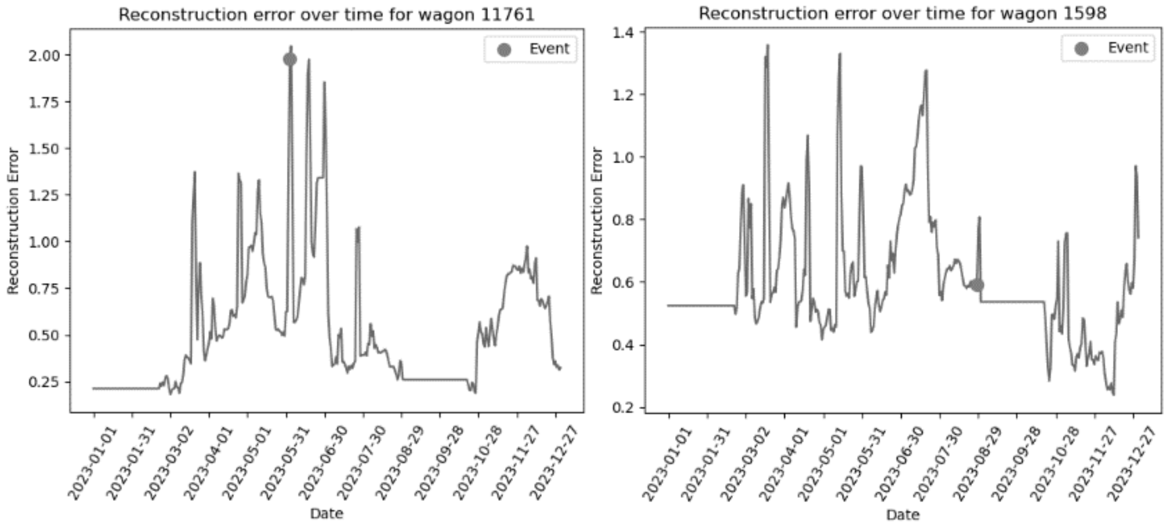

Figure 3 Comparison of simulated and real data results.

on this project can include model optimisation techniques such as experimenting with the exact architecture, as well as employing more complex frameworks such as Variational and Long-Short Term Memory autoencoders. It is also possible to expand this model into the wider class of braking related faults, thus increasing the prevalence of events within the data and therefore building a new profile for anomalous data.

5. Acknowledgements

I would like to express my gratitude to my family for their unwavering support throughout this academic journey. My heartfelt thanks also go to my academic supervisors, Adriano, Debora, and Melinda, for their invaluable guidance. I am deeply appreciative of my client mentors, Brad McCarthy, John McCarthur, Matthew Wojas and Milan Pareek, for their insights and encouragement. Lastly, I acknowledge the contributions of Jeremy Leggoe and Kimberlie Hancock from CEED for their dedication and organizational support.

6. References

Jennings, S. M. (2021). Unsupervised anomaly detection with LSTM autoencoders using statistical data-filtering. *Applied Soft Computing, 108*(1568-4946), 107443. doi:https://doi.org/10.1016/j.asoc.2021.107443

Schmidl, S. a. (2022). Anomaly detection in time series: a comprehensive evaluation. *Proc. VLDB Endow., 15*(2150-8097), 1779–1797. doi:10.14778/3538598.3538602