# Ore Car Bogie Dynamics Assessment with Wayside Condition Monitoring Data Analysis

Callan Brogan

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

Andrew Stevens, John McArthur
CEED Client: Rio Tinto

**Abstract**

*The project attempts to define performance groups within Rio Tinto condition monitoring data to describe characteristic train bogie dynamics. As there is no defined response variable this constitutes an unsupervised learning problem. Characteristic groups within the data are defined using a clustering algorithm coded in python. This defines similarity between bogie pairs using the dynamic time warping similarity metric and applies agglomerative hierarchical clustering to the resulting similarity matrix. Results describe several characteristic bogie dynamics in bogie geometry and wheel profile.*

## 1. Introduction

Rio Tinto operates over 13,000 ore cars over 1,700km of rail in Western Australia to ferry product from mines in the Pilbara to ports at Dampier and Port Headland. The bogie suspension sets are due for overhaul in coming years, with wear on friction shoes cited as the major criterion for overhaul. This project characterises bogie dynamic performance from an analysis of wayside condition monitoring data. This is to aid the development of an index to advise on the preference of bogies to be overhauled. The aim of this project is to define, from an unsupervised learning position, characteristic performance groups within condition monitoring data, and associate the corresponding bogie dynamics with friction wedge and wheel profile wear. The project clusters time series data using dynamic time warping and, with the aid of the Rio Tinto rolling stock team, interprets the resulting clusters as bogie performance groups. The intended result is a taxonomy of bogie performance groups and relationships between groups and flange wear.

## 2. Process

### 2.1 Condition Monitoring Data

The condition monitoring data used in this project describes the bogie geometry and the profile of the wheelsets. Measurements are taken at the Automatic Roll By (ARB) 29.6 site in Dampier using a machine vision system, which takes high speed photos of the bogies during roll by. The images are processed to collect component measurements. These measurements include axle tracking, angle of attack, flange width and height and friction wedge height. Car geometry data is taken for each of a bogie's two axles, and wheel profile data is taken for each of the four wheels. Discrete data, such as train direction and ore car lagging or leading, and categorical variables, such as bogie age and position within the train, describe the context of each measurement.

Condition monitoring data recorded by Rio Tinto's automatic roll by system is poorly conditioned for data analysis. The post processing of the collected machine vision images relies on reference points, which often slip and result in a significant variance within the data. The bogie components also wear very gradually, and the combination of high variance and minimal trend complicates analysis. Data is taken at irregular intervals as train schedules are subject to change and ARB data is only taken at the site outside Dampier. Bogie condition is therefore represented by a multivariate time series with irregular time intervals.

## 2.2 Determining Similarity

Establishing performance groups relies on quantifying the similarity of bogies from available data. Dynamic time warping is a metric which is used to calculate the similarity between time series pairs. It is particularly popular in literature (Moser & Schramm, 2019a) as it may be applied to whole time series of different lengths without first extracting features. Time sequences to be considered in clustering were formed with a sliding window of length 30 days and overlap of 50%. Dynamic time warping was then applied to every pair of data windows to form a 2D similarity matrix for each signal. The project used the dtaidistance (DTAIDistance 2.3.9 Documentation, 2018) Python library in the processing script to create signal similarity matrices.

Of interest to this project is the multivariate nature of the data. Ideally, the similarity of bogie pairs would be calculated from multiple data signals, so that similar bogies have a similar set of patterns in all available data signals. This may be achieved by summing independent signal similarity matrices elementwise to form a global similarity matrix. Clustering with all signals, however, yielded clusters which were messy and unintelligible. Instead, the project considered signals of the same measurements assembled into feature groups. The tracking feature group, for example, is composed of the tracking data of the two axles for the bogie. The result is a set of feature similarity matrices, each describing the similarity between all pairs of time windows for the relevant feature data. This approach assumes that the patterns and dynamics of feature groups are independent, and relationships between feature performance groups may be considered separately. This assumption is a simplification of reality but a necessary step at this early stage of analysis.

## 2.3 Clustering Feature Groups

To establish bogie clusters the resulting similarity matrix is input into an agglomerative hierarchical clustering algorithm, which treats each object as a singleton cluster and systematically forms new clusters from pairs most similar to each other. Each formed cluster is then assigned a similarity score calculated from those of its parents, and the process repeated until all objects are collected in one cluster. The output of this process is a dendrogram which describes the cluster formation process, as shown in Figure 1. Objects being clustered are listed on the y-axis, and the x-axis is the dissimilarity between connected clusters.

Clusters are defined with an appropriate similarity threshold, informed by the dendrogram. Sets of objects connected to each other with dissimilarity lower than this threshold are defined as clusters. The advantage of hierarchical clustering over other methods is that it does not assume the number of appropriate clusters and force objects to assign themselves to a cluster, as is common with density-based clustering mechanisms. This allows outlier objects to remain as singleton clusters, unassigned to any group.

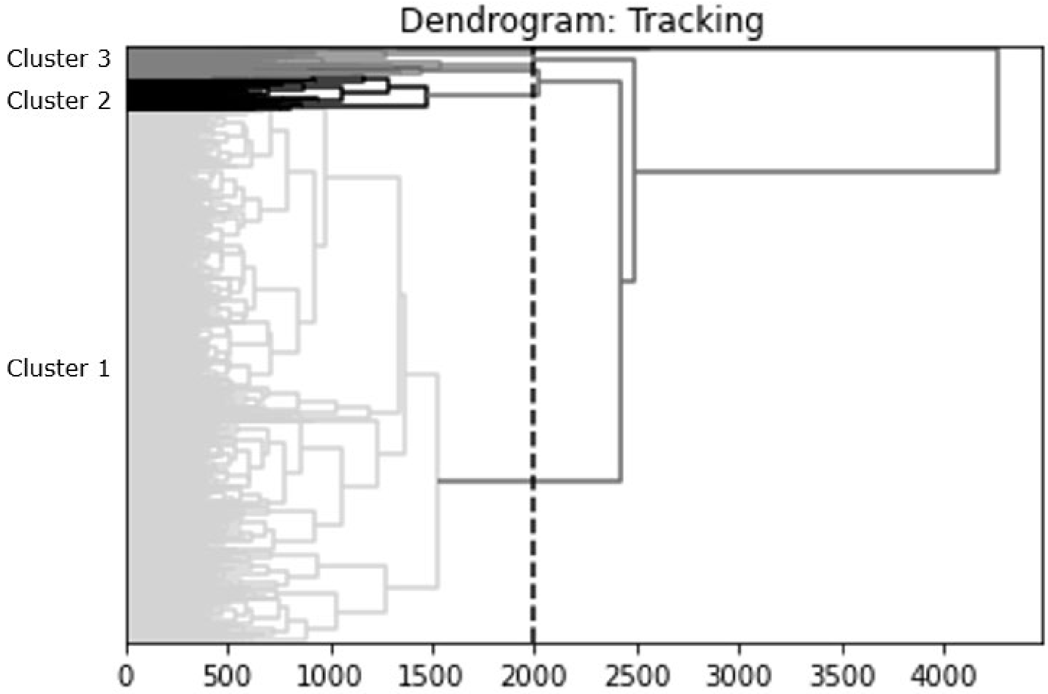


**Figure 1** Example Dendrogram. Three clusters formed from a similarity threshold of 2000

# 3. Results and Discussion

## 3.1 Established Clusters

The clustering algorithm was run with ARB data from January to April of 2024 for 500 ore car pairs. This data was split into 2000 bogies and 10000 data windows, and clustered by feature groups as described in section 2 above. The algorithm had difficulty classifying bogie pairs with the same data shape or pattern as similar if there was a significant difference in the mean of the two datasets. Instead, the dendrograms were purposefully cut at an early similarity threshold and the resulting clusters were sorted into performance groups with the author's supervision. The clustering threshold and number of clusters per feature group are listed in Table 1 below.

The following is a description of a selection of the identified performance patterns and a brief interpretation. Examples of bogies from each performance group are shown in figures 2, 3 and 4. In each case, the first of the supervised clusters for each feature group refers to normal functioning of the bogie, and the last group collects all unassigned singleton clusters from dendrogram as an anomaly group.

| Feature Group | No. Signals | Similarity Threshold | No. Raw Clusters | No. Supervised Clusters |
|---|---|---|---|---|
| Tracking | 2 | 1000 | 29 | 4 |
| Angle of Attack | 2 | 200 | 28 | 4 |
| Flange Height | 4 | 700 | 13 | 4 |
| Flange Width | 4 | 800 | 14 | 4 |
| Wedge Hight | 4 | 2000 | 16 | 6 |

**Table 1** List of established feature group clusters

## 3.2 Tracking

Tracking refers to the offset left or right of an axle with respect to the centreline of the rail. Data representing regular functioning has a mean of zero and low variance. The tracking feature group produced 3 significant groups. Groups 2 and 3 are of groups of data windows where the mean of the tracking of one axle is significantly different to the other. The group shown in Figure 2 appears to represent bogies which sit at an angle with respect to the rail, possibly due to asymmetrical wear on wheel flanges.

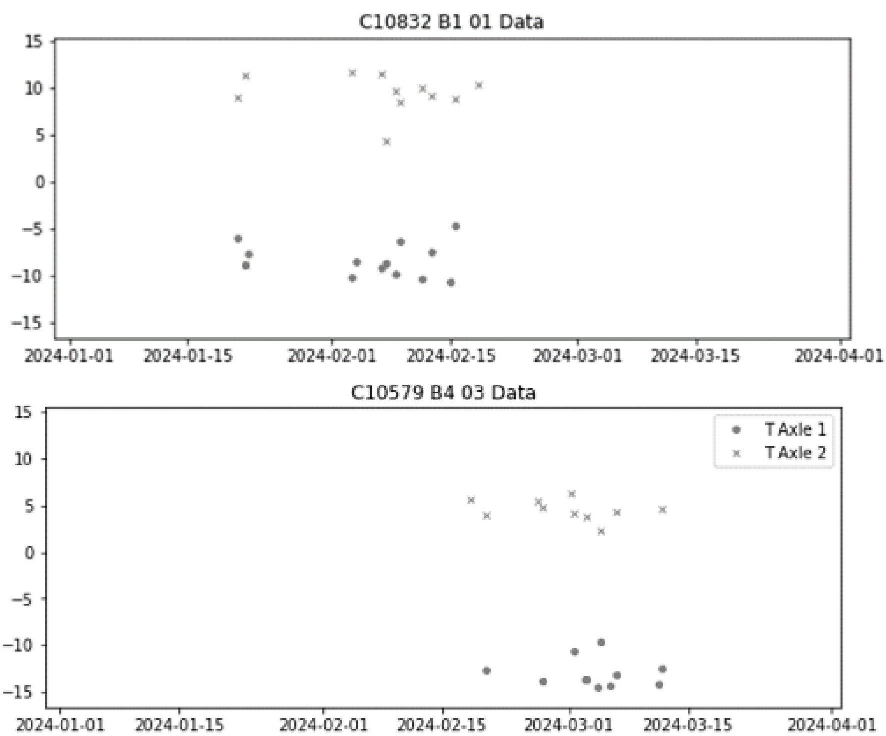


**Figure 2** Tracking data windows for bogie 1 of ore car C10832 and bogie 4 of ore car C10579, members of Tracking performance group 2.

### 3.3 Average Angle of Attack

Angle of attack refers to the angle an axle makes with the centreline of the rail. Data representing regular function dynamics has a mean of zero and minimal variance. The clustering algorithm identified three significant performance groups for the average angle of attack data. Groups 2 and 3 have one axle with two distinct means, which sit on either side of the data of the second axle, which behaves normally. These groups, captured in Figure 3, appear to represent bogies with one axle oscillating left and right, pivoting against the second axle. This may be indicative of hunting, a common train bogie dynamic.

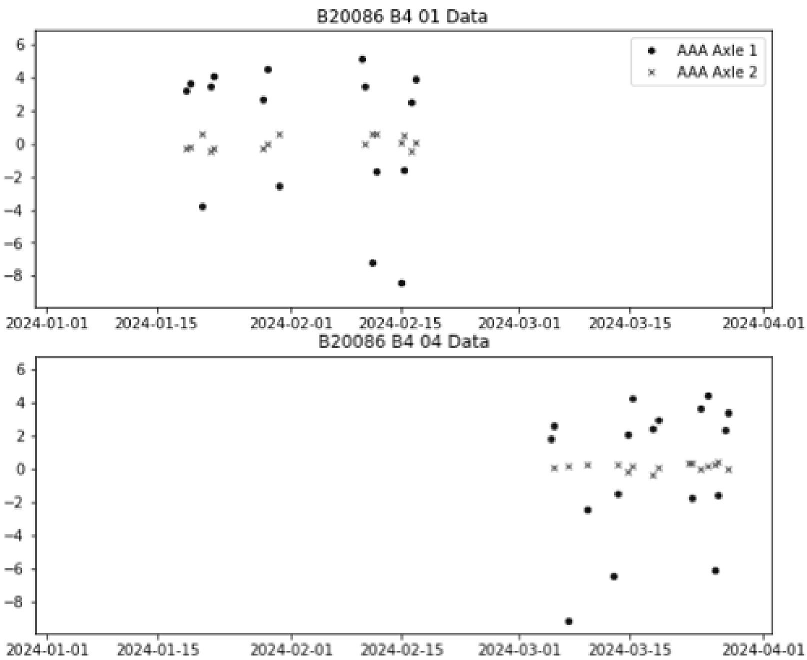


**Figure 3** Average angle of attack data windows for bogie 4 of the B20086 ore car, member of average angle of attack performance group 3.

### 3.4 Flange Height

The profile of the bogie wheels is described by the height and width of the flange. Data representing regular functioning has all four flange heights with a similar mean and minimal trend. There are three significant flange height groups identified from the output of the clustering algorithm. The second group is characterised by a sharp transition, and represents changeout of the wheel sets, as seen in Figure 4. Other groups are of combinations of wheel flange height means, which represent either asymmetric wearing or poorly paired wheelsets.

## 4. Conclusions and Future Work

The project has been successful in defining feature performance groups with wayside condition monitoring data. The project has made use of dynamic time warping to define similarity between bogie datasets, and agglomerative hierarchical clustering to form characteristic groups. Results from considering all signals in the clustering process proved to be too noisy, instead the project clustered feature groups independently.

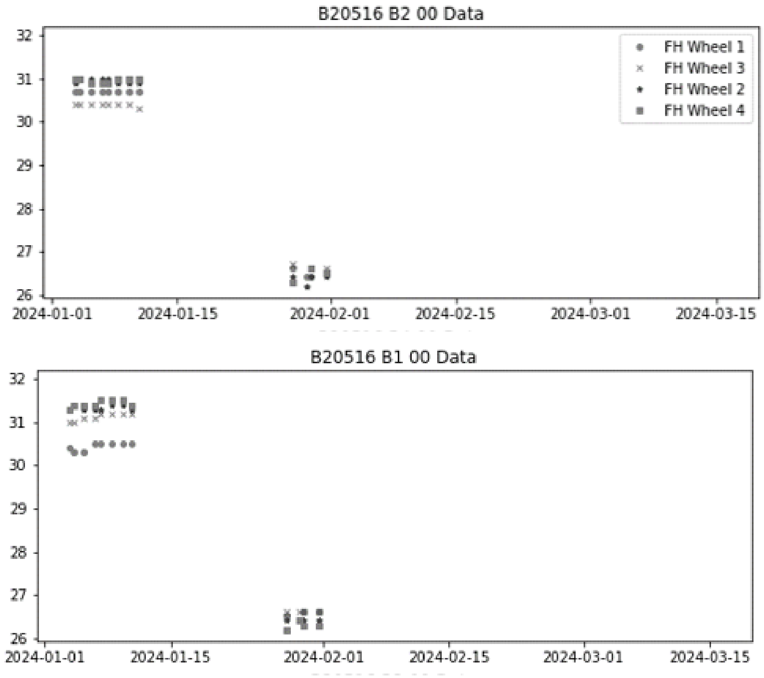


**Figure 4** Flange height data for bogies 1 and 2 of the B20516 ore car pair, member of flange height performance group 2.

Future work for this project will involve attempting to better describe and characterise the defined feature performance groups. The groups so far described are assumed independent of each other, and the degree to which the groups are interrelated will need to be evaluated. This will include considering relationships between groups, likely through the analysis of a conditional probability matrix of cluster membership. Future work will also involve defining the relationship between flange wear rate and the membership to performance groups, as will be necessary for describing the risk associated with bogie membership in the various performance groups.

# 5. Acknowledgements

I would like to acknowledge Rio Tinto for supplying the project and data, and in particular Andrew Stevens and John McArthur, who represented Rio Tinto and liaised with me to deliver results. I would also like to acknowledge academic supervisors who supported this project, Melinda Hodkiewicz, Adriano Polpo and Debora Correa.

# 6. References

Moser, U., & Schramm, D. (2019a). Multivariate dynamic time warping in automotive applications: A review. Intelligent Data Analysis, 23. https://doi.org/10.3233

Welcome to DTAIDistance's documentation! — DTAIDistance 2.3.9 documentation. (2018). Readthedocs.io. https://dtaidistance.readthedocs.io/en/latest/