Extreme Weather Performance Impact

Group Project - Data Science Fundamentals - Lancaster University

DIMITRIOS PRIOVOLOS, BOYUAN WANG, REKA LANTOS, SAYANTEE DUTTA, JOYCE WAMWEA, and DIMITRA GIORKATZII

Project name: `Weather trends against railway on-time performance`
Project number: `10`
Company involved: `Avanti West Coast`

1 INTRODUCTION

Railway transportation has always been a useful means of transportation for both people and goods over both short and long distances. Most nations' transportation networks are vulnerable to a wide range of weather extremes, including flooding, thunderstorms, temperature, and high winds [5]. Many weather phenomena can have a significant impact on public transportation, such as causing delays, which in turn lowers the quality of service. Extreme weather events are so prevalent in the UK and may have a significant influence on railway infrastructure. As a result, infrastructure owners and operators must monitor current weather impacts and establish appropriate actions to prevent railway delays and any other consequences that the weather phenomena may have [7]. By looking at historical data, the way that the railway performance is affected by different weather conditions can be determined, and appropriate measures can be predicted in order to minimise the consequences of those weather phenomena.[6]

This research focuses on the effects that extreme weather (flooding and high winds in particular) have on railway transportation lines. Specifically, the lines that were examined are the Euston – Scotland via West Midlands and Euston – Scotland direct (via Trent Valley). Data from Avanti West Coast, the project's collaborating organization, was used to conduct the analysis. The Avanti West Coast rail services serves a vast portion of the nation, and thus weather-related decisions often impact a lot of passengers and also affect overall railway performance on any particular day. One of those decisions, which is the main measure that will be examined in this report, is the enforcement of blanket speed restrictions. Blanket speed restriction is used when it is necessary for trains to run slower over a large area, and are vital for people safety and for the avoidance of collisions. This means that it is very important that the enforcement of blanket speed restrictions are predicted accurately and in time.

The questions we sought to address were as follows:

(1) What is the relationship between the extreme weather (flooding and high winds) and the Avanti West Coast North of Preston railway performance? What has changed in terms of performance?
(2) How does the introduction of blanket temporary speed restrictions north of Preston in response to extreme weather affect performance?
(3) What was the benefit of the emergency timetable work that took place from October 27th to October 30th, 2021, in response to weather-related temporary speed restrictions?
(4) How do we know whether and when NR should implement a weather-related emergency speed restriction?

This study focuses on the region north of Preston. The data provided contained information about operational decisions, effects of the actions taken and weather conditions.These data were used to evaluate the objectives listed above.

The following section presents the procedure and methods used to answer the research questions. Steps followed in conducting the analysis as suggested by Wickham, and Grolemu [12] includes aggregating the data, pre-processing, exploratory analysis and modelling. We began by aggregating data from the different files. Then, we pre-processed the data, only preserving the information that was relevant to our analysis. In order to make analysing the data more manageable, the data was divided into four segments, each addressing a specific question. Finally, we summarised the research's findings using plots and charts. In the last section, we discuss future work that could lead to a better understanding of the research questions.

2 METHODOLOGY

2.1 Data Selection

Data selection is the selection of appropriate training data and can be as important as the choice of algorithm[10]. Having engaged with the AWC team multiple times to understand the problems well and through exploring the data provided, we narrowed down the data and number of features. We were particularly interested in two types of data: weather and railway performance, especially weather-related delays, as well as the performance at times when emergency and blanket speed restrictions were introduced. Additionally, due to the nature of the research questions, we needed to identify different sets of data that could be used for the analysis and the modelling, as well as to cover specific locations and time periods.

The selected data included data from different weather stations which had features such as wind speed, air temperature and pressure, rainfall etc., and any data pertaining to delays caused by extreme weather (flooding and high winds) or other factors (high rail temperature, flooding not due to extreme weather etc.)

2.2 Data Pre-processing

Data pre-processing was an essential phase of this research since if the data is not cleaned and formatted properly, the analysis and the models could result in misleading conclusions. The fundamental goal of pre-processing is to end up with clean, meaningful data so that more accurate findings may be obtained[3].

The following steps were followed when pre-processing the data.

(1) Firstly, feature selection was implemented, thus we chose to investigate two routes, A and F of Scotland in more detail (filtered by their Bugle profit centre code), as well as specific highlighted periods such as the 27th and 30th of October 2021 (selected based on the journey dates).
(2) The next step was to convert all of the string values that represented dates into date-time types so that they could be handled like dates.
(3) Next, all data sets of weather were integrated into a single set and likewise, all data sets of delays were integrated into a single one. The combined set, which included all delay data sets were then sorted by dates and some of the features renamed for the purpose of the analysis.
(4) In the data delays dataset, all values of the feature 'Memo' were converted to uppercase. This was done to make dividing the data set into two distinct sets easier
(5) Observations with keywords like 'blanket', 'restrictions' and 'ESR' were merged into a new data set that contains all of the instances when blanket speed restrictions were enforced
(6) Finally, the remaining observations were merged into a new data set that contained all the instances where blanket speed restrictions were not applied

The pre-processed data, made it easier to analyze the data.

2.3 Research Strategy

The first 3 questions were examined and analyzed using graphics. This allowed us to compare multiple datasets visually and establish trends. Three different types of graphs were utilized. The first type is a box plot which can display the maximum, minimum, median, and upper and lower quartiles of a set of data [9]. The second type is a histogram which shows the frequency of each delay. A scatter plot is the third type, it depicts the relationship between two variables in the data set [1].

Python was used to build a few functions that returned some of these three graphs depending on the question we were trying to focus on. With the goal of examining either specific dates or all dates at once and the journey of specific people or all people at once, the functions returned some of these three graphs.

To address Question 4, we utilised a classification algorithm known as a logistic regression classifier. A logistic regression classifier was used to predict a dichotomous dependent variable. It utilized the maximum-likelihood ratio to determine the significance of the independent variables. The logistic regression model for p predictor variables is defined as follows:

$$P(Y = 1) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+...+\beta_p x_p)}}$$

where $P(Y = 1)$ denotes the probability of a data point belonging to a particular class and β_0,$\beta 1$,...,β_p are regression coefficients[11]. Generally, a threshold value is selected based on which the data points are segregated into different clusters.

For binary classification with a balanced dataset, as is our case, the threshold chosen is usually 0.5 [4]. Our model considered weather data with various features such as Wind Speed, Air Pressure, Relative Humidity etc. and predicted the value of a new variable, which indicated whether restrictions should be applied (0/1). To make the weather data comparable with the periods with restricted and unrestricted performance, we used re-sampling (on an hourly basis in this case). After this process, we merged the weather data and corresponding delay data and then created a new variable that indicated whether there was a restriction applied. The value 0 for the dependent variable corresponded to not applying any speed restrictions and 1 meant that speed restrictions should be applied. We then split the merged data into two sets- 70% of the data was used as a training set and the rest of the data was used as a test set.

Additionally to the logistic regression, we created a model using neural networks (NN) to improve the accuracy of possible predictions. NN can be well applied for classification, especially if the classes cannot be linearly separated and prediction problems, which made this application suitable for our research question[2]. To answer question 4, we constructed a sequential model, which is a multi-layer deep learning model built from a linear stack of layers, as it is useful for binary classification as well as time-series data, especially when the predictions can be time sensitive[8]. During the model construction, multiple layers were added on as well as drop-out layers to prevent overfitting. The data was transformed, re-sampled and merged in the same way as it was for the logistic regression, but in this case 80% of the data was used for training, as this split delivered the most optimal result.

3 RESULTS

3.1 Results for Question 1

We applied the procedure described above for Question 1 and we chose to compare the results of the delays based on all dates. From the box plot we obtained the following results:

- **Bad Weather delay (mins)**:
  - mean: 5.537
  - standard deviation: 10.958
  - media: 3.0
- **Other Speed delay (mins)**:
  - mean: 4.642
  - standard deviation: 6.798
  - media: 3.0

We demonstrate the results by using Fig. 1, in which the left box plot describes the delays of extreme weather and the right one the other speed delays, both based on all journeys and dates.

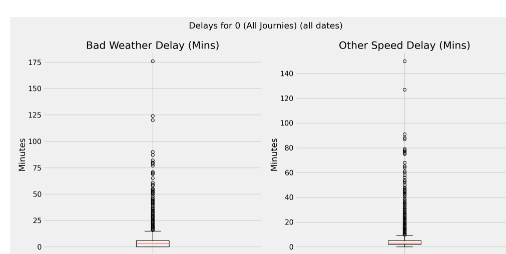

Fig. 1. Boxplot of bad weather delays vs other delays relying on all journeys and dates.

Here we can see that delays due to severe weather have a bigger range than the normal weather delays. Specifically, delays related to bad weather can surpass 50 minutes, while the other delays have a maximum value of 30 minutes. The same results can be derived from Fig. 2 as well, where we can see that normal delays were fluctuating at 0-15 minutes and the maximum delay was 30 minutes. On the other hand, most extreme weather delays were in the range of 0-15 minutes too, but there is a minor number of delays of 15-30 minutes and the maximum delay is over 50 minutes. In Fig.2, we can observe in a more clear way the frequency of the delays. It is shown that even though the delays due to severe weather are bigger, there are fewer of them. However, this could be due to severe weather occurring with less frequency.

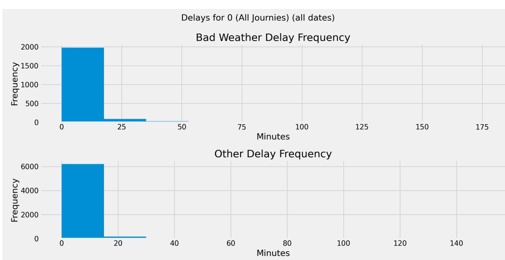

Fig. 2. Histogram of bad weather delays vs other delays relying on all journeys and dates.

3.2 Results for Question 2

For Question 2, we compared the cases where blanket speed restrictions were implemented, and blanket speed restrictions were not implemented. The comparison of these two cases was made by examining all train lines in two different time periods (December 2020, October 2021)

From the scatter plot in Fig. 3, on December 2020, it can be seen that the blanket speed restriction greatly improved the railway's performance. For the cases that speed restrictions were introduced, it is shown that most of the delays were under 15 minutes, and the overall number of delays was significantly lower than the cases with no speed limits. Moreover, the biggest delays that happened with speed limits enforced, were two cases of 50-55 minutes. For the non-speed restriction cases, there were a large number of delays longer than 15 minutes, and the longest delays lasted nearly 80 minutes.

Overall, trains with temporary speed restrictions had an average delay time of 3.632, a standard deviation of 4.178, and a median of 3. Trains

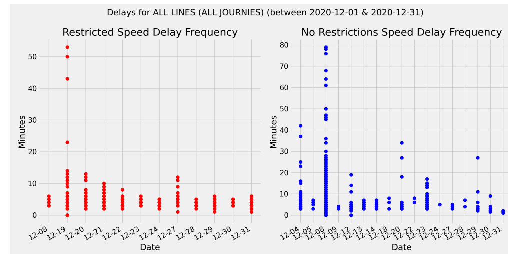

Fig. 3. Scatter chart showing all line delays for December 2020

without temporary speed restrictions had an average delay time of 6.076, a standard deviation of 6.479, and a median of 4. The results showed that trains with temporary speed restrictions had an average delay of more than two minutes shorter than trains without temporary speed restrictions. In addition, from the standard deviation, there was a significant difference between the average delay time of trains without temporary speed restrictions, and their stability is poor. It is also visible from Fig. 3 that the blanket speed restriction greatly reduced the number of delay cases. Overall, the blanket restriction measures for this time period managed to improve the performance of the railway successfully.

From the scatter chart in Fig. 4, we observe the temporary speed restrictions implemented in response to extreme weather in October 2021. Here it is shown that speed delay measures managed to ensure that the train left within 25 minutes of delay (except one case), and from the chart, we observe that the frequency of train delay was relatively high from October 27th to 30th. Compared with trains without temporary speed restrictions, trains with temporary speed restrictions rarely had serious delays. However, trains that did not implement the temporary speed restrictions experienced serious delays for several days. For example, one train was delayed for more than 140 minutes. It is also observed, as can be seen in the chart, that trains without temporary speed restrictions had a high frequency of train delay in a few days.

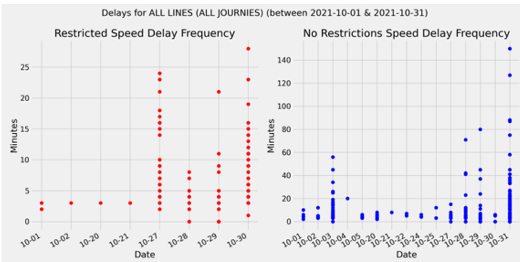

Fig. 4. Scatter chart showing all line delays for October 2021

Overall, trains with temporary speed restrictions had an average delay time of 4.865, a standard deviation of 5.321, and a median of 3. Trains without temporary speed restrictions had an average delay time of 5.785, a standard deviation of 13.725, and a median of 0. The results show that trains with temporary speed restrictions had an average delay of nearly a minute faster than trains without them. In addition to that the standard deviation of days without speed restrictions was higher. We suspect that this was caused by outliers. This showed that temporary speed restrictions in October 2021 were effective.

The conclusion was that temporary speed restrictions imposed in response to extreme weather in October, 2021 as well as in December, 2020 had a significant impact on reducing delays.

3.3 Results for Question 3

For Question 3 we investigate the benefits of the emergency timetable. We perform a comparative analysis between the 27th-30th of October 2021 with the 29th of October to 1st of November 2020, because the same dates were not available for 2020.

As the data showed, during the 27th-30th of October 2021 period, an emergency timetable was implemented. Blanket speed restrictions were enforced between the 29th of October and the 1st of November 2020, making these periods fairly comparable in order to identify if there was a benefit to the emergency timetable.

As shown in Fig. 5, during 27th-30th October 2021, the railway service was sometimes able to ensure the punctual departure of trains by initiating emergency timetable work, despite the impact of extreme weather at the time. From 29th of October to 1st of November 2020, the delays were fewer, but in the event they occurred, they lasted longer. Additionally, it was observed from the chart that the percentage of delaying trains departing within 10 minutes of delay during 27th-30th October 2021 is higher than that during 29th of October to 1st of November 2020. While there were some train delays from 29th of October to 1st of November 2020, the frequency of delays was much lower overall than from the 27th-30th October 2021.

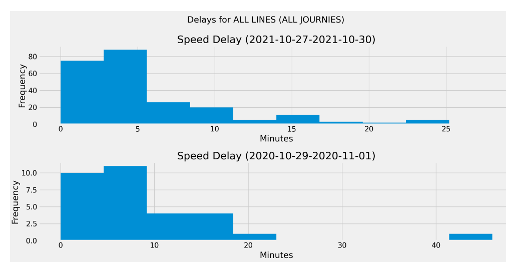

Fig. 5. Bar chart showing delays for lines A and F (ALL JOURNEYS)

The mean value for the delays for the 27/10/2021 - 30/10/2021 period is 5, the standard deviation is 5.543 and the median is 3. For the 29/10/2020 - 01/11/2020 period, we observed a mean value of 8.84, a standard deviation of 8.37 and a median of 6.5. The 27th-30th October 2021 window had the worst delays in terms of frequency, while the other one had higher average delay minutes. Despite having a high delay in minutes, the 29th of October to 1st of November 2020 period had a substantially lower frequency than the previous one. As a result, the October 2021 measures were significantly less effective than the compared measures.

3.4 Results for Question 4

For the final question, we investigated whether it was suitable for the Northern Railway to impose a weather-related emergency speed restriction based on current weather conditions. We implemented a logistic regression classifier that would predict either 0, which means no restrictions should be applied or 1, which means restrictions should be applied. We considered 70% of the data as a training set, and the rest was considered as a test set. After creating the model, we were able to obtain an accuracy of 76%. However, this model could be improved if the model was trained on a larger volume of data.

We also tested a NN model for the same question with and 80% train and 20% test split and in delivered results with a 91.7% percent accuracy, which is considerably higher. Additionally, similarly to the logistic regression, this accuracy potentially could be improved with the use of larger volume of data.

	Accuracy	Train-test split
Logistic regression	76%	70-30%
Neural Network	91.7%	80-20%

Fig. 6. Summary of the prediction model results

4 LIMITATIONS AND FUTURE WORK

Both the analysis and the models can have a number of validity issues and potential biases. One of the issues is that some of the periods we examined in the analysis were times when the UK had nationwide COVID restrictions and lockdown, which likely affected the operation of the railway (such as a decrease in the number of passengers), which could mean that the performance measured in this period is more likely to be an outlier value rather than a norm. Another issue might be that due to the limited weather data available and even after the re-sampling, the models are more likely to overfit than they would have done with a larger volume of training data. Finally, even though the models can predict fairly accurately if emergency speed restrictions are needed, the way they work doesn't make it possible to identify a certain threshold (such as volume of rainfall or wind speed) when they should be introduced.

One area of future work is improving the accuracy of our classification model by training and testing it using more data. For example, the weather data was not available for all delay instances, so we could not include all of them. Additionally, there is room to create a real-time model which consumes weather data in real-time and predicts expected delays. This could lead to better planning in response to sudden unexpected changing weather.

5 CONCLUSION

From the analysis, we can conclude that severe weather conditions have a negative impact on the performance of the railways. We observed that bad weather conditions increased the average number of delays and also resulted in higher variance as shown in higher values of standard deviation. We then explored the effect of the blanket restrictions that were imposed as a response to the severe weather conditions. We saw a positive influence of the blanket restrictions in terms of both the frequency and value of delays. Thus, we can infer that imposing blanket restrictions was effective in slightly improving the railway performance during severe weather conditions. We also found out that the emergence timetable work implemented between 27th-30th October 2021 performed worse than the blanket speed restrictions in most of the cases we analysed. However, we had limited data for this study case and exploring a larger amount of data about different time periods may result in different conclusions. Finally, we implemented a Logistic Regression Classifier that predicted whether a speed restriction should be applied based on the weather data and achieved an accuracy of 76% and a Neural Network model and achieved an accuracy of 91.7%. Again there is a scope of improving these models by introducing more data or exploring other algorithms.

ACKNOWLEDGMENTS

We utilised these libraries to write the code in Python: Numpy, Sklearn, Matplotlib, Tensorflow, Keras and Pandas.

REFERENCES

[1] 2022. https://byjus.com/maths/scatter-plot/

[2] D. Spiegelhalter D. Michie and C. Taylor. 1994. *Learning, Neural and Statistical Classification.* "Ellis Horwood".

[3] Carl French. 1996. *Data processing and information technology.* Cengage Learning EMEA.

[4] Imran Kurt, Mevlut Ture, and A Turhan Kurum. 2008. Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. *Expert systems with applications* 34, 1 (2008), 366–374.

[5] Geoff Lovea, Alice Soaresa, and Herbert Püempela. 2010. *Climate Change, Climate Variability and Transportation.* Elsevier. https://www.researchgate.net/publication/241124151_Climate_Change_Climate_Variability_and_Transportation

[6] Enikő Nagy and Csaba Csiszár. 2015. Analysis of delay causes in railway passenger transportation. *Periodica Polytechnica Transportation Engineering* 43, 2 (2015), 73–80.

[7] Irina Stipanovic Oslakovic, Erika J. Palin, Kenneth Gavin, and Andrew Quinn. 2019. *Implications of climate change for railway infrastructure.* WIREs. 1,2 pages. https://wires.onlinelibrary.wiley.com/doi/full/10.1002/wcc.728

[8] Jin Park and Dee W Edington. 2001. A sequential neural network model for diabetes prediction. *Artificial Intelligence in Medicine* 23, 3 (2001), 277–293.

[9] Chick H. Pierce, R. 2013. Workplace statistical literacy for teachers: interpreting box plots. https://doi-org.ezproxy.lancs.ac.uk/10.1007/s13394-012-0046-3

[10] Sebastian Ruder, Parsa Ghaffari, and John G Breslin. 2017. Data selection strategies for multi-domain sentiment analysis. *arXiv preprint arXiv:1702.02426* (2017).

[11] Strother H Walker and David B Duncan. 1967. Estimation of the probability of an event as a function of several independent variables. *Biometrika* 54, 1-2 (1967), 167–179.

[12] Hadley Wickham and Garrett Grolemund. 2016. *R for data science: import, tidy, transform, visualize, and model data.* " O'Reilly Media, Inc.".