A data model for heat-related rail buckling: implications for operations, maintenance and long-term adaptation

Jaroszweski, David; Fu, Qian; Easton, John

A data model for heat-related rail buckling: implications for operations, maintenance and long-term adaptation

David JAROSZWESKI[1], Qian FU[1], John EASTON[1]

[1]University of Birmingham, Birmingham, United Kingdom
Corresponding Author: David JAROSZWESKI (D.J.Jaroszweski@bham.ac.uk)

Abstract

Heat-related rail buckling is a significant operational and safety issue for the railways of Great Britain (GB). Although continuously-welded rail is pre-stressed to a stress-free temperature of 27°C, degradation and local topographic and microclimatic factors can lead to failures occurring at lower temperatures. These buckle events can cause widespread knock-on delays. As a common risk mitigation approach, speed restrictions are imposed for forecasted heat events, with these measures also being associated with non-trivial delays and disruption. The recently published UKCP18 climate projections indicate that the frequency, duration and magnitude of heat wave events will increase in the coming decades. This paper presents a data model to explain the occurrence of heat-related disruption incidents on GB's rail network. This model is built on historical delay data from Network Rail – the owner of GB's railway infrastructure, given explanatory variables for important meteorological and infrastructure features from published research in the literature. The model is implemented at two scales. Firstly at the national scale, including all Network Rail operational routes, and secondly for the South East of England (climatologically the warmest part of the country and hence at perceived greater risk) including the operational routes of Anglia, Wessex and South East.

Keywords: Rail, high temperatures, logistic regression, resilience, disruption

1. Introduction

Heat-related rail buckling is a safety and operational hazard which impacts the rail system during periods of high temperatures. Continuously welded steel track is particularly vulnerable to this hazard as lateral forces can build up as the track expands due to heat, with the passage of a train providing the additional force required to cause a buckle, potentially leading to derailment. The impact in terms of disruption can be significant, for instance, the hottest July day on record in the UK [1] caused 12,800 heat-related delay minutes, with an additional 23,700 minutes caused by preventative emergency speed restrictions.

Track is 'pre-stressed' to a stress-free temperature (SFT) to prevent buckling. For Great Britain (GB) the STF is typically 27°C [2]. In reality, once the track is laid, this resilience can reduce as the ballast moves and settles. STFs can be 3°C lower within a year of the track being laid, hence maintenance (particularly tamping) is essential to maintain resistance. Although determining track temperature can involve many factors, such as ambient temperature, exposure, cloud cover, relative humidity, track orientation and precipitation [2], a commonly-used empirical conversion gives the rail temperature as 3/2 greater than ambient temperature [3].

In terms of the causes of buckle events, as well as meteorological, topographic and asset variables, an important concept is that of 'buckle harvesting'. This concept is used to explain the observation that critical rail temperatures (CRTs) early in the summer period tend to be associated with a greater risk of buckling than those later in the year. This is due to the stock of at-risk assets being higher at the beginning of the year. These are 'harvested' during early heat events, and replaced or maintained, in effect increasing the resilience of the railways incrementally as the summer progresses, meaning that a particular high temperature value poses a higher risk the first time it is encountered in a season [4].

A previous study by the authors [5] used a logistic regression approach in a test region (the Anglia route of GB's rail network) to explore the relative importance of various explanatory variables, especially the maximum temperature within a predefined critical period of time prior to an incident. This study suggested that:

- All the temperature-related variables, except the maximum temperatures below 25°C, were statistically significant at the significance level of 0.05, indicating that they play a positive role in the causation of heat-related incidents;
- The difference between the minimum and maximum temperatures within the predefined time period is also a key variable, with the odds of a heat-related incident increasing 20% with a 1°C increase in diurnal range;
- The 'buckle harvesting' phenomenon may mean that lower temperatures earlier in the season cause at-risk sections of track to buckle, so that as higher temperatures are reached in the later summer months, the track has already become more resilient through reactive maintenance;
- A North-South orientation acts to reduce risk compared to an East-West orientation (although not statistically significant); and
- Although there was not strong statistical evidence that the other weather-related variables, such as wind and relative humidity, had significant impact on the occurrence of heat-related incidents; the signs associated with those variables reflect reasonable expectations. For instance, precipitation shows evidence for reducing buckling risk.

The previous study used an aggregated category of 'heat' defined by Network Rail – the owner of GB's railway infrastructure, where incidents from several 'reasons codes' were included, hence the contribution and strength of specific infrastructure and operational failures (e.g. for rail or points faults) were not assessed. Additionally, it was limited in application to the Anglia route in the East of England. The present study expands this research to include data for all regions of GB's rail network, as well as a regional comparison for several routes (i.e. Anglia, Wessex and South East routes) in the southeast of the country.

2. Data

The data was available for a period ranging from the start of financial year 2006/07 to the 30th of September 2018; the data model however was built on a subset of the data restricted from the beginning of May to the end of September for each year, so as to minimize the potential influence of cold weather-related incidents on the analysis.

2.1 Delay Incident data

Historical delay incident data was from Network Rail TRUST (Train RUnning Systems on TOPS) system, which records detailed information about every rail incident on the network, including the date, time, location, reason, the main actions taken in response to it, and the resultant delay (in minutes). In the present study, the analysis was limited to a selection of reported reasons (each abbreviated to two upper-case letters) for the heat-related delays:

- Broken/cracked/twisted/buckled/flawed rail (*IR*)
- Points failure (*IB*)
- CRT speeds (other than buckled rails) (*JH*)
- Severe heat affecting infrastructure the responsibility of Network Rail (excluding *JH*) (*XH*)

2.2. Meteorological data

Gridded data, supplied to Network Rail by MeteoGroup, were included for the study period:

- Temperature, given in 1 degree Celsius (°C) increments from 24 degrees. The present study considered the highest temperature prior to a reported heat-related delay as a categorical variable, labelled 'maximum<24°C', 'maximum=24°C', 'maximum=25°C', 'maximum=26°C', 'maximum=27°C', 'maximum=28°C', 'maximum=29°C', and 'maximum≥30°C', respectively. Note again, as suggested by existing studies, that 27°C is often deemed a critical temperature at or beyond which the risks of experiencing track buckling would dramatically increase.
- Diurnal range of the temperature (i.e. the absolute difference between the maximum and the minimum temperatures in a 24-hour period prior to the time of the incident). Existing evidence shows that rapid changes in temperature can be an additional risk factor for rail buckling.

- Total precipitation – precipitation can act to cool rails and is associated with greater cloud cover, assuming that a higher level of rainfall might help mitigate harm of high temperature.
- Whether it was the hottest day of year – evidence suggests that temperature-related rail failure follows a 'harvesting' pattern, with subsequent occurrences of temperature being associated with lower risk. This is labelled 'hottest heretofore' in results.

2.3. Other data

Track orientation

- The data model considered only the incident locations, such as stations, junctions and the stretches of track between these, which had ever experienced at least one heat-related incident. The track orientation for each incident location in this study was identified only by the straight line between the start and end of the incident location. This was due to lack of detailed geographical information of the tracks at the time of the analysis. For all the incident records, their track orientations were divided into four categories: east-west, north-south, northeast-southwest, and northwest-southeast.

The analysis of the above data viewed the likelihood of occurrence of heat-related incidents as the outcome variable of interest, given a selection of explanatory variables (or predictors), including the highest temperature and total precipitation prior to a reported heat-related delay, diurnal range of the temperature, and track orientation. In this case, a logistic regression model was used to explore the feasibility of establishing a data model. All of the predictors, except the diurnal range of the temperate and total precipitation, were categorical variables.

3. Data integration

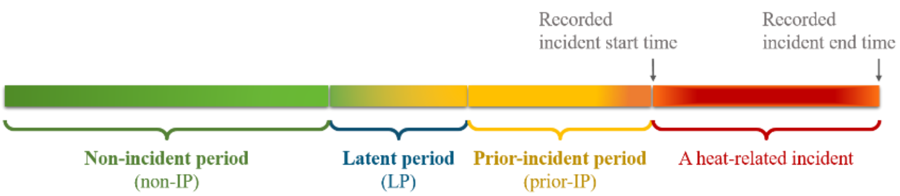

Figure. 1: Illustration of 'non-incident period' (non-IP), latent period (LP), and 'prior-incident period' (prior-IP) in the context of heat-related incidents

The integration of the above-described data was in both the spatial and temporal contexts was based on defining a 'prior-incident period' (prior-IP). As illustrated in Figure 1, the prior-IP should start a certain number of hours (e.g. 24 hours) before the recorded start time of an incident, and last until the occurrence of the incident. That way, weather observations during the prior-IP are considered to be the prevailing conditions in the causation of the incident. Conditions during the course of an incident are irrelevant in this case. On the other hand, a 'non-incident period' (non-IP) must be defined against each prior-IP. In contrast to the prior-IP, it would be assumed that the conditions (e.g. weather conditions) observed during the non-IP would be unlikely, or much less likely, to result in a heat-related incident (e.g. track buckling).

A logistic regression model was then employed to analyse the integrated data. The analysis was conducted for six iterations, featuring different reason codes in isolation and combination to discern the relative influence of the explanatory variables for different impact type. The iterations were: 1) *IR*, *IB*, *XH* and *JH*; 2) *JH* only; 3) *IR*, *IB* and *XH*; 4) *IR* only; and 5) *IB* only. These were repeated both at the national scale and for the southeast region (covering Network Rail routes of Anglia, Wessex and South East), which is climatologically the warmest part of the country and hence at perceived greater risk.

4. Results

4.1. Test 1 – IR, XH, IB and JH.

For all reason codes in combination at the national level, the estimated coefficient associated with the maximum temperatures up to 30°C were all positive and statistically significant, with odds ratios peaking at 1.7 at 27°C (i.e. the STF). The estimate for the hottest day heretofore was positive and significant, with an odds ratio of 1.5. For the southeast, maximum temperatures between 26°C and 29°C were significant with positive estimated coefficients, where odds ratios rose from 1.5 to 2.1 at 28°C and then fell to 1.8. Hottest day heretofore also had a statistically significant, positive coefficient, with an odds ratio of 1.3. All the other variables were either not significant or had odds ratios near to 0.

4.2. Test 2 - JH only.

The testing model estimated from the data of the reason code *JH* only showed the greatest significance for the explanatory variables among all the tests in this study. Nationwide, the coefficients associated with all temperature-related variables were statistically significant, with the highest odds ratio of 8.7 at 27°C. The track orientation of North-South was significant, with an odds ratio of 1.2. Hottest day heretofore was also significant, with an odds ratio of 2.9. Temperature change was positive and significant, with an odds ratio of 1.24. The coefficient of precipitation was negative and significant, with an odds ratio of 0.89.

Table 1: Estimation result (Southeast region)

Latent period = 5 × 24 hours						No. observations = 1391
						Degree of freedom = 13
Log-likelihood ≈ –741.31						
Variable	**Coef. est.**	**Std. err.**	***z*-statistic**	***P* > |*z*|**	**[95% conf. interval]**	**Odds ratio**
(Intercept)	-2.3191	0.300	-7.729	0.000	[-2.907, -1.731]	0.0984
Temperature (°C)						
Change within prior-IP	0.1944	0.031	6.332	0.000	[0.134, 0.255]	1.2146
maximum < 24 (ref.)	0.0000	-	-	-	-	1.0000
maximum = [24.0, 25.0)	0.4563	0.200	2.279	0.023	[0.064, 0.849]	1.5782
maximum = [25.0, 26.0)	1.6100	0.300	5.364	0.000	[1.022, 2.198]	5.0028
maximum = [26.0, 27.0)	2.2527	0.411	5.476	0.000	[1.446, 3.059]	9.5131
maximum = [27.0, 28.0)	2.0823	0.541	3.851	0.000	[1.023, 3.142]	8.0226
maximum = [28.0, 29.0)	1.4545	0.633	2.296	0.022	[0.213, 2.696]	4.2825
maximum = [29.0, 30.0)	2.2094	1.048	2.108	0.035	[0.155, 4.264]	9.1106
maximum ≥ 30.0	1.6909	0.803	2.105	0.035	[0.117, 3.265]	5.4244
Track orientation						
East – West (ref.)	0.0000	-	-	-	-	1.0000
Northeast-Southwest	0.0164	0.190	0.086	0.931	[-0.357, 0.390]	1.0165
Northwest-Southeast	-0.1762	0.210	-0.840	0.401	[-0.587, 0.235]	0.8384
North-South	0.1428	0.207	0.689	0.491	[-0.263, 0.549]	1.1535
Hottest heretofore	0.7280	0.166	4.381	0.000	[0.402, 1.054]	2.0710
Total Precipitation (mm)	-0.0869	0.024	-3.640	0.000	[-0.134, -0.040]	0.9167

For the southeast region (see Table 1), temperature showed positive significance throughout the range. Odds ratios were all above 4 for temperatures higher than 26°C, with values over 9 at 27°C and 30°C. Temperature change was positively significant with an odds ratio of 1.2. Total precipitation had a significant negative influence with an odds ratio of 0.9. Hottest day heretofore had a significant positive influence with an odds ratio of 2.1. The ROC curve for the *JH* model showed good predictive ability against random guess for the southeast region (see Figure 2), Figure 3 shows the actual delay minutes versus predicted likelihood of incident occurrence from the model, with a threshold of 0.51, capturing a majority of the high magnitude incidents.

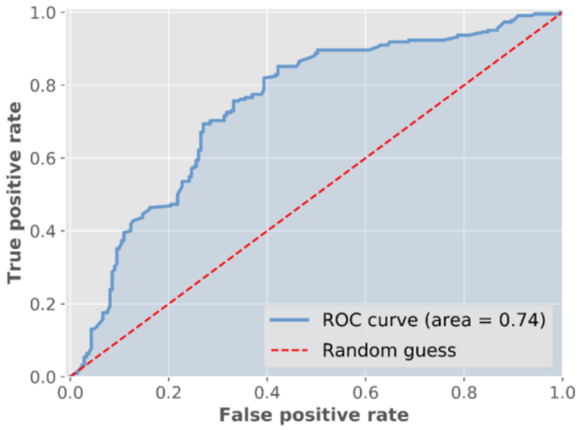

Figure 2: ROC plot for JH incidents in the southeast

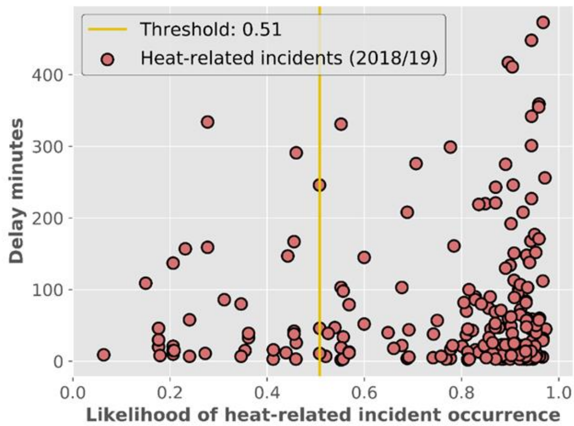

Figure 3: Actual delay minutes vs. predicted likelihood of the incident occurrence.

4.3. Test 3 – IR only.

The estimated coefficients associated with the maximum temperatures between 27°C and 30°C were positive and showed statistical significance in the nationwide context, with odds ratios rising from 1.5 to 2 within this range. For *IR* in the southeast region, only the maximum temperatures of 28°C and 29°C showed significance with an odds ratio of 1.5 and 2.9, respectively.

4.4. Test 4 – IB only.

Nationwide, temperature had significance for point failure between the maximum temperatures of 25°C and 29°C, with odds ratios rising between 1.2 and 1.4. Hottest day heretofore also showed significance with an odds ratio of 1.3. In the southeast region, the maximum temperatures of 25°C and 27°C showed positive significance in the causation of point failures, with odds ratios rising from 1.4 to 1.6.

4.5. Test 5 – IR, XH and IB.

This combination showed statistical significance between maximum temperatures of 25°C and 30°C, with odds ratios ranging between 1.1 and 1.4 at 29, then falling to 1.3 at the national scale. For the southeast region, this combination showed positive significance between 26°C and 29°C, with odds ratios increasing from 1.3 to 1.8 at 28, then falling to 1.6.

5. Discussion

5.1. Influence of explanatory variables on heat-related rail incidents

The results indicate that the specific impact categories or reason codes associated with heat-related failure have a less clear relationship with the explanatory variables than the operational measure of temporary speed restrictions. It is also apparent that for the infrastructure-related reason codes, the significance and odds ratios of the maximum temperature variables tends to reduce at relatively low temperatures e.g. 27°C to 28°C, whereas for *JH* peaks at comparatively higher temperatures and remains high and significant above 30°C. This suggests that the precautionary measures taken to reduce speeds at higher temperatures work well to reduce the impact of heat at these temperatures. As *JH* is an operational measure and governed by predefined temperature thresholds, this also explains the strength of the predictive power in Figures 2 and 3 for Test 1, the greater significance of Test 2 compared to 3-5, and the fact that the models were weaker in Tests 2-5.

The regional analysis shows that the southeast region generally have higher odds ratios associated with equivalent temperatures than the national picture. Further research is required to ascertain the reason for this, but the southeast is generally subjected to more frequent and longer duration heat events, with an additional component of latitude (lower latitudes being associated with greater incident solar radiation) potentially being a component [4]. The variable of hottest day heretofore was significant in several instances, again giving evidence to support the buckle harvesting theory [1]. Track orientation played very little part in the in terms of explaining incident occurrence. However, incident counts for each orientation indicate that east-west is by far the most frequent, however, these figures are not normalized, with further analysis of exposure (i.e. length of track orientation) required.

5.2. Further analysis and improvements to model

Additional interrogation of the delay data would be beneficial to ensure that the incidents are heat-related. This could involve analysis of keywords in the incident description for those commonly associated with heat-related incidents. Further sources of meteorological and asset information can be added, such as incoming solar radiation, cloud height, humidity, shading, latitude and asset condition. Furthermore, there is potential to incorporate interactions between variables in the statistical modelling approach, as well as modification of the prior incident period length to capture antecedent conditions.

Improvements to the allocation of weather cells for incidents can also be made. As incidents which occur between two stations or junctions do not have an exact location, a simple midpoint approach was used, with that location being used to select the weather cell. This does not account for the potential deviation of the track between these two points. The start and end locations are typically >10km part. An interim approach to deal with this uncertainty (both the exact location of the incident and the deviation of the line) is to create a buffer, formed by the radius between the midpoint and the end locations, within which the weather cells will be averaged.

5.3. Implications for industry

There is potential for the results to be integrated with output from the Natural Environment Research Council-funded project “Weather-induced single point of failure assessment methodology for railways”, which produced a criticality map of the GB rail network. Such a risk mapping approach would demonstrate how targeted interventions can be made at locations with the greatest risk of buckling and the greatest potential network disruption. The approach can also be improved with integration of high resolution topographic data currently being produced by the industry, such as LIDAR surveys of vegetation and topography.

The results demonstrate the potential for an operational forecast tool which can inform decisions for extreme event preparation and response through processes such as Network Rail’s Extreme Weather Action Team (EWAT) procedure. In the longer term, the approach is suitable for climate impact analysis. The frequency of rail buckling events is predicted to increase due to climate change [6]. The resolution of the model lends itself to the input of climate data from weather generator tools, which produce synthetic weather time series representative of future climates, such as those included in the recently released UKCP18 and H++ climate scenarios, the later providing projections for low probability, high impact extreme heat events. This will allow for the potential frequency and spatial distribution of future heat-related impacts to be assesses.

6. Conclusion

This is the most comprehensive data model for heat-related delays and presents a step change in temporal and spatial resolution from that of previous studies, presenting a clear pathway for further development by the meteorological services sector. The study indicated that although operational measures to reduce the risk of rail buckling and generally effective at higher temperatures, those temperatures around the STF (27°C) are associated with higher odds ratios for the specific infrastructure-related reason codes. The results have significance in the present day, and can also be coupled with recently released climate change projections (UKCP18 and H++) to gauge the potential impact of future climates and to prepare adaptation responses.

Acknowledgment

David Jaroszweski is supported by the NERC (Natural Environmental Research Council) fellowship NE/N012941/1 under the National Productivity Investment Fund Programme. Network Rail data supplied through the University of Birmingham’s data sharing agreement.

References

[1] E. Ferranti, L. Chapman, S. Lee, D. Jaroszweski, C. Lowe, S. McCulloch, and A. Quinn, "The hottest July day on the railway network: insights and thoughts for the future," *Meteorological Applications,* 2017, doi: 10.1002/met.1681.

[2] L. Chapman, J. E. Thornes, Y. Huang, X. Cai, V. L. Sanderson, and S. P. White, "Modelling of rail surface temperatures: a preliminary study," *Theoretical and Applied Climatology,* vol. 92, no. 1, pp. 121-131, 2008, doi: 10.1007/s00704-007-0313-5.

[3] G. A. Hunt, "An analysis of track buckling risk," British Railways Internal Report. RRTM013, p31, 1994.

[4] E. Ferranti, L. Chapman, C. Lowe, S. McCulloch, D. Jaroszweski, and A. Quinn, "Heat-related failures on southeast England’s railway network: insights and implications for heat risk management," *Weather, Climate, and Society,* vol. 8, no. 2, pp. 177-191, 2016, doi: 10.1175/WCAS-D-15-0068.1.

[5] Q. Fu, D. J. Jaroszweski, and J. M. Easton, "A prototype model for understanding heat-related rail incidents: a case study on the Anglia area in Great Britain," presented at the The 8th International Conference on Railway Engineering, IET London: Savoy Place, 16-17 May, 2018, Conference paper. [Online]. Available: http://digital-library.theiet.org/content/conferences/10.1049/cp.2018.0071.

[6] E. J. Palin, H. E. Thornton, C. T. Mathison, R. E. McCarthy, R. T. Clark, and J. Dora, "Future projections of temperature-related climate change impacts on the railway network of Great Britain," *Climatic Change,* vol. 120, no. 1, pp. 71-93, 2013, doi: 10.1007/s10584-013-0810-8.