The causes and consequences of reactionary delay

Data Sandbox: Improving Network Performance

Authors:

Dr Aidan Slingsby, City, University of London

Dr Cagatay Turkay, City, University of London

Dr Helen Wilkinson, Director, Risk Solutions

Jonathan Hyde, Director, Risk Solutions

Chris Rees, Director, Risk Solutions

Submission date: 29/03/2019

Executive summary

Aims and objective

Reactionary delay is a significant cause of overall train delay; it has grown over recent years and represents approximately 70% of all train delay minutes. The knock-on effects of reactionary delay appear chaotic and are difficult to predict.

This RSSB funded feasibility study was focussed on predicting and avoiding reactionary delay. Our objectives were to:

- Create a set of railway performance modelling tools and demonstrate how they can be used by Train Operating Companies and Network Rail to explore the causes and consequences of reactionary delay.
- Demonstrate how the tools can be used to test a range of interventions designed to improve performance, predicting the service performance that can be achieved through each intervention.

Outputs delivered

The study has created two prototype tools working in tandem; an agent-based model to simulate train services, and a set of interactive visualisations to explore the complex interactions between modelled train services.

These Rail Performance Tools have been demonstrated on a significant part of GWR's routes, simulating all train services using any part of the route from London Paddington to Bristol. The tools have been used to identify the root causes of reactionary delay and predict the effectiveness of some example interventions designed to improve performance.

This demonstration provides evidence that the tools work, and will be a valuable addition to the rail industry performance improvement tool kit.

Conclusions and value of outputs

The feasibility study has demonstrated that the combination of agent-based modelling of rail services, together with interactive visualisations of the complex interactions between trains, is a powerful way of exploring the complex causes and consequences of reactionary delay. The insights gained are valuable to Train Operating Companies and Network Rail as they can help identify ways to combat train service delays and improve service performance for rail customers. For example, to identify where intervention might provide most performance improvement.

The tools are also valuable as they can be used to test the impact of performance improvement strategies before committing to investment. The tools quantify the performance improvement that might be achieved, and tests the reliability of these improvements across multiple scenarios.

These benefits were planned from the outset of the project, but as the project ideas developed it became clear that the tools can be used for other valuable purposes, including to:

- stress test timetable changes to spot problems before introducing new services, and simulate how minor changes could alleviate problems.
- simulate specific worst-case scenarios to test contingency plans.
- understand the wider effects of speed restrictions and extended engineering access.

Next steps

High priority actions following the feasibility study are to:

- Promote the study findings and the availability of the tools in their current form to interested rail stakeholders, particularly Train Operating Companies who would be interested in applying the tools to their operational areas.
- Identify short term opportunities to fund improvements in the accuracy of the tools and make them easier to set up and use – this will encourage early adoption by rail partners. For example, model a larger area, prioritise trains at junctions and stations, allow some services to be cancelled, change station dwell times, improve the visualisation of delay results.
- Apply for longer term funding opportunities to develop the tools by adding additional features to increase their value to Train Operating Companies and Network Rail. For example, automate setting up the model for a new area, calculate the financial impact of reducing delays, associate rolling stock and crew with train services, add passenger loading.

Contents

1 Introduction

1.1 Background to the research

The railway industry collects information about the initial incidents that cause delay (**primary** delay) to help improve the performance of rail services and keep passengers informed. Primary delay has remained stable over recent years (300-350k minutes per year[1]), but the delay caused by an initial incident often cascades through the rail network, causing **reactionary** delay, which is proving difficult to understand and control.

Reactionary delay is a significant cause of overall train delay and has grown steadily over recent years (from 600k to 800k mins[1]). While the effects of primary delays on the punctuality of train journeys is generally straightforward, knock-on effects of this leading to reactionary delay is less easy to understand, because the knock-on effects are dependent on the interaction between multiple trains and the network characteristics of the track where the delay happened. There is also often positive feedback whereby reactionary delay can cause more reactionary delay in cascading reactionary chains affecting multiple trains.

GWR, in common with other train operating companies, are experiencing worsening delays on some of their key routes, and want to better understand the causes of this.

> "GWR is the primary operator on the Western Route, running over 1,700 trains a day across 270+ stations. Over the last couple of years, we have embarked on one of the largest railway investment programs in the industry, upgrading and electrifying our infrastructure along with undergoing a complete fleet transformation. Though this is aimed at improving customer journey times and overall experience, we constantly suffer from unforeseen incidents. There is a clear random nature to these, and the complexity makes it hard to understand at an operational level.
>
> Specifically, we have seen a worsening performance in our key London area service groups; London-Cotswolds (EF03), Outer London Thames Valley (EF05), Inner London Thames Valley (EF06), Reading Oxford Suburban (EF07) services, which overall on average have worsened by 1.2 % points since last year. Year to date, each of them has also declined."
>
> *Sanat Misra, Performance Strategy Manager, GWR*

To look at how industry data can be used and analysed to improve network operational performance RSSB set up some feasibility studies[1]. The aim of these studies was to develop tools that can support and optimise performance management. This report is

[1] Call for research - data sandbox: Improving Network Performance, December 2017

the final report of the research study carried out by City University of London and Risk Solutions, working with GWR.

1.2 Aims and objectives of the research study

Our study was focussed on the suggested research theme A: predicting and avoiding reactionary delay. Our objectives were to:

- Create a set of railway performance modelling tools and demonstrate how they can be used by Train Operating Companies and Network Rail to explore the causes and consequences of reactionary delay.
- Demonstrate how the tools can be used to test a range of interventions designed to improve performance, predicting the service performance that can be achieved through each intervention.

The proposed tools consisted of an agent-based model to simulate train services running across a significant part of GWR's routes, and a set of interactive visualisations to explore the complex interactions between modelled train services.

Agent based model (ABM): The model simulates timetabled train services, characterised using current data describing the railway system obtained from Network Rail. The model introduces incidents that cause primary train delay, characterised by an analysis of Network Rail attributed delay data. The model then simulates the train conflicts and subsequent reactionary delays as they develop across the rail network, each agent (train) affected by other trains and random events as they occur.

Interactive visualisations: The ABM produces thousands of delays across hundreds of model runs. In order to interpret these, they need to be statistically aggregated. High-level summaries are useful, but hide important detail. Information visualisation is capable of showing **hundreds of statistical values** for many locations, trains or times, **in interpretable ways**. It can also provide the context needed to interpret the data. Visualisation is also very well suited for showing comparison, giving the potential to help understand the effects of modelled interventions to be understood. **Interactions** allow the user to obtain more details, or to compute statistics for specific situations of interest, on-demand. Our visualisations provide visual summarises of the results needed to identify problematic locations and trains, but also allows users to dig into the details to understand the reasons for these delays and help plan for interventions.

Once the possible causes of reactionary delay are identified, interventions can be suggested that aim to reduce these delays. The modelling and visualisation tools can then be used to test the effectiveness of these interventions, by re-modelling the timetabled trains but this time with an alternative set of input information describing what the interventions are designed to achieve. For example, reducing the number of track based primary incidents, or reducing a range of incident durations. The visualisations then show the impact of these interventions, and the resulting improvement in service performance.

Improvements could be expressed in various ways. GWR expressed interest in the ability to test various performance measures in preparation for the change from the current Public Performance Measure (PPM) to Right-Time measures, and to understand the economic impacts of delays.

This research lays the groundwork for modelling and visualisation of more routes, over longer periods of time, with additional performance measures (such as passenger delay and the economic impact of delays). For this initial feasibility research study, however we agreed with GWR to focus on a route from their London service groups, London Paddington to Bristol, shown in Figure 1.

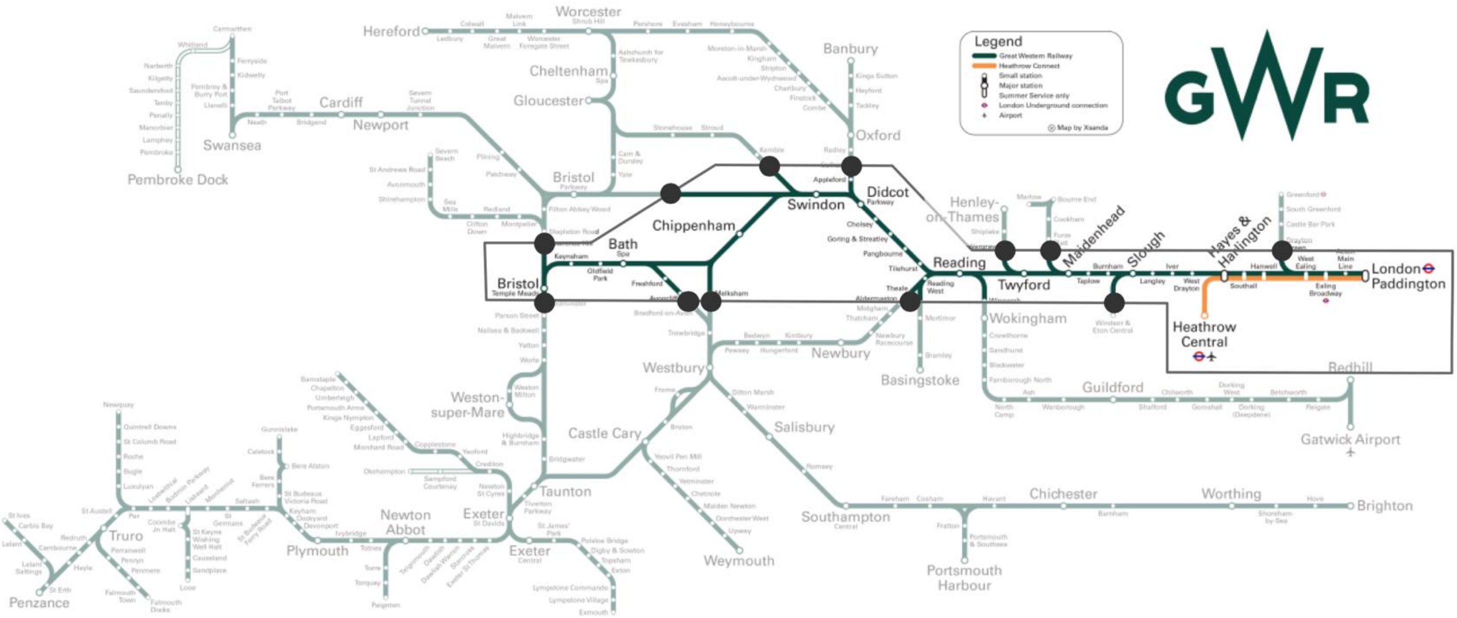


Figure 1: Modelled route

If the research is successful, we anticipate that the railway performance modelling tools could be developed into market ready products and become valuable components of the performance improvement arsenal used by Train Operating Companies and Network Rail.

2 Approach

This part of the report describes our approach to developing useful performance management tools for the railway industry:

- **Partnership with railway stakeholders**: Consulting with GWR and project steering group throughout the project on the modelling approach, assumptions and model outputs.

- **Railway performance modelling (agent-based modelling)**: Building an agent-based model, how it models train services, the data inputs it uses (delays, rail network and the timetable), the delay results produced, multiple model runs.
- **Characterising train service delays**: How incidents and delays are characterised and applied by the model
- **Calibrating the model**: How the model was calibrated to provide confidence that it can model real world train services
- **Exploring reactionary delay using interactive visualisations**: how we developed visualisations and how these were used to explore the causes and consequences of reactionary delay.

The **Findings** section of the report describes how the railway performance tools can be used to identify potential areas for intervention and quantify their effectiveness at improving performance.

Throughout this section we have identified areas where we think additional work would be useful - these are brought together in the **Recommendations** section of the report.

2.1 Partnership with railway stakeholders

This sandbox research project was designed as a partnership between academic understanding, complexity modelling and railway knowledge and expertise. The work was iterative, designing Rail Performance Tools in phases. Each design phase being reviewed by our rail partner GWR and the project steering group of industry experts. Feedback from industry stakeholders directed development of the next phase of work.

Steering the project: We held regular design discussions with GWR and the project steering group (members are listed in Appendix 1 – our stakeholders). These discussions proved valuable opportunities to test our assumptions, emerging findings and direct our resources to the most valuable developments.

Providing information and current data: Network Rail provided rail network topology (tracks, locations), the national timetable (trains, locations, timing), and historic attributed delay data (delay location, type and impact).

GWR provided operational control strategies (defining the rules for routing trains), the factors that affect delays (e.g. passenger loading, weather, crew availability, train dispatch), and real-time train timing and delay data.

Expert opinion: GWR provided expert opinion to inform the project, particularly the types of delay that can occur and the configuration of the Paddington to Bristol route. Members of the project steering group provided expert opinion about the way to interpret the timetable.

Analysing results and selecting interventions to model: We invited our rail industry partners to comment on how we used the project tools to produce delay results, and explore and interpret results.

Exploitation of research work: We discussed exploitation plans with RSSB, and towards the end of project shared emerging findings with other sandbox teams to look for ways of exploiting each other's work.

2.2 The railway performance model

2.2.1 Agent-based modelling

Understanding the root causes of performance issues is not easy. The railway system…

- has complex interactions and dependencies between individual components; passengers, trains, staff, stations, timetables, junctions, weather
- can be sensitive to small variations in inputs that can cause an escalating chain of events; 30 seconds late departure at a particular station / time can cause cascading delays across the network
- can be affected by rare combinations of events.

These characteristics create emergent behaviour and agent-based modelling is well suited to representing these characteristics.

Risk Solutions were responsible for the design, implementation and demonstration of an agent-based model (ABM) to simulate all train services that use any part of the route from London Paddington to Bristol. Train services are individual agents in the model, moving along routes defined by a timetable, following predefined rules which sometimes prevent them from achieving their journey on time. These rules define when and how trains can travel between locations, and stop/pass at locations, and create the conflicts that result in reactionary delays.

Trains are also delayed in the model due to incidents that are randomly applied to trains and locations across the modelled network. The incidents are characterised to be representative of typical delays from March 2018, using Network Rail's attributed data and GWR's Aegis train movement data. This process is described later.

Appendix 2 provides more information about how the agent-based model represents the rail network, the timetable, and the way that trains can conflict, creating reactionary delay.

The model can simulate many hundreds of days of rail services from the same timetabled day, producing a distribution of many delay propagation results, representing many more possible combinations of incidents than historic delay data. Filling gaps in current delay data sets.

Therefore the model can simulate a multitude of possible outcomes that can occur in a given day. It explores the range of possible train conflicts and subsequent reactionary delays that can cascade from a timetabled day of train services, producing detailed output data describing the many ways in which train services can be delayed.

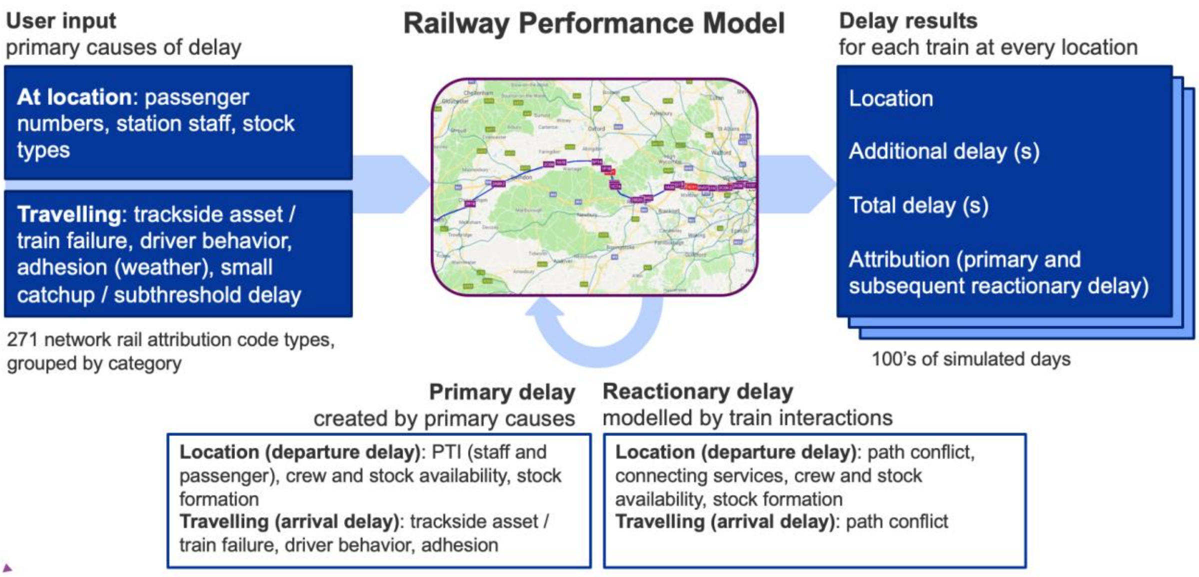


Figure 2: Railway Performance Model

The model produces detailed delay results at several levels of aggregation; individual train services, station stopping points, generating a large amount of detailed granular delay data, measured to the second. But the most significant value of these results is that they include the delays resulting from **all** train service interactions, not just the known attributed delays of over three minutes. Therefore, a full picture of what causes reactionary delay emerges, with an unparalleled level of detailed delay attribution, recording the combination of events and other factors that have influenced these modelled delays. These include:

- any train conflicts inherent in the normal timetable running, without the presence of any incidents
- trains suffering subthreshold delays (those that are under 3 minutes and not attributed to a cause)
- trains delayed due to incidents that can be attributed to a cause (delays over 3 minutes).

2.2.2 Modelling data inputs

The agent-based model is data driven, using input data files to characterise any part of the railway system. This means the same model can model any part of the UK rail network, or if required, the whole network, by changing the data input.

The data input files used by the model describe:

- **the railway network to be modelled**: rail reference lines, TIPLOC (locations), lines between locations
- **a timetable of train services**: automatically extracted from the national CIF timetable for the required day for the selected route(s)
- **incidents causing primary delays**: a set of incident types each with characteristics describing their frequency, where they can affect train services, their duration, and the duration of delays caused to services. Sixteen types of primary attributed incidents were characterised for our project demonstration, derived from Network Rail attribution data. Subthreshold delays were also characterised from train movement data provided by GWR from their Aegis system, and added to the input data.

To demonstrate the agent-based model, we characterised the route between London Paddington and Bristol, using a Wednesday timetable from March 2018, resulting in data input files describing:

- 90 locations (stations, depots and sidings), and 800 km of track
- 900 train services/journeys, with 4,600 station stops, from 23,000 timetable entries

These data input files can be edited in the model, or replaced with alternative data, describing other parts of the railway network, selecting different times of the year, and alternative incidents that can cause delays.

A short data preparation process is required to characterise the part of the rail network to be modelled, and automatically extract the relevant train services from the national timetable. Our original intention was that this process would be fully automated, exploiting any pre-existing data describing the railway network, specifically Network Rail's 'Network Model'. After some effort to track down data sets that link the timetabled services with the Network Model, and characterise the stations, depots and lines, we were not able to find existing data defining the necessary relationships between data sets, in particular, the relationship between TIPLOC locations used in the timetable and the Network Rail Network Model.

The last part of data preparation is characterising the incidents to be introduced in model runs, delaying train services and triggering reactionary delay.

An important next step is to automate the entire data preparation process, this would require some additional effort to create the necessary relationships between the data sets. When this is achieved the railway performance tools can be quickly applied to other parts of the UK rail network and more easily used by Train Operating Companies and Network Rail.

Modelling interventions

The model needs to be able to simulate the effect of interventions to improve service performance. These are achieved through changes to the input data, specifically the incidents that cause primary delays, based on assumptions of the performance benefits the interventions would bring.

For example, investing in more station staff could be modelled as reducing the frequency and duration of station based primary delays, or improving the time taken to resolve track-based incidents can be modelled as reducing the incident duration time for these types of incidents.

An example of intervention modelling was created as part of our demonstration and is described later in this report.

2.2.3 Model runs

The model can be tasked with running a single **simulation**, which might be a partial or complete day of timetabled train services. Or it can run a set of multiple simulations, one after another, using the same input data, but exploring the different ways that train services could interact. The input data might describe a typical day (a baseline) or a particular intervention intended to improve performance. This run of multiple simulations is referred to as a **scenario**. The demonstration model completes each simulation of 24 hours of train services in around 20 seconds.

A scenario can contain many hundred simulations so that a range of possible outcomes can be observed. The stochastic nature of the agent-based modelling process, with a range of randomly placed incidents causing train service delays, means that every simulation is different. The more simulations that are in a scenario, the more confident we can be that the range of performance improvements observed will bound those that are likely to occur in real life.

The model can also be set up with a **batch** of alternative scenarios, so that a range of interventions can be modelled, allowing the user to identify the best intervention (in terms of delay reduction) from the batch. The model will automatically run each scenario in turn, until all scenarios are complete.

2.2.4 Modelling outputs

Train service delay results: Every primary and reactionary delay that occurs in the model (including off-network primary delays and subthreshold delays), are saved in a results file for every simulated day along with data describing the train service delayed, where it was delayed, and the attribution of the delay. This therefore can result in hundreds of train service delay results files for each scenario.

Attributed delay results: The model also saves the train service delay results in the same format as the Network Rail attributed delay data, so that it can be analysed using any tools designed for this purpose, and compared with real-world data.

Train journey results: The model records every train service arrival and departure times at every TIPLOC location, and saves these in a results file together with the timetabled journey arrival and departure times.

Scenario results: The model calculates summary results for each simulation, for each type of incident and delay. These include counts of incidents and delays, the minimum, mean, maximum duration of incidents and delays. The model also calculates these statistics for all simulations in a scenario. These results are saved in a scenario results file, for each scenario in a batch.

The model also provides a summary of each scenario in the form of a chart, which is also saved as an image file together with the other results.

2.3 Characterising delays to train services

As previously mentioned, modelled train services can be delayed in four ways in the agent-based model:

1. **Attributed primary delays:** allocating delays to trains as a result of incidents added randomly to locations and lines throughout the modelled day. These delays are all over three minutes, and for the purposes of our demonstration, have been characterised using national attributed delay data from March 2018, provided by Network Rail. The delay characteristics can be altered by the user as required.
2. **Subthreshold primary delays:** allocating small delays to trains as they travel between every pair of locations in their journey. These delays are all under three minutes, and include negative delays so that late trains can in some cases catch up with their timetabled schedule. Trains arriving early at their station stops are not permitted to leave early. For our demonstration these delays have been characterised using Aegis train movement data provided by GWR, and can be altered if required.
3. **Reactionary delays:** calculated by the model as a result of trains conflicting at stations, sidings, depots, junctions and along lines. These delays cannot be altered using input data as they are a result of the modelling process.
4. **Off-network delays:** around 59% of train services modelled in our demonstration route join the modelled network part way through their journey (the locations with black dots in Figure 1). The model randomly adds a delay to approximately one third of these train services, so that they are not all joining the model according to the timetable. This delay represents all three types of delay to a train that is part way through its journey and we have called this 'off-network' delay as it has occurred off the modelled network. These delays have been characterised using Aegis train movement data provided by GWR, are part of the incident input data file and can be altered by the user if required.

Appendix 3 contains more detailed information about how these four types of delays have been characterised.

2.4 Calibrating the railway performance model

The agent-based model needs to be calibrated to provide confidence that it represents real rail services adequately. We calibrated the model to ensure that the model properly moves train services through the modelled network, and to test that the input data is correctly describing the railway system; the locations, platforms, lines, and junctions that make up the rail network, the timetable of train services to be modelled, and the incidents and delays that are to be applied during modelling.

The calibration process needs to be repeated if there are significant changes made to the model's input data, e.g. if a new area of the rail network is modelled.

We formed a four-step calibration process, described in Appendix 4 and summarised below. This process demonstrated both the impact of model assumptions on results, but also the impact of the complexity of the railway system on performance. The railway system has many interdependent components, that are highly coupled (dependent on each other), small changes in one component of the system (for example a subthreshold delay to a service) can result in large changes in the outputs (the punctuality of train services elsewhere in the system). This behaviour, observed in real life, was also observed in the model.

Step 1: Calibrate with timetable running (no incidents, no off-network delays)

In this initial step the model was run without any incidents causing delays to trains, to simulate purely the correct operation of timetabled services. This step was used to identify, and as far as possible remove or reduce, limitations in the modelling assumptions or input data. For example, in some places it is critical that the number of platforms and lines available for train services to use is closely matched to the timetabled services. If this was not the case we would observe rapid build up of delays. We worked with GWR and steering group partners to identify and remove these model artefacts.

These calibration model runs resulted in 4.3% of train services arriving at stopping locations later than one minute (first bar in Figure 3), over a third of these are within 1-3 mins (first bar in Figure 4). Ideally, we might expect in this step to see no trains arriving late. These delays may be due to modelling artefacts we were unable to identify or small delays inherent in the timetable. The project steering group advised that there are likely to be small delays due to imperfections in the timetable.

Step 2: Calibrate with subthreshold delays only

In this second step we introduced subthreshold primary delays.

The second bar in Figure 3**Error! Reference source not found.** shows that an additional 0.7% of trains arrive late, resulting in 5.0% of station stops over one minute late.

Figure 4 displays the same simulation results, showing that many of the additional late train services are arriving between 1 and 3 minutes late and that there are additional train services arriving between 5 and 10 minutes late.

It is not possible to identify the root causes of subthreshold delays in the current modelled delay results. But, it may be that in the real-world some of the subthreshold delays grow in size during a train's journey, until they are large enough to be included in the attribution process and are attributed as threshold delays. We believe that there will be value in modifying the approach to subthreshold delays, to model some of these potential precursors to attributed threshold delay. This approach might help to improve the modelling of subthreshold train service delays, and at the same time allow users to explore the possible causes of subthreshold delay.

Step 3: Testing the generation of attributed primary delays

This step involved the model artificially creating a large number of virtual primary incidents, sampled from the distributions defined in the input data. These are saved as a delay test file and can be analysed to check that a representative distribution of incidents and delays are created from the input data.

We tested the incidents generated by the demonstration model, and they were a reasonable reflection of the national attributed delay data, and are appropriate for a demonstration of the agent-based model.

It is possible to fine tune the incident input data to represent a specific region, or other part of the railway network. The model can then be calibrated to represent a specific part of the network.

Step 4: Calibrate with attributed primary delay input data (creating a baseline)

The third bar in Figure 3 shows the modelled train service punctuality with all of the delay types in place; subthreshold primary delays, threshold primary delays, off-network delays, and any resulting reactionary delays. In our demonstration, this represents a typical days of train services, a baseline.

The result is that 18% of train services are modelled as arriving late at their station stops. An additional 13% of trains arrive late when this is compared with the timetable only simulation (i.e. without any primary delays).

Figure 4 displays the same simulation results, showing that the majority of late arriving train services are arriving between 1 and 15 minutes late.

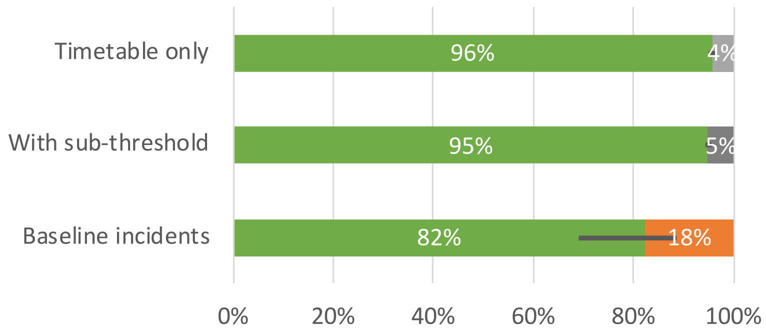


Figure 3: Model calibration, train service punctuality (all stations)

Note: the horizontal line in the lower bar of Figure 3 denotes the minimum and maximum values of train service punctuality across all simulations.

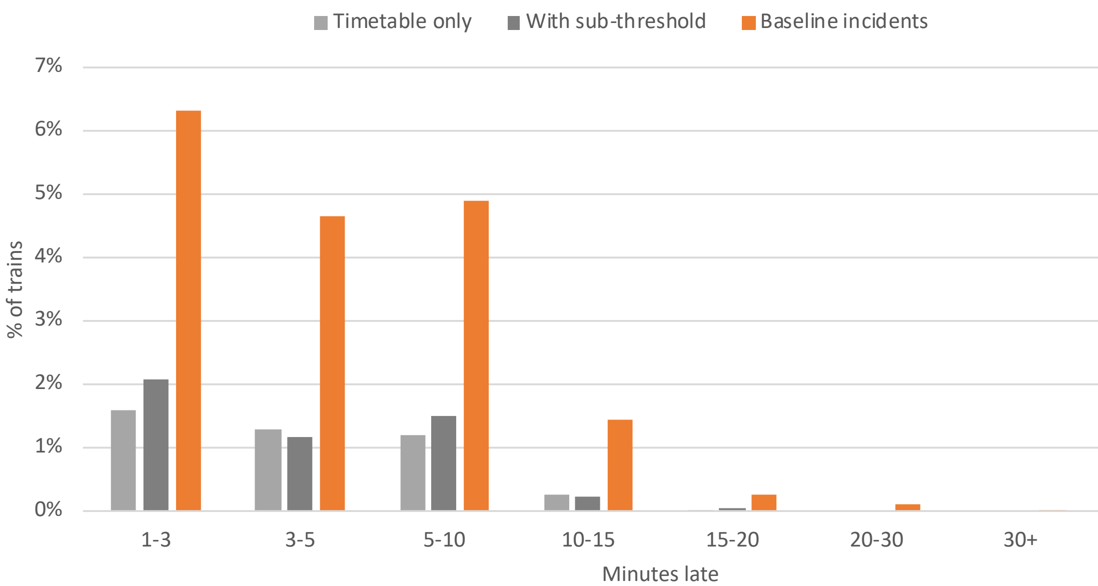


Figure 4: Model calibration, train services arriving late (all stations)

Figure 5 is using the model's heat map display to show train service punctuality along the demonstration route. The size of the circles represents the total amount of train service delays at each location and line.

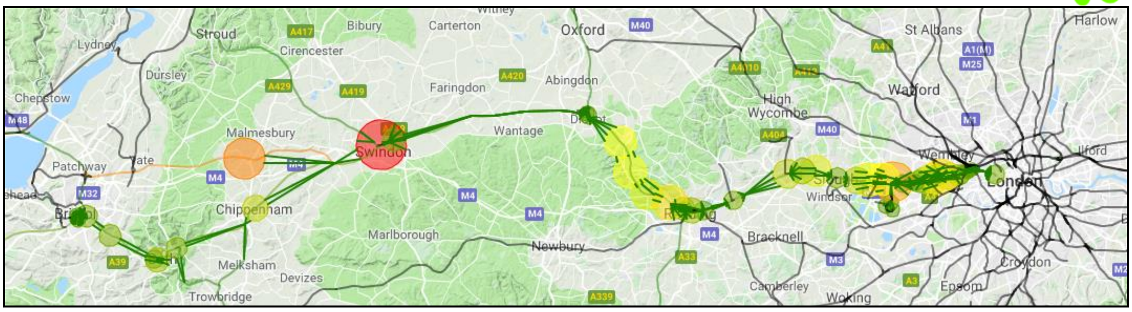


Figure 5: Model calibration, punctuality heat map

The proportion of train service delay minutes due to the different types of delays calculated by the model is shown in Figure 6. Sixteen percent of delay minutes are due to reactionary delay, with an equivalent proportion due to subthreshold delays. Twenty four percent of delay minutes are due to attributed primary delays, and 44% are due to off-network delays.

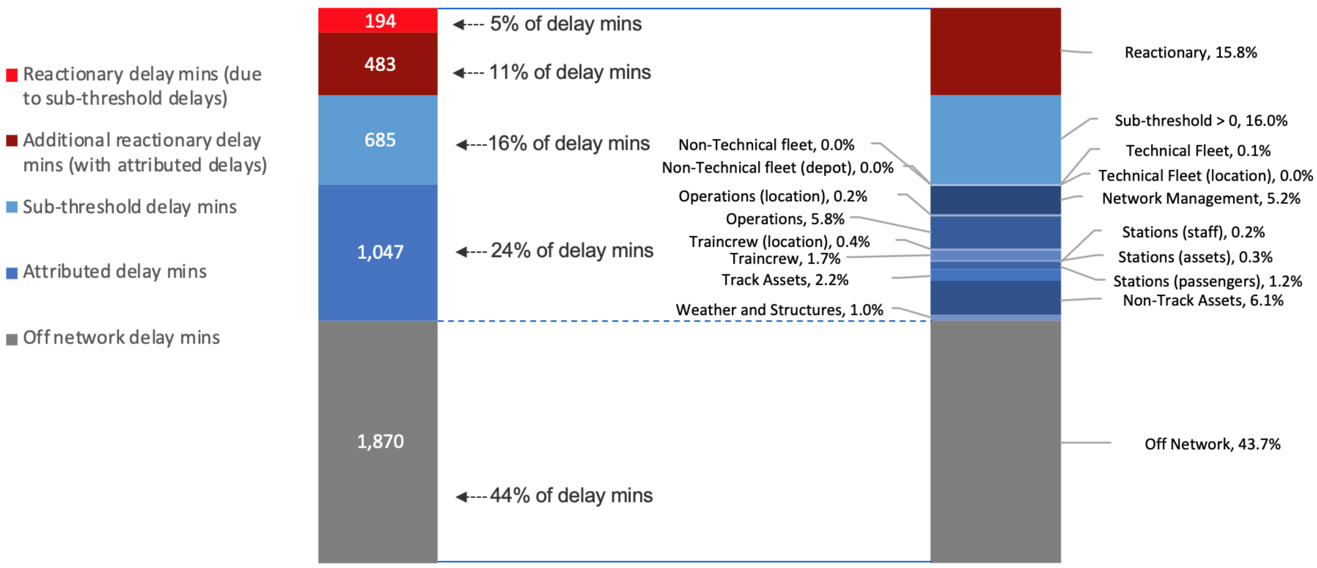


Figure 6: Baseline delay minutes (average per day)

The modelled reactionary delay includes any delay greater than 0 seconds, and is triggered by both subthreshold delays and attributed incidents. Every cause of reactionary delay is recorded and available for analysis through the interactive visualisations. This provides a wealth of information about the relative importance of the type and size of causing delays, to direct finding effective interventions to reduce reactionary delay.

We have been able to spend a short time looking at some causes of reactionary delay in our demonstration model runs, but the relationships between reactionary delay and the causes of reactionary delay could be explored further.

Off Network delays are a significant proportion of the delays being modelled. This is due partly to the simple assumptions used to model off-network delays in the current model, which may be artificially exaggerating the delays – these assumptions should be re-assessed. The effect of the simplifications is exaggerated because a large number of train services enter the modelled network from the many spurs off the Paddington to Bristol route. If the modelled area was widened then this effect would decrease.

The characterisation of off-network delays needs to be improved, both by increasing the modelled area to reduce the number of train services affected by this uncertainty, and improving the characterisation of off-network delays.

The attributed delay minutes are split between the types of primary incidents, with the highest proportions of threshold primary delay minutes due to non-track assets (6.1%), Network Management (5.2%), and Operations (5.8%). These might be candidates for interventions, but it would be premature to assume this before fully understanding the causes of the reactionary delay by using the interactive visualisations to explore the modelling results in more detail.

2.5 Exploring reactionary delay using interactive visualisations

Interactive visualisation is an integral part of the approach used in this project. The model outputs comprise **thousands of primary and reactionary delays**, generated from multiple model runs. These need to be considered from different perspectives and different levels of abstraction (i.e. by train service, by location and by time-slice) in order to verify, interpret and take actions from the model results.

The high-level statistical summaries produced by the model are useful, but these hide important detail. Information visualisation is capable of showing **hundreds of statistical values** for many locations, trains or times, **in interpretable ways**. For example, a map of delays is more interpretable than a table of numbers for each location. This is because when appropriately represented graphically – e.g. using bar length (preferably) or colour lightness – humans can instinctively read the values and can make comparisons. **Interacting with visualisations** allows the user to obtain more details or to compute statistics for specific situations of interest. An example of a visualisation is shown in

Figure 7.

Our approach

The visualisation tools were developed in close collaboration with Risk Solutions and GWR were updated on progress throughout. We deliberately configured the way we worked so that we could work in parallel, by establishing an agreed file exchange format. The visualisations supported the model calibration and in doing so, established the information we thought TOCs would need to use the tools.

Development of the interactive visualisation tools was iterative. We did some initial paper sketching, but very quickly built **high-fidelity prototypes** that could be demonstrated and shared so others could test them. We tracked the versions of the prototypes and the comments made. This was mainly between Risk Solutions and City, but GWR were able to provide feedback during face-to-face sessions and video demonstrations. We used these comments to improve or design new prototypes.

The eventual tool (summarised below) is developed from one of our later prototypes that considers data from a location perspective. We developed it beyond a simple prototype and incorporated some of the functionality from previous earlier prototypes. Four prototypes were produces in total: (1) an interactive delay explorer for historical attributed data, (2) an interactive diagram that shows how individual trains interactive with each other by location and in time (this was incorporated into the eventual tool – see the end of Appendix 5), (3) an interactive visualisation that helps identify trains with different delay properties and (4) the location based visualisation that developed into our final tool.

The prototypes and eventual tools were written in Java and are – as such – cross-platform, working on platforms with Java Runtime Environment installed. These can be installed on Windows, Mac and Linux platforms.

The eventual visualisation tool

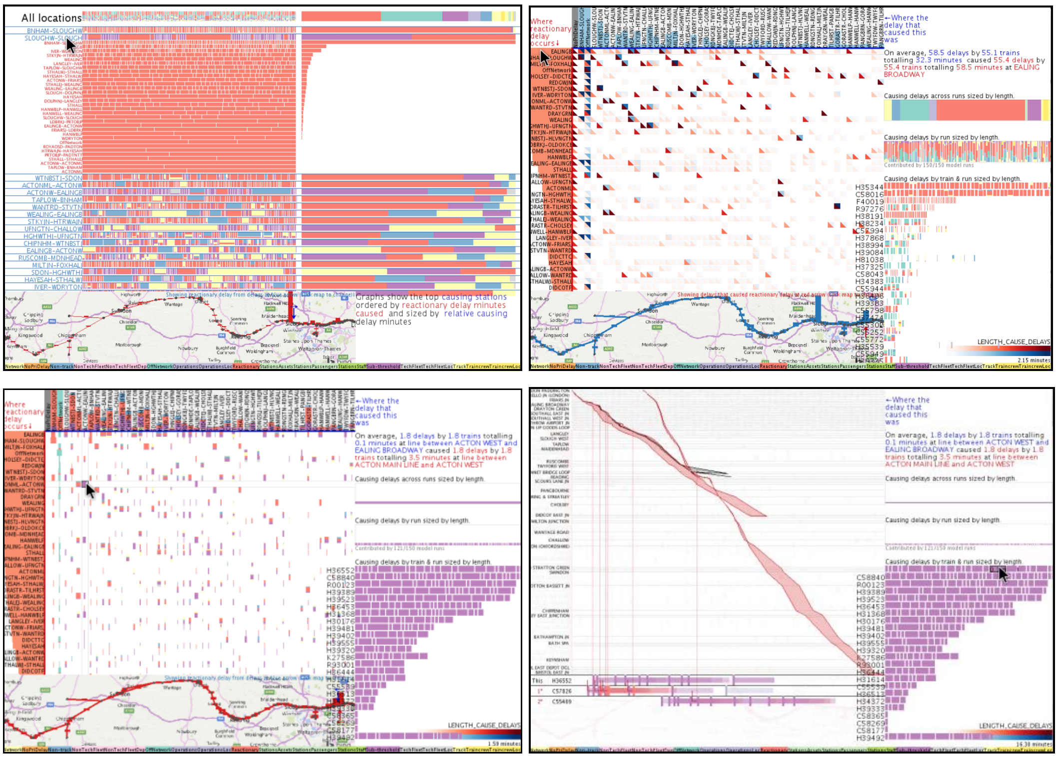


Figure 7: Screenshots from the eventual interactive visualisation tool. See Appendix 5 for explanations.

We have designed the interactive visualisations to help in the following ways:

- **End-user TOCs identifying problematic trains, locations and times**: The delay data needs to be considered **by location**, in order to identify problematic locations at which interventions might be appropriate. The variation between model runs indicates the **range of possible consequences of delays**.

  Our visualisation tool *identifies the locations that cause the most reactionary delay. It then links the location at which the reactionary delay occurs to these locations. For each of these, it:*

  - *summarises the relative lengths of the primary delays by primary delay type, and how this varies between simulations.*
  - *it shows the top five locations at which the reactionary delay occurs.*

  This allows the user to identify the most problematic locations (those that cause most reactionary delay), the locations where the reactionary delay occurs and therefore the locations at which it might be worth trying to reduce primary delays.

  Future work: *for this feasibility study, we only do this by location, but doing it by train and time of day might also be valuable.*

- **Model development and involvement of TOCs**. Modellers need to verify the **assumptions** and **mechanisms** used by the model by studying delay information produced by the model for **specific trains, locations and times**, particularly for those giving unexpected or extreme results. Interactive methods allow full details of locations, trains and times to be displayed in an interpretable way. This can be used as a basis for discussion with the TOC and other stakeholders and helps the model developer identify bugs in the models and supports verifying whether these have been fixed.

  Our visualisation tool *allows a detailed examination of the mechanisms of delay, where causing and reactionary delay is summarised for every pair of stations. Full details of the mechanism by which trains are delayed, in a specific model run. can be inspected with a simple mouse-over. The mechanism is displayed with a location vs time graph, on which all the involved trains are plotted.*

  Future work*: we would also like to adapt this tool to work with real attributed delay data and train movement data, so that the mechanisms of delays in a previous period can be explored.*

- **An end-user TOC designing interventions.** Although the above visualisation and interactions would need some training to use, it is likely to be helpful to TOCs who are designing interactions. Once the problematic locations and trains have been identified, obtaining the details of the causing delay types, the specific trains that cause the delays and those that pick up the resulting reactionary delays can be investigates for each model run.

  Our visualisation tool *would support this, as outlined above.*

  Future work*: We would like to do some dedicated work with a TOC to establish which statistics, visualisations and interactions would best support designing interactions.*

- **An end-user comparing the effect of an intervention.** Once possible interventions have been identified and coded, the model produces delay results for this intervention. Interactive visualisation can be used to understand the impact of these interventions and the details of this by location, train and model run. It is likely that the effect of the intervention will vary by model run, so that some interventions may have more consistent effects than others.

  Our visualisation tool *allows side-by-side direct comparison, but does not currently directly encode the difference between model runs.*

  Future work: *We would like to develop tools to compare different interventions.*

Appendix 5 describes how the visualisation tools can be used to explore the causes and consequences of reactionary delay. This detailed walk-through is summarised as part of our demonstration of how to use the Rail Performance Tools, described in the next section of this report.

3 Findings

Our research objective was to explore whether creating agent based modelling and interactive visualisation tools are useful in understanding rail service delays and performance. We have created tools for railway stakeholders and propose that:

- Agent based modelling is a powerful tool to model the complex and sensitive interactions between dependent train services in a railway system.
- Interactive visualisations are helpful to identify underlying patterns, the causes and consequences of delay in complex data
- Modelling interventions designed to improve service performance can demonstrate how effective each intervention can be

The following pages in this section describe how the tools can be used to identify the causes and consequences of reactionary delay, and how the benefits of interventions can be quantified in terms of improved punctuality and reduced delay minutes for train services.

3.1 A demonstration of how to use the railway performance tools

The following process describes how the Rail Performance Tools can be used by GWR, other Train Operating Companies, and Network Rail. This is a demonstration of the potential to model train services on a route not unlike Paddington to Bristol. Given a properly characterised input data set, the model and visualisations could be used to simulate a real-world route in any part of the UK rail network.

The process demonstrates how the project tools can be used to identify where reactionary delay is occurring, prompt suggestions for potential interventions to improve service performance, and quantify the performance improvements that are likely to be the result.

Step 1: Model typical train service interactions and establish a baseline of rail performance

The first step would be to choose and design a baseline scenario (e.g. a typical day of train services), and represent this as input data for the Rail Performance Model. Run the model using this scenario, at least 150 times, to produce a wide range of possible train service interactions. The result data represents the performance baseline against which improvements can be measured, and a rich source of information about how trains services can conflict, causing reactionary delay.

The detailed train service delays are recorded by the model for analysis and visualisation, together with high level service performance statistics. This represents the baseline performance.

DEMONSTRATION: We used the calibrated input data files for the Rail Performance Model, previously described, as our demonstration baseline scenario. This produced detailed delay data representing service performance for train services using parts of a route not unlike that from London Paddington to Bristol.

Step 2: Identify the root causes of service delays

The second step would be to use the interactive visualisations to explore the modelling results from the baseline scenario, to identify what is causing service delays. **This process is described in detail in Appendix 5**, in summary:

1. The initial summary view shows the types of delays that are causing the most reactionary delay, and a list of locations where these delays occur, sorted by those that cause the most reactionary delay. The proportions of causing delays are indicated, as well as variation between model runs. *The user is looking for the locations where they might consider an intervention to reduce these delays, and the types of delays to target.*

2. Interacting with this view allows the user to explore the amount of reactionary delay that is caused by types of delays at the (causing) locations, and where the train services are affected by this reactionary delay. *The user is looking for the causes of significant reactionary delay at multiple locations.*

3. The view also shows the number of simulations affected by the (causing) delays and the amount of reactionary delay that results. *The user is interested in the common causes of reactionary delay that feature in many of the model runs, or perhaps the causes of the worst performing days.*

4. The user also has the option to explore the results in additional detailed views showing:
    - Every possible pair of locations in the modelled network and how delays at one are affecting reactionary delays at the other
    - Every train's journey in each simulation to see how they have been affected by other trains
    - Study the mechanism of the delay and whether this is consistent across model runs.

DEMONSTRATION: Using the interactive visualisations and our demonstration baseline results we confirmed that; reactionary delay causes further reactionary delay, subthreshold delays cause the most reactionary delay, and after this, a number of the attributed primary incidents are causing most reactionary delays. In our demonstration these were:

- Non-track incidents
- Operations incidents
- Network Management incidents

These incidents caused the most reactionary delays when they occurred between specific locations, particularly between:

- Chippenham and Wootton Bassett Junction and Swindon
- Acton West and Ealing Broadway, and West Ealing and Ealing Broadway
- Wantage Road to Steventon
- Uffington to Challow

Reactionary delays occur along the whole route between Paddington and Bristol, but particularly on the approach to Paddington (Southall to Ealing Broadway). However, it was reactionary delay between Burnham and Slough West that caused the most cascading reactionary delay at other locations.

The worst day was simulation number 116: An operational delay at Taplow delayed 40 train services, by 3-9 mins, causing reactionary delay to 41 other train services, along the whole route. These delays were concentrated at Ealing Broadway, Burnham to Slough, and Wantage Road to Stevenson).

Step 3: Design interventions aimed at reducing delay

The next step is to design and characterise interventions intended to reduce reactionary delays. During this step railway stakeholders would design interventions targeted at reducing or preventing the causes of significant delays that they were interested in resolving.

DEMONSTRATION: We characterised three sets of interventions, and explored what service performance improvement these could achieved using the agent-based model.

Intervention 1: Reduce number and duration of all primary incidents:

- halve the probability of incidents occurring
- halve the incident duration
- halve the delay duration

This is probably not achievable in practice; a very significant effort would be required to achieve this level if impact. But we wanted to demonstrate what a significant performance improvement would look like when modelled and visualised.

Intervention 2: Reduce number and duration of just the types of primary incidents that are causing the most significant amounts of reactionary delay (halving as above):

- network management incidents
- non-track incidents

- track incidents
- off-network incidents

In practice an intervention to achieve this impact would also require a considerable effort, but we wanted to demonstrate the effect of targeting interventions at the most significant type of delays.

Intervention 3: Reduce duration of all incidents by 50%

This intervention was designed to demonstrate the effect of focussing efforts to significantly improve recovery times from incidents.

Step 4: Model the impact of interventions to predict service performance improvement

Having designed some candidate interventions, in this final step in the process users would adjust the model input data to represent each of their interventions. The model is then used to simulate the interventions, and saves both the detailed results for further analysis using the visualisation tools, and the overall punctuality and delay performance. The model can run a series of scenarios, one after the other, so that results can be compared and the best intervention selected.

DEMONSTRATION: We used the model to run the three scenarios described in step 3. The next section of this report describes the performance results of each scenario, compared with the baseline performance.

3.2 A demonstration of the results of modelling interventions

Figure 8, Figure 9 and Figure 10 compares the performance of the three demonstration interventions with the baseline model runs created during the calibration phase of the project.

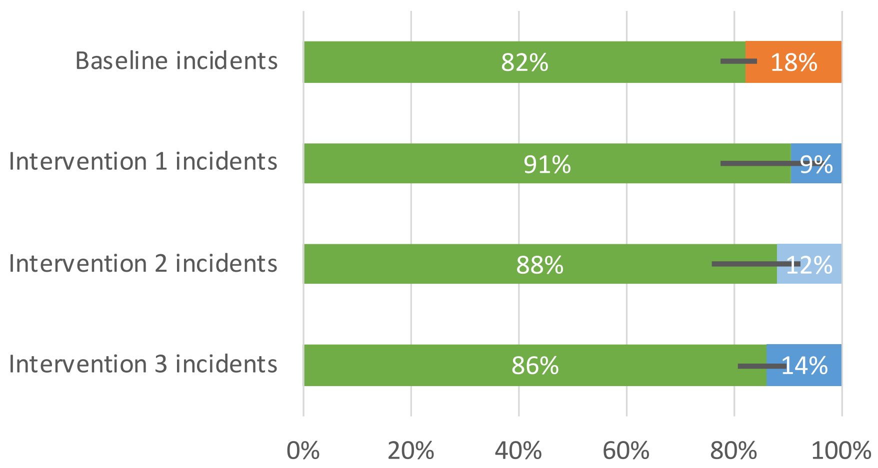


Figure 8: Train service punctuality (% trains arriving within 1 min, all stations)

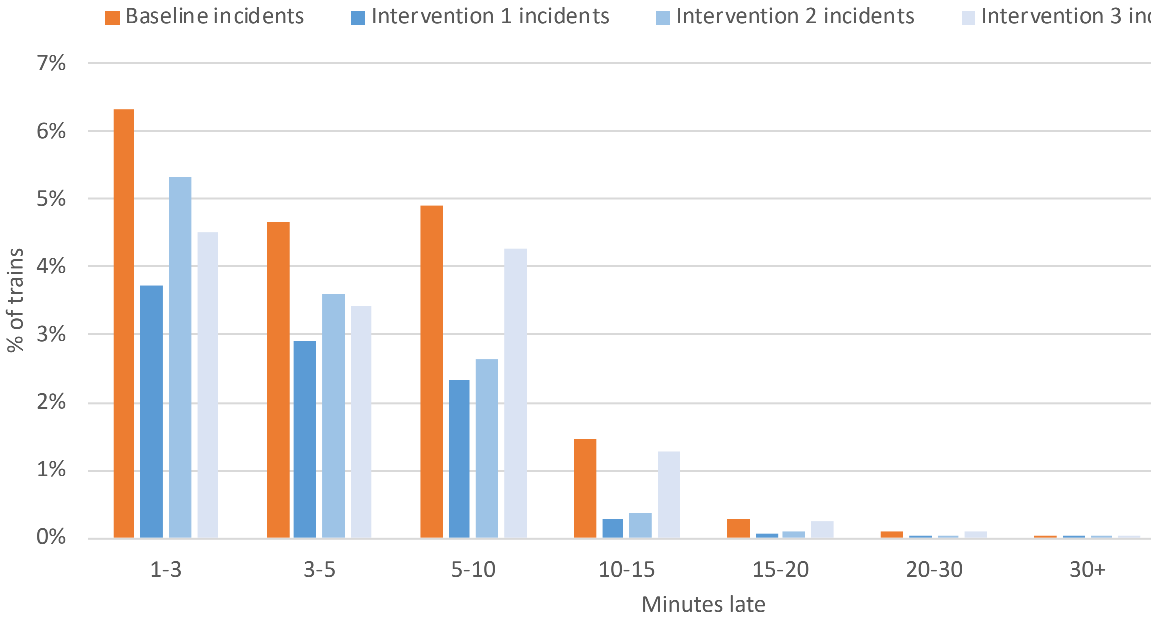


Figure 9: % of train services arriving late (all stations)

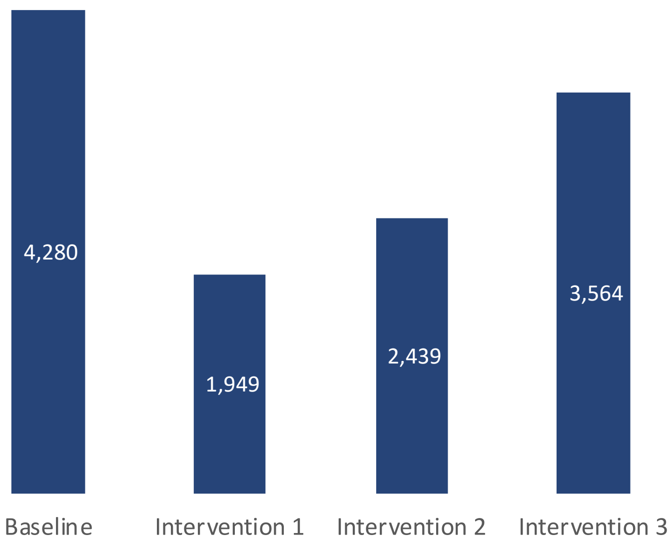


Figure 10: Train service delay minutes (average per day)

Intervention 1 (reducing incidents by 50%) results in halving the number of train services arriving more than a minute late (Figure 8), and halving the delay minutes to train services (Figure 10). At first glance this appears linear, but we need to model more scenarios to establish the actual relationship between incident rates/duration and punctuality.

However, **intervention 2** (reducing only 3 targeted high impact incident types by 50%) also results in significantly fewer train services arriving a minute or more late (43% in our demonstration), and a significant reduction in delay minutes (37% in our demonstration).

It is worth exploring this with some more modelling of how targeted interventions can improve punctuality, using the interactive visualisation tools to help identify the principle causes of reactionary delays and hence target interventions. This approach will help focus scarce resources where they can count the most.

It was also interesting to note that **intervention 3** (speeding up recovery from incidents) also results in materially fewer train services arriving a minute or more late (22% in our demonstration) and a material reduction in delay minutes (17% in our demonstration).

Therefore, this confirms that a faster incident recovery can reduce overall reactionary delay. Again we need to explore in more detail what improvements in recovery are possible and at what cost for different types of incident.

Conclusion: We have demonstrated in this project that the Rail Performance Model can be used to test different types of intervention, and quantify the resulting performance in terms of train service punctuality and delay minutes. The model records the delays to all train services at every at each TIPLOC and between every pair of TIPLOCS in their journey. Therefore, it would probably be possible to calculate a wide range of alternative performance indicators if required.

We were only able to test three intervention scenarios, but the process of setting up the model for a new scenario was not time consuming. Therefore, there is plenty of scope for the tools to be used by Network Rail and/or Train Operating Companies to estimate the benefit of many more alternative interventions.

The original proposed purpose of the project was to test whether we could create tools to help identify the causes and consequences of reactionary delay, and this has been achieved. Over the course of the project we have also identified some other ways to use the tools and to improve their accuracy.

3.3 Additional value emerging from the project

An additional finding is that there are other reasons why it is beneficial to model train services over a large area, not just to explore reactionary delay. Through discussion with the project steering group and GWR, we have identified some ways in which the current tools could be used outside the proposed remit of our project, and include:

- **Timetable changes**: modelling the impact of adding or removing train services from the timetable, or stress testing a completely new timetable before introduction

- **Speed restrictions**: quantifying the impact of speed restrictions and extended possessions, on other parts of a route
- **Joint Performance Improvement Plans**: providing quantified evidence of expected performance improvements
- **Contingency plans**: testing existing operational contingency plans to large scale disruption, and improve their effectiveness.
- **Dwell times:** The potential impact of reducing dwell times

3.3.1 Dwell time analysis

To demonstrate one of these ideas, we took the opportunity to alter the dwell time used by the model, to explore the effect that this could have on service performance. The results are shown in Figure 11 and Figure 12.

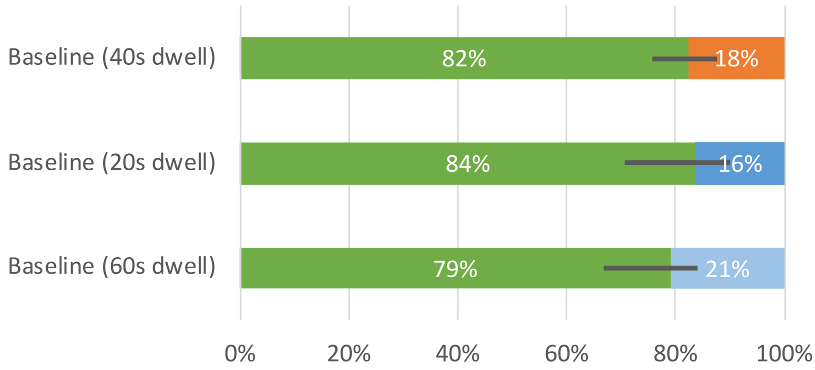


Figure 11: Train service punctuality (all stations) affected by dwell time

The model currently applies a fairly crude approach to dwell times (this was not the focus of our research). Reducing dwell times for late train services from 40s to 20s allows trains to catch up with their timetabled services if they are delayed. This reduces the number of train services arriving late at station stops from 18% to 16%. Increasing dwell time from 40s to 60s for late train services increases the number of train services arriving late from 18% to 21%.

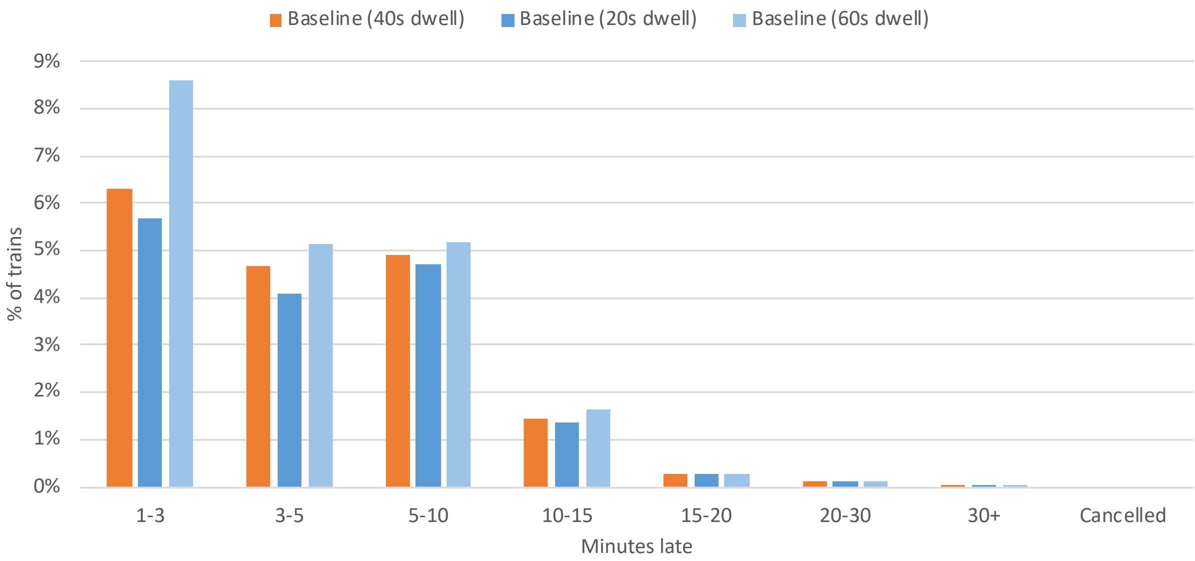


Figure 12: Train services arriving late (all stations) affected by dwell time

Figure 12 shows that these changes mainly affect train services arriving between 1 and 3 minutes late.

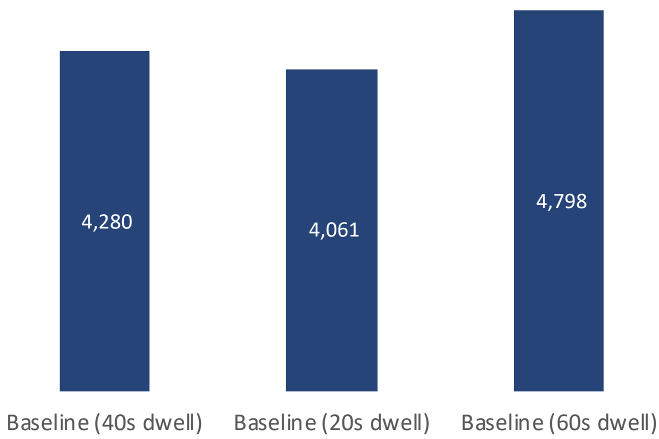


Figure 13: Train service delay minutes (average per day), affected by dwell time

A companion sandbox research project, conducted by Southampton University, examined the characteristics of dwell times. This project may be able to provide valuable information to improve the dwell time assumptions used by our Rail Performance Model.

4 Impacts and benefits

As mentioned before, **primary** delay to train services has remained stable over recent years, but **reactionary** delay has grown steadily, and has been difficult to understand and control. The effects of reactionary delay are dependent on the interaction between multiple trains and the network characteristics of the track where the delay happened. Reactionary delay can cause more reactionary delay in cascading reactionary chains affecting multiple trains. These affects are difficult to predict.

This feasibility study has created tools that can be used to understand some of these characteristics, and generate the following impacts and benefits:

Gain insight into the underlying causes of reactionary delay: The Rail Performance Tools can be used to explore the complex causes and consequences of reactionary delay. This helps identify new ways to improve performance.

Quantify performance improvements: the tools also quantify the likely impact that interventions might have on service performance. This is of great value to Train Operating Companies and Network Rail:

- **Avoid abortive trials:** users can quickly test various options for improving performance without the expense of time-consuming trials.
- **Use scarce resources carefully:** users can model a wide range of possible interventions to improve performance, then select the ones that demonstrate the best performance outcomes.

Resilient solutions to improve performance: the tools test interventions across a wide range of possible combinations of incidents and train service delays. Users can be confident that the interventions they choose to implement will be resilient and reliable.

Using the tools in this way to improve performance was the original concept proposed for the sandbox research. And the tools in their current form are available to GWR, Network Rail and RSSB to use and develop further.

Through project discussions Network Rail, GWR and the project steering board, we have identified that the current tools could also be used in a number of other ways, with some additional benefits:

- **Model timetable changes:** avoid unintended consequences following changes to the timetable by modelling these along with typical incidents and train service delays to check their robustness. It would be beneficial to test the performance impact of adding new services to the timetable, to avoid overcommitting pathways. And, to remove services from congested pathways to reduce reactionary delay.
- **Model specific incident/situations:** improve responses to a range of situations, for example, develop the best ways to deal with seasonal issues (leaves, broken or

buckling rails), test contingency plans for the most disruptive days, how to deal with the impacts of extended possessions, or reveal the wider impact of introducing temporary speed restrictions.

5 Recommendations

The Rail Performance Tools are at a prototype stage, and can be used in their current form, but require some development so that they can be used readily by rail industry stakeholders. There is now an opportunity to develop the current tools to higher levels of technology readiness and provide more evidence of their value. Appendix 6 describes the Railway Industry Readiness Levels (RIRL) and Technology Readiness Levels (TRL) achieved by this project, and how the recommended next steps and longer-term development ideas could deliver higher levels.

Our project recommendations are to:

1. Improve accuracy of modelling train services to increase confidence in results
2. Improve the usability of interactive visualisations so that they can be more easily adopted by railway stakeholders
3. Automate data preparation processes
4. Scale up the tools so that they can be applied to any part of the rail network
5. Explore how tools can be used for other purposes
6. Enhance the tools so that they can be used for other valuable purpose.

The following paragraphs outline the actions required to achieve these recommendations.

5.1 Improve accuracy of the modelling

There are a number of ways that the current agent-based model could be quickly improved, without adding sophisticated new features.

Recommended actions:

- **Model more routes in a larger area** covered by a train operating company to reduce the uncertainty associated with modelling off-network train services (see actions in recommendation to scale up the tools).
- **Improve modelling algorithms**. For example; resolving conflicts at junctions and stations using train service priorities, add rules to cancel train services, model station and line capacity with more granularity.
- **Improve modelling of incidents and delays**. For example; being able to introduce specific delays at specific locations and to specific train services, model subthreshold

delays as emerging attributed delays, improve characterisation of off network delays, consider if delay duration varies with train service type.

- **Improve method of modelling dwell times** to take account of passenger loading at different times of day / weekend, station characteristics, other parameters identified by the feasibility study conducted by Southampton University.

Benefit: improved confidence in the accuracy of model results when testing alternative scenarios aimed at improving performance.

5.2 Improve interactive visualisations

Recommended actions:

- **Use the visualisation tools to review historical attributed delays.** Our tools have focussed on modelling output, but they could be adapted to work with any historical attributed delays. ***Potential benefits:*** *to* ***help understand the mechanism of existing delays*** *and to assess the* ***real-world impact of an intervention*** *after it has been implemented.*
- **Additionally present the results by train and time.** We focussed on identifying **places** at which delays happen and for which interventions should be designed, but we would also like to support identifying **times** and **trains**. ***Potential benefits****: some interventions may be better focussed on particular* ***times of day, day of week*** *and for* ***trains with particular characteristics****.*
- **User-centred work for model calibration.** The visualisation tools helped us check that the ABM was working as expected, but we were **not formally focussed** on this for this project. ***Potential benefits:*** *modelling may become more important in future and is it important that we produce good models.*
- **User-centred work with TOCs.** Although we aimed to run from user-centred approach with TOCs, we did not have the space we needed to do this formally. We would like to formally work with a TOC to:
    - Establish the best information, statistics and visual representation needed by TOCs to produce high-level summaries of the model outputs
    - Establish how TOCs design interventions and what information, visualisation and interactions are needed to support this.
    - Establish the best information, statistics and visual representation needs by TOCs to assess the effect of interventions on delays

  Potential benefits: *This will provide specifications that can be used to develop operational tools*

5.3 Automate data preparation processes

There is scope to automate the manual parts of the process required to prepare input data for the Rail Performance Tools, so that they can be quickly applied to another part of the UK rail network.

Recommended actions:

- Automatically derive the input data that describes the route to be modelled, selecting to use either the Network Rail Network Model, or the network model used by the Train Planning System (TPS).
- Characterise a particular set of incidents and delays required for modelling scenarios (derived from Network Rail attributed delay data), automatically creating the model's incident input data.

Benefit: Less effort required to characterise the Rail Performance Tools to a new part of the UK rail network, and less effort required to create a scenario for modelling from source delay data.

5.4 Validating the Rail Performance Tools

The Rail Performance Tools have been calibrated and demonstrated as part of this research feasibility study. We recommend that the tools are now validated, by characterising a larger part of a TOC's operational area, and characterising modelled incidents that are likely to be affecting services in this area. The resulting baseline performance can be validated using TOC performance data.

We have designed the tools to be data driven, so can be applied to any part of the UK rail network, when provided with the suitable input data files. Table 1 lists the activities required to tailor the tools and apply them to a part of the UK, and the source data that is required.

Table 1: activities required to apply Rail Performance Tools to a region or route

Activity	Source data required
Manually characterise routes and automatically extract timetable from CIF: there is potential to develop an end to end automatic process to do this.	CIF timetable Signalling route diagrams with TIPLOC and platforms
Characterise incident and delays: choose periods of interest within the last 12 months, characterise delays for these periods. There is potential to automate the process of characterising delays.	Network Rail historic attributed delay data

(Optional) **Characterise the subthreshold delays:** for the routes to be modelled (to improve accuracy of model)	Remote Condition Monitoring system GPS and TIPLOC timing data, e.g. Aegis, Nexala
(Optional) **Characterise the off-network delays:** for the routes to be modelled (to improve accuracy of model)	TBD
Calibrate model: do model runs and adjust modelling parameters where necessary	None

Recommended actions:

- Partner with one or more Train Operating Companies (TOCs), characterise the Rail Performance Tools to represent a larger portion of their operating area
- Identify key performance issues facing the TOC, and how the tools can help explore solutions
- Train key users how to use the tools and use them to model scenarios aimed at improving performance

Benefit: Validate the rail performance Tools and at the same time help the TOC to identify the best ways to deal with real world performance issues. This will also create additional evidence of the usefulness of the Rail Performance Tools and encourage further take-up. Through discussions with several TOC's we know there is interest in the following uses for the tools:

- Understand the causes and consequences of reactionary delay suffered by a TOC
- Create evidence to support Joint Performance Improvement Plans
- Analyse the impact of speed restrictions
- Test and improving contingency plans

5.5 Use the existing Rail Performance Tools for other purposes

Calibrating the agent-based model revealed a small amount of residual reactionary delay when there were no primary incidents, due to train service conflicts inherent in the timetable. Train Operating Companies have told us they recognise this as a feature of the timetable, citing days where their train services suffer small delays even when there are no incidents.

Recommended actions:

- Use the tools to stress test new timetables before they are introduced, providing evidence of the impact of new timetables on train service performance

- Use the tools to test the impact of adding new services to the timetable, and removing problematic services
- Explore how to enhance the tools to identify timetable 'white space' where there is potential to add train services, and congested pinch points where the timetable could be adjusted to reduce the likelihood of delays.

Benefit: Fully exploit the funding to date in the sand-box project to help improve the process of changing the timetable.

5.6 Enhancing the Rail Performance Tools

There is an opportunity to enhance the Rail Performance Tools through developments to increase their scope and features, and provide more value to rail stakeholders.

Recommended actions:

The following list of tasks should be considered for further research funding (the list is ordered from quick wins to more difficult/time consuming tasks):

- **Economic impact:** Calculate the financial impacts of delays in each modelled scenario
- **Rolling stock:** Model the impacts of the availability of rolling stock and the knock-on dependencies between timetabled train services
- **Train crew:** Model the impacts of the availability of train crew and the knock-on dependencies between timetabled services
- **Passenger demand**: Model passenger demand and adjust dwell times accordingly
- **White / dark space**: Identify where there is space in the timetable for additional services and highlight existing pinch points / congestion
- **Alternative infrastructure/timetables**: test alternative rail network topology, e.g. the impact of HS2 and other infrastructure investment
- **Machine learning goal seek**: Use machine learning to suggest the best interventions to reduce delays
- **Real-time**: use real time train movement data to create a starting point for modelling train services into the next few hours, identifying the critical train services or locations where delays could easily cascade, passing this information in real time to staff to help maintain service performance
- **Real-time passenger information**: Provide real-time predictions for through network journey timings for passengers

6 Next steps

RSSB should publicise this report and the value of the Rail Performance Tools widely to rail industry stakeholders, e.g. through the Better Operations Programme Board, and Operators Performance and Planning Forum.

Train Operation Companies and Network Rail should consider applying the current Rail Performance Tools to larger areas of rail services, and consider if the tools could be used to test timetable changes.

RSSB should consider short term development in the Rail Performance Tools; to improve modelling accuracy, the usability of interactive visualisations, and automation of the data preparation process. This will help take-up of the tools by Train Operating Companies. For example, as part of the Enabling Better Performance Research Challenge (PERFORM).

RSSB and Network Rail should consider longer term development in the Rail Performance Tools, prioritising the more advanced features that could be added. For example, the PERFORM programme, the Network Rail CP6 R&D fund and the proposed ORR £40 million performance innovation fund.

City University of London and Risk Solutions should:

- Review outcomes from the other RSSB feasibility studies and identify ways of combining the results (an example of which might be the dwell time analysis by University of Southampton, and the subthreshold delay analysis done by the University of Brighton)
- Study the Network Rail roadmap for Traffic Management Systems to identify how the Rail Performance Tools can evolve in line with operational strategy
- Look for opportunities to demonstrate the potential uses of the Rail Performance Tools to Train Operating Companies and other rail stakeholders (Rail industry conferences, meetings with TOC performance teams and candidate TOCs through the franchise programme, through their networks of academic and engineering partners)
- Apply for additional sources of funding to supplement investment from railway industry stakeholders for any of the potential development activities. This could include European funding, Innovate UK funding.
- Promote the project outcomes through journals, conferences, magazines, seminars, website case studies, social media.

Appendices

Appendix 1 – Project stakeholders

Appendix 2 – Modelling train interactions

Appendix 3 – Characterising delays to train services

Appendix 4 – Model calibration process

Appendix 5 – Using the interactive visualisations

Appendix 6 – Technology Readiness Levels

Appendix 1 – Project stakeholders

Steering group

Richard Raine, Network Rail

Karl Butler-Garnham, Network Rail

Nicola Curran, ScotRail

Tim Jones, RfL

Giulia Lorenzini, RSSB

John McMorrow, RSSB

Melissa Frewin, RSSB

Hassan Khalil, RSSB

Karen Lane, BAE systems

Justin Willett, RSSB

Liz Davies, RSSB

GWR representatives

Project team, City University of London and Risk Solutions

GWR

Sanat Misra, Performance Manager, Great Western Railway

Malgorzata Hodgkinson, Project Planning and Delivery Manager, Great Western Railway

Michelle Grimshaw, Project Planning and Delivery Manager, Great Western Railway

Simon Greenwood, Performance Analysis Manager, Great Western Railway

James Linley, East Regional Performance Manager, Great Western Railway

Mark Salisbury, Performance Analyst, Great Western Railway

Project team

Dr Aidan Slingsby, City, University of London

Dr Cagatay Turkay, City, University of London

Dr Helen Wilkinson, Director, Risk Solutions

Jonathan Hyde, Director, Risk Solutions

Chris Rees, Director, Risk Solutions

Data providers

Andrew Gosney, Head of Delivery – Data Management, Network Rail

Stephen Wiggins, Geospatial Analytics Manager, Network Rail

Kathryn Upson, Network Rail

Thomas Gallagher, Network Rail

Appendix 2 – Modelling train interactions

Trains are modelled to move along routes defined by the timetable input data. The timetable defines the expected arrival and departure times (to the second) at all intermediate TIPLOC locations between the train's origin and destination locations. The national CIF timetable does not include all TIPLOC locations, or timings, so these are added by the modelling tool during a CIF extraction process that needed to be created.

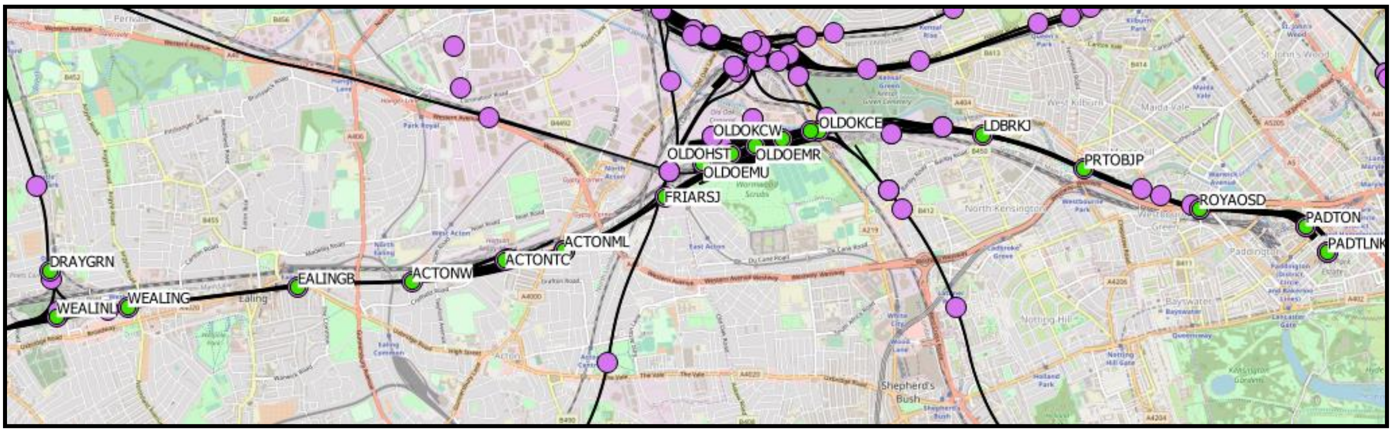


Figure 14: Example of infilled TIPLOCs used for modelling (green)

Trains can be delayed and conflict either at **locations** or on **lines** between locations.

Locations

There are four types of modelled locations, based on TILOC codes; stations, sidings, depots and junctions. Trains can stop (if timetabled) or pass, and occupy specific platforms, (these are virtual in the case of sidings and depots, and do not exist at junctions). Locations and their platforms are defined in the model input data.

Platforms are split into multiple segments (platform A, B, and sometimes C), which can be occupied by trains of various lengths of a 'multiple unit'. A multiple unit length occupies a single A, B or C platform segment. This means that shorter trains can occupy a single platform and is important so that trains can enter and leave stations from opposite directions (e.g. Birmingham New Street). A train will attempt to stop/pass the specific platform identified in the source CIF national timetable.

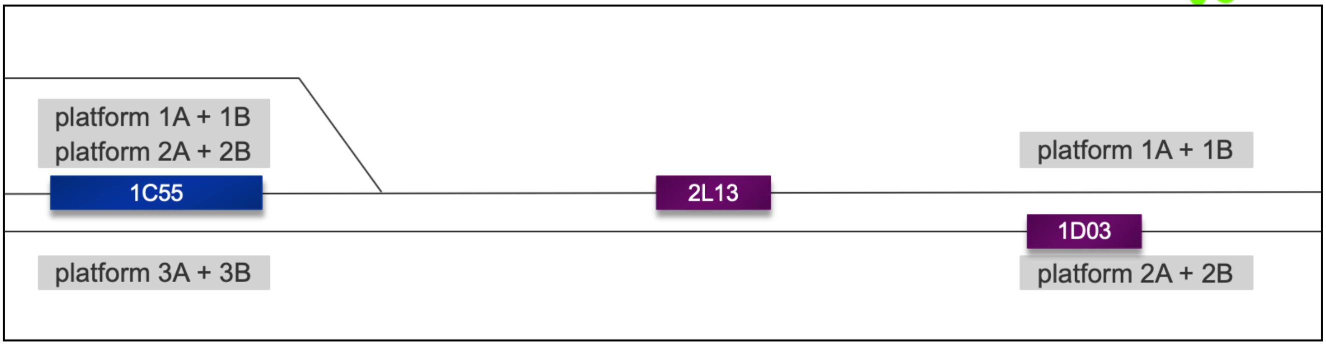


Figure 15: Station locations

Location delays: Trains can be delayed at locations by the types of incidents listed in Table 2 of the main report. These are defined in model input data, and can be altered if required.

Delays are measured in seconds and applied to trains while they are at locations, affecting their departure time. Multiple delays can occur to the same train at locations, but only the maximum delay is used to set the departure time and flagged in the delay results as the delay that contributes to the train's overall journey time. This means that some shorter delays are effectively masked by longer delays, and interventions aimed at reducing the longer delay may not return trains to their timetabled departure times because the short delay still occurs.

Dwell times: If a train is timetabled to stop at a location then it does not leave the allocated platform until the timetabled departure time, or if it is late, after a minimum of a 40 second dwell time (a default value chosen for our demonstration modelling). If the train is not timetabled to stop then it passes through without any dwell time, but the platform it requires has to be available!

There is potential to improve the modelling by adjusting dwell times to take account of time of day / passenger loading, type of station, or other factors. We are recommending working with the feasibility study led by Southampton University who have looked in more detail at dwell time.

We have also considered that dwell times could be modelled to take account of passenger loading on trains and at stations. This would add complexity but with suitable passenger demand source data this might be possible.

Reactionary delay at locations is recorded in the model's delay results in the following situations:

- Departure delays: when a train cannot leave its platform because the line ahead is occupied
- Arrival delays: when a train cannot enter the platform allocated in the timetable because the platform is fully occupied (this is actually calculated while the train is travelling between locations (see below)

We hoped to be able to model a train's dependency on there being available stock and train crew to run a service. This would create additional reactionary delay at locations while a train service waits for its forming stock, or train crew; both of which may be formed by a previous train service due to arrive at the location. We were not able to add this to the model within the budget and data constraints of the feasibility study, but are recommending that this is considered as a priority for further funding.

Lines between locations

Trains travel between locations on lines, defined in the model input data. The train's travel time on a section of track is defined by the time between consecutive TIPLOC locations in the national CIF timetable source data. If the model has had to add TIPLOCs that were not in the CIF timetabled journey then it calculates arrival times at intermediate TIPLOCs that are proportional to the length of each section of line between timetabled TIPLOCs. The modelled timetable is defined in input data, and can be altered for specific train journeys if required.

The accuracy of train service timing between consecutive locations could be improved by modelling stock characteristics such as acceleration, braking and top speed.

The model can use multiple lines between consecutive TIPLOC locations, if these are defined in the input data, matching the actual lines available in the real railway system. A modelled train will attempt to use the specific line identified in the source CIF national timetable, if it is occupied by a slower train then an alternative line will be selected by the model.

A more sophisticated routing algorithm could be considered for modelling, designed to reflect operational decisions made by Network Rail. This is one of our study recommendations.

Multiple trains can occupy the same line between locations, up to the number of signalling berths that exist on the line (fixed black signalling), defined in the input data. Trains cannot overtake each other unless they are on a different line.

Line delays are measured in seconds and applied to trains while they are travelling between locations, affecting their arrival time at the next location. Trains can be delayed between locations by the specific types of attributed primary incidents described in Table 2.

Trains can also be delayed, or catch up time, between locations. The model applies a subthreshold delay to every train while in between locations, which can be both +ve seconds or -ve seconds (recovery time built into timetable). This characterisation is described in the next section of this report.

In addition, trains can be delayed while travelling between locations if there is no spare capacity available at the next location.

Multiple delays can occur between locations, and unlike at locations, these all contribute to the train's journey time.

Junctions

Junctions are special cases of locations, and do not have any platforms for trains to stop at. Trains always pass through a junction, unless there is no capacity in the next section of line that they are timetabled to use. In these cases, an alternative line is considered to the next timetabled location, or the train waits in the previous section of line until the junction is clear.

The prioritisation of trains at junctions is currently first-come-first-serve. The accuracy of the model would be improved by prioritising trains through junctions based on their headcode priority, or other suitable mechanism.

Off-network delays

Around 59% of trains modelled in our demonstration route join the modelled part of the network part way through their journey (the locations with black dots in Figure 1). The model adds a delay to approximately one third of these trains, so that they are not all joining the model according to the timetable. This delay represents both attributed threshold delay and subthreshold delay.

Our recommendations include improving this aspect of delay modelling.

Model map representation of locations

The agent-based model includes a GoogleMap representation of the modelled network, which shows locations and lines, and can show trains moving while the model is running. This representation draws locations as circles (diameter showing the number of platforms) and straight lines between locations, one for each modelled track. The actual track is shown in grey lines. Figure 16 shows a portion of the modelled route with some trains (purple is a moving train, blue is a train at a station stop, red are trains that have reactionary delay).

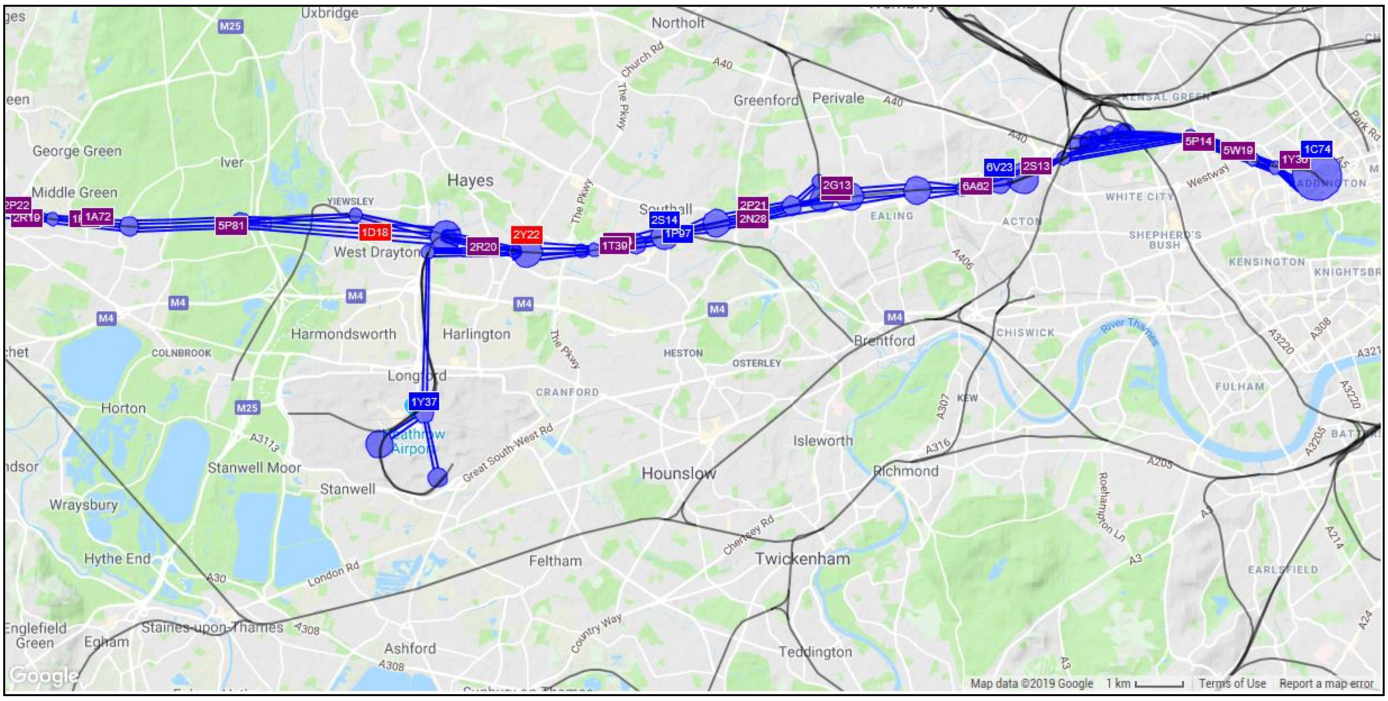


Figure 16: Model map

Processing train movements

The model maintains a list of trains that are currently being modelled, introducing new trains to this list as their timetabled departure time occurs. It also maintains a list of the next trains that need to either depart from a location, or arrive at the next location. This second list is sorted to identify the next train to move.

For each train that needs to move the model:

- Calculates the train's expected journey time to the next location according to the timetable,
- Adds any delays from incidents that are at the location, or line currently occupied by the train
- Checks for any conflicts with other trains, causing reactionary delays
- Calculates which delays contribute to the train's overall journey time
- Adjusts the expected arrival time at the next location
- Records delays and their attribution

After reviewing all of the trains that needed to move, they are all resorted to reveal the set of trains that need to move next.

Appendix 3 – Characterising delays to train services

Characterising attributed delays

Threshold delays of over three minutes are attributed by Train Operating Companies and Network Rail to identify action plans to improve operational performance. The Delay Attribution Board[2] is an industry body set up to 'Lead, Monitor and Advise' the Rail Industry on the attribution of train service delays and cancellations.

Attributed delay data is a key resource to characterise the known incidents that affect rail services, introduce train service delays, which can then interrupt other train services through train conflicts, resulting in reactionary delay. The agent-based model uses information in an input data file to characterise incidents and apply them to the modelled rail services. There is no limit to the number of incident types that the model can theoretically apply, but for the purposes of our demonstration we decided to reduce the 271 attribution codes to a manageable number for modelling, and with the help of GWR, ended up with 16 types of incidents, shown in Table 2.

Table 2: Types of incidents causing delays

Type of incident	location	Delay attribution codes[3]
Track Assets	anywhere	JA, JB, JS, IR, IS, IT
Non-Track Assets	anywhere	IB, ID, IP, I0, I1, I2, I3, I4, IA, IC, IE, IF, IG, IH, IK, II, IJ, IL, IM, IN, IO, J0, J2, J3, J7, JT
Technical Fleet	anywhere	DD, M0, M1, M8, M9, MB, MC, MD, ME, MF, ML, MN, MO, MR, MT, MV, MW, NA
Technical Fleet	stations/depots	M7 doors, MY coupling
Non-Technical fleet	anywhere	MS short stock
Non-Technical fleet	depots	MU depot issue
Network Management	anywhere	I5, I6, I7, I8, IQ, IZ, J4, J5, J8, J9, JG, JL, JP, JX, OA, OB, OC, OD, OH, OJ, OK, OL, OM, ON, OP, OQ, OS, OU, OV, OW, OZ,

2 http://www.delayattributionboard.co.uk/about.htm
3 Incident types and their associated attribution codes were chosen by GWR

Type of incident	location	Delay attribution codes[3]
		Q1, QA, QB, QM, QN, QP, QT, ZS, ZU, ZW, ZX, ZY, ZZ
Operations	anywhere	DB, FA, FH, FJ, FK, FL, FM, FN, FO, FS, FU, FX, FZ, RL, TB, TF, TM, TN, TO, TP, TR, TS, TU, TX, TY, TZ
Operations	stations and depots	RD, RK, TA, T8
Traincrew	anywhere	FC, FG, FI, FP, TG, TH, TJ, TK, TW, T2
Traincrew	stations/depots	FE not available, T3 waiting connections other transport, TI rostering
Stations (staff)	stations	R1, R2, R3, R4, R5, R6, R7, R8, RI, RJ, T4
Stations (passengers)	stations	RB, RC RO, RP, RQ, RR, RT, RU
Stations (assets)	stations	RE, RH, RV, RW, RY, RZ
Adhesion	lineside	OE no sandite, QH leaves, QI leaves
Severe Weather, Autumn & Structures (excluding adhesion)	anywhere	IV, IW, J6, JD, JH, JK, OG, Q3, QJ, X1, X2, X3, X4, X9, XH, XT, XW

As GWR had chosen Wednesday 21 March 2018 as a typical timetable day to model, we analysed a month of national attributed delay data from March 2018, to define these 16 types of incidents and the train service delays that typically result. Sundays were excluded due to engineering works, and we did not include reactionary delays. The dataset consisted of 26,000 incidents, resulting in 46,000 primary delays.

Each type of incident was characterised with:

- **Incident frequency**: the daily rate at which incidents occur (per km, per train journey, per location)
- **Incident duration**: a distribution of incident duration measured in minutes, with various types of distribution available, e.g. normally distributed, triangular distribution)
- **Delay duration**: a distribution of the duration of delays caused to train services as a result of the incident, with various types of distribution available.

A normalisation process was used to transform national incident frequencies to those that are more appropriate to the London Paddington to Bristol route modelled.

We recognise that any inherent bias in the attribution process will be present in the modelled delays.

It is possible to group incident types in a different way, or represent incidents for an alternative time of year to tailor the model for specific situations. It is also possible to filter incidents by route or location to accurately represent a particular part of the UK rail network.

Our analysis identified many incidents with a duration of zero minutes, e.g. incidents affecting a train service at a single point in time, such as the Technical Fleet incidents shown in Figure 17. We separated these from the other incidents that remain in place over a period of time, affecting multiple train services, and created two incident frequencies for the model to use; frequency of incidents with zero minute duration, and frequency of incidents with non-zero duration. This allowed us to create a reasonable distribution of the non-zero delay durations (e.g. Figure 18).

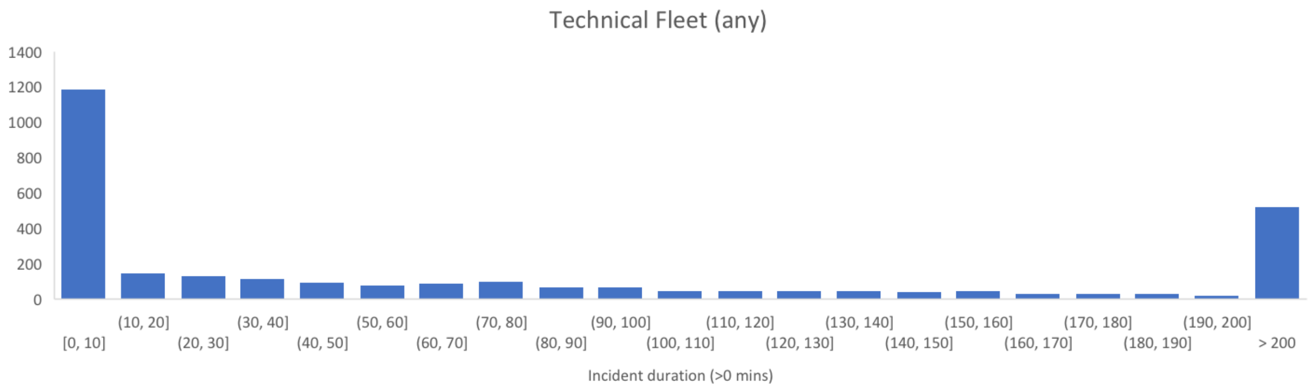


Figure 17: Technical fleet incident duration

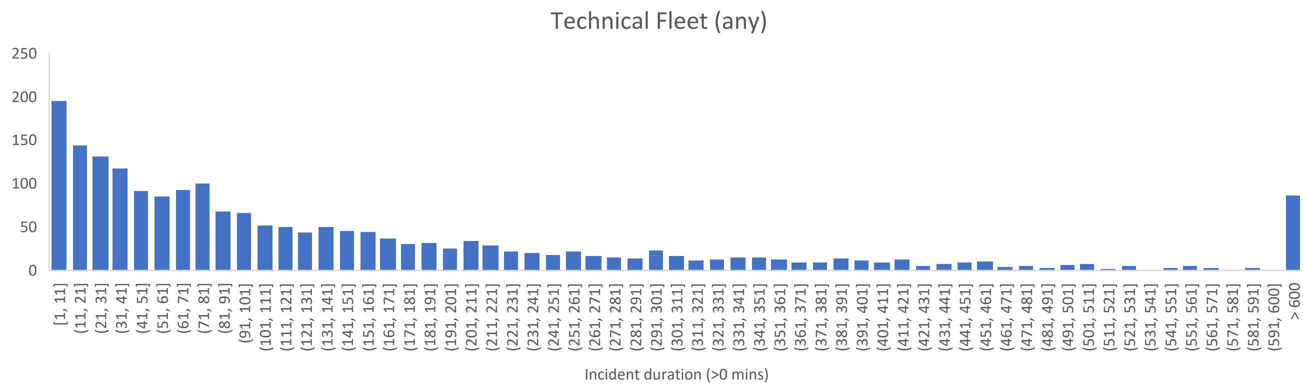


Figure 18: Technical fleet incident duration (>0 mins)

Characterising subthreshold delays

The model also needs to apply delays to train services of under three minutes, and these subthreshold delays are not yet attributed to a particular incident type by railway stakeholders. GWR provided train movement data from their Aegis system for trains travelling between London Paddington and Bristol (other train movement systems or data providing GPS based train timing information could be used). We were able to remove delays over three minutes from this data, compare train service timings at TIPLOC locations with timetabled arrival and departure times, and determine the characteristics of subthreshold delays.

Subthreshold delays were characterised as follows:

- **Incident frequency**: every train service is affected by a subthreshold delay between locations, so the frequency is 1.
- **Incident duration**: this was not relevant to subthreshold delays as the incidents are continuous.
- **Delay duration**: a normal distribution was derived from the Aegis data to describe delay per km, with a mode of -0.037 and a standard deviation of 0.062 (Figure 19). The model multiplied the sampled delay per km by the distance between locations to produce a subthreshold delay value.

When the model calculates subthreshold delay it applies a maximum of three minutes, any delay over three minutes is represented in the attributed delay (see above). Many subthreshold delays are negative, allowing train services to catch up with the timetable, so the model also applies a minimum value, to ensure that a train service could never have a large enough negative delay so that it was traversing a section of track in less than the time it would take at line speed (line speed and length of each section of track is included in the input data).

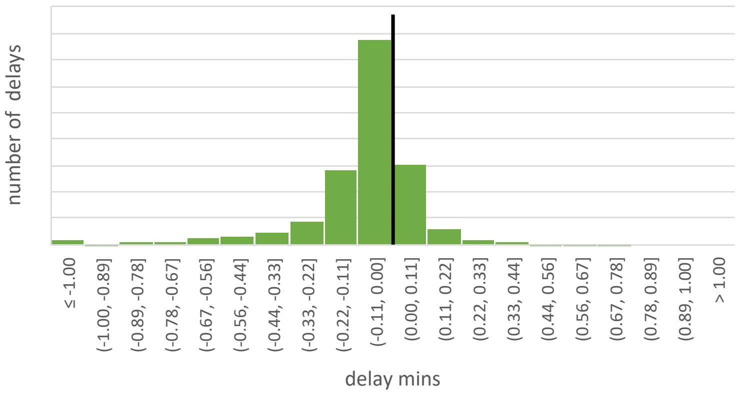


Figure 19: Subthreshold delay distribution

It would be beneficial to use different sources of train movement data when the model is applied to other parts of the UK rail network. This would tailor the model to fit the particular subthreshold delay characteristics of another part of the network.

Characterising off-network delays

Every train service joining the modelled network from another route was allocated a random duration of 'off-network' delay (which could be zero). The locations where trains could join our demonstration route are shown in Figure 20 as black dots. We proposed that the delay duration would have a random distribution, but could also be affected by the type of train service and the distance it had travelled from its journey origin.

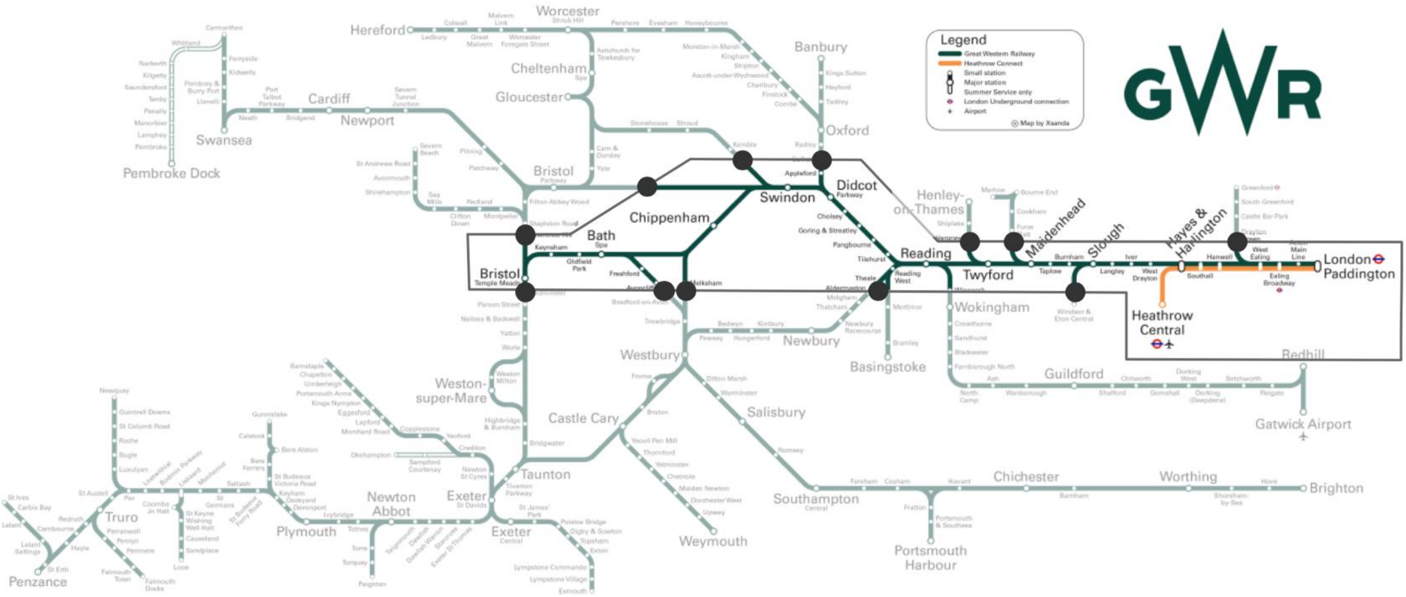


Figure 20: Extent of demonstration modelled route (Paddington to Bristol)

Off-network delays were characterised as follows:

- **Incident frequency**: every train service joining the modelled network is potentially affected by an off-network delay, so the frequency is 1.
- **Incident duration**: this was not relevant to off-network delays as the incidents are instantaneous.
- **Delay duration**: a four stage triangular distribution was derived from the Aegis data to describe delay.

We did not have access to other operators train timing at the points where they entered our demonstration network, but GWR provided us with Aegis train movement data for their services. From this we calculated the delays to GWR train services as a function of distance from their origin as a [proxy for other operators train services. Unfortunately this data only contained main line services, so we were not able to adjust off-network delay by type of train service to include commuter services, freight, etc. Also, the journey from London to Bristol did not represent the considerable distances that some train services travel before joining the modelled network.

Our analysis suggested that we could use a triangular delay distribution with a minimum of -6.61, a mode of 2.53, p98 of 16.38, and a maximum of 24.91 minutes, together with a multiplication factor of 0.024 per km from journey origin. Our initial trials with this characteristic created very significant delays for Cross Country train services with distances of 500km from their origin locations, so we chose not to use the distance factor for our demonstration modelling.

We recommend off-network modelling is improved by:

- obtaining more data
- including more routes in the area modelled in any analysis to reduce the number of off-network train services included in the modelling.

Appendix 4 – Model calibration process

The agent-based model needs to be calibrated to provide confidence that it represents rail services. There are two aspects that need calibration:

1. To check that the model is properly moving trains through a modelled network.
2. Test that the input data is correctly describing the railway system; the locations, platforms, lines, and junctions that make up the rail network, the timetable of train services to be modelled, and the incidents and delays that are to be applied during modelling.

These two aspects are difficult to separate, as they depend on each other, but it is possible to independently test that the incident characteristics are correct before they are used in modelling.

Testing incident input data

An input data test feature was created as part of the model, to generate incidents according to the input data characteristics of any incident data file. The model creates an output file, containing incidents and delays, generated using the characteristics defined by the input data file. These incidents would be applied by the model to train services during a simulation. The results file can be analysed to test that the incidents and delays are represented as intended.

Testing the model and remaining input data

The remainder of the calibration process simultaneously tests the agent-based modelling functions and the remainder of the input data. Early versions of the agent-based model used a simplified timetable of four trains to test that the basic modelling concepts were practical. The following calibration process was created to both test the accuracy of the model, and the input data being used.

Step 1: Calibrate with timetable running (no incidents)

This process tests that the model will run the timetable, with no incidents or delays being introduced, with the expectation that most trains would run to time and there would be little reactionary delay. This tests that the timetable extracted from the CIF input data matches the network to be modelled represented locations and lines in the input data.

For this test multiple simulation runs are required, at least 50, over the full 24 time period in the timetable, in 'timetable only' mode (no primary delays). Trains that might be delayed are permitted to catch up to their scheduled stop times, using randomly generated negative subthreshold delays (a normal distribution of catch up minutes), so multiple runs are required. Any trains with long delays are identified and the causes of the delays explored.

The project steering group advised that there are likely to be a small amount of reactionary delay due to imperfections that may be in the timetable. We were also expecting some delay due to approximations in the modelling process.

These tests resulted in several changes to the numbers of platforms at locations, the numbers of lines between locations and the number of signalling berths allocated to lines. There were also some trains in the timetable that appeared to be duplicates, which were removed. These alterations were discussed with GWR before the input data was updated.

After these data input improvements the model runs still result in a small amount of reactionary delay. In timetable mode the model shows that 4% of trains arrive at stopping locations late, shown in Figure 21. Late = a train arrives later than 1 minute after it's timetabled arrival time.

There are lots of small reactionary delays from timetable only runs, impacting on low level punctuality, resulting in 4% of trains being late, almost half of these are within 1-3 mins. These may be conflicts between timetabled trains. It might be possible to reduce delays by removing conflicts from the timetable.

Step 2: Calibrate with subthreshold delays only

This process tests that the model will recreate low levels of delay (non attributed). It runs a normal timetable, and only applies subthreshold delays to trains.

For this test multiple simulation runs are required, at least 50, over the full 24 time period in the timetable. Trains can be delayed by randomly generated subthreshold delays (up to a maximum of three minutes), so multiple runs are required. Trains with long delays are identified and the causes of the delays explored. Train punctuality can be compared with national statistics for trains arriving up to three minutes late (national performance statistics for period 10 2018/19 show that 65% of trains were early or on time, and 35% were early). And, train punctuality can be compared with a train operator's performance statistics.

Figure 21 shows the results of these steps, an additional 0.7% of trains arrive late after introducing subthreshold delays.

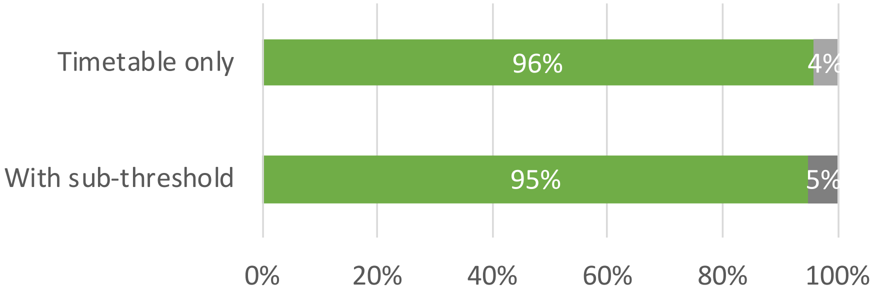


Figure 21: Model calibration, train punctuality (at all stations)

Figure 22 displays the same simulation results, showing that many of the additional late trains are arriving between 1 and 3 minutes late. And, that there are additional trains arriving between 5 and 10 minutes late.

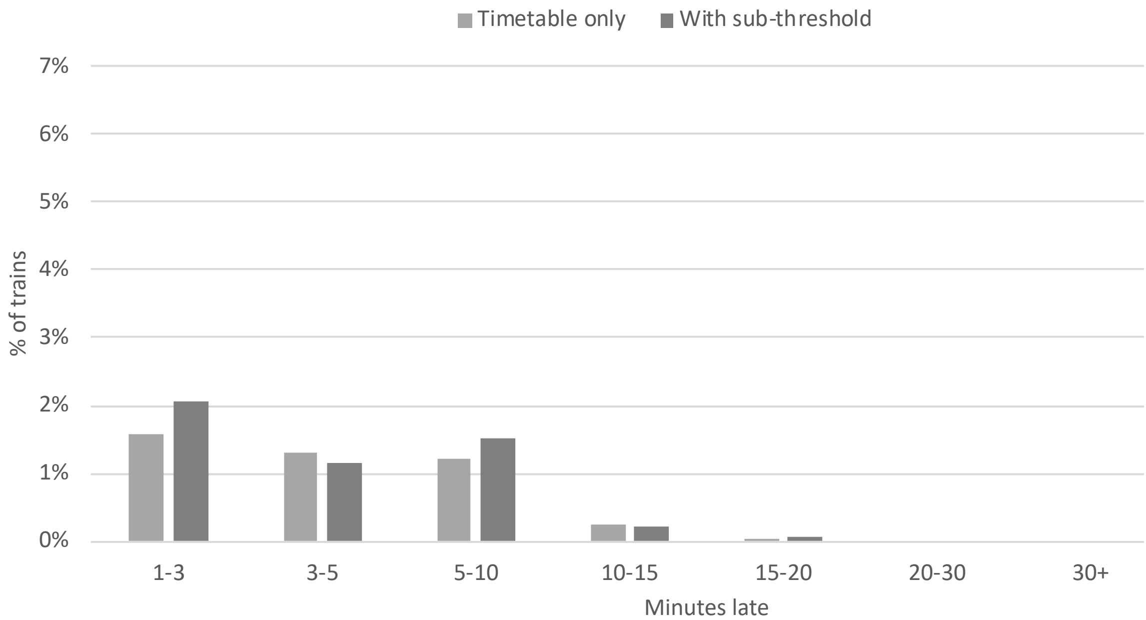


Figure 22: Model calibration, trains arriving late (all stations)

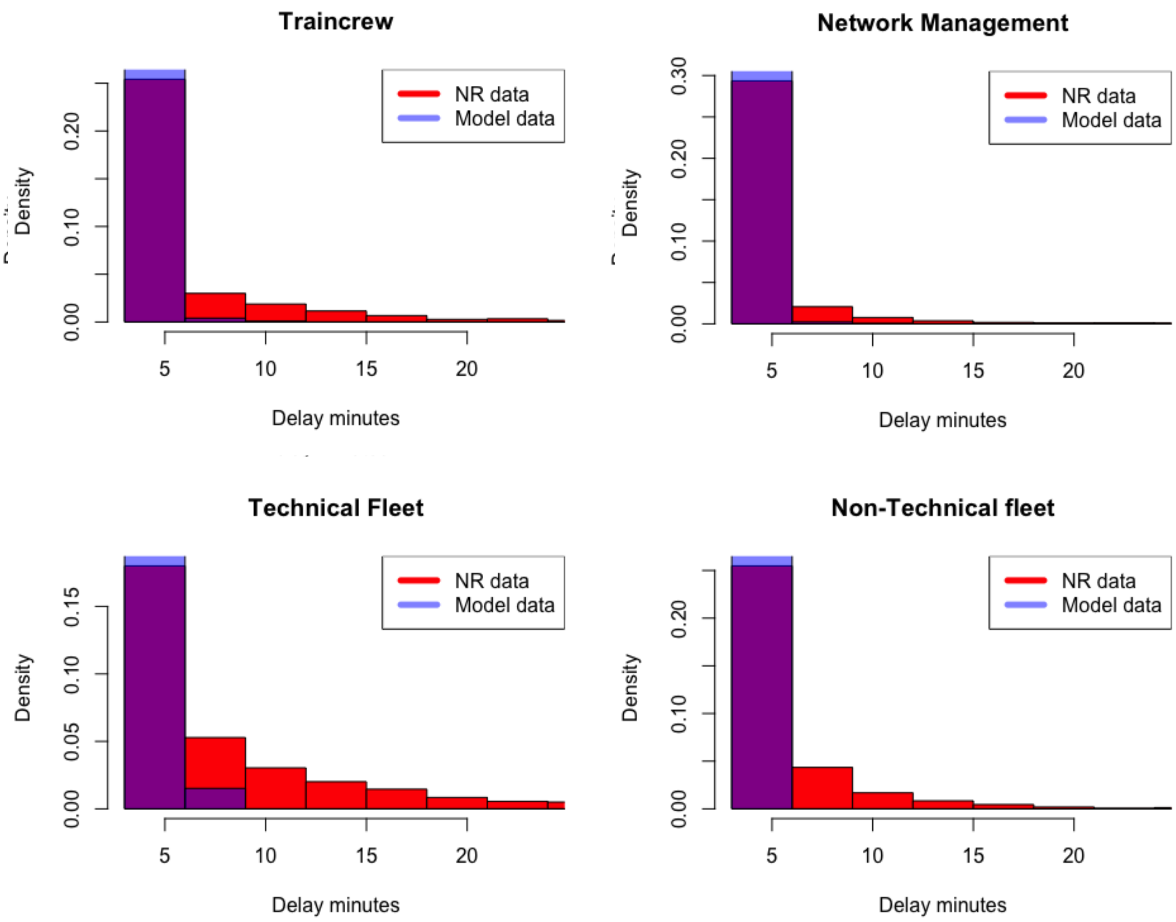
The timetable and subthreshold delays cause considerable punctuality issues. It is not possible to identify causes of subthreshold delays, a proportion of which will become threshold delays. Modelling may help identify these.

Step 3: Testing attributed delay input data

This stage of calibration involves running the model with attributed incidents. This process tests that the model will recreate all types of primary delay at appropriate levels. It runs a normal timetable, and applies both subthreshold delays and attributed delays to trains.

For this test multiple simulation runs are required, at least 150, over 24 hours. The train delay results files for each simulation are analysed to test that the appropriate number of delays of each incident type have been applied by the model to train services. The incident input data file can be adjusted as required, to modify the frequency and duration of incidents, and modify the duration of delays.

Figure 23 shows a selection of delay distributions produced by the model (red), comparing then with Network Rail attributed data delays (blue), showing where these match (purple). The delay data sets are very closely matched, but the Network rail delay data has longer tails in their distributions due to a small number of unusually long delays. These could be recreated in the model if required by extending the p98 point in the four stage triangular distributions used in the input data to characterise the delay duration distribution. We also tried a few alternative delay distributions shapes but these did not introduce a significant change in the punctuality results, e.g. Johnson SU



and Normal distributions.

Figure 23: Comparing modelled delay with attributed delay

Step 4: Calibrate with attributed delay input data (creating a baseline)

The final stage of calibration involves reviewing all model results from step 3, and comparing them with real world train service performance statistics. This process tests that the model will recreate all types of primary delay at appropriate levels. Once calibrated, these modelling results provide a baseline for modelled service performance, against which any interventions designed to improve performance can be measured.

For this test multiple simulation runs are required, at least 150, over 24 hours. Trains with long delays are identified and the causes of the delays explored. Train punctuality can be compared with national statistics. And, train punctuality can be compared with a train operator's performance statistics.

Figure 24 and Figure 25 show the results of these steps for our demonstration model, with 18% of trains arriving late after adding attributed delays, an additional 13.4% of trains arrive late compared with the timetable only simulation (without any primary delays).

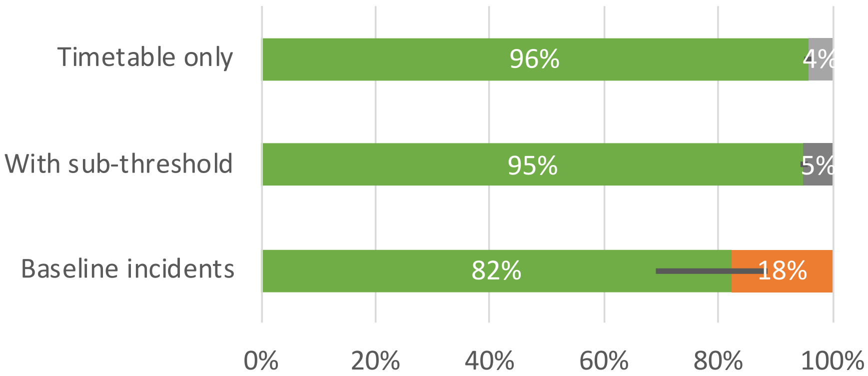


Figure 24: Model calibration, train punctuality (all stations)

Figure 25 displays the same simulation results, showing that a significant number of trains are arriving up to 15 minutes late.

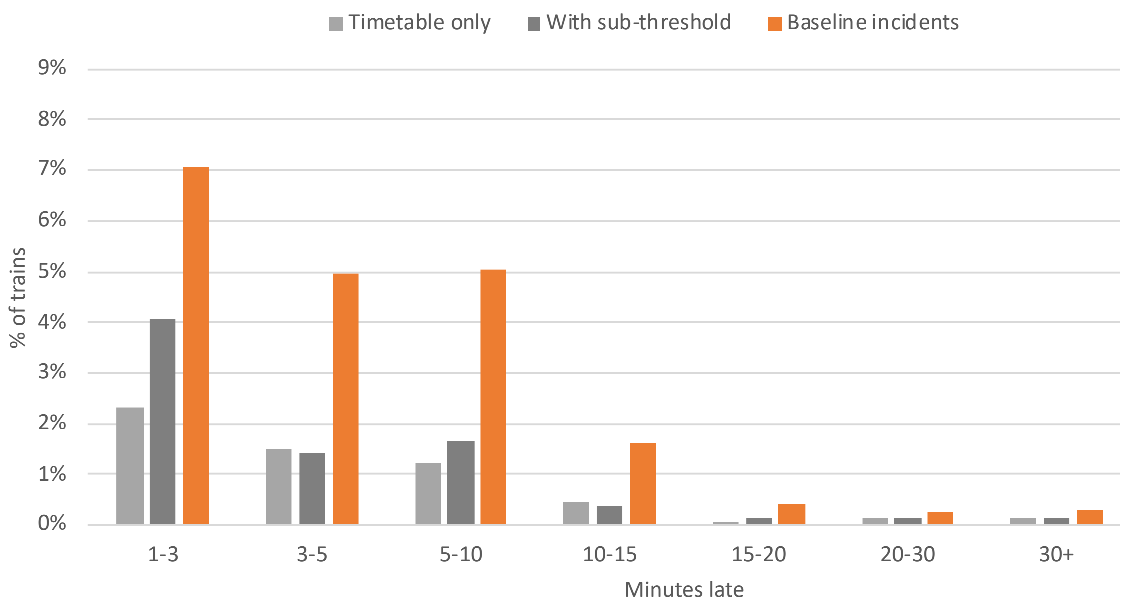


Figure 25: Model calibration, trains arriving late (all stations)

The proportion of delay minutes to trains due to the different types of incidents is shown in Figure 6. 16% of delay minutes are due to reactionary delay, with an equivalent proportion due to subthreshold delays. 24% of delay minutes are due to attributed delays, and 44% due to off-network delays.

The attributed delay minutes are split between the types of primary incidents, with the highest proportions of delay minutes due to non-track assets (6.1%), Network Management (5.2%), and Operations (5.8%). These might be candidates for interventions, but it would be premature to assume this before fully understanding the causes of the reactionary delay by using the interactive visualisations to explore the modelling results in detail.

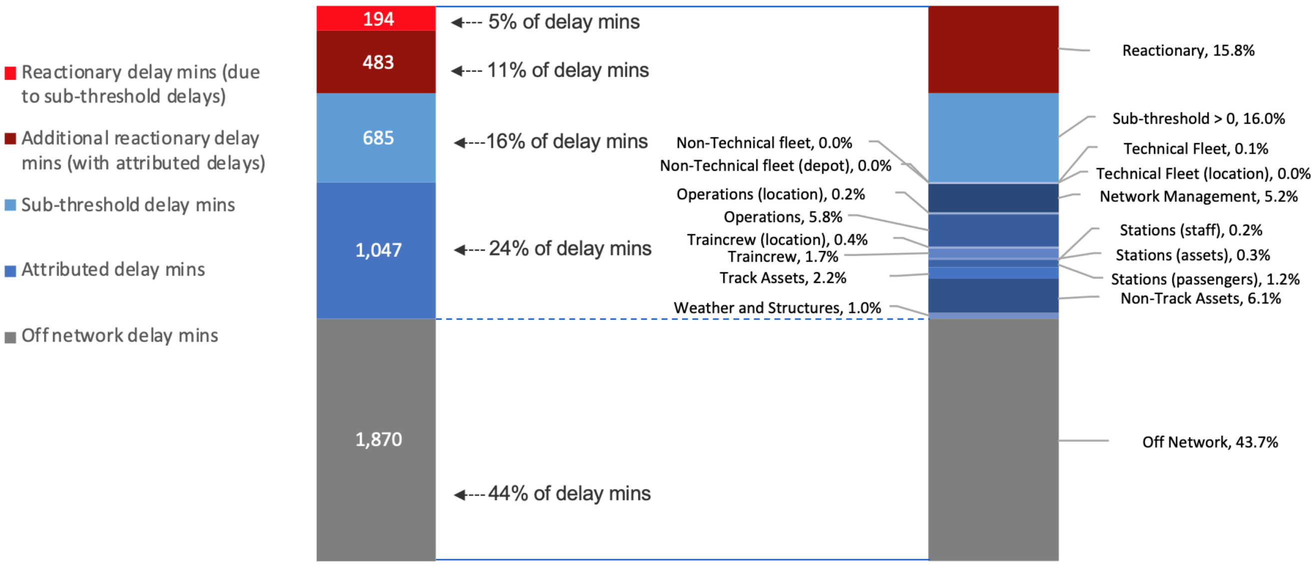


Figure 26: Baseline delay minutes

An attempt was made to compare these modelled results with national performance statistics, and with GWR performance statistics, but the results were not conclusive. Our demonstration used incident characteristics derived from the national attributed data from March 2018 (to ensure enough data for our analysis). These are unlikely to reflect the specific range of incidents that are present on the Paddington to Bristol route.

In summary, attributed delays (over 3 mins) reduce the number of trains arriving on time by 12.7%, both directly through the trains being delayed by incidents, but also as a result of the additional reactionary delay minutes caused. They represent 24% of arrival delay mins.

Subthreshold delays (under 3 mins) also contribute to delayed arrival times and are the cause of 16% of arrival delay mins, and cause an additional 5% of arrival delay mins through resulting reactionary delay. It would be worth understanding the causes and consequences of subthreshold delay.

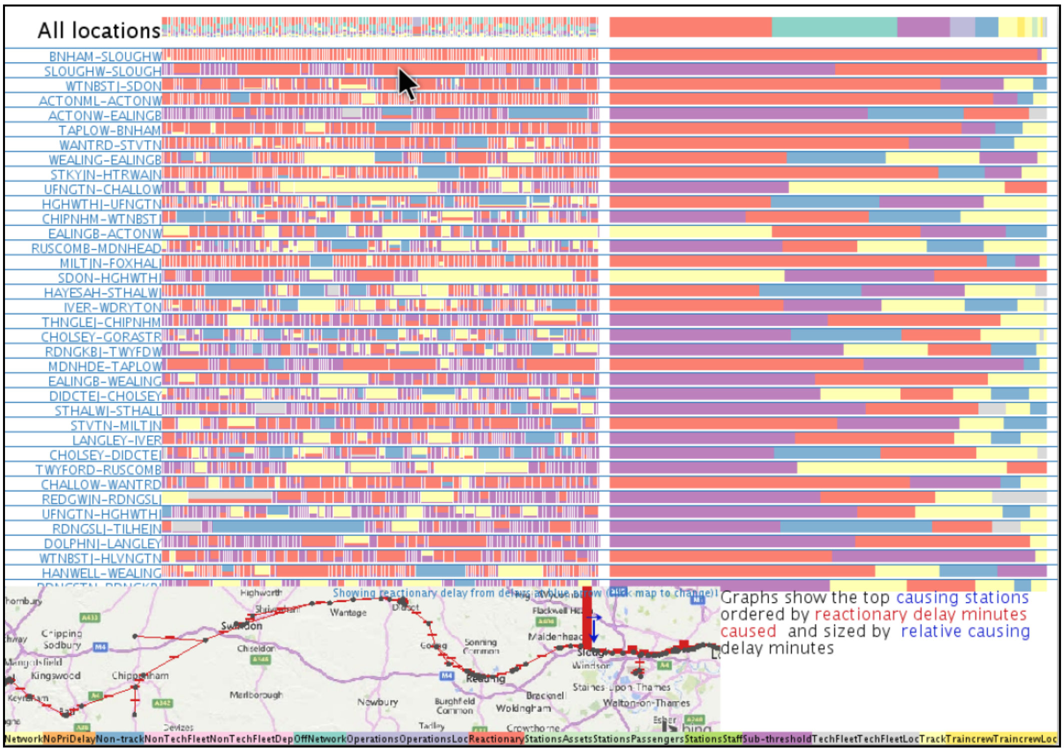
Off-Network delays are significant (the cause of 44% of arrival delay mins) these are made up of reactionary and attributed delays from outside the modelled network. 59% of trains in the model originate off the modelled network, a 1/3rd of them are delayed. This type of delay has been crudely modelled, there is a need to improve accuracy of modelling off-network delay. And, extend modelled area to encompass more routes/services to reduce off-network interactions.

Appendix 5 – Using the interactive visualisations

This section serves to illustrate how the visualisation presents the model results in a way that informs the interventions.

WT1: Identifying locations that cause most reactionary delay and locations where this reactionary delay occurs, using the summary view

The steps below demonstrate how the ***summary view*** *of the interactive visualisation tool helps* ***identify locations that cause reactionary delay*** *and helps* ***design interventions****.* ***What this enables:*** *This enables you to find the locations* ***that cause most delay****, find which are the* ***affected locations****, find which are the most* ***likely*** *(by looking across model runs), and helps* ***design interventions*** *by indicating the delay types that cause this (with locations).*



Whole screenshot

WT1a: Overall types of primary delay

Excerpt from above screenshot

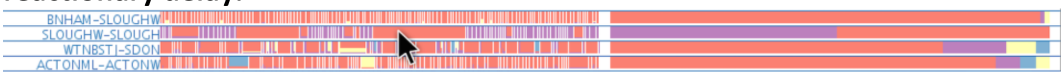
This stacked barchart (top right) shows the **relative amount of types of causing delay minutes** across **all locations**, with the top three delay types being reactionary delay (red), off-network delay (turquoise) and sub-threshold delay (purple). ***Headline****: It is likely that reactionary, off-network and sub-threshold delay are the most common reasons for reactionary delay.*

WT1b: Variation of overall delay types between model runs

Excerpt from above screenshot

This is a stacked barchart (top left) **per model run**. Although there are some differences between model runs, the **proportions are quite similar**. The **width** of the model run indicates the relative amount of causing delay minutes. This is **quite consistent across model** runs, but a few of the 150 model runs have longer causing delays and a few have shorter ones. ***Headline:*** *The causing delay types are consistent across model runs.*

WT1c: Identifying the nature of causing delays at the locations that cause most reactionary delay.



Excerpt from above screenshot

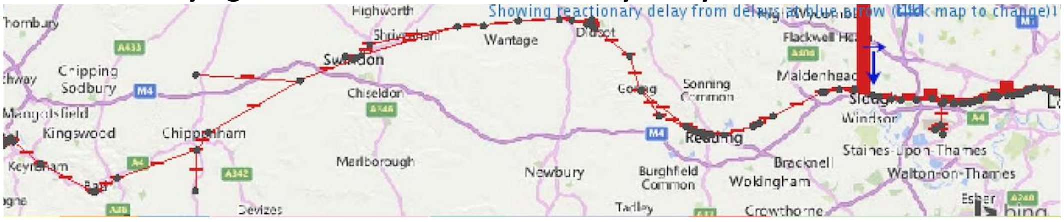
The list of locations or sections of track (blue names in left hand column) is ordered by the **reactionary delay minutes caused**, where the **top three locations that cause delay** are "Burnham – Slough West", "Slough West – Slough" and "Wootton Bassett – Swindon". Each location has two bar charts, the right hand shows the proportion of delays types causing reactionary delay at this location, and the left hand one shows how this varies across all model runs. In terms of **causing delay type**:

- Causing delays at "Burnham – Slough West" are **almost entirely** (other) reactionary delays (red) and this is consistent between model runs
- Overall, causing delays at "Slough West – Slough" are about **half-and–half** subthreshold delay (purple) and reactionary delay (red). However, the reactionary delay **only happens in some model runs**, and when it does happen, it is a longer delay.
- "Wootton Bassett – Swindon". Again, this is **mostly** reactionary delay (red), and it **occurs in all model runs**, but to different degrees.

Excerpt from above screenshot

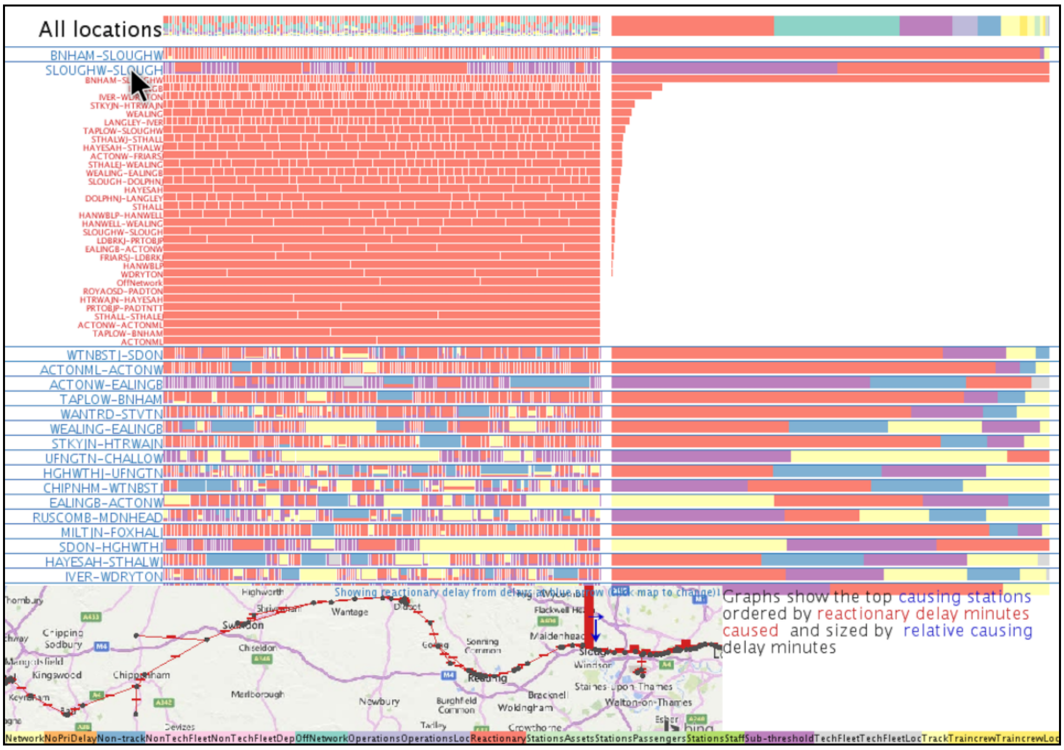
The **tenth top location/line** that causes reactionary delay is "Uffington – Challow". Overall, it has about equal amounts of sub-threshold (purple) and track (light yellow) delays types, but the track delay type only occurs in **a few model runs**, but when it occurs, it is lengthy.
Headline: *Reactionary and subthreshold track dominate the delay types that cause reactionary delay, but some other delay types are less consistent between model runs and relatively lengthy, indicating they are unlikely, but if they happen, they may have a large impact*.

WT1d: Identifying the nature of the reactionary delay



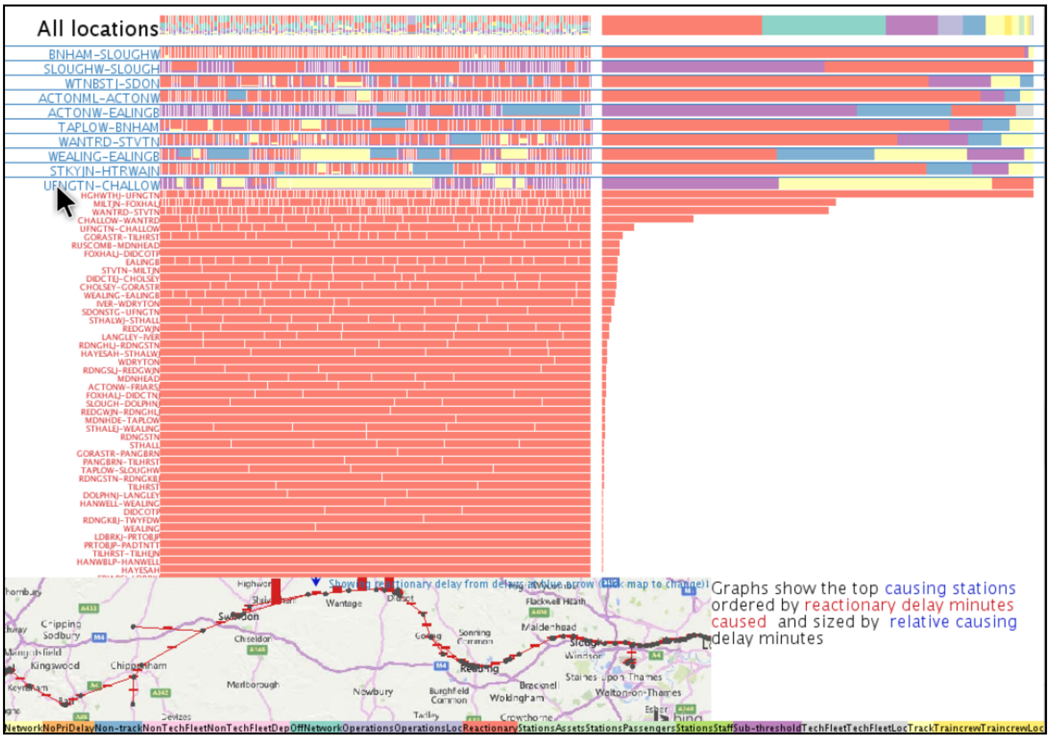
Excerpt from above screenshot

The map shows the **reactionary delay minutes** (red bar heights) for **the selected causing location** (see mouse pointer in main screenshot) indicated with a blue downward-facing arrow with the right-facing arrow indicates direction. In this case, the **overwhelming majority of reactionary delay occurs just before the causing delay** (the tallest red bar is to the left of the blue arrow for eastward-bound services), but it also causes lesser delays in **other parts of the network**, particularly for train ahead (eastwards, towards London).



Whole screenshot

By moving the mouse over the name of the causing location, we get a list of locations **at which the reactionary delay occurred** for the causing location. This list is ordered by the **amount of reactionary delay minutes**. Consistent with the map, the red bars on the right show that the top reactionary location ("Burnham – Slough West") has much more lengthy reactionary delay than any of the other locations.



Whole screenshot

Placing the mouse on the location with the tenth most reactionary delay ("Uffington – Challow" as mentioned above), indicates that the **top reactionary locations are further afield** from the causing location (see the proximity of the red bars to the blue arrow in the map). Also, note that the **top reactionary locations are represented in many model runs** (the white vertical lines separate the model runs), but the bottom reactionary delays only occurs in one or two model runs.

Headline: *Reactionary delay is often close to the causing delay, but also occurs in other part of the network in both directions. Where long delays only occur in few model runs, they are unlikely but may cause problems in the rare cases where they occur.*

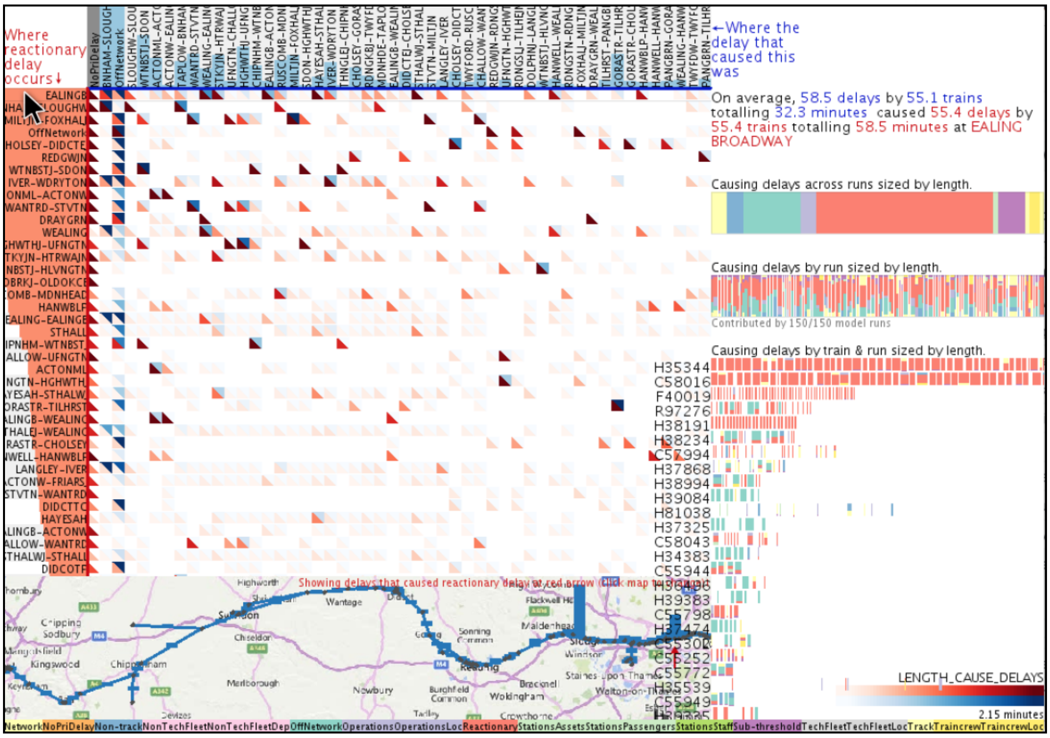
WT2: Identifying locations that are subject to most reactionary delay and locations where this reactionary delay occurs, using the detailed view

The summary view (previous walkthrough) is configured to identify locations that cause reactionary delay, and where the reactionary delay for each of these locations is. The ***detailed view*** *allows you to look at* ***both****, and we will use it to identify locations that are* ***subject to reactionary delay****.*

The steps below demonstrate how the ***detailed view*** *of the interactive visualisation tool helps* ***identify locations that are subject to reactionary delay*** *and helps* ***design interventions****.*

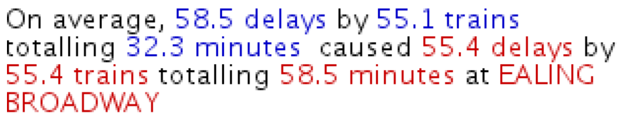
What this enables: *This enables you to find the locations that are* ***subject to*** *most delay, find which* ***locations ultimately cause this****, and help* ***design interventions*** *by indicating the delay types that cause this (with locations).*

WT2a: Identify locations that are subject reactionary delay



Whole screenshot

This is the detailed view, where locations that **cause** reactionary delay are **columns** (blue) and locations that are **subject to** reactionary delay are **rows** (red). (Note that this is just the top left of the matrix – other locations are currently off-screen.) Both are ordered by reactionary delay minutes caused (columns) or subjected to (rows). The length of the blue bars in the columns is the causing delay minutes. The length of the red bars in the rows is the reactionary delay minutes. The mouse pointer is on the location (row) **subject to** most reactionary delay ("Ealing Broadway") across all causing locations. A **text description** of what this rows presents is written at the top right. This is a **high level summary average for the reactionary delay that happens at this location**.



Excerpt from above screenshot

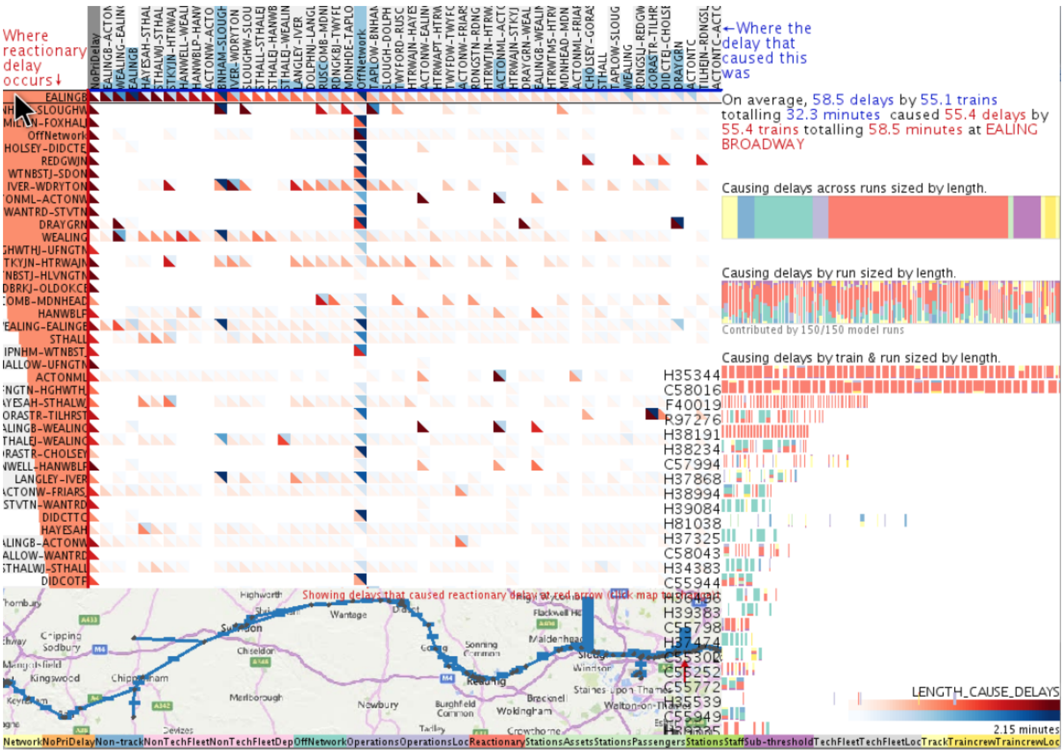
The top **three locations subject to reactionary delay** (across all causing locations) are "Ealing Broadway", "Burnham – Slough West" and "Milton Junction and Foxhall Junction".

WT2b: Identify causing the delay types

Excerpt from above screenshot

With "Ealing Broadway" selected as the location subject to most reactionary delay, the graph on the right shows the relative length of causing delay minutes just **for the reactionary delays at location** ("Ealing Broadway"). Off-network (green) is about a fifth of this (trains that are delayed before they join the modelled network) and reactionary delay is about half of this. These vary in (causing) length between model runs, but not significantly. The **number of model runs** that produce a delay here is written below (all model runs, in this case).

WT2c: Identify causing the top causing locations



Whole screenshot

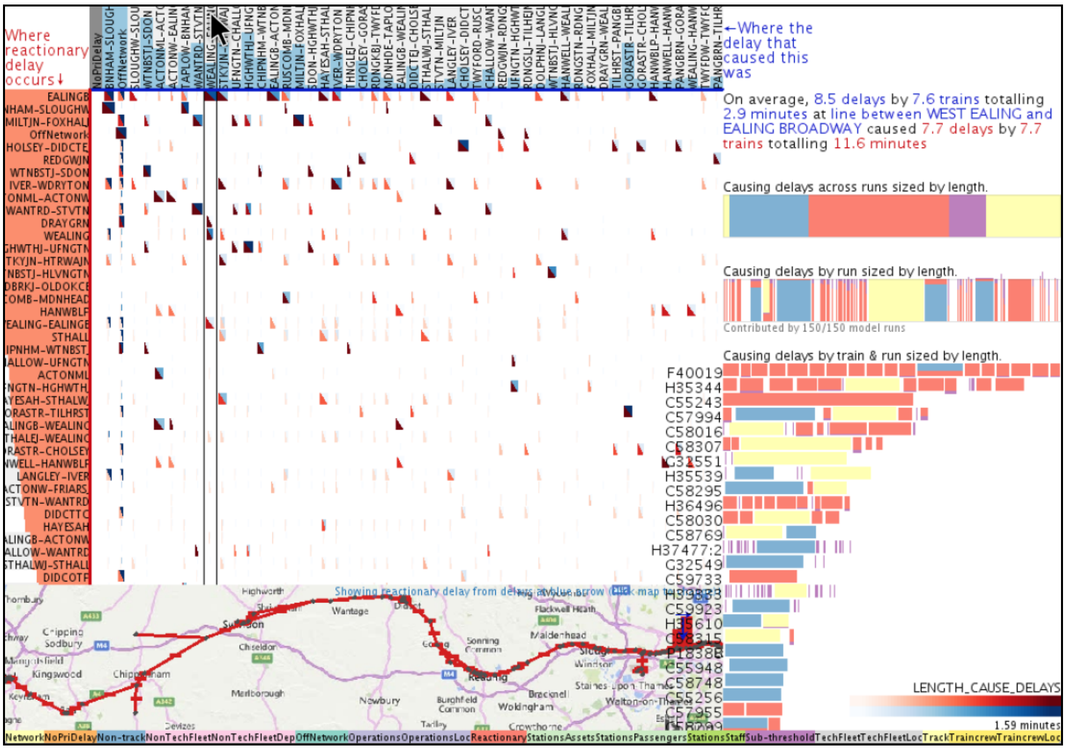
The column (causing locations; blue) have been sorted by the amount of reactionary detail caused to the selected (top) row (reactionary locations; red; "Ealing Broadway"). The cells are diagonally split into triangles, where the darkness of the **blue** corresponds to the delay minutes of **causing delay** and the darkness of **red** corresponds to the delay

minutes of **reactionary delay** (you'll see that the red shading for the columns in the top row gradually lightens towards the right; see legend bottom right). Thus, the top locations that cause delay to Ealing Broadway are "Ealing Broadway – Action", "West Ealing – Ealing" and "Ealing Broadway" (itself). This is reflected in the map.
A reason why "Ealing Broadway" is subject to more delay is that it is **affected by more locations**. The matrix in the screenshot shows that many of the top causing locations do not cause delays at other stations.
Headline: Ealing Broadway is the location at which there is most reactionary delay, but this was not picked up earlier. This is because many locations contribute to delays there.

WT3: Identify which types of causing delay are commonly associated with reactionary delay, where, and how likely these are to occur

What this enables: *This enables you to find the locations where certain delay types are more likely, and thus could be targeted by interventions.*

WT3a: Identify delay types associated with reactionary delay



Whole screenshot

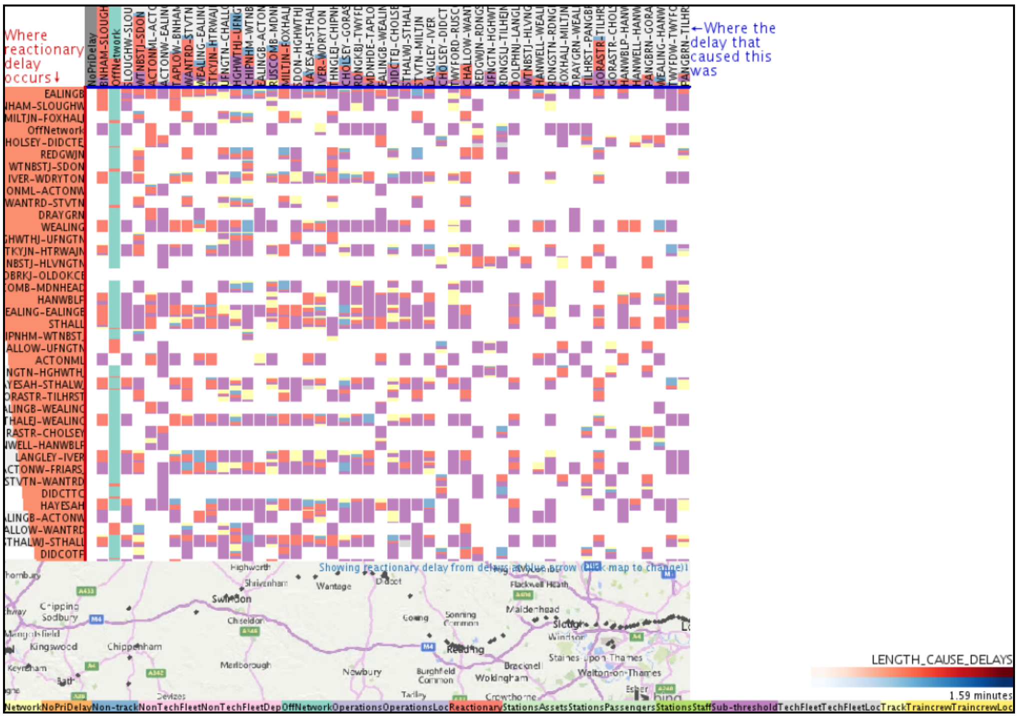
In the matrix, columns (causing locations) are **sorted by the amount of reactionary delay minutes they cause**, and the blue bars indicate the amount of causing delay minutes.
Moving the mouse over from left to right indicates the delay types. This one ("West Ealing – Ealing Broadway") is the first that has significant amounts of non-reactionary

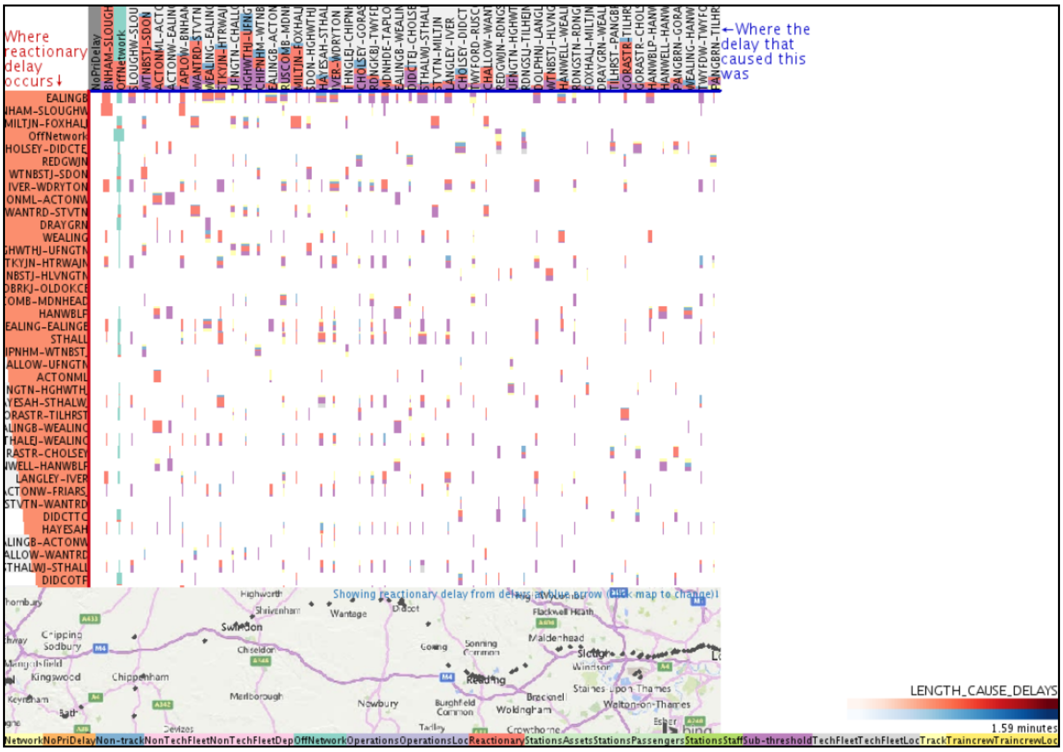
delay (the ones to the left are dominated by reactionary delay which it is hard to intervene to directly reduce).

Excerpt from above screenshot

However, from the model runs graph, The non-track (blue) is only present in 4 model runs and the weather (pale yellow) is only present in one model run. This is because these primary incidents are not so common.

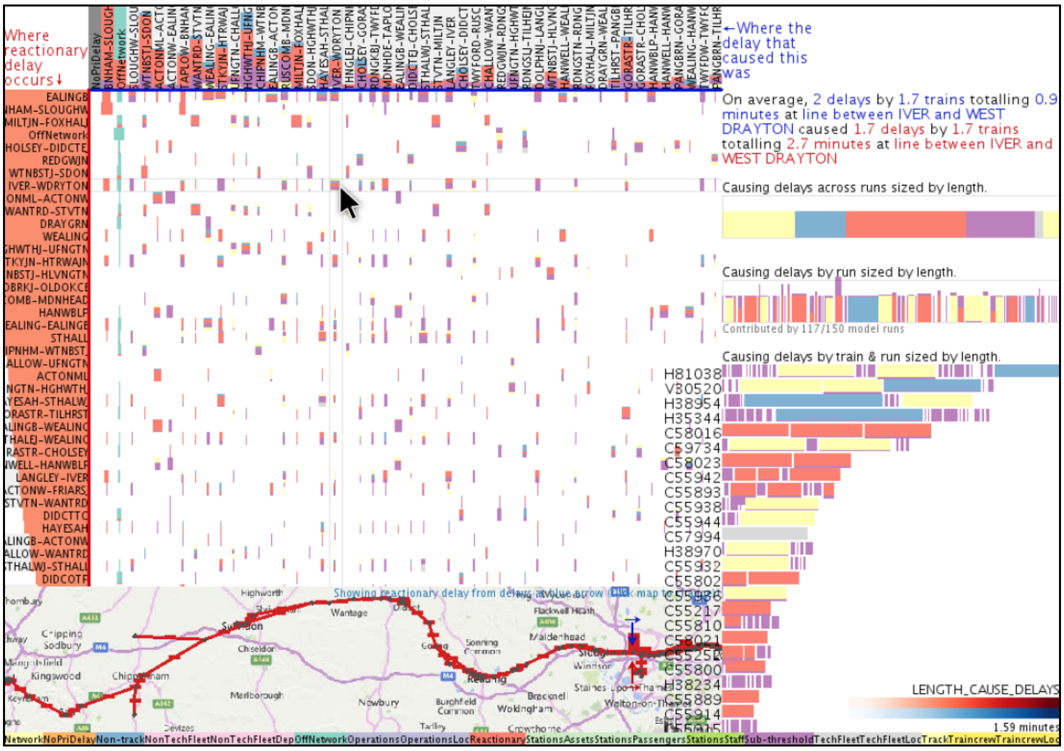
WT3a: Identify location pairs where delays are more likely





Whole screenshots

Another way in is to look at the **pairs** of locations (the one that **causes** the reactionary delay and the one that **is subject to** the reactionary delay). As before, each cell is a pair of locations. But instead of drawing triangles inside, the proportion of causing delays are shown. **In the bottom example**, the width of these represents the number of model runs that produce delays here. So the **fattest ones** represent location pairs in which all the model runs produce delays, and the **thinnest ones** are those locations where only one model run produces delays.
Notice that many of them are thin, indicating that many delays are not so likely.



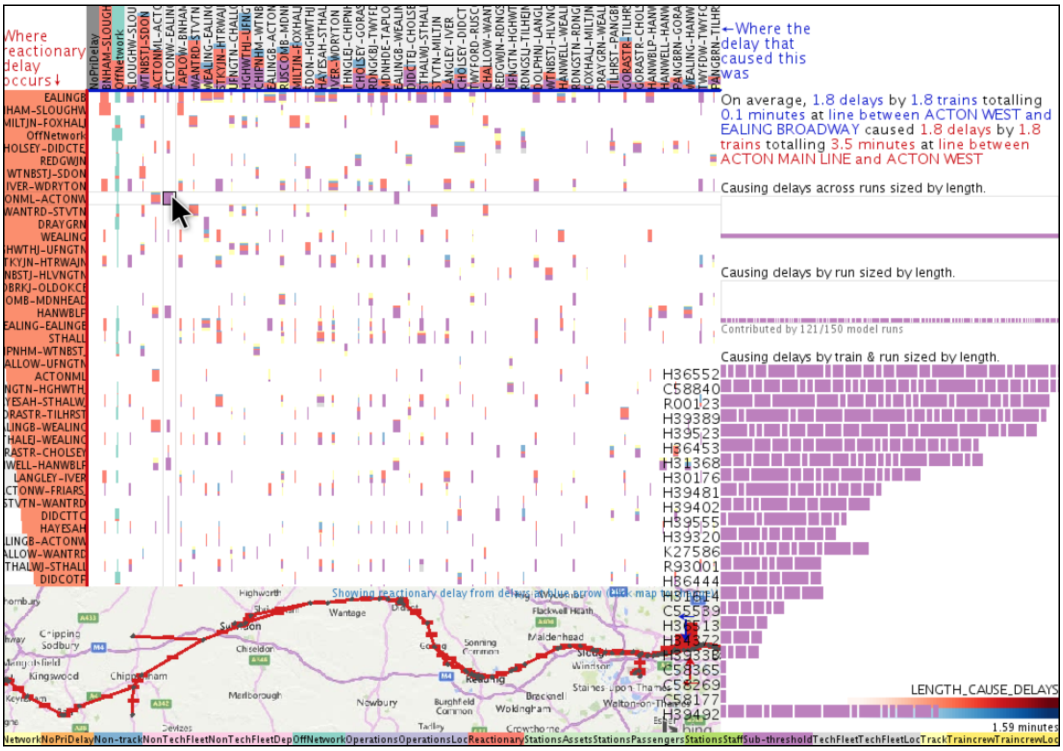
Whole screenshot

Here, we've found a pair of locations, in which "Iver – West Drayton" is the location that both **causes** and **is subject to** reactionary delay. This pair of locations (cell) has been selected. These delay types are represented in many model types, so are more likely than in some other situations.

WT4: Investigate the importance of sub-threshold delays and study the mechanism by which it causes delays

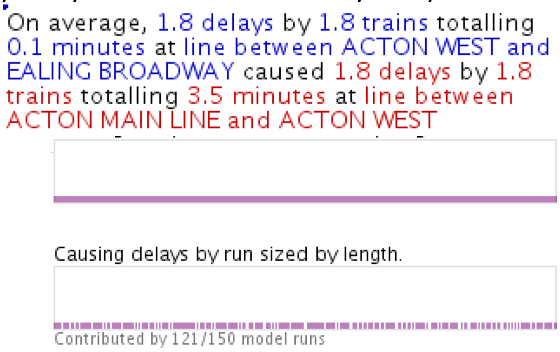
What this enables: *Sub-threshold delay is currently not recorded nor attributed, so is "under the radar". Our interactive visualisations let us investigate these from the modelled results.*

WT4a: Identify location pairs where sub-threshold delay occurs in many model runs



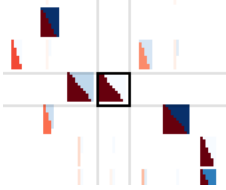
Whole screenshot

Here, we've selected a pair of locations with a "fat" purple square, indicating that many model runs produce sub-threshold primary delays that cause reactionary delay.



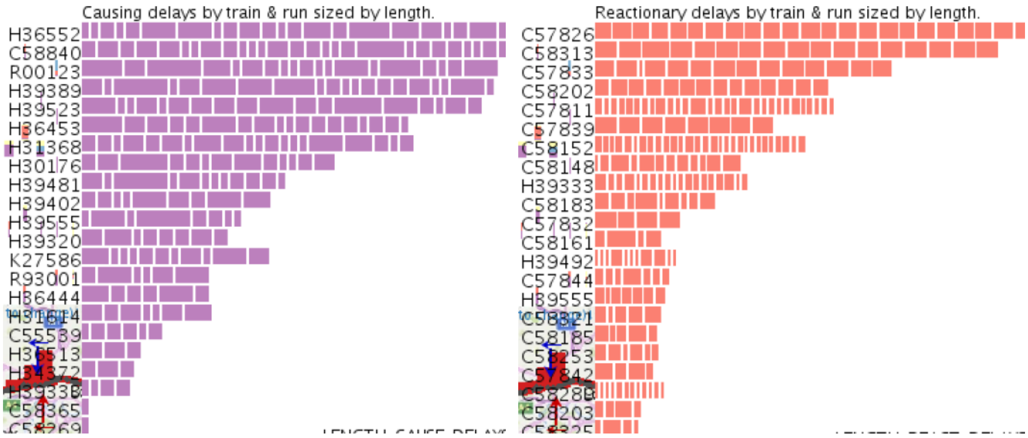
Excerpts from above screenshot

The text shows that on average, 0.1minutes of delay is causing 3.5 minutes of delay. The bar chart shown here are not very tall, because the amount of causing delay is small (by definition; these are sub-threshold delays), but the amount of reactionary delay is significant in comparison.



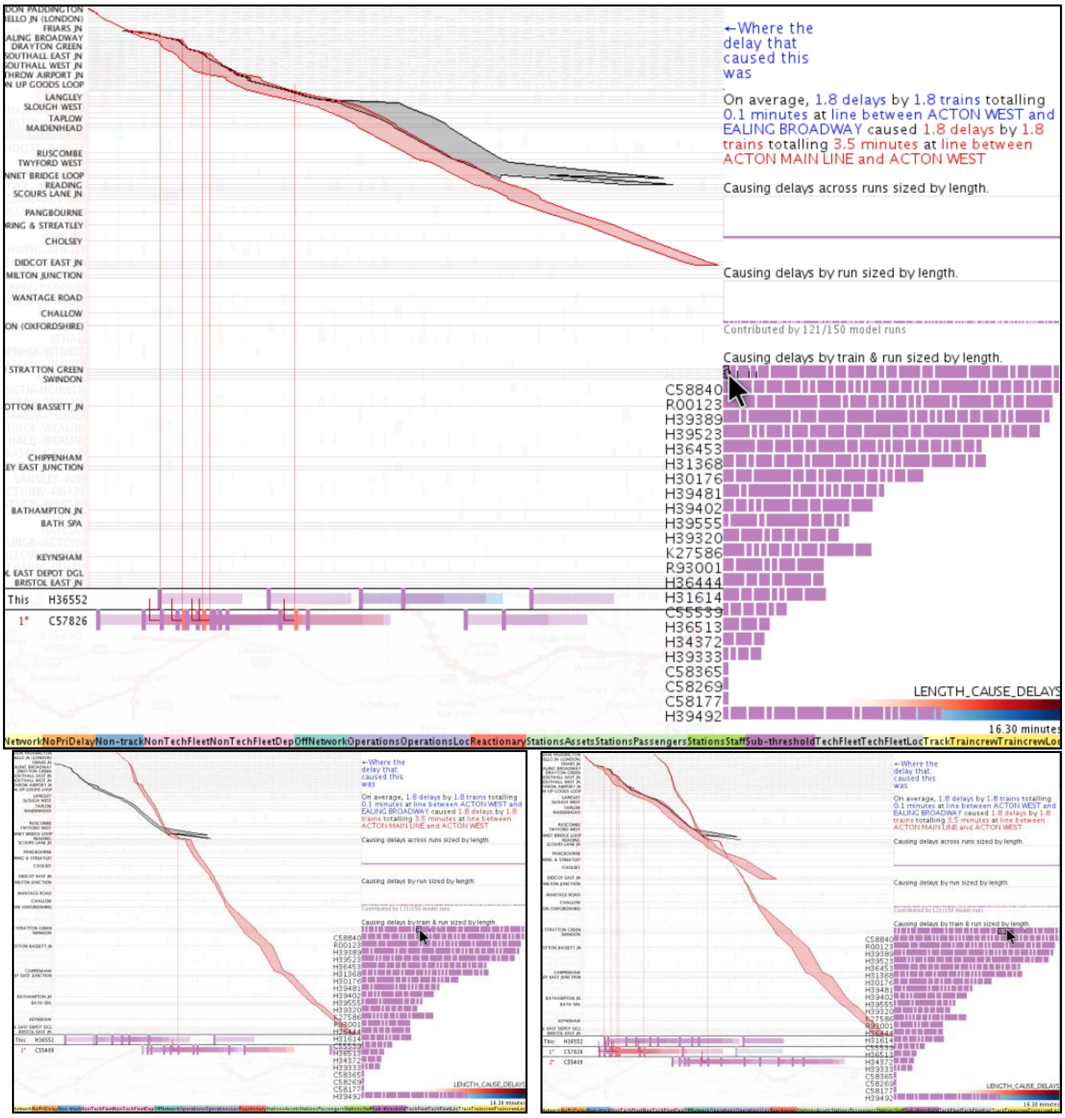
Excerpt from above screenshot, but showing the red/blue triangles instead, where there's an obvious imbalance between causing delay and reactionary delay.

This is interesting, as it may indicate that sub-threshold delays, in some circumstances have more significant effect than one might expect.



***LEFT:** Excerpt from above screenshot, showing **causing trains**. **RIGHT**: As left, but showing trains affected by reactionary delay.*

These parts of the whole screenshot show the **trains that caused the delay on the left** (at the column location), and **the trains that were subject to the delays on the right** (at the row location). In both case, this is produced by a fair few model runs. Now we can look at the mechanism of this.

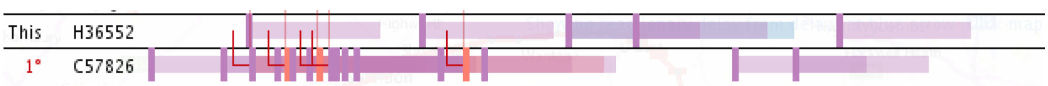


3 whole screenshots (the first is bigger, so you can see the detail)

When we select a specific run of a specific train, we see a graph **showing how the train interacts with other trains**. Above, we have selected three different model runs for the same train. The first of the three screenshots should be big enough to show the detail. The **vertical axis is the location**, with London at the top and Bristol at the bottom and various points along the route marked. The **horizontal axis is time**.

The train that is **selected** is **black/grey** with the leftmost line being the timetabled times of arrival, the right line being the time of actual arrive, and (correspondingly) the width of the grey area between the lines is how late it is.

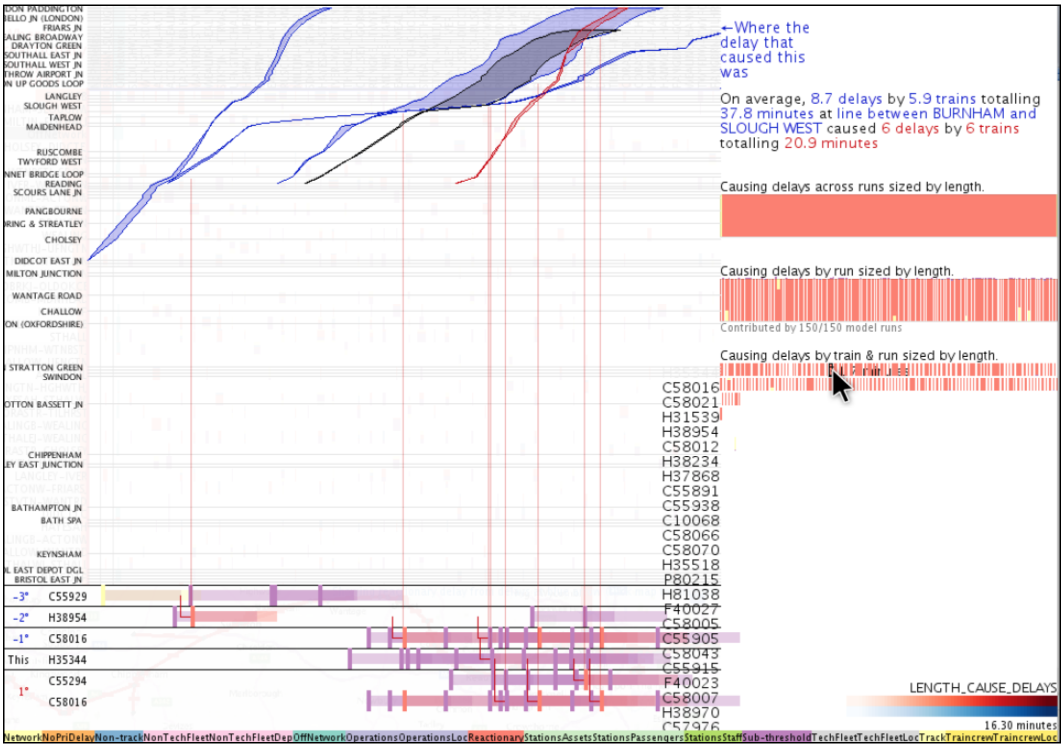
Trains that it delays are in **red**, and the vertical red lines indicate where there was reactionary delay.



Excerpt from the first (of the three) screenshots

At the bottom is a graph showing the selected train (labelled “this”) and the train that is causes a delay to (labelled “1°”). The bars show the delay – mostly purple, as most are reactionary delays.
These graphs show the exact mechanism of how reactionary delay happens and one can quickly browse to compare how this differs by model run.

WT5: Investigate the propagating reactionary delay



Whole screenshot

In this example – as before – the selected train is black/grey. There is a chain of trains that caused this to be late (labelled in blue as “-3°”, “-2°” and “-1°”) and a subsequent train that is delayed by the selected train (labelled in red as “1°”)

Appendix 6 – Technology Readiness Levels

Table 2 describes the Railway Industry Readiness Levels (RIRL) and Technology Readiness Levels (TRL) achieved by this project, and how the recommendations will deliver higher levels.

Table 3: Technology Readiness Levels

RIRL	TRL (MoD)	
Level 3 (project objective): proof of concept	TRL 1: Basic principles observed and reported. TRL 2: Technology concept and/or application formulated. TRL 3: Analytical and experimental critical function and/or characteristic proof-of-concept.	feasibility study
Level 4: industry specification	TRL 4: Technology basic validation in a laboratory environment.	feasibility study
Level 5: prototype, scale up/out issues fixed, improved calibration/accuracy, more complete testing	TRL 5: Technology basic validation in a relevant environment.	next steps
Level 6: adoption of prototype by other party(s)	TRL 6: Technology model or prototype demonstration in a relevant environment.	next steps
Level 7-9: adopted by more rail partners	TRL 7: Technology prototype demonstration in an operational environment. TRL 8: Actual Technology completed and qualified through test and demonstration. TRL 9: Actual Technology qualified through successful mission operations.	longer term development