Understanding the Impact of Driving Styles on Reactionary Subthreshold Delays on a Fixed Block Signalling System

Oliver Bratton [a], Giorgio Medeossi [b],
[a] MTR Corporation Ltd
[b] trenolab

Abstract
Train punctuality in the UK is focussed on measuring the time trains are booked to pass a fixed point and when that event occurs. What is not considered in this measurement of performance is whether the capacity of the system is being optimised. It is posited in this paper that performance needs to consider how closely the delivered train service matches the minimum time signals should be red for that pattern of train services. Any changes to the operation of the system that cause the signals to be red for longer than necessary will decrease system capacity and this will have a detrimental effect on *delay per incident*.

This paper compares on-train data recorders (OTDR) from 2002 and 2018. It shows that average braking rates have declined from 4%g to 3.5%g. This will typically add 4 seconds per stop. Train lengths in the UK have also increased in this time, with a typical train length increase being from 8 to 10 cars. If the slower braking curves and longer trains are combined, and a hypothetical block joint positioned 300m from a stopping point, it can be shown that the signal in rear will take 8 seconds longer to clear on average today than in 2002.

While the impact of these changes on time at destination can be easily demonstrated using distance/time graphs, the effect on the signalling system is more complex. The simulation system *trenissimo* has been used to show that the effect on a system of longer trains and slower braking curves is [x], with the system responding in a non-linear way to very small changes in train operations. It is posited that this is key reason for the increase in delay per incident currently being seen in the UK.

Keywords
Braking, Capacity, Performance

1 Introduction

A railway is a balance of journey time, intensity of service (given infrastructure constraints) and service reliability. While the general principles of how these interplay are axiomatic to operations management, the precise mechanism by which one affects the other is less well understood. In the UK, the primary focus has been 'PPM' - the per cent of trains that arrive at destination within 5 minutes of booked time (10 minutes for long distance operators) having called at all booked stops.

PPM does not measure how effectively a train uses the available capacity of the system. One way to improve the PPM metric is to increase the planning time between stations; this increases the probability that a single train will arrive at the timing point 'on-time', even if it has been subject to a delay en-route. While this works for individual trains, it does not consider the interaction of train movements with a fixed-block signalling system. A fixed-block signalling system is designed to operate with trains at a given speed; the greater the difference of actual train speed against this optimised speed, the longer signals will display a red aspect.

When the network is considered as a system (rather than as a collection of individual trains moving against the timetable), performance becomes not only the ability of trains to cover a distance between two points in a given time, but the interaction of trains. This is becomes increasingly important as the system approaches capacity, or when there is an incident and trains operate to the signalling system rather than to a timetable.

UK railways have seen an increase over the past 7 years of two metrics - 'Delay per Incident' (DPI) and 'sub-threshold delays'. This means that delays affecting passengers have increased, even though the number of incidents has declined. This has largely been attributed to increase in the number of services in operation and an increase in passenger numbers (affecting dwell times). There are parts of the UK system, however, where there have not been significant timetable changes and where passenger numbers have decreased in recent times—yet performance has still declined, even with fewer infrastructure incidents. This suggests that the increase in DPI and sub-threshold delay is not only caused by passenger numbers and/or increased numbers of trains but that other factors are affecting how the system performs.

1.1 Background to train operations in Wessex

The Wessex route in the UK (operated by First-MTR South Western Railway) connects London with the south west of England, as shown in figure 1. The terminal at London Waterloo is the busiest station in the UK; 8 tracks connect it with Clapham Junction, from where four continue to Reading, and four — the South Western Main Line — towards Southampton, Bournemouth and Weymouth. Additional branches connect the line with other centres in the region, such as Portsmouth and Exeter.

This intercity traffic uses two of the four tracks available between Woking and London, with the other two dedicated to local services: in peak hours the same fast track is used by 24 trains per direction. Despite the high density of block sections, with such high traffic density the probability of unplanned braking actions due to restrictive signal aspects is quite high; the lower speed cause a further increase of the blocking time, which propagates to the following trains.

The timetable on the Wessex route in the UK has been largely unchanged since 2004. Although the rolling stock in use is principally the same, the formations have changed with trains being typically increased from 8 cars to 10 cars on suburban routes. The infrastructure layout has remained the same with the exception of Waterloo which saw some remodelling in 2017. Despite considerable efforts by Network Rail and the train operator, performance has declined steadily since 2010 (figure 2).

Passenger numbers increased on Wessex have increased steadily up to 2017 but recent data from the Office for Rail and Road shows a decrease in journey numbers by 7.9% in the past year to 212 million per annum (the lowest level since 2012/13) yet there has been no corresponding improvement in performance.

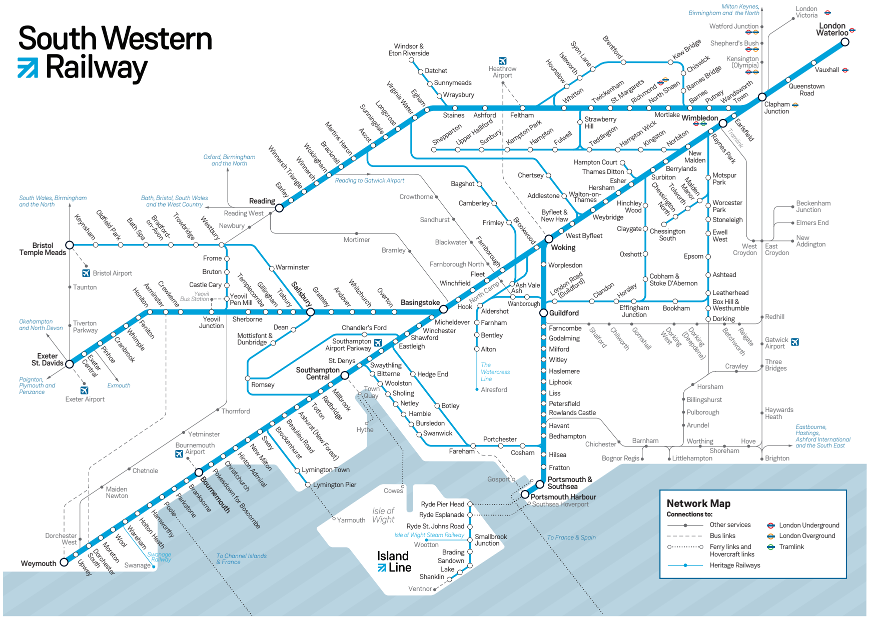


Figure 1: South Western Railway routes

1.2 Identifying changes to operations

The changes in performance in Wessex have been subject to numerous recent reports but these focus on the reliability of the infrastructure [ref to Holden report] rather than on how the train operator is delivering the service. This is partly because asset failures are readily identifiable (typically these cause delays that the system can easily identify) whereas train operations and station delays are often much smaller and harder to identify, despite being more prevalent. While there is a more work taking place to identify where these 'sub-threshold' delays are occurring, there is very little historical data to show if these have changed.

In this study, Hasler TELOC on-train data recorders have been analysed from files extracted in 2002 and 2018 and the braking curves from the data sets compared. These changes have then been simulated to quantify how much of the decline in recent performance can be attributed to the changes in how a train operated uses the available capacity of the network. The focus of the study is on how drivers brake to a halt. This is because this variable is wholly within the control of the train operator and is an action repeated continuously on a train's journey. Even a small change in braking style is likely to manifest itself when repeated for every station stop or restricted aspect. Furthermore, the braking of rolling stock conforms to standards and therefore a step one application (i.e. 3%g) in 2002 will be consistent with a step one application in 2018.

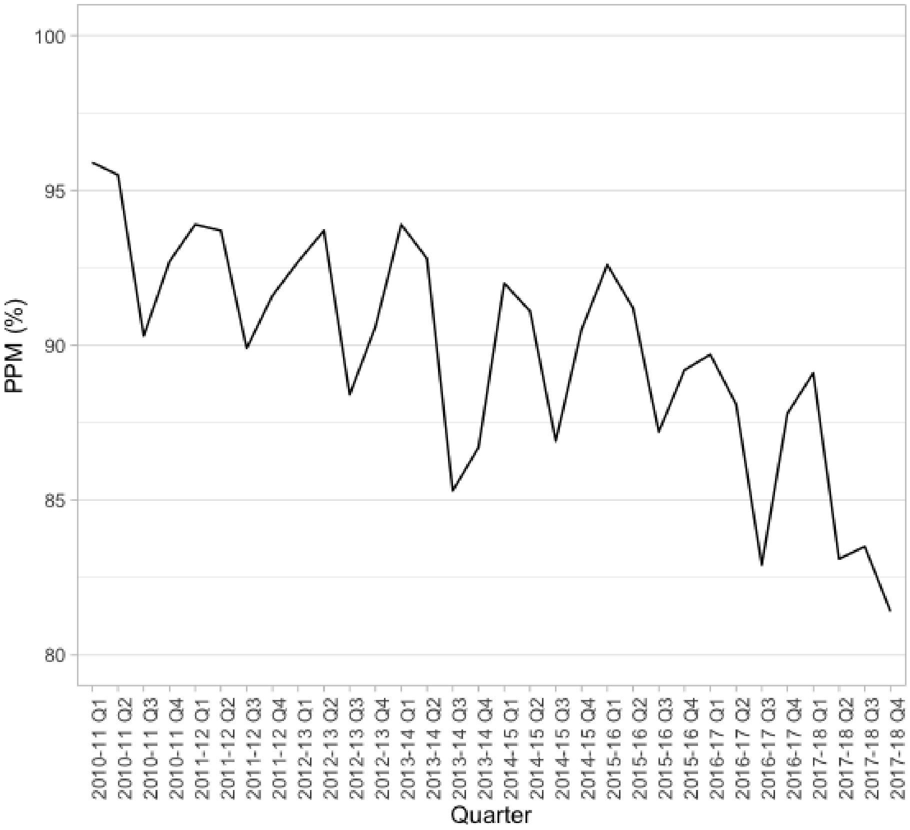


Figure 2: South Western Railway performance since 2010

2 Methodology

2.1 Sources of 2002 data

Routine down-loading of OTDR has only become common in recent years. In the early days of OTDR, it was necessary to physically connect a laptop to the train to obtain a file, and then save the data to a floppy disk, making data collection onerous. Most files were therefore obtained to investigate an operational incident or a unit defect. No policy existed to preserve these files. The author, however, as part of a previous role in 2002 working between the then Infrastructure Manager (Railtrack) and Transport Undertaking (South West Trains), had access to some files obtained as part of safety investigations. The files came from class 458 (Alstom Juniper) units introduced in Wessex around 2000 and one of the earliest new fleets fitted with OTDR as part of the build specification.

The files were uncovered as part of a review of archived data by the authors in 2018.

2.2 Processing of 2002 data

The recovered data archive included 6 OTDR files and a copy of Hasler TELOC 2.0 software. The files were processed in TELOC and saved as .LTM files (a native TELOC text-based format). These files were the processed in Unix using a bash script to remove extraneous lines of text in order that the files could then be saved as .CSV files. Hasler files were then limited to approximately 65,000 rows; this tended to equate to about 24 hours' worth of data per file.

2.3 Sources of 2018 data

MTR is the parent company of MTR Crossrail. MTR Crossrail currently operates the TfL Rail services in London on behalf Rail for London. The TfL Rail services will become the Elizabeth Line on the opening of the Crossrail tunnels through central London.

MTR Crossrail operates class 345 (Bombardier Aventura) units. These are equipped with Hasler OTDR and TELOC software. MTR Crossrail provided data from the class 345s from February 2018. The files were processed through TELOC and exported to Microsoft Excel in order that the outputs for speed and time could be converted to .CSV format. Due to a limitation of the software, it was only practical to parse a subset of the total available class 345 data to .CSV format.

Although 2018 and 2002 data come from different rolling stock and routes, it should be noted that they operate on similar suburban railways. Station distances and line speeds are comparable for the two railways.

2.4 Extraction of braking curves

The R data.table package was used to process the .CSV files to extract the braking curves with ggplot2 being used to produce graphics. The start of the braking curve has been defined as being the maximum speed in the last 90 seconds prior to stopping. Curves were only included where the maximum speed was between 50 and 65 mph. [The scripts are attached as an appendix?]

There is no geographical context to the class 458 braking curves. The OTDR has a distance column but there is no means of identifying precisely to where those distances apply. The class 458s were used on the Waterloo - Reading and the Waterloo - Alton services but there were multiple variations in stopping patterns that make it hard to interpolate where the train is from the data available. Each curve could therefore be subject to variables such as railhead adhesion, gradient and track curvature but these have not been factored into the work.

Location information is available from the class 345 OTDRs but, for consistency with the work on the class 458s, this has not been included.

Identifying the train stopping point

It is extremely difficult to identify the exact moment a train wheel stops rotating from either GPS or the OTDR. In the class 458 data in particular, the algorithm appears to hold the last known rotational velocity of the wheel rather than replace the value with a zero. Since all other metrics (such as distance and acceleration) are derived from the wheel rotation, the train will often never appear to become stationary. This makes it difficult to trust the data below 2.5 mph (1.1176 ms^{-1}); the train speed often stays above 0 mph even when it is known that the train is stationary.

A stopping point has therefore been defined as being (for the class 458 data) 3 seconds after the train speed has decreased below 3mph. For the 345s a combination of the position of the power brake and train speed has been used since the odometer readings show higher accuracy (but still not clearly reaching an absolute 0 mph), with the stopping point being assumed as being 1 second after the train speed decreases to 1 mph. GPS data also shows limitations in identifying the exact stopping and so has not been used for this work.

2.5 Changes to driver training policy

There have been changes in the driver training policies in the UK over the past 20 years in response to incidents involving drivers failing to stop at red aspects [ref to Ladbroke Grove, Southall,], and to accompany the fitment of Train Protection Warning System. The mitigation for preventing SPaDs including braking on sight of restricted aspects, not entering a platform at more than 30 mph, not exceeding 20 mph at 200 yards (approximately 200m) from the signal; and stopping 20 yards short of the signal. Furthermore, drivers were discouraged from using step three braking (i.e. 9%g) and taught to only brake in steps 1 & 2 (3%g and 6%g respectively). Whereas, on suburban systems, drivers used to drive at line speed on double yellows (it being possible to brake to a stand from line speed from sighting the single yellow), drivers are now required to start braking on sight of a double yellow. Despite the changes in driving styles, there has been no corresponding changes to signalling design specification to account for these changes.

2.6 Simulation of the network

Simulation tool

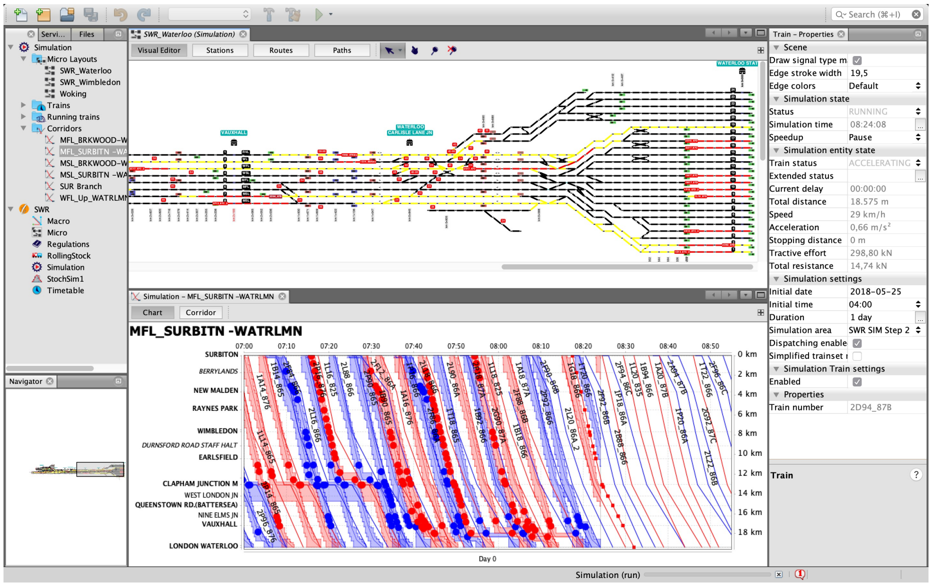
The *trenissimo* simulation programme (see 3) has been used to simulate the small changes in driver style. The tool (de Fabris et al., 2018) has been developed by trenolab. It is a synchronous, microscopic simulation tool, aimed at reproducing railway operations in the most accurate way, with a special focus on the representation of stochastic factors, such as the dwell times and the variability of running times. *trenissimo* is a Java application natively compatible with all operating systems. It was developed using the Netbeans Platform, a framework designed to create a very flexible and user-friendly environment.

The tool reproduces railway operations in a mixed continuous-discrete approach: it calculates the solution well-known motion equation (Wende, 2003) of trains in a continuous way, considering the discrete processes of signal states. At present, *trenissimo* features the Italian, French, British and Norwegian signalling systems, as well as the ETCS Level 1 and 2.

One of key strengths of *trenissimo* is that the dispatcher is simulated: as in the real world, while automatic block signals are automatically set to green, a dispatcher oversees the operation, opening the home and exit signals based on the planned and actual positions of trains. As a result, and similar to real operations, the dispatcher always controls operations: he can take decisions based on simple rules, or more complex algorithms. This principle allows implementing robust deadlock-prevention algorithms, as well as testing the effective impact of different dispatching strategies. Based on a set of rules, the dispatcher is also able to cancel train services, or short-turn them to reduce the propagation of delays.

Following the principles explained in (Medeossi et al, 2018), the key stochastic inputs for an accurate simulation are the initial delays, the dwell times, and the variability of running times. *trenissimo* implements the combination of stochastic dwell times and departure inaccuracies proposed in (Longo, 2012) to accurately represent the dwell times of the early- and late-arriving trains. The dwell time is considered as the stop time related with the exchange of passengers and this is applied to all trains stopping at a station, while the departure inaccuracy represents the departure variability of trains that arrive early at the stop, but do not depart on time due to an overlong departure process, or to passengers arriving at the last second.

In previous work (Medeossi et al, 2011) it was demonstrated that to represent accurately the running time a set of parameters is required, each representing the way drivers drive during one of its phases: acceleration, cruising, braking and coasting. Additionally, ideally braking at stops, signals and speed restrictions would be considered separately though this is not currently implemented. The work also proposed and tested a method to estimate the distribution of these parameters based on GPS or OTDR data, which has been used in this study as an input for the simulation.



Train – Properties	
Scene	
Draw signal type m	☑
Edge stroke width	19,5
Edge colors	Default
Simulation state	
Status	RUNNING
Simulation time	08:24:08
Speedup	Pause
Simulation entity state	
Train status	ACCELERATING
Extended status	
Current delay	00:00:00
Total distance	18.575 m
Speed	29 km/h
Acceleration	0,66 m/s²
Stopping distance	0 m
Tractive effort	298,80 kN
Total resistance	14,74 kN
Simulation settings	
Initial date	2018-05-25
Initial time	04:00
Duration	1 day
Simulation area	SWR SIM Step 2
Dispatching enable	☑
Simplified trainset r	☐
Simulation Train settings	
Enabled	☑
Properties	
Train number	2D94_87B

Figure 3: Screen shot of trenissimo system

2.7 Simulation of the system

The May 2018 morning peak hour timetable of the Woking - London Waterloo section of the South Western Main Line was simulated using the real distributions of input delays, dwell times and departure inaccuracy derived from track-circuit logs. The input delays, which are the distribution of departures from the first ocp within the simulation area, were filtered to remove secondary delays (Medeossi et al., 2011), while the distributions of departure inaccuracy and dwell times were respectively obtained considering only record of late- and early-arriving services, plus the variability of driving styles obtained from the analysis of OTDR data.

Three scenarios were simulated. The first represents 2018 operations, in particular considering the 2002 braking behavior and shorter train formations. The second scenario instead considers the 2018 braking behavior, while the third one combines it with the 2018 train formations.

The simulation of each scenario was repeated for 200 times using a Monte Carlo approach; the occupation time of selected block sections during the peak hour (08:00 – 09:00) and delay indicators (mean delay and punctuality at 5') at arrival at Waterloo in the were used as KPIs of each scenario.

3 Results

3.1 Comparison of braking curves

Figure 4 shows the spread of speeds of braking curves at three second intervals to a halt from the class 458s in 2002 and the class 345s in 2018.

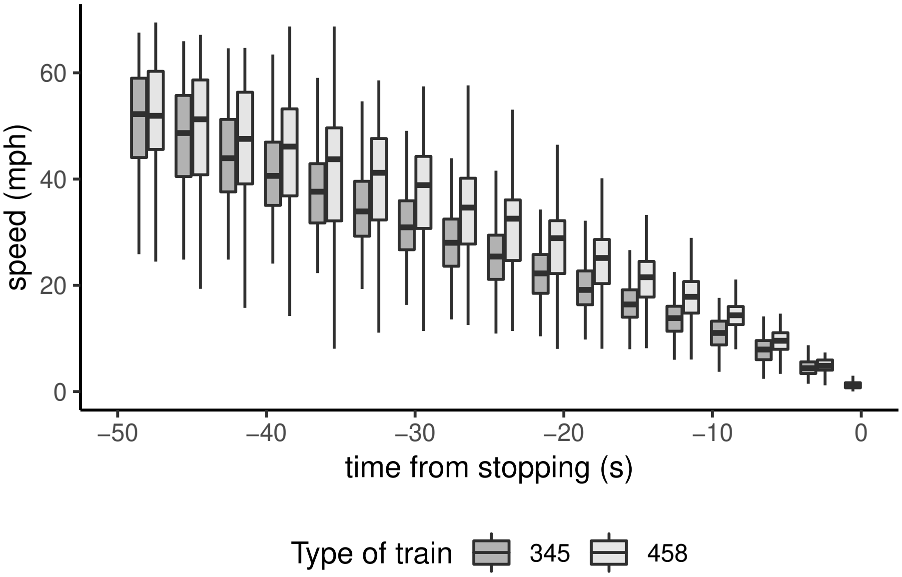


Figure 4: Braking curves from 2002 (class 458s) and 2018 (class 345s)

At 50s prior to stopping there is very little difference between the spread of results for the two data sets but the class 345s start braking sooner and, at -30s to stopping, the median for the 345s is 8mph slower than the 458s. When comparing the median for the two data sets, the class 345s take 4 seconds longer than the class 458s to halt.

3.2 Implications of changing train length

The class 458s were bought as four-car units. For peak services they would be coupled together to make 8 car trains, totalling 164m in length. In 2016 the trains were lengthened to five-car sets to provide increased peak capacity. When coupled, these are now 204m long.

It is easy to calculate the increase in time it takes for the longer trains to clear a given track circuit, and the commensurate time the signal in rear stays at red. This needs, however, to be combined with the changes in braking rate. If it is assumed that there is a block joint 300m from the stopping point, the slower braking curve and the longer train means that the protecting signal in rear will be red for 8 seconds longer than it would have been in 2002.

For a suburban railway with a planning headway of 150 seconds, this represents a near 6% loss in capacity.

3.3 Network Rail modelling assumptions

Network Rail has a standard for the modelling assumptions to be used in RailSys simulations. Braking of Electric Multiple Unit trains is required to be modelled at 0.588 ms^{-1} (8%g). It is apparent that this was not met in 2002, and the discrepancy has increased since then.

Figure 5 shows the braking curves of the class 345s in 2018 compared with the expected trajectories from step 1 braking (3%g) and step 2 braking (6%g). This shows that, even though the class 345s have a continuous brake, the drivers continue to operate within the range that they are used to driving stepped brakes. This shows the extent to which drivers *feel* braking.

The Network Rail assumption for RailSys of 8%g is met in only one instance. This is unsurprising given that drivers are taught to avoid step 3 braking. Since braking typically involves amending the braking curve at some point before stopping, and drivers must be assumed to be using less that step 3 braking, it is extremely unlikely that a driver will be able to average 6%g since that would apply using the same step 2 brake throughout the whole braking curve. Instead, we see the outcome of a mix of step 1 and step two braking, averaging considerably less than the 8%g assumed by Network Rail.

3.4 Impact of changing train length and braking curves on performance

textbf[Note that these simulations are to be re-run. Currently the simulation is a doubling in train length from 4 to 8, not from 8 to 10 car] Figure 6 shows the mean simulated arrival lateness at Waterloo for the up fast lines. These graphs show the cumulative effects of a loss in capacity caused by the longer trains and slower braking styles, and the exponential rate at which delay accumulates once trains start to interact. It also shows that the system is extremely senstive to changes in train length. Of particular note is the rapid increase in lateness around 08:30 in the morning; the peak time for arrivals into Waterloo. The delays then increase exponentially, only recovering at the end of the peak as the services start to thin.

Figure 7 shows the total delay on route for each train for each of the scenarios. It can be seen how the combination of driving styles and longer trains are in themselves alone sufficient to fundamentally alter the performance of a route. Since fast trains to do not stop at many stations, the impact of the driver styles is much lower than that of the train length.

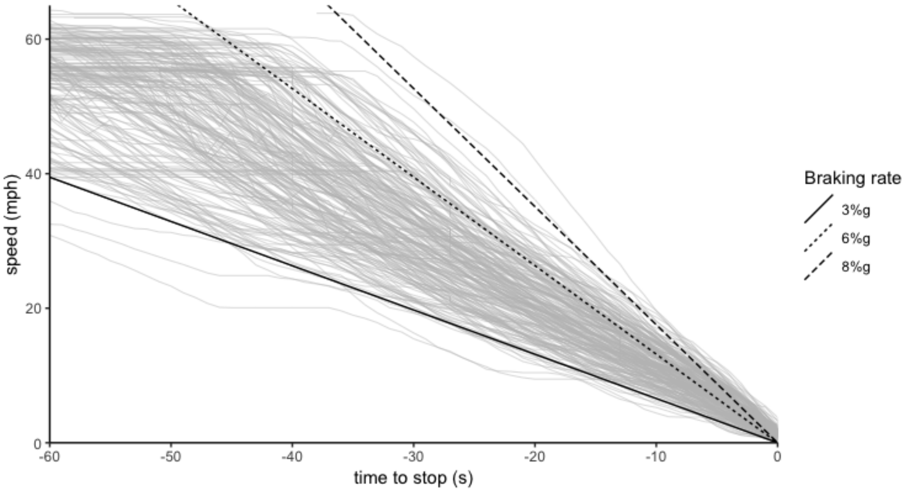


Figure 5: Comparison of 2018 braking curves with braking steps and NR modelling assumptions

The impact of driver styles is, however, almost 50% greater on long trains than it is on short trains.

The increases in overall lateness are shown in table 1 for each of the changes.

Increase from driving style	5.9%
Increase from longer trains	21.7%
Increase from longer trains and driving style	37.1%

Table 1: Changes in delay from 2002 to 2018

4 Discussion

4.1 Validity of findings

[Note of 1 Feb 2019 - these results are of the Up Fast only. The Up Slow results also need to be incorporated. Second, these results are based on doubling of train lengths where as the train formation changes on the fast are less substantial than on the main.]

A number of assumptions have been used for the purposes of the study. The consistency between the sets suggests that these assumptions do not invalidate the findings. The braking curve rates of 2018 are lower than 2002 from similar starting speeds; both data sets are lower than the modelling rates assumed by Network Rail for capacity and performance modelling; and the simulations show that even small changes in braking rates on a network with similar operating constraints leads to an increase in delay per incident and decline in overall performance of the system. It is acknowledged, however, that there are weaknesses of the study caused by not having the same base assumptions. In practice, however, there

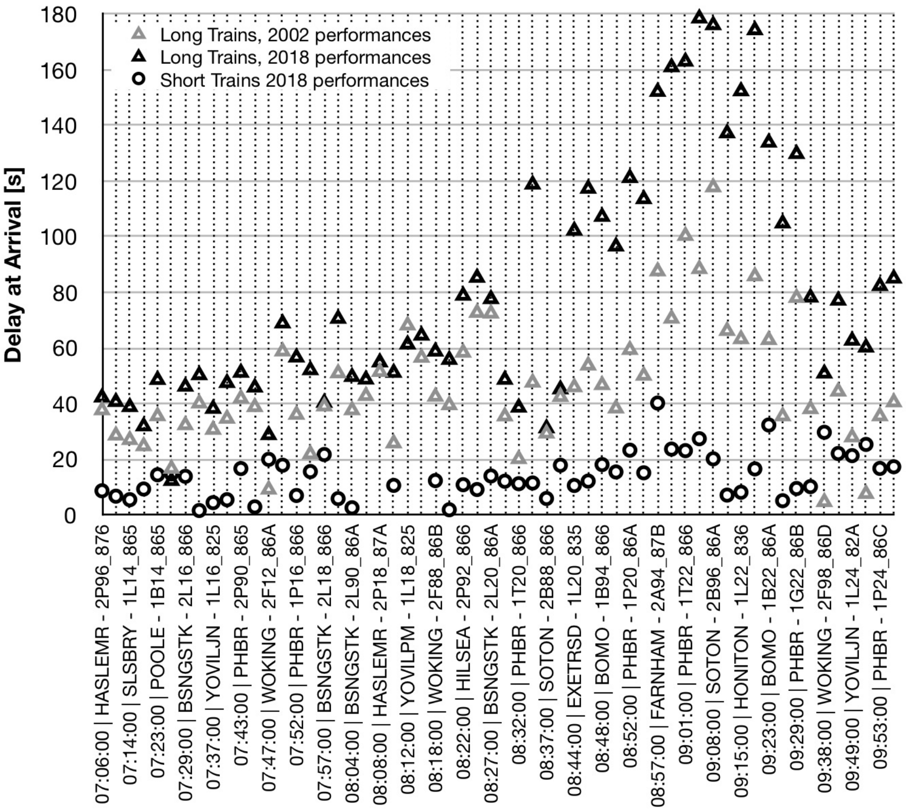


Figure 6: Simulated average lateness at Waterloo for changes to train length and braking style

is very limited data available from 15 years ago and, even where this does exist, it is very unlikely that the same units are operating on the route, or that the route has remained entirely unchanged.

The system shows considerable sensitivity to changes in train length. On an urban system, with relatively short track circuits, this is perhaps unsurprising since it takes the train longer to clear the track circuit at slower speeds.

4.2 Implications

The train operator has the capability to determine the rate at which delay dissipates across the network by varying from the optimum profile at which, for a given train service, signals revert from red to a proceed aspect. This effect is greatest when trains are constrained by the signalling system, rather than by the timetable. This typically occurs in congested routes, at busy junctions, and during disruption. Two factors that have implications for performance but which (in the UK at least) have not been assessed prior to their introduction are the

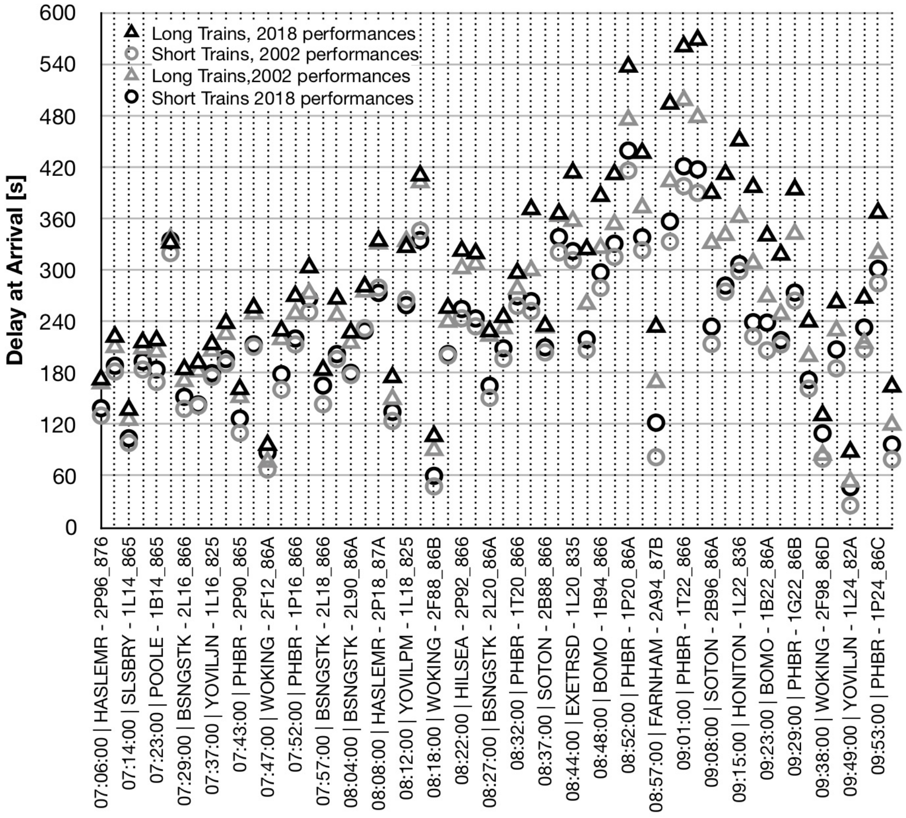


Figure 7: Total delay on route for each scenario

length of trains and speed at which they are driven. Both of these by themselves increase the time in section, but a combination of both greatly increases the delay with slower driving styles and longer trains being 50% worse than slower driving and short trains.

5 Summary of findings

The change in driver styles between 2002 and 2018 have resulted in it taking four seconds longer today to halt a train from 60mph than in 2002. When combined with longer trains, the signalling system takes longer to clear. The *trenissimo* simulations show the sensitivity of the network to these changes. On the up fast, the combined changes account for [x]% of trains being more than [x] minutes late at destination.

References to be completed

References

Albrecht T., Goverde R.M.P., Weeda V.A., Van Luipen J., 2006. “Reconstruction of train trajectories from track occupation data to determine the effects of a Driver Information System”, In: Allan J., Brebbia C.A., Rumsey A.F., Sciutto G., Sone S., (eds.), *Computers in Railways X*, WIT Press, Southampton.

Besinovic, N.; Quaglietta, E.; Goverde, R.M.P., 2013. “A Simulation-based Optimization Approach for the Calibration of Dynamic Train Speed Profiles”, *Journal of Rail Transport Planning & Management 3*, (2013).

Cui Y., Martin U., Zhao W., 2016. “Calibration of disturbance parameters in railway operational simulation based on reinforcement learning”, *Journal of Rail Transport Planning & Management*, vol 6, issue 1.

de Fabris S., Medeossi G., and Montanaro G., 2018. “trenissimo: Improving the microscopic simulation of railway networks”, In: *Proceedings of the 16th International Conference on Railway Engineering Design & Operation July 2-4 2018*, Lisbon.

Medeossi, G., Longo, G. and de Fabris, S., 2011. “A method for using stochastic blocking times to improve timetable planning”, *Journal of Rail Transport Planning & Management 1*, (2011), pp. 1-13.

Powell, J., Palacin R. (2015). “A comparison of modelled and real-life driving profiles for the simulation of railway vehicle operation”, *Transportation Planning and Technology*. 38. 10.1080/03081060.2014.976984.