Sub-Threshold Delays and GB Timetable Quality

John Armstrong [a,1], John Preston [a], Stephan Zieger [b], Nils Nießen [b]
[a] Transportation Research Group, University of Southampton,
Burgess Road, Southampton SO16 7QF, UK
[1] E-mail: j.armstrong@soton.ac.uk, Phone: +44 (0) 23 8059 9575
[b] Institute of Transport Science, Mies-van-der-Rohe-Str. 1,
D-52074 Aachen, RWTH Aachen University, Germany

Abstract

Performance on railways in Great Britain (GB) has improved with the reduction in traffic levels since the onset of Covid-19. However, if and when traffic levels return to (and potentially exceed) pre-Covid levels, performance is likely to deteriorate again, and an improved understanding of the relationships between timetable characteristics and performance will be needed. It is proposed that some recurring sub-threshold delays are related to timetable quality issues, particularly in relation to the allocation of dwell and running time allowances (or supplements) in the timetable planning rules. High-level analysis of performance data supports this hypothesis, indicating consistent levels of sub-threshold delay at busy locations on the network, with running time delays sometimes being recovered at major stations with extended scheduled dwell times. Detailed analysis of timetable planning rules and historic timetable and sub-threshold delay data indicates some clear and strong relationships between the dwell and running time allowances provided and resulting performance. However, these relationships can be complicated and obscured by the presence of multiple sources of secondary delays on busier sections of the network. The initial application of location-specific relationships to predict cumulative performance over longer route sections is demonstrated.

Keywords

Timetable quality, sub-threshold delay, performance data, allowances, cumulative lateness

Paper Type

Type A: Research paper (i.e. theoretical research)

1 Introduction

This paper describes work undertaken for the Data-Driven Robust Timetabling (DDRT) project, sponsored by Britain's Rail Safety and Standards Board (RSSB) and by Network Rail, the infrastructure manager (IM) of Britain's railway network. The aim of the project was to identify and quantify timetable-related performance risks, thus enabling the development of timetables that provide robustness without consuming excessive levels of capacity. The project proposal included four work packages (WPs):

- WP1: High-level, 'top-down' performance analysis
- WP2: Review of existing, alternative approaches to timetable performance analysis
- WP3: Detailed, 'bottom-up' performance analysis

- WP4: Application of findings as ’overlay rules’ and in timetable planning software

This paper focusses primarily on the work undertaken for WP3 and the results obtained; aspects of WP1 are also described and summarised to provide context for WP3.

Following this introduction, the background to and objectives of the research are first described. The methodology and data used (and their evolution) are then presented, together with the results obtained. Ongoing and further work are then briefly summarised, followed by conclusions, acknowledgements and a list of references.

2 Background

Prior to the outbreak of Covid-19, increasing traffic levels on Britain’s railways led to declining levels of performance (i.e. punctuality) and increasing levels of reactionary (or secondary) delay. This is illustrated in Figure 1, based on data from Network Rail for the years between 2010 and 2017 (Armstrong, Raine et al., 2019). The relatively constant blue line represents moving annual average (MAA) primary delay, the yellow line represents increasing MAA reactionary delay, and the red and purple lines respectively represent the increasing difference and ratio between the two.

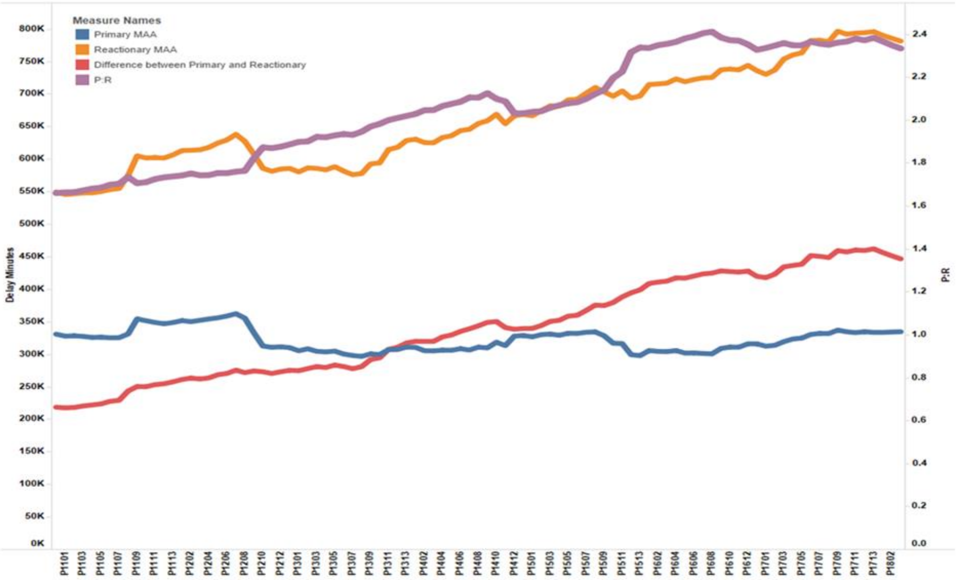

Figure 1: 2010-17 Delay Trends (Network Rail)

In a railway industry speech (Network Rail, 2019), Network Rail’s Chief Executive confirmed the operational challenges presented by increasing traffic levels and reduced levels of spare time in the timetable for recovery from perturbation. He noted that secondary delay then accounted for 70% of attributed delay, over twice as much as the underlying primary delays, and that growing levels of unattributed ‘sub-threshold’ delays (i.e. delays of less than three minutes) accounted for approximately 35% of all delay across the network. Because they are unattributed, these relatively small delays and their causes, and their potential to trigger larger secondary delays, are poorly understood.

As reported by the Office of Rail and Road (ORR, 2020a), the Covid-19 pandemic caused “a substantial reduction in train services on the network which led to improvements in punctuality and reliability.” On-time arrivals (i.e. within one minute of the scheduled arrival time) at stations between April and June 2020 increased by 17.1 percentage points relative to the year before to 86.4%. Similarly, the Public Performance Measure (PPM, the percentage of short- and long-distance trains arriving at their final destinations within 5 and 10 minutes respectively of their scheduled times) increased by 6.2 percentage points to 96.2%. These “were the highest quarterly percentages (i.e. best) since the time series began”, in 2014-15 and 1997-98 respectively. Cancellations were halved to 1.2% of scheduled services, “the lowest quarterly percentage (i.e. best) since the time series began in 2014-15.” Total passenger train km between April and June 2020 were 84.2 million, 39.6% less than for the equivalent quarterly period in 2019 (ORR, 2020b).

While some of this performance improvement may be due to the development of better timetables (the operation of fewer trains reduces the potential for conflicting train paths and the need for timetable compromises and scheduled delays), much of it is likely to have arisen from lower levels of secondary delay as a result of reduced traffic volumes. If and when traffic levels return to, and potentially exceed, their pre-Covid volumes, performance is likely to deteriorate again, re-emphasising the need for improved understanding and measures of timetable quality, in terms of the relationship between timetable characteristics and the resulting train performance. Previous consultancy work has been undertaken in Britain in an attempt to understand the sources of secondary delays in particular, but this has tended to focus on areas other than the details of the timetable and its relationship with the timetable planning rules, and their combined influence on performance.

In the course of the European ON-TIME project, Goverde and Hansen (2013) identified four broad measures and definitions of timetable quality:

- ***Timetable feasibility*** *is the ability of all trains to adhere to their scheduled train paths. A timetable is feasible if the individual processes are realizable [i.e. achievable] within their scheduled process times and the scheduled train paths are conflict free.*

- ***Timetable robustness*** *is the ability of a timetable to withstand design errors, parameter variations, and changing operational conditions [and] is one part of the overall robustness of the railway system.*

- ***Timetable stability*** *is the ability of a timetable to absorb initial and primary delays so that delayed trains return to their scheduled train paths. Hence, a timetable is stable if any train delay can be absorbed by the time allowances in the timetable without active dispatching [i.e. intervention in the form of re-routeing, re-scheduling and/or cancellation of trains].*

- ***Timetable resilience*** *is the flexibility of a timetable to prevent or reduce secondary delays using dispatching (re-timing, re-ordering, re-routeing) [and] can be viewed as the complement of robustness.*

They also observed (ibid.) that Network Rail’s timetable planning process provides timetable stability, but does not guarantee feasibility or robustness. This is consistent with the performance improvements seen with reduced traffic levels in response to Covid-19, meaning that a large proportion of the remaining primary delays can be absorbed without causing knock-on, secondary delays.

WP2 of the DDRT project reviewed a range of detailed, microscopic approaches to the assessment of timetable quality that are already in use elsewhere. These include:

- Recoverability quotients, i.e. the number of timetable cycles required to recover from a given level of initial delay (Pachl, 2018).
- Queuing theory applications (Nießen, 2014).
- Max-Plus Algebra (Goverde, 2014).
- Periodic Event Scheduling Problem (PESP)-based optimisation approaches (Cacchiani and Toth, 2018).

However, it was confirmed in WP2 that these approaches typically require more detailed data inputs than are – so far – routinely available in Britain, notably more accurate technical running time and headway values. This problem arises because the specified planning values for minimum dwell, running, headway and other times specified in the timetable planning rules (TPRs) for Britain's railways (Network Rail, 2020a) include dwell and running time supplements and buffer times, but do not specify their values or those of the underlying technical minimum time values. Further work is thus required to produce the required inputs and/or to modify the approaches to enable the use of the input data currently available. It was therefore concluded that, for the purposes of DDRT, and for WP3 in particular, the primary focus should be on the investigation of empirical relationships between timetable characteristics and performance.

3 Objectives, hypotheses and assumptions

As indicated in the Introduction, the overall aim of the DDRT project was to identify and quantify timetable-related performance risks, and enable the development of feasible timetables that provide robustness while making efficient use of the available capacity. In particular, the project aimed to test the following three research hypotheses:

(i) Timetable quality (in terms of feasibility and robustness) issues are reflected in at least some of the sub-threshold (i.e. < 3 minutes) delays experienced across the network. If a timetable complies with the TPRs, it is unlikely to cause threshold (≥ 3 minutes) primary delays, although this is not necessarily the case for any resulting 'knock-on', secondary delays.

(ii) Some of these delays reflect incorrect allowances in the TPRs. Some sectional running times (SRTs), headways, margins and dwell times may have insufficient supplements or buffer times, others may have too much.

(iii) Timetable quality could be improved by analysing and amending the allowances (and, ultimately, by adjusting the TPRs).

It was assumed that the timetable-related (i.e. small, sub-threshold) primary dwell time delays are due to insufficient scheduled dwell times, and that primary running time delays are due to insufficient scheduled running times. Dwell times should not be affected by secondary delays (i.e. delays caused by late-running preceding or otherwise conflicting train movements), unless a signal controls trains' departures from a platform, and clearance of the signal is prevented by delays to preceding trains and dwell times are extended as a consequence. Secondary running time delays are likely to be due to insufficient planning headway values (including buffer times) to absorb the effects of routine, small delays to preceding services (such delays may also occur on the approach to a station if a preceding/conflicting train's departure is delayed due to insufficient scheduled dwell time). These assumptions are summarised in Table 1.

Table 1: Sources of Dwell and Running Time Primary and Secondary Delays

Delay Types	Primary	Secondary
Dwell Times	Insufficient minimum dwell times	N/A (but possibility of inadequate 'downstream' allowances)
Running Times	Insufficient minimum running times	Insufficient planning headways and/or dwell times (for preceding trains)

4 Methodology and Findings

The Wessex Route of Britain's railway network, and the South West Main Line (SWML) in particular, were selected for the initial analysis in WPs 1 and 3. The work described here focussed mainly on the four-track section of the SWML between Basingstoke and Waterloo, where South Western Railway (SWR) is the primary passenger train operating company (TOC) and network user. SWR's network map is shown in Figure 2 (SWR, 2017-2020).

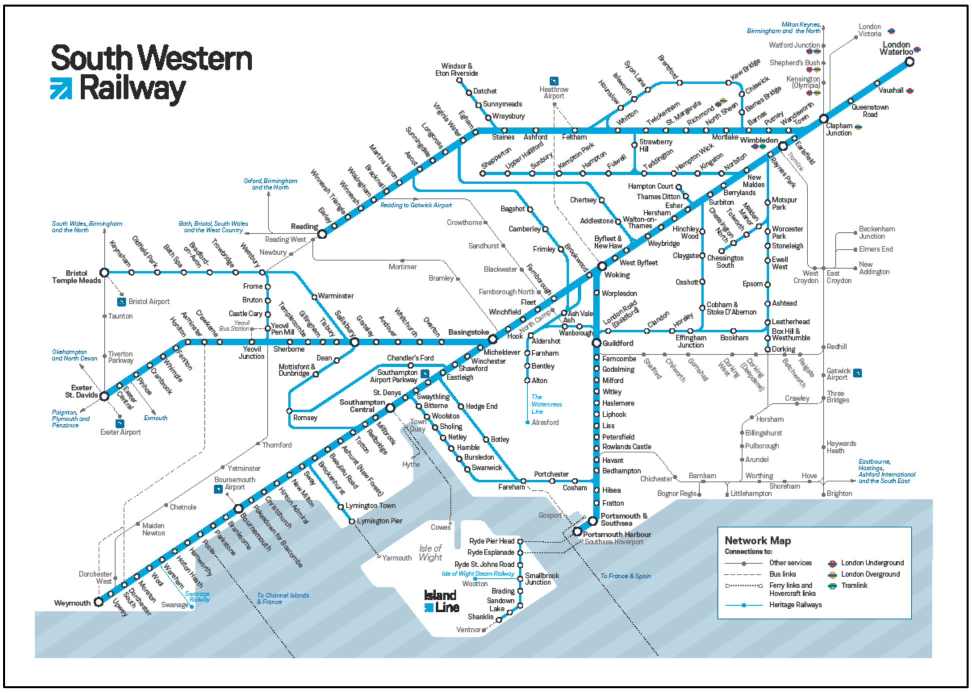

Figure 2: SWR network map

The Wessex route covers the area between London Waterloo and locations to the south and west, including Southampton, the choice being made for reasons of route, timetable, operations and rolling stock familiarity, and the mix of traffic using the route. For some of the work described below, the study area was extended to include the four-track route section between Worting Junction, where the double-track lines from Southampton and Salisbury converge, to Basingstoke. Between Worting Junction and Basingstoke, the route is shared with freight and CrossCountry-operated passenger services. The analysis was conducted at two distinct levels of detail, corresponding to WPs 1 and 3 described above.

4.1 WP1: high-level, ‘top-down’ analysis

Network Rail initially provided Lateness data for SWR for the 12 months between the December 2017 and December 2018 timetable changes. Lateness data records the differences between actual (truncated to the nearest minute) and scheduled train times. The initial results of the high-level, ‘top-down’ analysis of this data were described by Armstrong, Zieger et al. (2019). In summary, it was found that relatively small delays consistently occur at the busiest stations and sections on the network, and that these could potentially be addressed by amendments to the timetable and the TPRs. Such amendments would consume additional system capacity, requiring trade-offs, given the busyness of these locations. However, this was prior to the outbreak of Covid-19, since when demand levels have fallen dramatically and the timetable has been ‘thinned out’ with resulting performance improvements. This provides an opportunity to introduce timetable and TPR changes to ensure that improved performance can be maintained if and when demand returns towards pre-Covid levels. For validation purposes, Network Rail also provided Lateness data for Great Western Railway (GWR) train services between London Paddington and south Wales and the south and west of England. The GWR network map is shown in Figure 3 (GWR, 2020).

Correct as of 12 June 2020

Figure 3: GWR network map

The GWR Lateness data covered the 15 months following the December 2018 timetable change, and included the first three months following a major timetable change in December 2019 and prior to the outbreak of the Covid-19 pandemic, which caused a large reduction in passenger demand and a reduced timetable from 22 March 2020.

The equivalent high-level analysis of the GWR Lateness data produced similar findings to those for SWR, i.e. that trains were subject to recurring small delays at the busiest stations and sections, and it also identified less busy locations that were similarly subject to recurring delays. As was the case for SWR, locations with 'surplus' time allocations were also found, suggesting that there might be scope for some reallocation of scheduled times, to improve performance and make better use of available capacity. It was found that overall dwell time performance improved after the December 2019 timetable change, but that overall running time and total performance deteriorated, suggesting that dwell time allocations were partly compensating for insufficient scheduled running times. An analysis of running time performance between Bristol Temple Meads and London Paddington supported this, also indicating a deterioration following the timetable change. Subsequent detailed analysis of performance on the SWML, described below, also indicated that running time lateness was being recovered at stations with extended scheduled dwell times, effectively reversing the concept of 'zones of concentration' at stations and junctions, and 'zones of compensation' on intermediate line sections (Caimi et al., 2009).

4.2 WP3: detailed, 'bottom-up' analysis

Because of the non-availability of some of the detailed data required to apply the established methods of timetable quality assessment referred to in Section 2, an empirical approach was taken, making use of the available historic timetable, TPR and performance data. This type of data-driven performance analysis is advocated, in a related context, by Besinovic (2020).

Initial Approach

Based on results obtained from the analysis of SWR performance data in WP1, a single station and single route section were initially selected for detailed analysis of dwell and running time performance. The chosen station was Farnborough (Main), a busy station on the four-track section of the SWML in London's outer suburban railway 'commuter belt' with two platforms, one each on the Up (towards London) and Down Slow lines. The route section initially chosen was the Up Slow line from Wimbledon to Earlsfield, carrying stopping/local inner suburban commuter services to London Waterloo. Each was chosen on the basis of its position in the 'mid-range' of performance data, i.e. some but not all trains were subject to delay on average, providing a range of conditions and data points, and suggesting that it should be possible to identify the factors that were causing delays to some services, while other trains were unaffected. A further reason for the initial choices made was that both locations are on the four-track section of the SWML, the area of particular interest in the context of the DDRT project.

Initial analysis did not produce the desired results because of the absence of variation in scheduled dwell and running times, however: 65 out of 67 (97%) weekday trains stopping at Farnborough (Main)'s Platform 1 (Up direction) had a scheduled dwell time of one minute, the minimum value specified in the TPRs, and 281 of 284 (99%) weekday trains running from Wimbledon to Earlsfield on the Up Slow line had a scheduled running time of three minutes, again the minimum value specified in the TPRs. This meant that, while there was considerable variation in the dependent, performance variables, there was very little variation available in the independent dwell and running time variables to 'explain' the variations in performance, and this was reflected in initial graphical representations of the data, which showed little in the way of obvious relationships. This problem was exacerbated by the initial inclusion of threshold (i.e. ≥ 3 minutes) delays in the dataset, some of which were quite large and further obscured any relationships between timetable characteristics and sub-threshold delays.

Analytical Refinements

With hindsight, the lack of variation in planned timings was an obvious potential problem and challenge, given the nature of the timetable planning process and rules, and the comparative regularity and intensity of the timetable in these locations, which means that trains are likely to be scheduled at minimum dwell and running times. As an initial response, the running time analysis was extended to the 248 weekday trains scheduled on the Up Fast line from Wimbledon to Clapham Junction, which carries longer-distance services, of which very few stop at Wimbledon, but some stop at Clapham Junction, providing a range of stopping patterns and Pass-to-Pass and Pass-to-Stop running times, and thus a greater degree of 'explanation of performance'. The numbers of trains in these datasets are too large to present their schedules graphically here, but the range of scheduled SRTs for Wimbledon – Clapham Junction is shown in Table 2 for the diesel and electric multiple units (DMUs and EMUs) on the route, including some older British Rail (BR)-era suburban units.

Table 2: SRTs for the Up Fast line from Wimbledon to Clapham Junction

Rolling Stock	Movement	SRT (minutes)
Class 159 (DMU)	Pass-to-Pass	3.0
Class 159 (DMU)	Pass-to-Stop	4.5
Class 450/444 (Desiro EMU)	Pass-to-Pass	3.0
Class 450/444 (Desiro EMU)	Pass-to-Stop	4.5
Class 450/444 (Desiro EMU)	Stop-to-Stop	5.0
Class 455 (Ex-BR suburban EMU)	Pass-to-Pass	3.5

Discussion of the initial results with the project sponsors also led to a refinement of the process for identifying and calculating dwell and running time delays. For the initial calculations, dwell time delays were calculated simply as the change in lateness at a station between arrivals and departures (and running time delays were calculated as the change in lateness between departure from one location and arrival at the next). However, when an initial lateness value is negative (i.e. a train arrives at a station or enters a section early), this simple calculation approach tends to exaggerate lateness, and may indicate dwell or running time delays where none in fact exist. Dwell time lateness and delay calculations were therefore reviewed, and the process updated as summarised in Table 3.

Table 3: Dwell Time Lateness and Delay Calculations

Arrive\Depart	Early	On Time	Late
Early	No Delay	No Delay	Delay (+ive)
On Time	No Delay	No Delay	Delay (+ive)
Late	Delay (-ive)	Delay (-ive)	Delay (-/+ive)

If a train arrives at a station early or on time, and departs early or on time, it is assumed that no delay occurs: even in cases where an early arrival and on-time departure results in an extended actual dwell time, this does not necessarily indicate an insufficient scheduled dwell time. If a train arrives early or on time and departs late, a positive (i.e. > 0) delay is recorded. If a train arrives late and departs early or on time, a negative delay is recorded. Finally, if a train arrives late and departs late, a delay is recorded which may be positive or negative (or zero), depending upon the relative arrival and departure lateness values. While this initial analysis and classification is based on dwell times, a similar approach can be applied to running time lateness and delay, simply by transposing the Arrive and Depart lateness categories in Table 3, e.g. if a train departs early and arrives early, no delay is recorded.

As noted above, the initial results were also obscured by the inclusion of large, above-threshold delay values, and all threshold and larger delay values were therefore excluded from subsequent analysis to enable focus on sub-threshold delays only, which are more likely to be timetable-related, and also comprise the majority of the records obtained: typically more than 95%, and in excess of 99% for some locations, supporting the observation above that the timetable planning process in Britain may provide stability, but with no guarantee of feasibility or robustness. This focus on sub-threshold delays, combined with the revised approach to calculating delays, clarified the relationships between timetable characteristics and performance, even in the absence of running time variations and 'explanatory power'. The results of the revised (linear) regression analyses of the Up Slow and Up Fast lines from Wimbledon to Earlsfield and Clapham Junction respectively are summarised in Tables 4 and 5. The relationships between performance and additional (i.e. in excess of the minimum values specified in the TPRs) running time and preceding headways (all in minutes) were assessed individually and in combination, as shown.

Table 4: Statistics for Net Allowances and Performance: Wimbledon – Earlsfield

Explanatory Variable	**Intercept**	**t-Stat**	**Slope**	**t-Stat**	**R^2**
Additional Running Time	0.592	57.556	-0.133	-1.010	0.004
Additional Headway	0.635	46.721	-0.023	-4.685	0.083
Additional Running Time and Additional Headway	0.636	46.657	-0.153 -0.024	-1.212 -4.728	0.088

The R^2 values for Wimbledon – Earlsfield are very small, particularly for the Additional Running Time variable in isolation, indicating that the relationships involved are quite weak, i.e. that little of the variation in average delay is explained by the variations in additional running times and preceding headways. The signs of the slope values are, however, consistent with expectations: increases in running times and/or preceding headways correspond to reductions in average delay, as we would expect. Also, the t-statistic values indicate that the explanatory variables are statistically significant (marginally so, in the case of Additional Running Time). The intercept values are quite consistent for all three explanatory variable combinations, and are highly statistically significant.

The very low R^2 value for additional running times is likely to be due in part to the lack of running time variability, and possibly also to the approximate nature of the truncated lateness values used, although the impact of the latter may be reduced by the process of obtaining their averages. As additional running times and headways increase, they have a diminishing effect on performance, i.e. 30 minutes of additional headway is unlikely to provide performance benefits (in normal operating conditions) greater than those provided by, say, five additional minutes, and this too may affect the quality of the correlations obtained. Some small increases in R^2 were obtained by 'capping' the values of the additional running times and, particularly, headways to limit the size of their largest values (alternatively, a non-linear, negative exponential relationship between allowances and performance can be considered, as discussed below).

Table 5: Statistics for Net Allowances and Performance: Wimbledon – Clapham Jn.

Explanatory Variable	**Intercept**	**t-Stat**	**Slope**	**t-Stat**	**R^2**
Additional Running Time	0.762	33.715	-0.231	-12.416	0.370
Additional Headway	0.589	19.275	0.006	0.651	0.002
Additional Running Time and Additional Headway	0.840	27.663	-0.256 -0.031	-13.231 -3.740	0.402

The R^2 value for additional running time for Wimbledon – Clapham Junction is much larger than for Wimbledon – Earlsfield, perhaps reflecting the much greater variation in the additional running time values, and the slope value is again negative, as expected. The R^2 value for the additional headway values is very small, though, and the slope value is positive, contrary to expectations, and is not statistically significant. When both explanatory variables are included in the analysis, however, both slope values are negative, consistent with expectations, and the combined R^2 value is greater than the sum of its individual parts. However, the changes in slope signs and t-Statistic values for Additional Headway, when it is analysed in combination with Additional Running Time, suggest that there may be correlation between the two explanatory variables, and thus an issue of multi-collinearity requiring further statistical investigation.

The intercept values obtained are again highly statistically significant, and similar in value to those obtained for Wimbledon – Earlsfield, varying between 0.59 and 0.84, and suggest that there are consistent running time delays of between 0.5 and one minute when trains are scheduled to operate at the specified minimum planning values for running times and/or headways. Some small increases in R^2 were again obtained by capping the values of the additional running times and headways. It may also be worth disaggregating the analysis by stopping pattern and/or traction type, but it was decided that the analysis should first be extended to additional locations, rather than focussing exhaustively on the line sections from Wimbledon to Earlsfield and Clapham Junction.

The SWR timetable data was analysed to identify stations and line sections with high degrees of scheduled dwell and running time variation, and thus increased potential levels of ‘explanatory power’ in relation to performance. The locations with the greatest levels of variation are shown in Tables 6 and 7.

Table 6: Stations with High Levels of Dwell Time Variation

Station	**Platform**	**No. of Distinct Scheduled Dwell Times**
Woking	1	7
Basingstoke	3	6
Raynes Park	Not specified	6
West Byfleet	Not specified	6
Basingstoke	4	5

In some situations where the infrastructure configuration restricts trains on a particular route to the use of a unique platform at a given station, the timetable data does not explicitly state the platform number: such cases are listed as ‘Not Specified’ in Table 6 and, where required, the timetable data was subjected to further analysis to identify the platform numbers.

Table 7: Route Sections with High Levels of Running Time Variation

Section	Line	No. of Distinct Scheduled Running Times
Clapham Jn. – Waterloo	Main Fast Line	12
Raynes Park – Wimbledon	Slow Line	8
New Malden – Wimbledon	Fast Line	7
Basingstoke – Farnborough	Fast Line	4
Farnborough – Woking Jn.	Fast Line	4

Because of concerns about the approximations arising from the use of truncated Lateness data, as noted above, Network Rail also provided ‘Control Centre of the Future’ (CCF)-derived train describer (TD) data for SWR services, recorded to the second, together with the signal berth offset data required to convert the TD records to timings at stations and other timing point locations (TIPLOCs). In contrast to Lateness data, which directly reports the variations in lateness of (and thus delays to) scheduled services, it is necessary to compare the actual timings in TD records with the corresponding planned values in the timetable data to calculate variations in lateness and thus the delays along each train’s recorded journey.

Improved Dwell and Running Time Results

The refined analytical approach and the additional TD data were applied to some of the stations and sections shown in Tables 6 and 7. West Byfleet is an outer suburban, commuter station on the SWML, close to and on the ‘Up’ (i.e. towards London) side of the major station and junction at Woking, as shown on Figure 2. The Up Fast line between New Malden and Wimbledon carries outer suburban and longer-distance train services towards London, with no trains stopping at New Malden (the platforms on the Fast lines there, the centre tracks in the four-track alignment, are not in use) and very few stopping at Wimbledon, and then only outside peak operating periods. Both locations are sufficiently distant from London to be relatively unaffected by any secondary delays that occur as traffic accumulates on the approaches to Clapham Junction and London Waterloo.

Further analysis of the timetable data indicated that West Byfleet had six distinct scheduled dwell times for services calling at Platform 1 (towards London) and two for services calling at Platform 3 (from London, towards Woking). Platforms 1 and 3 are both on the Slow lines, and Platform 2, on the Down Fast line, is rarely used. Average sub-threshold dwell time delays are plotted against scheduled additional (i.e. in excess of the TPR-specified minimum value) dwell times in Figures 4 and 5, using performance results obtained from both Lateness (shown in blue) and TD (shown in orange) data for the December 2017 timetable (note the differences in the horizontal scales: in Figure 4 the scale ranges from 0 to 6, whereas in Figure 5 it ranges from 0 to 0.6).

In both cases, and for both the Lateness- and TD-based results, it can be seen that average sub-threshold dwell time delays decrease as scheduled dwell times increase, consistent with expectations. The Lateness and TD values for Platform 1, particularly for the lower dwell time values, are quite consistent and ‘closely-packed’, producing relatively high levels of correlation and R^2 values for linear regression of almost 74% and just over 49% respectively (the negative exponential equivalents are 63% and 47% respectively, although these are not strictly directly comparable). The two sets of values have similar intercept values, and the slope of the Lateness-based regression equation is slightly greater than for its TD-based equivalent.

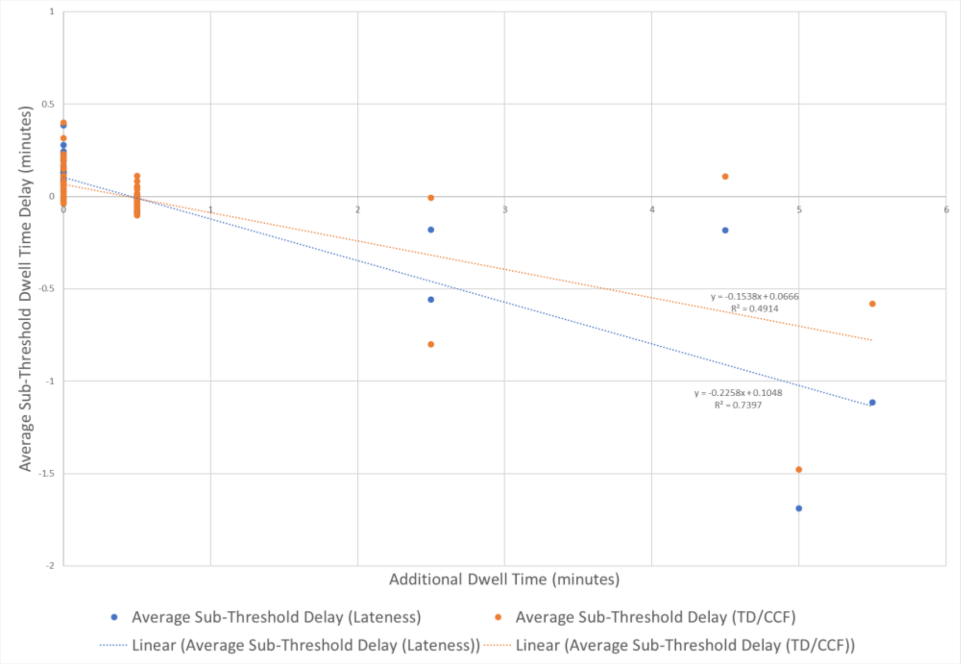

Figure 4: West Byfleet Platform 1 Dwell Time Performance

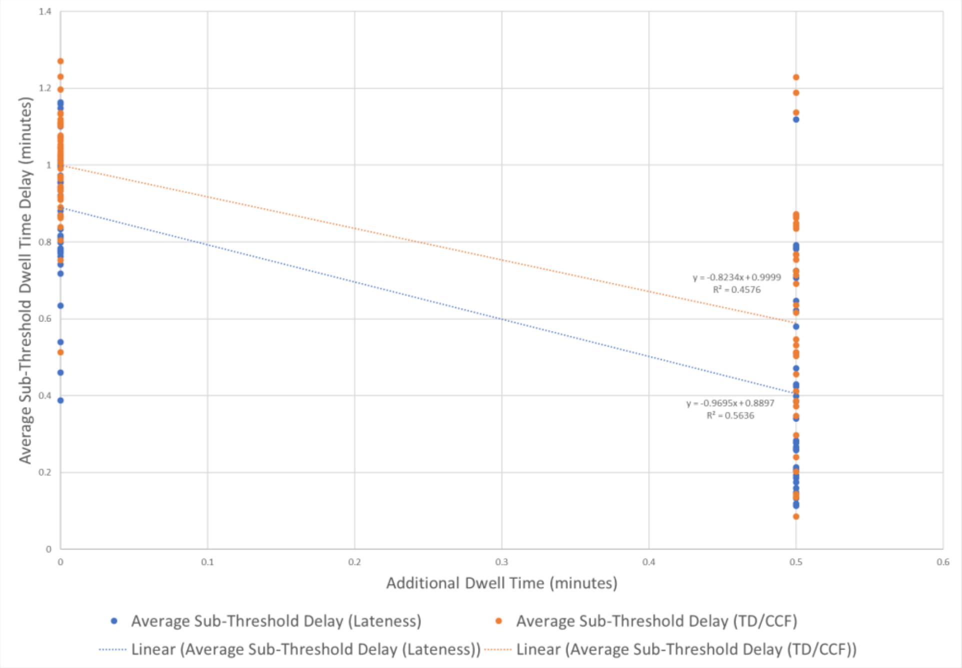

Figure 5: West Byfleet Platform 3 Dwell Time Performance

A greater spread of results can be seen in Figure 5 for both scheduled additional dwell times for Platform 3, and the R^2 values are correspondingly lower, although still reasonable, at 56% and 46% respectively for the Lateness and TD data. The average sub-threshold delay values are again reduced as the planned dwell time increases, but the delays are consistently higher than those for Platform 1. This is somewhat counter-intuitive, given that Platform 1 performance would be expected to be affected by London-bound morning peak traffic, but there may be significant conflicting boarding and alighting passenger movements due to ‘churn’ on the Woking-bound Platform 3, and there may also be a greater operational focus on on-time departures for busy London-bound services in the morning peak. All trains serving Platform 3 are delayed on average, whereas while most Platform 1 services with a minimum scheduled dwell time are delayed on average, the regression results indicate that 30s additional scheduled dwell time reduces average delays to slightly below zero.

In general, the regression equations suggest that TD-based delay values are higher than their Lateness-based equivalents. This is consistent with the truncation of the recorded Lateness values, but the situation is again complicated by the calculation of delay values as the difference between two Lateness/TD values, which, combined with the calculation of average values, is likely to reduce the influence of truncation. An example is provided by a train that arrives in a station on time at 12:30:00 and is scheduled to depart at 12:30:30. If the train departs 29s late, at 12:30:59, truncation will result in a recorded departure time of 12:30:00, indicating an early departure and zero actual dwell time; if it departs at 12:31:00, no truncation will occur, and the late departure will be correctly recorded, but a 1s difference in actual departure time results in a 60s difference in the recorded departure times and any subsequent delay calculations.

Combined plots of average sub-threshold Lateness and TD values (again shown in blue and orange respectively) vs. additional running time values are shown in Figure 6 for the Up Fast line of the SWML between New Malden and Wimbledon.

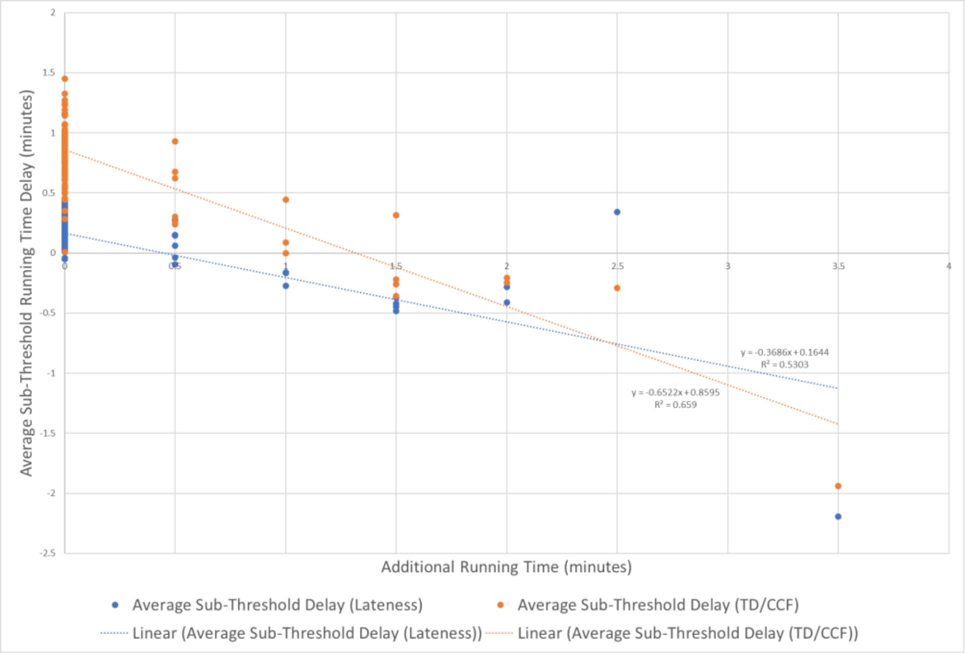

Figure 6: New Malden – Wimbledon Running Time Performance

Both the Lateness and TD datasets provide quite high levels of correlation with average sub-threshold running time delay values, with R^2 values of 53% and 66% respectively (and 49% and 69% respectively for negative exponential regression) – the indicated relationships are again as expected, with average delay declining as planned running times increase. The intercept value for the TD data is over 0.5 minutes larger, but the regression equation has a larger slope, i.e. performance is predicted to improve more rapidly with increases in running time values. The data shows that most trains scheduled at their minimum running times experience some delay; the Lateness data regression indicates that trains should achieve the planned running time on average if 0.5 minute is added, but the TD data suggest that this will still result in an average delay of approximately 0.5 minute, and that almost 1.5 additional minutes of running time are required to achieve punctuality on average.

Longitudinal Analysis: Southampton Central to London Waterloo

The analysis described in the preceding paragraphs provides a 'transverse' view of the relationship(s) between the timetable and performance at a specific location throughout an operating weekday, reflecting the operational characteristics of the station or section in question as well as the timetable there. However, the schedules of most trains on the national network cover multiple stations and sections, with trains' progressive and overall performance reflecting the cumulative operational characteristics and performance effects of the stations and sections that they serve and traverse.

A 'longitudinal' analysis of performance along the route from Southampton Central to London Waterloo was therefore undertaken, using Lateness data (a full TD dataset for the route was not available) and covering longer-distance services with relatively infrequent stops, routed via the Up Fast line of the four-track section of the SWML between Worting Junction, Basingstoke and London Waterloo. The results of transverse analyses of individual stations and sections were combined and ordered sequentially for this purpose, and cumulative historic and predicted lateness values were calculated. Train stopping patterns vary among services on this route, but the dominant service pattern for most of a normal (i.e. pre-Covid) weekday consists of half-hourly semi-fast services calling alternately at Southampton Airport (Parkway), Winchester and Woking, and at Southampton Airport (Parkway), Winchester, Basingstoke and Clapham Junction. These two service patterns were used for the initial longitudinal analysis, and their route includes the Up Fast line between New Malden, Wimbledon and Clapham Junction, which was the subject of previous analysis, as described above. Historic cumulative lateness values, based on sub-threshold delay data for the December 2017 and May 2018 timetables, are shown for both stopping patterns in Figure 7. The blue line represents the services that call at Woking, and the red line represents the alternate trains that call at Basingstoke and Clapham Junction. The two lines are almost identical from Southampton Central to Worting Junction (WRTINGJ), between which there is no difference in the stopping patterns, but diverge considerably thereafter.

Consistent with some of the WP1 findings described above, lateness recovery is particularly evident for dwells at Woking and Basingstoke (BSNGSTK), and to a lesser extent at Clapham Junction (CLPHMJM), all of which have minimum scheduled dwell times of 1.5 minutes. The effects of allowance sizes and locations are also apparent: trains stopping at Woking have a total of five minutes additional dwell and running time after Worting Junction, and terminate at Waterloo just over 0.5 minute late on average, whereas the Basingstoke/Clapham stopping services have 2.5 minutes of additional allowances, and terminate almost two minutes late on average. The allowances for the Basingstoke/Clapham trains are provided between Basingstoke and Hampton Court Junction, towards the centre of the graph, where their average cumulative lateness can be seen to decline to below zero on approach to Woking, and to increase steadily thereafter, apart from some recovery at Clapham Junction. For the trains stopping at Woking, the allowances are provided at Woking and between Wimbledon and Waterloo, with three minutes provided between Clapham Junction and Waterloo, reducing the terminating lateness value considerably.

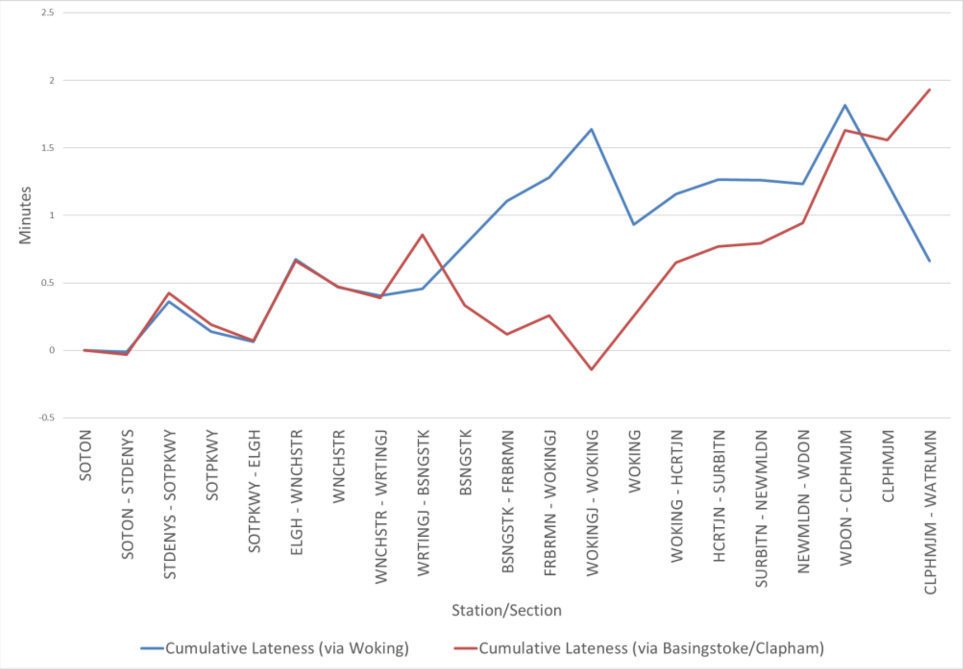

Figure 7: Southampton Central – Waterloo Cumulative Lateness

Although these services are compliant with the TPRs, the graphs indicate that they are subject to considerable sub-threshold delays and lateness accumulation and dissipation, and that they consistently terminate late, especially those that stop at Basingstoke and Clapham Junction. Despite this, their performance in terms of sub-threshold delays (i.e. ignoring the effects of larger delays) and traditional performance metrics is quite good, with no cumulative lateness values over two minutes, and trains easily achieving their PPM target (for longer-distance services) of arrival at their terminus within 10 minutes of schedule. By the standards of the newer ‘on time at all stations’ (i.e. within one minute of the scheduled time, and less than one minute late, at all stations) metric (Network Rail, 2020b), the Basingstoke/Clapham services depart from Clapham Junction and terminate at Waterloo more than 1.5 minutes late on average, while the Woking services are technically on time there, departing just under one minute late on average. The recorded increases in lateness between New Malden and Wimbledon are consistent with the results shown in Figure 6.

Performance Prediction: Worting Junction to London Waterloo

One of the primary objectives of the DDRT project was to enable the prediction of performance using historic timetable and performance data and the relationships between them, and particularly on the basis of the relationships between sub-threshold delays and additional or ‘net’ allowances in excess of the TPR-specified minimum values, of the type shown in the regression equations in Figures 4 to 6 above.

Equivalent ‘transverse’ analyses were therefore undertaken of all stations and sections on the Up Fast line between Worting Junction and London Waterloo (re-using previous analyses where available), using the available Lateness data. Given the relatively high levels of correlation previously obtained, linear regression was again used to identify relationships between additional dwell and running times and performance where possible; in the case where only a single scheduled dwell or running time was available (i.e. Basingstoke – Farnborough, Start to Pass), the mean associated sub-threshold delay value was obtained. The results (again in minutes) of these additional analyses are summarised for the December 2017 data in Table 8, together with the t-statistics for the intercept and slope values, which indicate that the results are all statistically significant.

Table 8: Relationships between Additional Allowances and Sub-Threshold Delays

Section/Station	**Movement**	**Intercept**	**t-Stat**	**Slope**	**t-Stat**	**Mean**
Worting Jn. – Basingstoke	Pass to Pass	0.339	8.643	-0.524	-9.939	
	Pass to Stop	0.558	25.833	-0.407	-3.513	
Basingstoke	Dwell	-0.108	-2.371	-0.482	-9.513	
Basingstoke – Farnborough	Pass to Pass	0.718	22.162	-0.582	-7.923	
	Start to Pass					0.094
Farnborough – Woking Jn.	Pass to Pass	0.173	7.298	-0.487	-4.072	
Woking Jn. – Woking	Pass to Pass	0.143	4.266	-0.475	-9.171	
	Pass to Stop	0.450	26.698	-0.525	-12.394	
Woking	Dwell	-0.451	-7.502	-0.206	-1.812	
Woking – Hampton Ct. Jn.	Pass to Pass	0.743	10.143	-0.362	-4.435	
	Start to Pass	0.230	15.935	-0.195	-11.827	
Hampton Ct. Jn. – Surbiton	Pass to Pass	0.088	10.395	-0.158	-3.308	
Surbiton – New Malden	Pass to Pass	0.029	2.767	-0.508	-24.253	
New Malden – Wimbledon	Pass to Pass	0.165	18.614	-0.369	-16.716	
Wimbledon – Clapham Jn.	Pass to Pass	0.862	22.012	-0.266	-10.306	
	Pass to Stop	0.754	30.048	-0.382	-7.128	
Clapham Jn.	Dwell	-0.148	-9.991	-1.231	-3.870	
Clapham Jn. – Waterloo	Pass to Stop	0.848	14.499	-0.523	-18.509	
	Start to Stop	0.640	11.911	-0.455	-16.840	

The regression relationships and average values obtained from the December 2017 data were then used to predict the cumulative lateness development of two trains from the May 2018 timetable: 1W20, calling at Woking; and 1W60, calling at Basingstoke and Clapham Junction. The results are shown in Figures 8 and 9 respectively, with the dashed lines representing the predicted cumulative delay values.

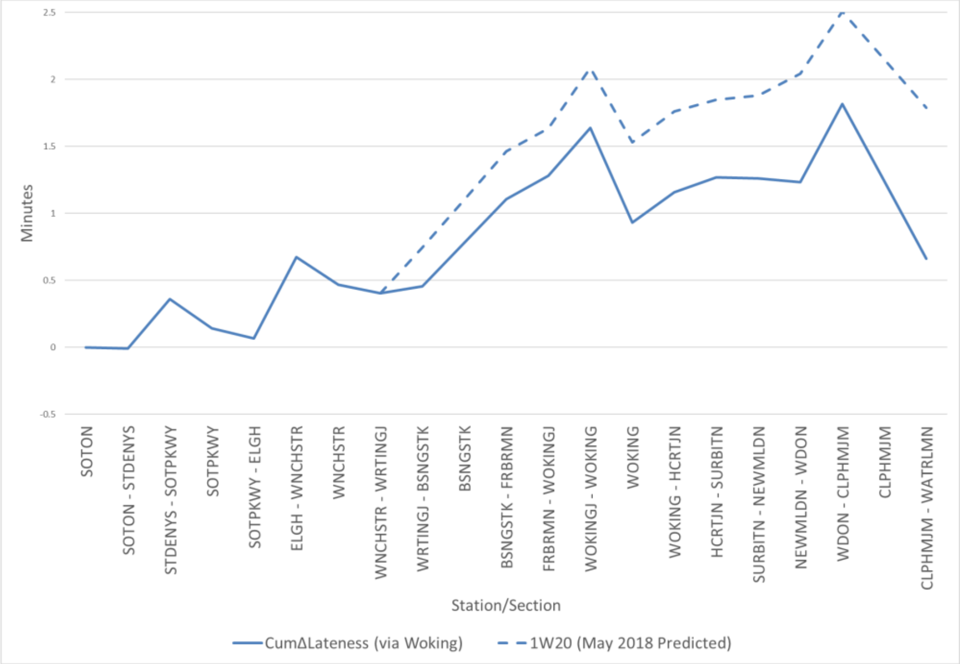

Figure 8: Predicted Cumulative Lateness for 1W20

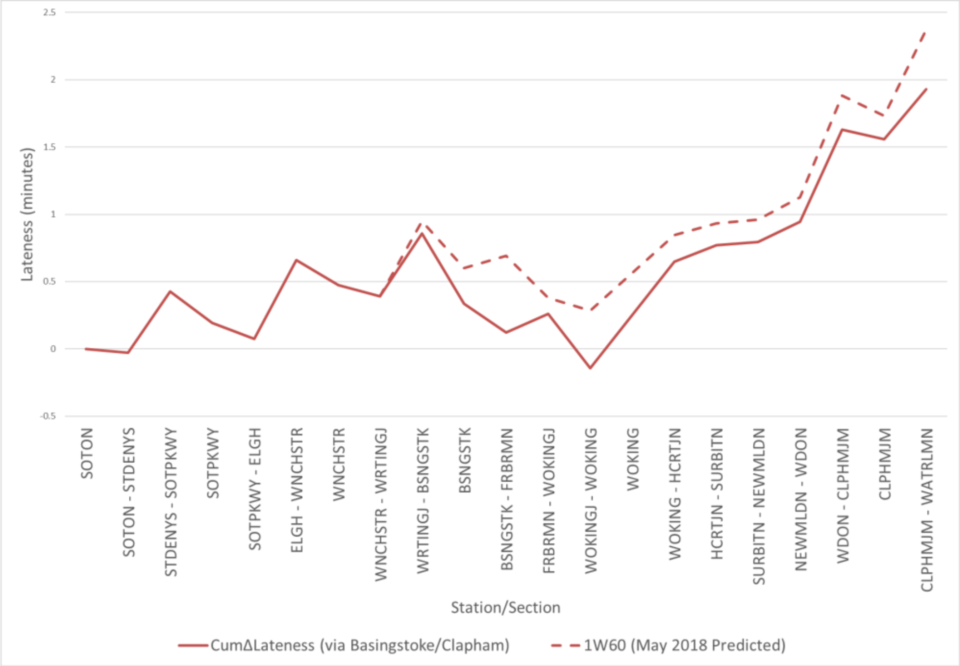

Figure 9: Predicted Cumulative Lateness for 1W60

For 1W20, the delay between Worting Junction and Basingstoke is over-predicted by approximately 30s, and this is carried through the rest of the journey, with the cumulative level of over-prediction increasing slightly, so that it is just over one minute at the terminus at Waterloo. However, the general shape of the prediction curve otherwise matches the cumulative delay pattern for the Woking trains quite well. For 1W60, there are also some deviations along the route, but the shape of the predicted curve again matches the historical data quite well, and deviates by less than 15s at Clapham Junction and by approximately 30s on termination at Waterloo. The over-prediction is likely to be due in part to the fact that the regression analysis includes all trains, peak and off-peak, and so may over-predict the delays incurred by off-peak services.

Further analysis of 1W60 was undertaken, comparing the predicted cumulative lateness with the actual, recorded values for the service for the duration of the May 2018 timetable, and also with the recorded performance data for a similar, later train, 1W66, scheduled to arrive in Waterloo during the evening peak, when the system is busier and trains are likely to be subject to additional delay. The results are shown in Figure 6, with the dashed red line representing the predicted performance for 1W60, the solid red line representing its recorded performance, and the green line representing the recorded performance of 1W66 across the December 2017 and May 2018 timetables.

Figure 10: Predicted and Recorded Cumulative Lateness for 1W60 and Recorded Cumulative Lateness for 1W66

It can be seen that the cumulative lateness of 1W60 is over-predicted between Basingstoke and Woking, but that the model provides a very accurate prediction between Woking and Clapham Junction, and then over-predicts terminating lateness by approximately one minute at Waterloo, again probably reflecting the inclusion of peak period performance data in the prediction model. Comparing the predicted lateness curve for 1W60 with the actual lateness curve for 1W66, it can be seen that the predicted interpeak lateness is generally less than the actual evening peak lateness, as would be expected (the reduction in lateness for 1W66 between Clapham Junction and Waterloo is due to an additional three minutes included in the train's scheduled running time over that section of the route, reflecting the high levels of traffic and potential for delay during the peak). These results mark a significant step towards meeting the project objectives, but further work is clearly needed to improve the accuracy and extend the coverage of the outputs.

Additional Running Time Analysis: Raynes Park to Wimbledon

In contrast to the relatively 'clean' results seen in Figures 4 and 6 (and similarly to the dwell time results shown in Figure 5), stopping train services are subject to a wide range of delay effects as they converge towards their termini, and the secondary effects of preceding headway and junction margin values become increasingly influential. This can be seen in Figure 11, showing the relationships (or the apparent lack thereof) between preceding headways/junction margins and running time delays (obtained from Lateness data only) for stopping services on the Up Slow line from Raynes Park to Wimbledon (there is a converging junction at the 'London end' of Raynes Park station, meaning that trains departing from the station are affected by both headways and junction margins between them and preceding services).

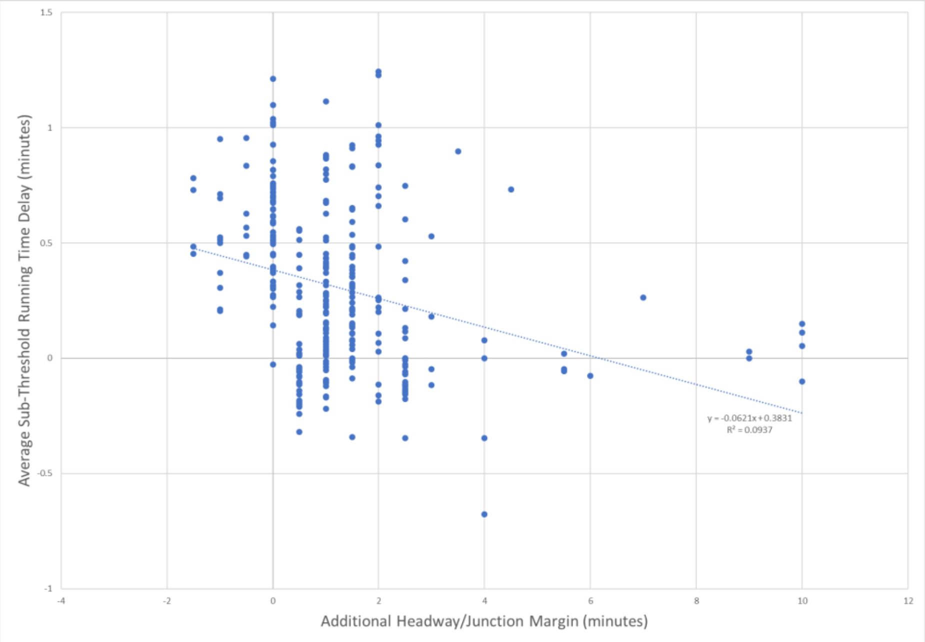

Figure 11: Raynes Park – Wimbledon Margins and Headways

While there is an overall trend of decreasing delay with increasing margins and headways, the correlation is quite low (R^2 = 9.4% for linear regression, and 10.0% for negative exponential), and there is considerable variation within each margin/headway 'band', illustrating the complexity of the influences on performance, and requiring further investigation. It can, however, be seen that all trains with 'negative additional margins' (trains operating at less than the minimum TPR-specified junction margin) are subject to delay, as shown in the top-left quadrant.

5 Ongoing and Further Work

Work is ongoing to identify, investigate and improve the understanding of the ‘micro’ influences on dwell and running time performance, producing ‘scatter’ of the type illustrated in Figures 5, 10 and, to a lesser extent, 6. In parallel with and beyond this, there is also a need to further automate and scale up the processes used, increasing the speed and scope of analysis as well as the accuracy and quality of the results. This should include the use of Big Data analysis techniques at the network level to identify locations prone to sub-threshold delays and which also have significant variations in scheduled running and dwell times and headways and/or junction margins. These should be complemented by the use of machine learning tools to analyse and identify the detailed relationships between timetable characteristics and performance at individual locations.

Preliminary analysis of GWR timetable data between December 2018 and March 2020 indicates a good range of dwell and running time values for stations and sections between London and Bristol, suggesting that it should be suitable for developing the relationships needed to quantify and predict timetable-related performance, and thus for extending, validating and improving upon the work presented in this paper.

6 Conclusions

Performance on Britain’s railways has improved with the reduction in traffic levels since the onset of Covid-19. However, as and when traffic levels increase again (as they should, if the wider priority of modal shift from road and air to rail is to be met), performance problems are likely to re-emerge in the absence of actions to improve punctuality and reliability.

The high-level analysis undertaken in WP1 of the DDRT project indicates consistent historic exceedances of scheduled dwell and running times. Although the major pre-Covid December 2019 timetable change on GWR improved dwell time performance, this was at the expense of deteriorations in running time performance, contrary to the approach where line sections are used as zones of compensation and stations/junctions are used as zones of concentration, and similar findings emerged from WP3 for SWR operations.

The detailed analysis undertaken in WP3 demonstrates an approach to identifying the historic relationships between timetable characteristics and dwell and running time performance, and its findings support the project hypotheses. The application of these findings to the assessment and prediction of overall train service performance is also demonstrated, providing the basis for a means and method of scoring timetable quality in terms of predicted lateness and delays. Significant further work is required to complete this, in terms both of understanding the details of secondary delay causation and transmission, and of scaling up the overall process for validation and wider application.

Acknowledgements

Stephan Zieger’s contributions to the work described above were funded by the Deutsche Forschungsgemeinschaft (German Research Foundation) – 2236/1 as well as Erasmus+ STT and Erasmus+ Traineeship. The authors are grateful to RSSB and Network Rail for project coordination, funding and performance data provision, to Bellvedi for the provision of software, and to Phil Dargue of Winder Phillips Associates for his support for the project.

References

Armstrong, J., Raine, R., Draper, S., Preston, J., 2019. "Developing Objective Timetable Quality Metrics", In: *Proceedings of the 12th World Congress on Railway Research (WCRR2019)*, Tokyo, Japan.

Armstrong, J., Zieger, S., Preston, J., Nießen, N., 2019. "Identifying and Reducing Performance Uncertainty in UK Railway Timetables", In: *Proceedings of International Railway Symposium Aachen 2019 (IRSA 2019)*, Aachen, Germany.

Bešinović, N. (2020). "Resilience in railway transport systems: a literature review and research agenda", *Transport Reviews*, vol. 40:4, pp. 457-478. DOI: https://doi.org/10.1080/01441647.2020.1728419

Cacchiani, V., Toth, P., 2018. "Robust Train Timetabling", In: *Handbook of Optimization in the Railway Industry* (Borndörfer, R., Klug, T., Lamorgese, L., Mannino, C., Reuther, M., Schlechte. T., eds.). Cham: Springer.

Caimi, G., Fuchsberger, M., Burkolter, D., Herrmann, T., Wüst, R. and Roos, S., 2009. "Conflict-free train scheduling in a compensation zone exploiting the speed profile", In: *Proceedings of the Third International Seminar on Railway Operations Research, (ISROR RailZurich2009)*, Zurich, Switzerland.

Goverde, R.M.P., 2014. "Timetable Stability Analysis", In: *Railway Timetabling & Operations* (Hansen, I.A., Pachl, J., eds.). Hamburg: Eurailpress.

Goverde, R.M.P., Hansen, I., 2013. "Performance Indicators for Railway Timetables", In: *Proceedings of the IEEE International Conference on Intelligent Rail Transportation (IEEE ICIRT 2013)*, Beijing, China.

GWR, 2020. Network and Accessibility Map – Great Western Railway [online]. Available from https://www.gwr.com/~/media/gwr/pdfs/maps/accessibility-network-map-poster-june-2020.pdf [Accessed 25 November 2020]

Network Rail, 2019. The central importance of the operating function [online]. Available from https://www.networkrail.co.uk/wp-content/uploads/2019/05/Andrew-Haines-speech-the-central-importance-of-the-operating-function-25-1-19.pdf [Accessed 26 November 2020]

Network Rail, 2020a. Operational Rules (zip folder download – 144 MB) [online]. Available from https://www.networkrail.co.uk/wp-content/uploads/2020/10/Operational-Rules.zip [Accessed 30 November 2020]

Network Rail, 2020b. Railway Performance [online]. Available from https://www.networkrail.co.uk/who-we-are/how-we-work/performance/railway-performance/ [Accessed: 8 December 2020]

Nießen, N., 2014. "Queuing", In: *Railway Timetabling & Operations* (Hansen, I.A., Pachl, J., eds.). Hamburg: Eurailpress.

ORR, 2020a. Passenger Rail Performance 2020-21 Quarter 1 [online]. Available from https://dataportal.orr.gov.uk/media/1808/passenger-performance-2020-21-q1.pdf [Accessed 17 November 2020]

ORR, 2020b. Passenger Rail Performance 2020-21 Quarter 1 [online]. Available from https://dataportal.orr.gov.uk/statistics/usage/passenger-rail-usage/table-1243-passenger-train-kilometres-by-operator/ [Accessed 11 March 2021]

Pachl, J., 2018. Railway Operation and Control (4th ed.). Mountlake Terrace: VTD Rail Publishing.

SWR, 2017-2020. Download our network map [online]. Available from https://www.southwesternrailway.com/~/media/Images/InteractiveMap/NetworkMap.ashx [Accessed 25 November 2020]