Optimal Real-time Line Scheduling for Trains with Connected Driver Advice Systems

Ajini Galapitage [a,1], Amie R. Albrecht [a], Peter Pudney [a,b], Peng Zhou [a]
[a] Scheduling & Control Group, University of South Australia
[b] Future Industries Institute, University of South Australia
[1] E-mail: ajini.galapitage@mymail.unisa.edu.au

Abstract
On most rail networks, if a train is delayed then following trains will not know about the delay until they encounter a trackside signal that tells the driver that the next section of track is still occupied. The train will usually have to slow significantly, which causes delays to propagate back along the track. By using in-cab Driver Advice Systems connected to centralised scheduling systems, train delays can be detected as they happen, and new schedules can be calculated and issued to following trains so that additional delays are avoided. It is impossible to re-schedule the whole rail network at once in real time as the problem is too large. An alternative, more practical approach is micro-scheduling to independently optimise small sections of the network.
We describe and illustrate a method that can be used to ensure adequate and energy-efficient train separation. The method can be used during timetable planning to ensure robust timetables or can be used in real time to prevent trains from encountering restrictive signals, smoothing the flow of trains along a corridor.

Keywords
Optimal train control, dynamic rescheduling, line scheduling

1 Introduction

Energy-efficient driving strategies are often disrupted by train separation constraints, particularly when there are short time headways between trains and when some trains are delayed. When a train encounters a restrictive signal it will usually have to slow significantly, which disrupts efficient driving and introduces delays that can propagate back through the network.

Driver Advice Systems (DAS) can help trains follow a schedule precisely, and save energy at the same time [Scheepmaker et al., 2017, Panou et al., 2013, Albrecht et al., 2016a,b]. Connected Driver Advice Systems (C-DAS) extend this capability by adding communication with a central control system, which can provide real-time updates to individual train schedules in response to disruptions on the network.

Previous work has described how C-DAS can be used in real time to smooth the flow of trains through junctions, by adjusting the target arrival times of trains approaching the junction to avoid conflicts [Galapitage et al., 2018, Chen et al., 2015]. The on-board DAS ensures that the revised targets are achieved.

Luan et al. [2018a,b] discuss the integration of real-time traffic management and train control. Part 1 gives a good overview of various approaches, and develops mixed integer programming solutions to the problem of determining optimal sequences, routes and arrival times for trains. Part 2 discusses optimal scheduling and energy efficiency, but use simplified speed profiles and assume constant gradient, curve and speed limits on each block section.

In this paper, we show how measurements of train movements can be used to identify locations and times where trains are delayed along a line without junctions, and we use examples from a long-haul freight line and from an intercity passenger line to show how small adjustments to train schedules can be used to ensure safe separation of trains while minimising energy use.

2 Measuring Train Delays

Many railways in the UK use the Energymiser[1] driver advice system, developed by Australian company TTG Transportation Technology and based on train control methods and software developed by the Scheduling and Control Group at the University of South Australia [Albrecht et al., 2016a,b]. As well as giving train drivers advice on how to drive efficiently, these units collect data that includes the position and speed of each train at 10-second intervals. This data can be used to analyse the performance of a railway. In this section we use journey logs from Chiltern Railways to investigate delays on the rail network. In particular, we use data collected on trains travelling from Princes Risborough to London Marylebone via High Wycombe during August 2016. Figure 1 shows two of the Chiltern routes to the west of London Marylebone. Princes Risborough is three stations south of Aylesbury, just south of a junction where trains from London can either head north to Aylesbury or continue west.

Our data from August 2016 includes 2172 "up" trips from Princes Risborough to London Marylebone via High Wycombe. Figure 2 shows the measured speed profiles of trains for the "up" direction, highlighting both the different stopping patterns and the considerable variations in speed.

2.1 Variation in Section Durations

Energymiser journey logs can be used to determine how long it took trains to drive between stops, and how much variation there was in these section durations.

Table 1 shows the durations of stop-to-stop journey sections for the measured train journeys. The columns are:

- the origin of the trip section
- the destination of the trip section, which is not necessarily the next station along the route
- the number of trips that did this section
- the median section duration, in seconds; half the trips had a section duration less than this value, and half had a section duration greater than this value
- the first quartile section duration, in seconds; one quarter of the trips had a section duration less than this value

[1] http://www.ttgtransportationtechnology.com/energymiser

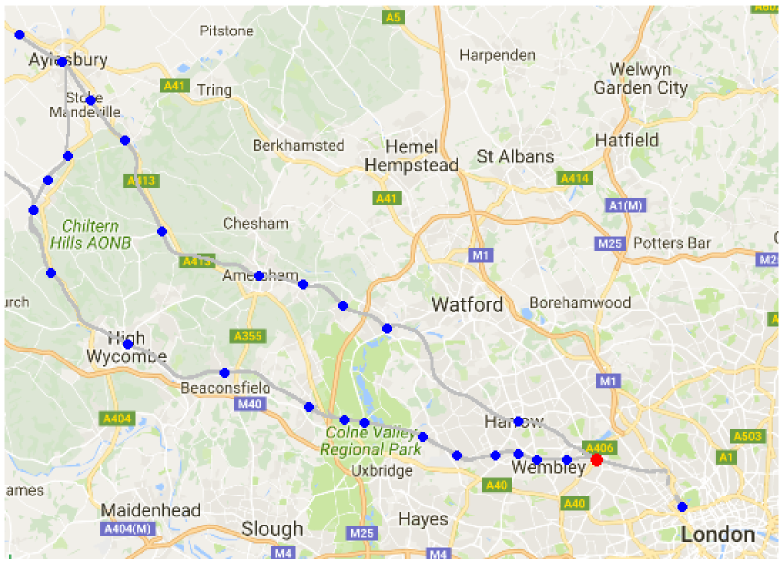

Figure 1: The Chiltern rail network west of London Marylebone. The background map is from Google.

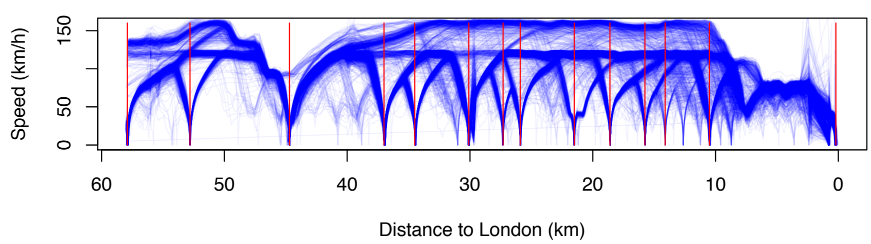

Figure 2: Speed profiles of trains travelling from Princes Risborough to London Marylebone via High Wycombe, August 2016.

- the third quartile section duration, in seconds; one quarter of the trips had a section duration greater than this value
- the inter-quartile range (IQR), in seconds; this is the difference between the third quartile value and the first quartile value, and is a measure of the variation in section durations.

Sections with wide variation, where the duration IQR is more than 10% of the median section duration, are indicated with a '*'.

Some of the variation in section running durations could be due to variations in driving advice provided by the Energymiser units. However, many trains run late; for these trains, the advice is to drive as quickly as possible and so the variations are due to other factors.

2.2 Slowing Between Scheduled Stops

We are particularly interested in times and locations where trains are slowed between stops by the signalling systems. We did not have access to signalling data, but if a train slows significantly between scheduled stops then it is almost certainly because of traffic issues.

Figures 3 and 4 show times and locations where trains travelling in the "up" direction slowed to less than 40 km/h. The darker dots indicate speeds less than 20 km/h. The horizontal black lines represent train station locations. We expect trains to travel slowly when arriving at a stop and departing from a stop, but low speeds away from stops indicate a traffic problem.

The delays at 8 km are near Neasden Junction; trains travelling in the "up" direction through the junction are often delayed by train movements on other paths through the junction. Galapitage et al. [2018] describe a method for real-time rescheduling of trains at junctions.

There were several other sections where trains slowed to less than 40 km/h. There are no junctions on these route sections.

3 Line Scheduling

A train following another train along a track will be delayed if it gets too close to the leading train. Once a train has encountered a restrictive signal, it will have to slow; this can introduce further delays on the corridor.

In this section we describe how small adjustments to individual train schedules can be used to ensure adequate separation between trains to avoid encounters with restrictive signals. The method can be used in the timetable planning stage to develop robust timetables, or in real time to ensure smooth running on a corridor.

We will illustrate the method using four simulated but realistic examples.

3.1 Example 1: Long-haul Freight

The Dedicated Fast Freight Corporation in India is building two new rail corridors, in the east and west of the country. These corridors will each carry a mix of freight train types over long distances, with headways between trains as low as six minutes. Crew change locations are fixed on each corridor. To maximise capacity, the running time between any given pair of adjacent crew change locations will be the same for every train. However, differences

Table 1: Section durations for trains travelling from Princes Risborough to London Marylebone via High Wycombe, August 2016.

			section duration				
origin	destination	trips	median	Q1	Q3	IQR	
Princes Risborough	Saunderton	433	203	197	211	14	
Princes Risborough	High Wycombe	466	454	436	465	30	
Saunderton	High Wycombe	430	279	273	284	11	
High Wycombe	Beaconsfield	1207	270	262	281	19	
High Wycombe	Seer Green & Jordans	25	338	330	347	17	
High Wycombe	Gerrards Cross	72	490	477	504	28	
High Wycombe	West Ruislip	11	745	728	774	46	
High Wycombe	Wembley Stadium	75	943	884	1066	182	*
High Wycombe	London Marylebone	165	1545	1440	1679	239	*
Beaconsfield	Seer Green & Jordans	524	100	96	106	10	*
Beaconsfield	Gerrards Cross	644	247	241	254	13	
Beaconsfield	London Marylebone	16	1247	1174	1347	173	*
Seer Green & Jordans	Gerrards Cross	547	162	157	169	12	
Gerrards Cross	Denham Golf Club	197	99	93	102	9	
Gerrards Cross	Denham	208	145	139	149	10	
Gerrards Cross	West Ruislip	86	289	280	302	22	
Gerrards Cross	South Ruislip	223	362	354	369	16	
Gerrards Cross	Wembley Stadium	166	613	604	628	24	
Gerrards Cross	London Marylebone	221	1180	1119	1257	138	*
Denham Golf Club	Denham	196	51	47	54	7	*
Denham	West Ruislip	120	179	174	184	10	
Denham	South Ruislip	173	251	247	257	10	
Denham	Wembley Stadium	64	502	495	510	15	
Denham	London Marylebone	28	1041	994	1062	68	
West Ruislip	South Ruislip	68	145	140	150	11	
West Ruislip	Northolt Park	63	241	233	247	15	
West Ruislip	Sudbury Hill Harrow	13	297	295	315	20	
West Ruislip	Wembley Stadium	63	405	392	417	26	
South Ruislip	Northolt Park	144	119	116	123	7	
South Ruislip	Wembley Stadium	220	297	290	309	19	
South Ruislip	London Marylebone	49	841	817	871	54	
Northolt Park	Sudbury Hill Harrow	24	69	67	73	6	
Northolt Park	Wembley Stadium	154	198	191	207	16	
Northolt Park	London Marylebone	11	705	699	816	117	*
Sudbury Hill Harrow	Wembley Stadium	12	137	136	148	12	
Sudbury Hill Harrow	London Marylebone	12	841	814	897	83	
Wembley Stadium	London Marylebone	409	563	523	610	87	*

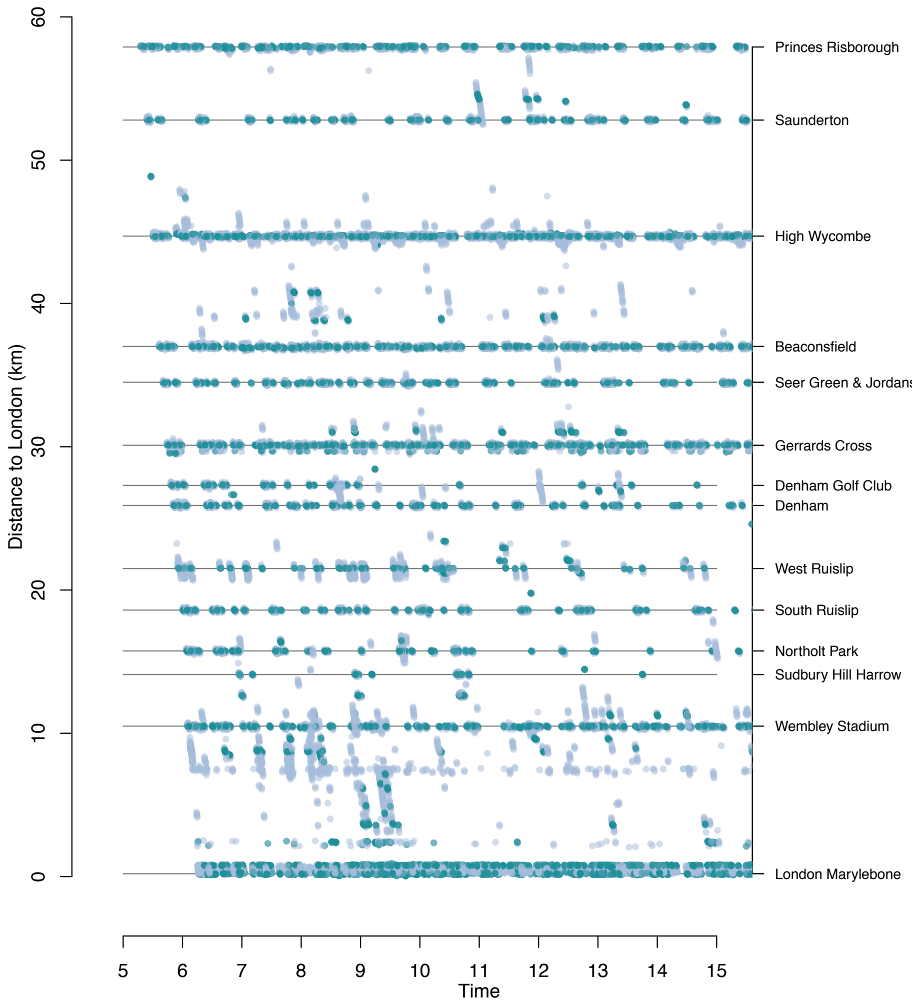

Figure 3: Times and locations where up trains slowed to less than 40 km/h, between 05:00 and 15:00. The darker dots indicate speeds less than 20 km/h.

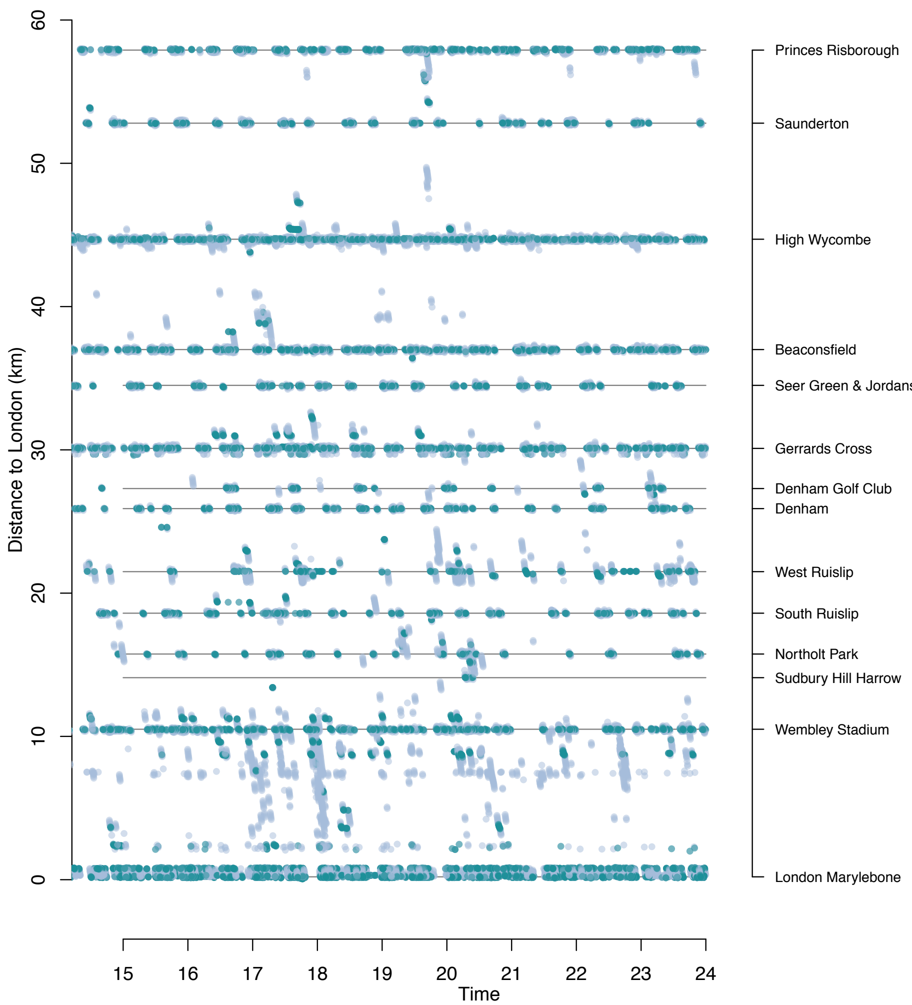

Figure 4: Times and locations where up trains slowed to less than 40 km/h, between 15:00 and 00:00. The darker dots indicate speeds less than 20 km/h.

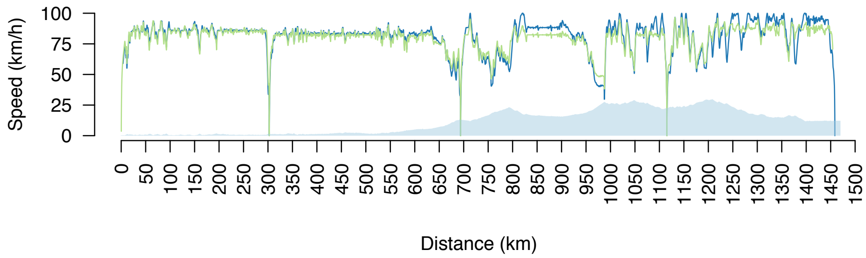

Figure 5: Speed profiles for a bulk train (blue) and a container train (green).

in train performance will mean that the separation between trains will vary as trains drive between crew change locations. For example, a heavy train may be slowed on hills more than a following light train.

The track will have three-aspect signalling with a spacing of 1.5 km. This means that a following train must be at least 3 km behind a leading train, otherwise it will encounter a yellow signal and have to slow.

The separation between two trains traveling along a line depends on the relative speeds of the two trains, which in turn depends on the locomotive performance and trailing load of each of the trains, and on the gradients, curves and speed limits. Figure 5 shows energy-efficient speed profiles for a bulk train (blue) with a trailing mass of 6500 tonnes and for a container train (green) with a trailing mass of 4500 tonnes. The optimal speed profiles were calculated using our Energymiser software. The shaded region at the bottom of the graph indicates the elevation profile of the track. Both trains have the same section running time for each of the four sections of the trip, but the speed profiles are different because of the different train characteristics. For example, the laden bulk train is slowed more by the hills near 985 km than the lighter container train, and so has to travel faster elsewhere in the journey to make up time.

Trains will normally follow each other with a headway of six minutes at each stop. Figure 6 shows the two journey paths with time on the horizontal axis. The container train starts six minutes behind the bulk train.

Figure 8 shows speed profiles $v_1(t)$ and $v_2(t)$ for the bulk train and container train on the third section of the route. Each speed profile has been optimised independently to meet the overall section duration of 5H45M with minimum energy. The heavier bulk train is slowed more by the hills than the lighter container train.

The distance between the two trains at any time is the train *separation*. Figure 7 shows separation as a function of time. We can see from Figure 7 that the trains are too close near times 03:53, 09:01, 10:45, 13:16 and 15:01. The low separation near times 03:53, 09:01 and 15:01 occur because the leading train is stopping for crew changes, and the following train catches up while the leading train is slowing to a stop. In these situations we can allow the following train to get close because it is also going to stop at these locations, and there is space at the crew change locations for more than one train. We are more interested in low separations that occur between stops, at times 10:45 and 13:16.

One way to prevent a following train from getting too close to a leading train is to insert timing points for the following train that will slow it at certain places on the track. The lowest separation occurs at time 10:45, where the distance between the trains is 1438 m.

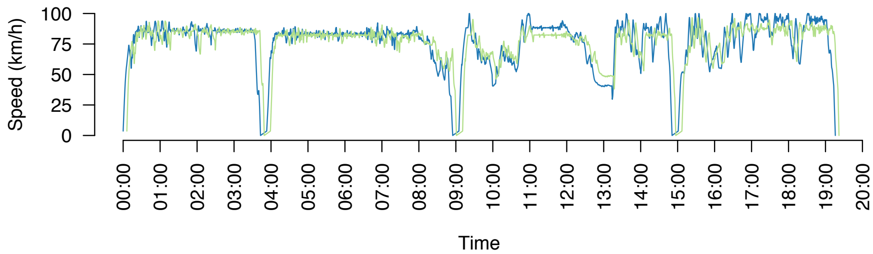

Figure 6: Speed against time for the bulk train and the container train.

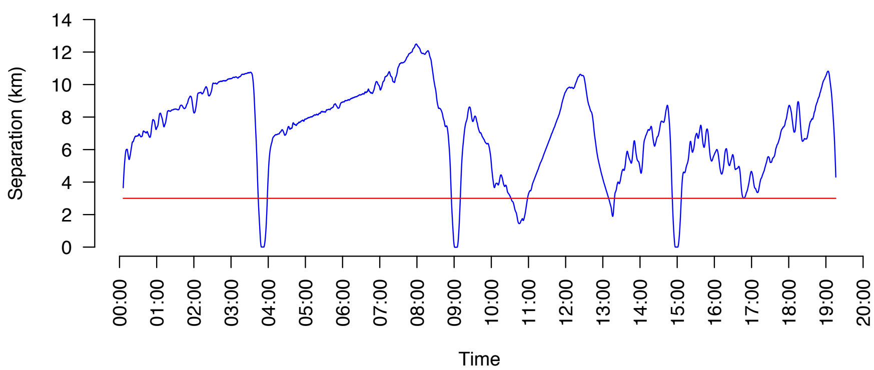

Figure 7: Separation between the two freight trains.

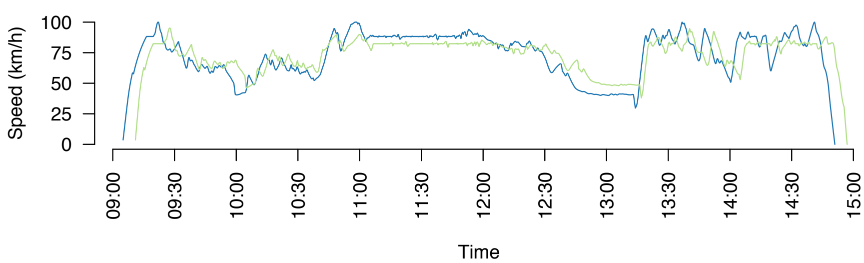

Figure 8: Speed profiles of the two freight trains on the third journey section.

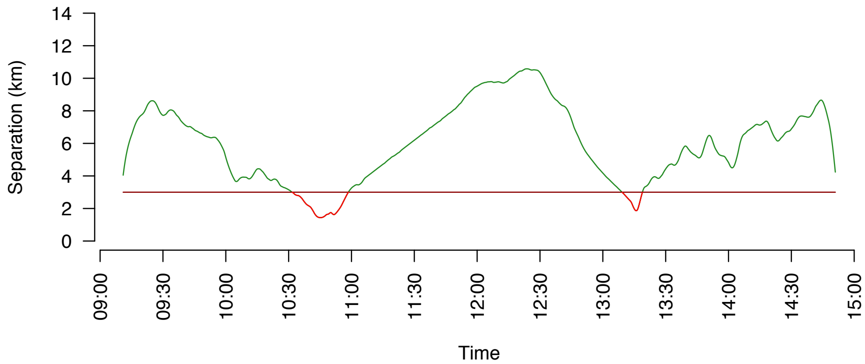

Figure 9: Separation between the two freight trains on the third journey section.

The minimum separation required is 3 km. Let $x_1(t)$ and $x_2(t)$ be the locations of the leading and following trains at time t. The separation between the trains at time t is $\delta(t) = x_1(t) - x_2(t)$, and the rate of change of this separation is $\delta'(t) = v_1(t) - v_2(t)$. Figure 9 shows the separation δ. The green regions are where the separation between the trains is less than 3 km.

One approach to resolving the separation violations is to place a timing constraint at times with minimum separation. In each of the regions with low separation, we search for the time $\tau = \arg\min_t \delta(t)$ at which the separation between the two trains is minimum, then set a timing constraint

$$t_2^*(x_1(\tau) - h) \geq \tau$$

for the rescheduled Train 2, where h is the minimum allowable distance between the trains. This constraint ensures that, at time τ, Train 2 will not have passed the location that is distance h behind the location of Train 1. The path of the rescheduled Train 2 is described by the distance profile x_2^* and the speed profile v_2^*.

The original train paths have minimum separation of 1.438 km at time τ = 10:45:10, with zero derivative. The new profile for Train 2 has $x_2^*(\tau) = x_1(\tau) - h$, because the timing constraint is active, and $v_2^*(\tau) \leq v_2(\tau)$, because Train 2 is now travelling slower at time τ. The new separation is

$$\delta^*(\tau) = x_1(\tau) - x_2^*(\tau) = h$$

with

$$\delta^{*'}(\tau) = v_1(\tau) - v_2^*(\tau) \geq v_1(\tau) - v_2(\tau) = 0$$

and so $\delta^*(\tau - \epsilon) \leq h$ for small ϵ; that is, the new separation is slightly less than h immediately prior to time τ. Figure 10 shows more detail around time τ.

The separation constraint is still violated after introducing a single timing constraint. To resolve this, instead of adding one timing point at the minimum separation point we can add timing constraints throughout the journey. In practice, we add timing constraints at closely spaced discrete points in regions where the minimum separation dips below 3 km. We use

$$t_2^*(x_1(k\Delta t) - h) \geq k\Delta t, \qquad k \in \{0, 1, \dots\},\ \delta(k\Delta t) < h$$

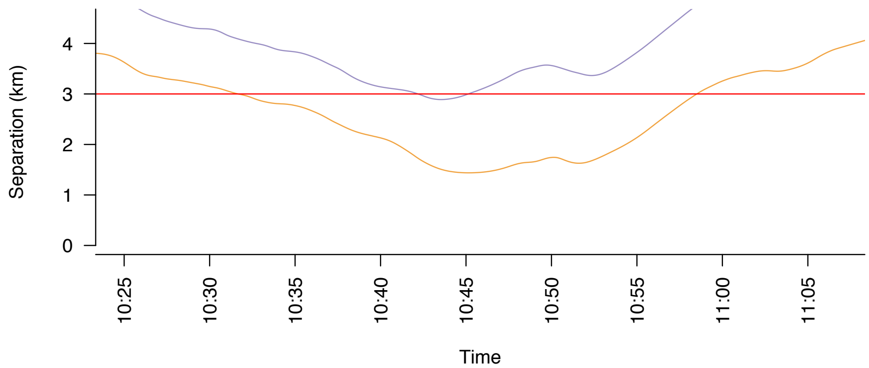

Figure 10: Detail of the original freight train separation δ (orange), and separation after rescheduling the second train (purple), around time τ.

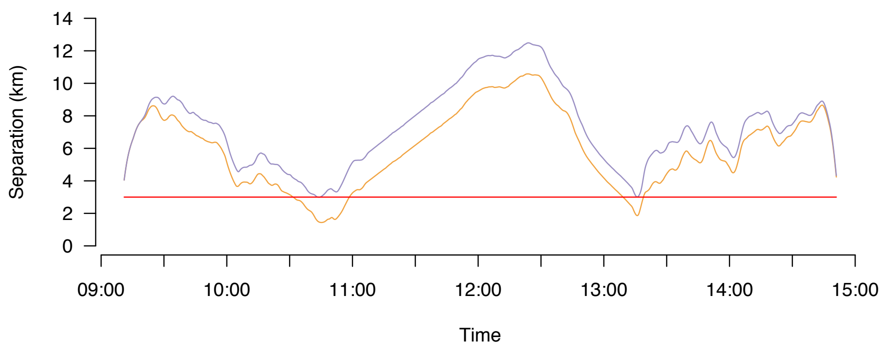

Figure 11: The original freight train separation δ (orange), and separation after rescheduling the second train (purple).

where Δt is a time step chosen to suit the problem; in this case, we used $\Delta t = 25$ seconds. Figure 11 shows the original separation δ (orange), and separation after rescheduling the second train (purple) to increase the separation throughout both regions where the separation drops below 3 km.

The rescheduled container train is still able to finish its journey on time. We expect it to use more energy, since its path has been constrained. In this case, the extra energy use is negligible—just 0.0133% more than without the extra timing constraint.

The particular example does not require a trade-off between speeding up the leading train and slowing down the following train, as suggested by Albrecht et al. [2018].

3.2 Example 2: Express from London

In this next example, we simulate the motion of two express passenger trains running from London Marylebone to Princes Risborough. This is a 58 km journey taking 25 minutes. The

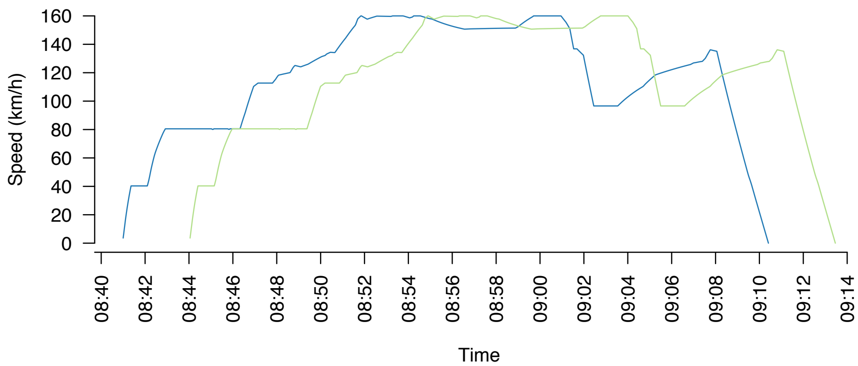

Figure 12: Speed profiles for London – Princes Risborough trains.

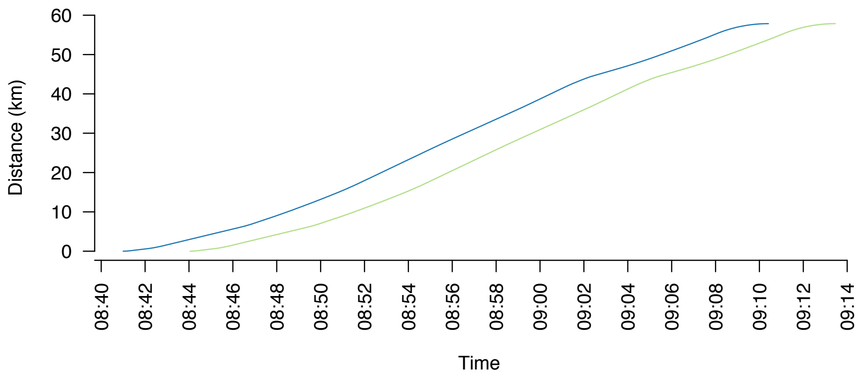

Figure 13: Train graph with two identical passenger trains running express from London.

two trains have identical characteristics and identical optimal journey profiles. Figure 12 shows the optimal speed profiles.

As with the previous example, we assume that the required minimum separation is 3 km. We start the second train as soon as the first train has travelled 3 km, to maximise the likelihood of interaction between the trains. Figure 13 shows the distance that each train has travelled at a given time.

Figure 14 shows the separation between the two trains. The relatively low speed limits leaving London mean that the first train speeds up while the second train is still travelling slowly. This increases the separation between the trains, and the separation remains above the critical 3 km for the remainder of the journey. There is no need to intervene.

3.3 Example 3: Approaching London

Our example trains travelling away from London never got too close due to the initial speed limits. Next we simulate two trains running towards London, from Wembley Stadium to

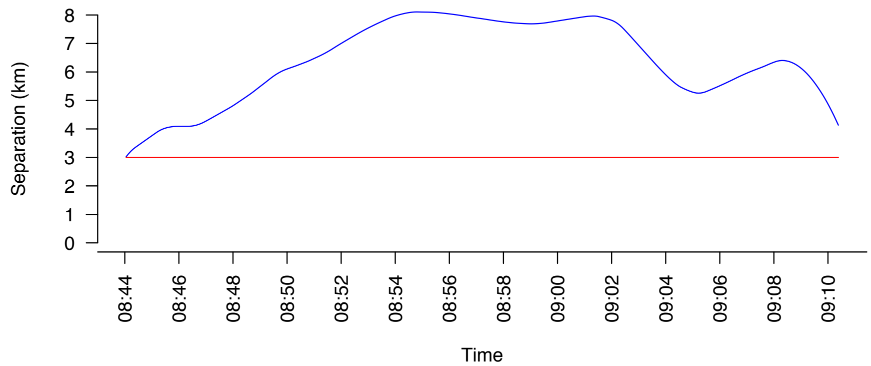

Figure 14: Separation of the two trains from London.

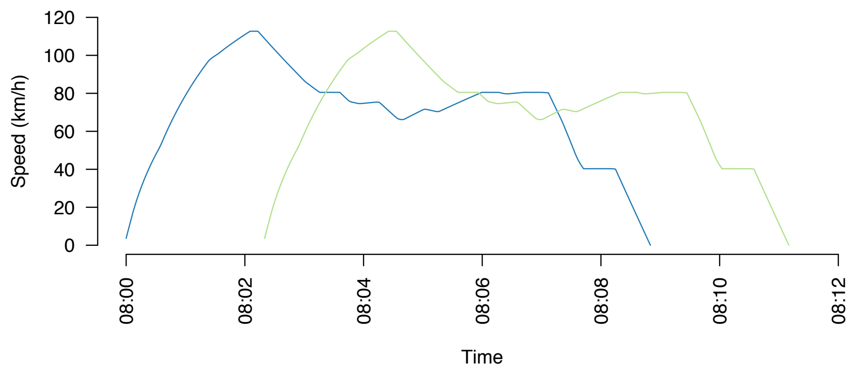

Figure 15: Optimal speed profiles of two trains approaching London.

London Marylebone. This is a 10.5 km journey which takes just under 9 minutes. Once again, we start Train 2 as soon as Train 1 has travelled 3 km. Figure 15 shows the optimal speed profiles of the two trains when separation is not considered.

Figure 16 shows the separation between the two trains (orange). The two trains are closest together at 08:08:49 when Train 1 is arriving at London Marylebone. We add timing constraints for Train 2 to keep the 3 km separation from Train 1, using the same method as in our first example. Figure 16 shows the separation after adding the timing constraints for Train 2 (purple).

The rescheduled Train 2 finishes its journey 48 seconds late, and consumes 1.76% less energy than the original journey.

If we want Train 2 to arrive on time then we need to speed up Train 1. So next we run the Train 1 fast as possible and adjust Train 2 to meet the minimum separation requirement. Figure 19 shows the separation after adding timing constraints for both trains (purple). After making Train 1 as fast as possible, it arrives 29 s early and Train 2 is still 17 s late at the destination. Together, the trains use 23% more energy than the optimised journeys. Because of the low speed limits near the end of the journey, it is not possible to meet the separation

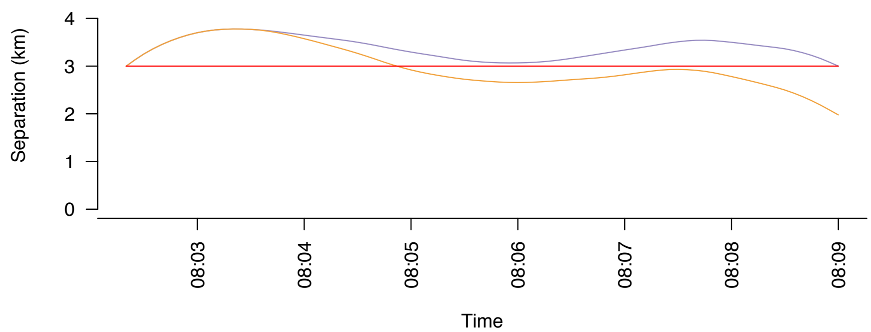

Figure 16: The original separation (orange) for the two trains approaching London, and separation after rescheduling Train 2 (purple).

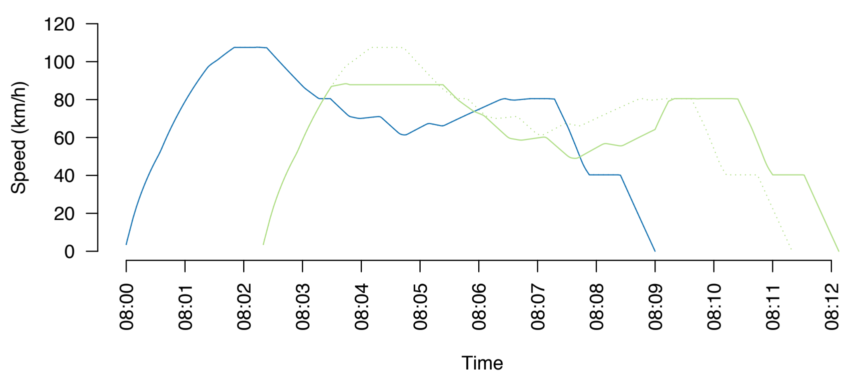

Figure 17: Original speed profiles of the two trains approaching London (blue and dotted green lines) and speed profile of the rescheduled second train (green).

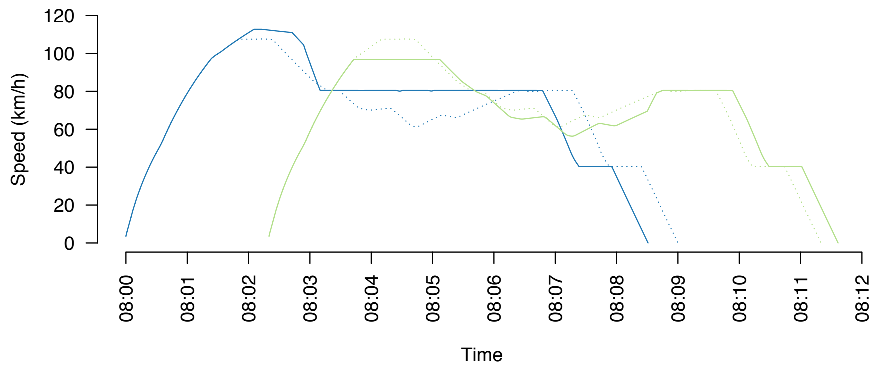

Figure 18: Original speed profiles of the two trains approaching London (dotted lines) and speed profiles of the rescheduled trains (blue and green).

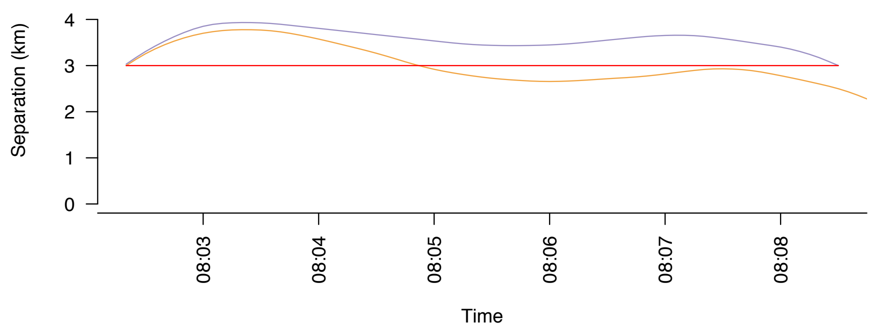

Figure 19: The original separation of the two trains approaching London (orange), and separation after rescheduling both trains (purple).

constraints without changing the time between arrivals at London Marylebone.

3.4 Example 4: Mid-journey Speed Restriction

None of the examples so far demonstrate a scenario where both trains arrive on time and the optimal strategy is a compromise between speeding up Train 1 and slowing Train 2. Our final example does this, using a scenario where the minimum separation occurs in the middle of the journey.

In this example, we simulate two trains running from London Marylebone to Princes Risborough, which is a 57.86 km journey with a duration of 36 minutes. The trains get closer together in the middle of the journey as they encounter a low speed limit that we have imposed to demonstrate the principle.

Train 1 starts its journey at 08:41:00 and finishes at 09:17:00. Train 2 starts and finishes three minutes after Train 1. Each train consumes 630 MJ energy. Figure 20 shows the speed profiles of the two trains, and figure 21 shows the separation between the two trains.

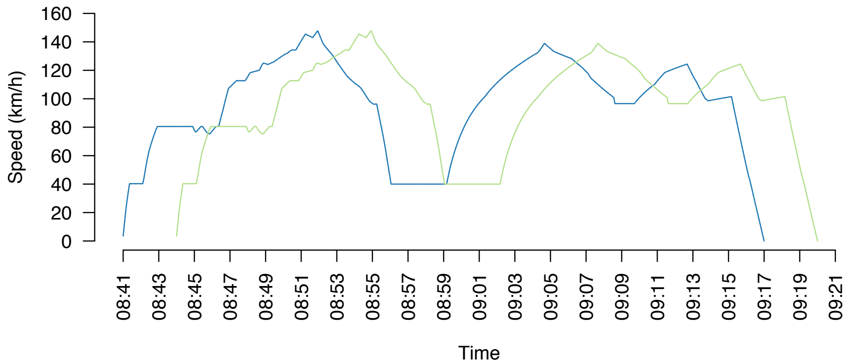

Figure 20: Original speed profiles of the two trains with a mid-section speed restriction.

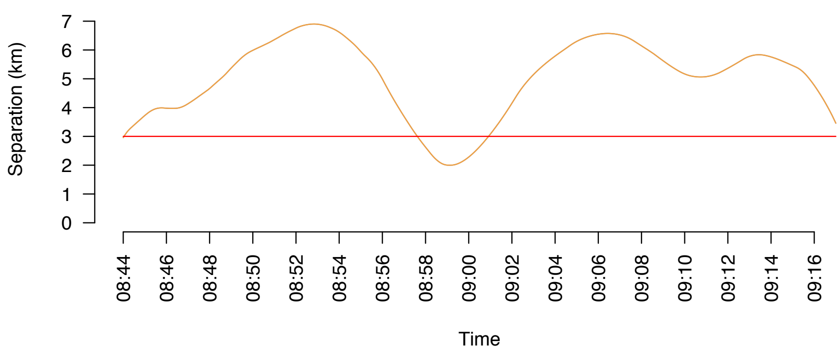

Figure 21: Original separation of the two trains with a mid-section speed restriction.

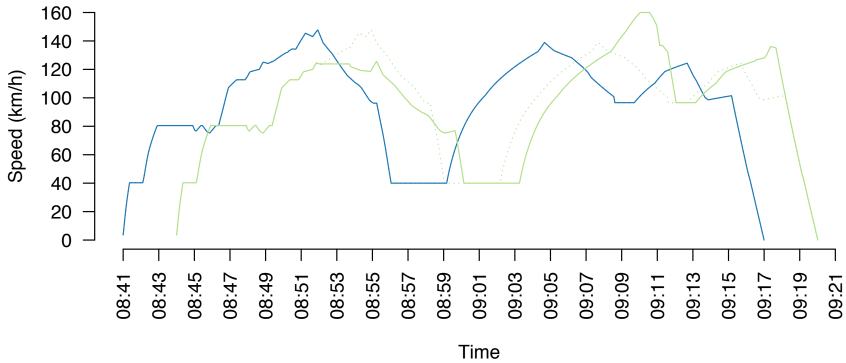

Figure 22: Speed profiles of the two trains with a mid-section speed restriction, after rescheduling Train 2.

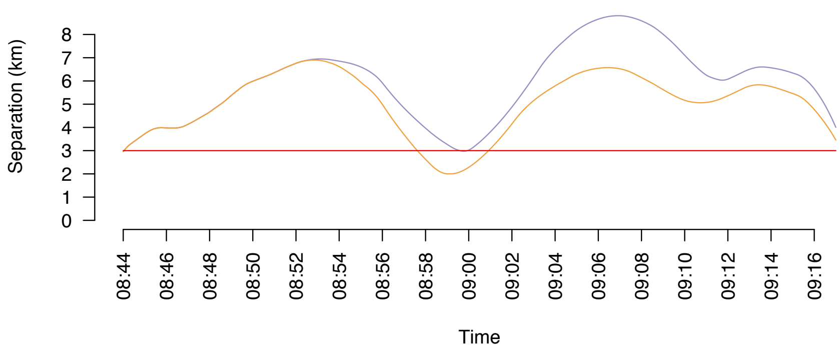

Figure 23: The original separation of the two trains with a mid-section speed speed restriction (orange), and separation after rescheduling Train 2 (purple).

The minimum separation occurs at 08:59:05 when Train 1 is at 27.022 km and Train 2 is at 25.022 km. Since the minimum separation occurs near the middle of the journeys, we can change the speed profiles of either train without compromising their ability to finish on time. We can either speed up the first train or slow down the second train to meet the separation requirement, or do a combination of both.

Slowing Down the Second Train

First, we simulate the optimal journey for Train 1 and slow down Train 2 to meet the separation constraint near 08:59:05. Figure 22 shows the speed profiles of the two trains after adding a timing point for Train 2. The dotted green line represents the original speed profile of Train 2. Both trains still arrive at the destination on time, but together consume 6.8% more energy than the original optimal journeys. Figure 23 shows the separation before and after rescheduling Train 2.

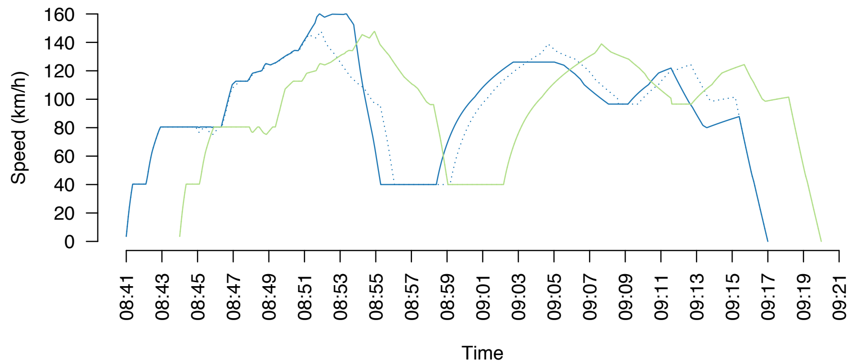

Figure 24: Speed profiles of the two trains with a mid-section speed restriction, after rescheduling Train 1.

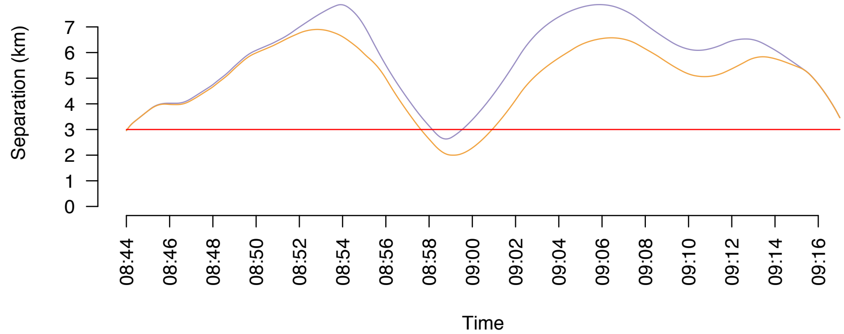

Figure 25: The original separation of the two trains with a mid-section speed restriction (orange), and separation after rescheduling Train 1 (purple).

Speeding Up the First Train

Next, we simulate the optimal journey for Train 2 and speed up Train 1 to meet the separation constraint near 08:59:05. Figure 24 shows the speed profiles of the two trains after adding a timing point for Train 1. The dotted blue line indicates the original speed profile of Train 1. Both trains still arrive at the destination on time, but together consume 5.1% more energy. Figure 25 shows the separation before and after rescheduling Train 1.

Adding Timing Points to Both Trains

We can find the optimal compromise between slowing Train 2 and speeding up Train 1 by imposing a latest arrival time τ for Train 1 at $x_1 = 28.022\,\text{km}$, and then driving Train 2 to avoid getting too close to Train 1.

The earliest that Train 1 could arrive at the timing point x_1 is $\tau =$ 08:59:20 and the latest it could arrive at the timing point is $\tau =$ 09:00:06. We vary the time τ in this interval then run Train 1 with this constraint and Train 2 to avoid Train 1. We calculate the total energy

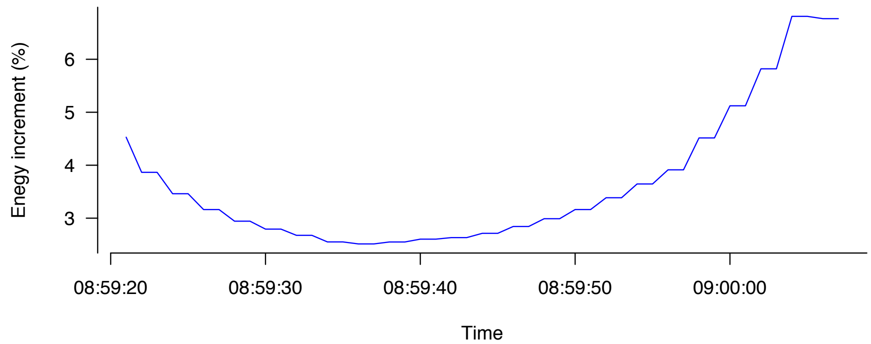

Figure 26: Overall energy increment for the two journeys with a mid-section speed limit, for different values of the timing constraint τ.

consumption for each τ. Figure 26 shows the total energy increment for each value of τ.

The graph is “lumpy” because of the numerical precision of the Energymiser software used to calculate optimal journeys. Nevertheless, the graph shows that we can meet the separation constraints and minimise the overall energy use by setting $\tau \approx$ 08:59:37.

4 Conclusion

Trains will be delayed if they get too close to the train ahead. These types of delay can be reduced by designing robust timetables with adequate train separation, and then by using Driver Advice Systems to ensure that trains are driven to the timetable.

Nevertheless, when a train is delayed, the delay can propagate to following trains if they encounter restrictive signals, which introduces further delay.

We have described a method that can be used while planning timetables to ensure adequate separation between trains, but also in real time to make small adjustments to individual train schedules so that restrictive signals are avoided. The method can simply adjust the schedules of following trains to maintain the required separation, or can find the energy-optimal trade-off between speeding up a leading train and slowing down a following train.

References

A. Albrecht, P. Howlett, P. Pudney, X. Vu, and P. Zhou. The key principles of optimal train control-Part 1: Formulation of the model, strategies of optimal type, evolutionary lines, location of optimal switching points. *Transportation Research Part B: Methodological*, 94:482–508, 2016a.

A. Albrecht, P. Howlett, P. Pudney, X. Vu, and P. Zhou. The key principles of optimal train control—Part 2: Existence of an optimal strategy, the local energy minimization principle, uniqueness, computational techniques. *Transportation Research Part B: Methodological*, 94:509–538, 2016b. ISSN 0191-2615. doi: http://dx.doi.org/10.1016/j.trb.2015.07.024.

A. Albrecht, P. Howlett, P. Pudney, X. Vu, and P. Zhou. The two-train separation problem on non-level track—driving strategies that minimize total required tractive energy subject to prescribed section clearance times. *Transportation Research Part B: Methodological*, 111:135 – 167, 2018. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2018.03.012.

L. Chen, C. Roberts, F. Schmid, and E. Stewart. Modeling and solving real-time train rescheduling problems in railway bottleneck sections. *Transactions on Intelligent Transportation Systems*, 16(4):1–9, 2015.

A. Galapitage, A. R. Albrecht, P. Pudney, X. Vu, and P. Zhou. Optimal real-time junction scheduling for trains with connected driver advice systems. *Journal of Rail Transport Planning and Management*, 8(1):29 – 41, 2018. ISSN 2210-9706. doi: 10.1016/j.jrtpm.2018.02.003.

X. Luan, Y. Wang, B. D. Schutter, L. Meng, G. Lodewijks, and F. Corman. Integration of real-time traffic management and train control for rail networks–Part 1: Optimization problems and solution approaches. *Transportation Research Part B*, 115:41–71, 2018a. doi: 10.1016/j.trb.2018.06.006.

X. Luan, Y. Wang, B. D. Schutter, L. Meng, G. Lodewijks, and F. Corman. Integration of real-time traffic management and train control for rail networks–Part 2: Extensions towards energy-efficient train operations. *Transportation Research Part B*, 115:72–94, 2018b. doi: 10.1016/j.trb.2018.06.011.

K. Panou, P. Tzieropoulos, and D. Emery. Railway driver advice systems: evaluation of methods, tools and systems. *Journal of Rail Transport Planning & Management*, 3:150–162, 2013.

G. M. Scheepmaker, R. M. Goverde, and L. G. Kroon. Review of energy-efficient train control and timetabling. *European Journal of Operational Research*, 257(2):355–376, 2017.