How and Why Freight Trains Deviate from the Timetable

Evidence from Sweden

Palmqvist, Carl-William; Lind, Anne; Ahlqvist, Victoria

Published in:
IEEE Open Journal of Intelligent Transport Systems

DOI:
10.1109/OJITS.2022.3160546

2022

Document Version:
Peer reviewed version (aka post-print)

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

How and Why Freight Trains Deviate from the Timetable: Evidence from Sweden

Carl-William Palmqvist, Anne Lind, Victoria Ahlqvist

Abstract European infrastructure managers (IMs) create annual timetables for trains that will run during a year. Freight trains in Sweden often deviate from this by being *added*, *cancelled*, *delayed* or *early*, resulting in increased costs for IMs and railway undertakings (RUs). We investigate the frequency of and causes for these deviations, using one year of operational data for 48,000 trains, and 15 stakeholder interviews. We find that about 20% of freight trains are *added* once the timetable has been created, and that *cancellations* occur for about 35% of freight trains, mostly at the RUs' initiative. *Delays* are common: some 40% of departures, 30% of runtimes, and 20% of dwell times are delayed. Running *early* is even more common: 80% are ready to depart early, and 60% do so, while 40% of runtimes and 75% of dwell times are shorter than scheduled. We find links and feedback loops between the root causes for these deviations and suggest that IMs reserve more of the capacity that is needed for freight trains and instead distribute it throughout the year. This could lead to more appropriate, attractive, and reliable timetables for freight trains, whilst greatly reducing the amount of planning effort.

***Index Terms*— Freight transportation, Rail transportation, Rail transportation reliability**

I. INTRODUCTION[1]

Railways are an important part of transport systems, and many governments aim to shift freight transport from road to rail with increasing degrees of intermodality (see e.g. [1], [2]). Such a transfer can reduce the congestion on roads and decrease the environmental impact from the transport sector. Utilizing existing capacity in a more efficient manner is important since development of railway infrastructure is very expensive. Timetables are crucial to getting the most use out of railway capacity [3]–[5]. To attract customers and increase the modal share of railways, timetables need to offer reliable, short transportation times.

This study analyses the capacity allocation process and deviations from the annual timetable for freight trains in Sweden. In Sweden and other European countries, annual timetables are made by the infrastructure manager (IM) in a process spanning years. This is regulated by EU directives, implemented in national law, and is therefore similar in the member countries [6]. However, there are often deviations from timetables, especially for freight trains.

This paper identifies and analyses four types of possible deviations from the timetable and the intended capacity allocation: trains can be 1) added, or 2) cancelled, 3) delayed, or 4) early. Discrepancies between scheduled and operated traffic are problematic: some trains may have been refused access because the timetable appears full [7], some trains have longer scheduled runtimes than necessary [8], and the revisions to the timetable requires extensive resources and increases costs for IMs and the railway undertakings (RUs).

To increase the reliability of operations and enable a higher utilization of capacity, it is important to better understand what causes the deviations described above. This paper thus aims to describe the frequency and explore the root causes for discrepancies between the annual timetable and executed freight train transports, and to identify potential areas for improvement, particularly in terms of capacity allocation.

II. BACKGROUND AND LITERATURE REVIEW

A. Capacity allocation for freight trains

The railway infrastructure in EU countries is to a large extent state-owned, while the trains are run by separate companies. IMs are responsible for allocating capacity in a fair and non-discriminatory manner based on applications from the RUs [9]. The allocation is done at a tactical planning level, up to 18 months in advance, by preparing annual timetables. Tactical planning is usually done for groups of products or resources (i.e. groups of trains with similar tasks and characteristics) to handle uncertainty [10],

First submitted on December 15th, 2021. This work was supported in part by the Swedish Transport Administration Grant TRV2020/119576.

C. W. Palmqvist is with the Department of Technology and Society, Lund University, Box 118, 221 00 Lund, Sweden. (e-mail: carl-william.palmqvist@tft.lth.se).

A. Lind is with the National Capacity Optimisation, the Swedish Transport Administration. (e-mail: anne.lind@trafikverket.se).

V. Ahlqvist is with the Department of Industrial Management and Logistics, Lund University, Box 118, 221 00 Lund, Sweden. (e-mail: victoria.ahlqvist@tlog.lth.se).

and does not consider individual activities (such as exact arrival and departure times of specific trains) [11]. This is not the case when railway capacity is allocated in the EU, as all trains are scheduled with a high precision in the annual timetable. Compared to other industries, the process for allocation capacity lacks a stage of production planning, where the previously made plans can be adjusted to changed conditions [12]. As a result, the annual timetable cannot handle uncertainty in the RU's applications. The case is very different for freight train operations in North America, which essentially operate without a timetable (see for instance [13], [14]) – the deviations mentioned in this paper would have an entirely different meaning in that context. To mix unscheduled freight trains with scheduled passenger trains in a constructive way is not easy, but may be an important research direction for the future.

B. Adding or cancelling freight trains

To cope with the inherent uncertainty without having a proper production planning step, trains are added or cancelled in a short term so-called ad hoc process. These are the first two types of deviations discussed in this paper. They are initiated by either the IMs or RUs, are handled on a first-come-first-served basis and must be answered within five working days. Existing train paths cannot be changed and are thus limitations when new trains are added.

Freight train paths from the annual timetable are often cancelled in the ad hoc process. To illustrate the magnitude: during 2016 about 18% of all scheduled trains were cancelled in Sweden [8], almost all of them freight trains. Several earlier studies [8], [15], [16] have analysed the occurrence of these changes, but not their causes. [17], however, identifies several problems in the capacity allocation that can be regarded as causes of deviations. For example, it is difficult for freight companies to know the demand when they apply for train paths, as the applications for the annual timetable are submitted up to 13 months in advance. Cancelled trains can thus in part be a result of RUs applying for more train paths than necessarily required, to enable flexibility in their own planning.

C. Delayed freight trains

Once freight trains run, it is not uncommon that they are delayed. This is the third type of deviation we discuss in the paper. These deviations are often monitored by measuring punctuality, a key indicator in railway operations. Delays can damage the RUs' reputation and increase their operational costs, and customers may need to increase their stock levels to handle late deliveries [18], [19]. In 2019, 20% of the freight trains were more than five minutes delayed, compared to 9% of passenger trains [20]. Typically, delays for freight trains are also larger than for passenger trains, and it is not uncommon for them to be measured in hours rather than minutes [21]. The method in [22] together with an average load of 638 tonnes [23], suggests that the value of time for a freight train in 2019 was around 2,200 SEK (roughly €220), and including operational costs would raise this figure higher. Expressed differently, some 73 million tonnes were moved by railway in 2019, with a value of time of 2.8 SEK per tonne hour.

The causes for delay are the subject of a range of research papers, but many of these focus on passenger rather than freight trains. Some of the causes identified in the literature include technical issues and communication problems [24], weather factors [25], the size and distribution of supplements [6], the length of the train run [26], heterogeneity in terms of speed differences, different stop patterns and varying headways [27], [28], and capacity utilisation [28], [29]. [30] specifically mention a need for studies mapping the underlying causes behind the delay cause attribution codes used by IMs.

D. Early freight trains

The last type of deviation, trains running before schedule, is more common for freight trains than passenger trains [31]. During 2018, 53% of all freight trains in Sweden departed more than 15 minutes ahead of their scheduled departure time [30]. Early trains are not captured by typical punctuality measures, and in contrast to delays, causes for earliness are not coded, nor have they been the topic of many previous studies. [30] describe an early departure as a reason for deviations, but do not analyse the reasons behind early departures. [17] only mentions that there are many explanations for early departures, without going into them.

Whether a train can run before or after the allocated train path is decided manually by individual dispatchers, who check the graphical timetable to see if any conflicts may arise and make changes with pen and paper. In Sweden, the basic rule is that trains running on time have priority, and exceptions are only allowed for special reasons. There is no IT support for these decisions, and the dispatcher will often not be able to check the entire (sometimes hundreds of km long) train path before allowing the train to depart [31]. Digital graphs are gradually being introduced in Sweden, which will eventually give dispatchers a better overview.

Once in the network, freight trains can often run faster than scheduled. To reduce the likelihood of delays, timetables must include sufficient margins and buffer times between trains. Unfortunately, this leads to longer scheduled travel times, unnecessary capacity utilisation, and increased costs. Train weight and running characteristics are also not always known when applications are made to the annual timetable, so a train may be lighter and faster than expected when it runs.

Excessive stops also contribute to long travel times. While some stops are commercially useful, so-called technical stops exist to accommodate other trains when there are meetings on single track lines or to let faster train pass on double track lines. Sometimes technical stops become unnecessary as train paths often include stops for interactions with trains that have been cancelled after the timetable was established or have been conflict regulated with trains that have another running day pattern. This is particularly common on single track lines and in Sweden, up to 35% of the stops are of this nature [32]. Since these stops are unnecessary and omitted in practice, they create a potential for trains to run faster and earlier than scheduled.

E. Fees and incentives

There are some fees attributed to causing delays, as an incentive mechanism. The following figures are from [33]. If a delay is attributed to the RU, the RU must pay 75 SEK/delay minute to the IM, and vice versa. Additionally, if the IM causes a delay in the range of 60-179 minutes to freight trains it must pay 13 000 SEK to the RU, or 20 000 SEK for delays of 180+ minutes. There is thus an incentive for RUs to avoid causing delays. Departing ahead of schedule reduces the risk of this, as fees for delays are calculated compared to the scheduled times, and any delays caused to other trains by the early freight train will be attributed to dispatching and be reimbursed by the IM rather than the RU.

There can also be fees for cancellations, depending on attribution and timing [33]. If the IM causes a cancellation attributed to infrastructure failure, new or urgent trackwork, or dispatching reasons, it must pay 1 000 SEK + 20 SEK/train km cancelled if the train was scheduled to run within 24 hours. If the cancellation is between 24h and 14 days ahead, the fee paid is 750 SEK + 15 SEK/train km cancelled, and 500 SEK + 10 SEK/train km cancelled for those that happen between 15 and 83 days ahead of departure. Cancellations by the IM or RUs that occur 84 days or more ahead of time do not lead to fees. The RUs pay a reservation charge of 10% of the train path charge if the path is cancelled 83-15 days ahead of departure, 20% for those 14 days – 24h before departure, or 500 SEK + 50% of the track access fee if it is cancelled within 24h. There is thus a small incentive against strategically overbooking paths, but not a strong one.

F. Incremental planning and TTR

Extensive research has been conducted in Sweden since the early 2000's on how to improve the timetabling process, much of this under the umbrella term of "incremental allocation". A summary of the concept, as well as a rough roadmap for implementation, can be found in [34]. In summary, delivery commitments are extracted from timetable applications in stage 0. These are the commercially important points of a timetable that should be known and fixed. In stage 1, not all production specifics need to be set in stone in the annual timetable, only these delivery commitments. In stage 2, trains are allowed to have different train paths on different days of the annual timetable. In stage 3 "trains that are marred by uncertainty" can be scheduled on shorter notice. In stage 4 a new planning method is introduced in the long-term planning process, which ensures that the delivery commitments can be fulfilled, without having to plan everything else in detail at the same time. In stage 5, the process is adapted to market needs by allowing applications to have different time frames, both far in advance and on short notice.

The importance of getting to stage 2 in the above process is highlighted by [35]. In the current process, timetables are scheduled such that there are essentially no conflicts between trains, but the timetable is planned for one "representative" day per year, rather than for all unique days. This means that trains can be scheduled to wait for another train at a specific station every day, even though the other train only runs once a week. In fact, [35] show that in the year of 2014, there were as many as 314 unique timetable days. They also developed a method to schedule detailed and unique timetables for all these days, while keeping the delivery commitments fixed, and show that these unique timetables are more efficient.

There is also some ongoing research into the concept of "reserve capacity". Article 48 of the SERA directive [9] states: "Infrastructure managers shall, where necessary, undertake an evaluation of the need for reserve capacity to be kept available within the final scheduled working timetable to enable them to respond rapidly to foreseeable ad hoc requests for capacity. This shall also apply in cases of congested infrastructure." [36] summarises a recent pilot-study on the topic, and discusses both how much capacity should be reserved, how the reserved capacity should be valued, and the "reservation price" that a given request must surpass to be allowed to use the capacity. These and similar ideas are to some extent present in the ongoing European TTR-project [37], particularly regarding rolling planning, but also with path optimisation requests.

III. RESEARCH DESIGN

This study aims to explore potential root causes behind deviations between the annual timetable and executed freight train transports. It does so through an iterative, single case study process inspired by the case study design by [38] and combined with elements from the Punctuality Improvement Method System (PIMS) by [39]. This is a collection of methods developed for punctuality improvement work, designed to be a continuous improvement system for practitioners and facilitators in the railway system, whilst the goal with the case study is to examine the phenomenon of freight train deviations holistically from multiple angles.

The first step in the study was to map the current process based on literature, process documents, and unstructured interviews with key informants. The second step included the collection and analysis of operational data from the annual timetable of 2019, covering 48,000 freight trains on a Swedish mainline railway, and data from the application systems used for the annual timetable and in the ad hoc process, to identify patterns and possible systematic deviations. The data analysis is described in section 3.1. The third step was to conduct and analyse 15 interviews with railway practitioners and experts at various RUs and an IM, see section 3.2. The last step was to conduct a root cause analysis, based on output of steps 2 and 3, allowing measures to be taken to reduce the problems.

A. Case study and operational data

The operational data was collected from the Swedish Transport Administration's monitoring system. The data covered the section between Malmö marshalling yard and Hallsberg marshalling yard via Lund, Hässleholm, Nässjö, Mjölby and Motala, pictured in Fig. 1. This section is part of the TEN-T Scan-Med corridor and one of the most important sections for freight traffic in Sweden, including a total of 70 stations of varying sizes. The first part of the section, from Malmö to Mjölby, is part of the double-tracked Swedish Southern mainline. The traffic on this section is a mix between freight and passenger trains, including local, regional, and long-distance passenger services, with a large variance in speed profiles. The capacity utilisation is high, especially during peak hours. Between Mjölby and Hallsberg, the line alternates between single and double track. The traffic consists mainly of freight trains, although there are local passenger trains from Mjölby to Motala and some regional passenger trains using the entire section from Mjölby to Hallsberg. In general, the capacity utilisation is lower here than south of Mjölby.

Fig. 1: A schematic illustration of the studied railway line. Double track from Malmö-Motala, mostly single track from Motala-Hallsberg. These are the eight main stations, there is a total of 70 stations along this line.

Two different data sets were collected, both covering the annual timetable of 2019. The data used to analyse the changes made in the ad hoc process include all freight trains that were scheduled to run at least on part of the line between Malmö and Hallsberg, also covering cancelled and added trains. In total, about 48,000 trains were included.

The data used for delayed and early trains covered all freight trains scheduled to pass at least 45 of the 70 stations on the selected line, which in total included about 10,500 train runs across 600 different train numbers. In addition to train numbers, timestamps etc., this data also contains information on delay attribution, manually coded by dispatchers when a delay increase of three or more minutes is registered. Similar data is available for cancellations, but not for added trains, or deviations where the train operates ahead of schedule. Please refer to [40] for more information on railway delay attribution in Sweden.

All data was used to corroborate, substantiate, and complement the statements made by the interviewees, and it helped inform the interview questions.

B. Interviews with practitioners

We conducted 15 semi-structured interviews, focusing on the processes of capacity allocation and planning, as a complement to the operational data discussed above. Nine of the interviews were with six different RUs and six with the IM, all in Sweden. All interviews were recorded and transcribed, and the interviewees had the opportunity to check and approve their answers. The RUs represented have between eight and 1,800 employees. Four of the RUs only have freight traffic, one has both freight and passenger traffic, and the last RU has mainly freight traffic but occasionally runs passenger traffic as well. All interviewees are presented in Table 1.

TABLE 1
INTERVIEWEES AND THEIR ROLES

Interviewee	Role	Interviewee	Role
RU1	Long term planner	IM1	Annual timetable planner
RU2	Quality manager	IM2	Annual timetable planner
RU3	Prod. planning manager	IM3	Ad hoc timetable planner
RU4	Operations manager	IM4	Ad hoc timetable planner
RU5	Long term planner	IM5	Operations manager
RU6	Production manager	IM6	Train dispatcher
RU7	Production manager		
RU8	Planning manager		
RU9	Planning manager		

When the respondents had confirmed the transcribed version of the interviews the transcriptions were rearranged and coded. In a first step the answers were coded based on the four different types of deviations (added, cancelled, late, early). This was followed by a more detailed coding on individual causes that were mentioned. In the interviews the effects of the four types of deviations were also explored, and these were coded in a similar process.

The answers were compared to identify and continue to explore the causes and effects that multiple respondents agreed upon and those that were contradicting. The data was coded and structured in cause-effect loops and relations between causes. Effects mentioned across the different deviation types were also explored. Notably, these analyses were based on the interviewees' *perception* of what constitutes a cause and an effect. The final step in the analysis iterated back to previous research on the subject to further explore potential causes that could be identified in the data as well as potential improvement actions.

IV. ADDING FREIGHT TRAINS

Sometimes, freight trains are added to the annual timetable, after it has taken effect. In our data, this applies to about 20% of trains, or 18 trains per day. Fig. 2 shows the number of additional trains paths per month (totaling 6,618 paths over the year). Unfortunately, no data is available for rejected applications, or about when the applications are made or processed. In the interviews, RUs say that they are unsatisfied with the ad hoc process, and instead choose to apply for extra trains in the annual timetable (see Cancelling Freight Trains, below).

Issues mentioned during the interviews include long response times and long runtimes for added trains. Several RUs also mention that their applications for adding trains are often rejected, sometimes only due to small errors, such

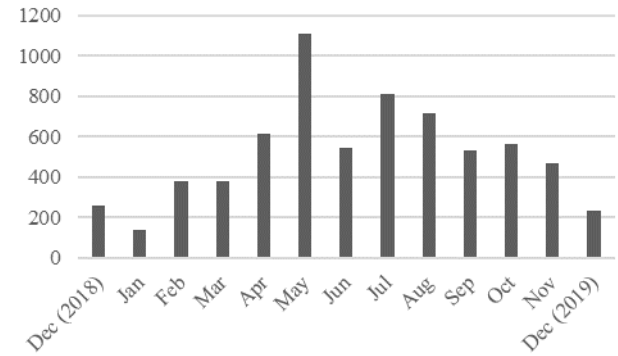

Fig. 2: Number of applications accepted for new train paths. These do not include applications for trains that have either been rejected or first accepted and subsequently cancelled, no data is available for such cases.

as typos. Many rejections result in new applications, as the transport still needs to be carried out, which means that the rejections increase the workload for both the RUs and the IM. New applications following on rejections also arise when applications are rejected due to lack of capacity. Today, the RUs can only state a certain preferred departure or arrival time, rather than an interval, which means that the timetable planner themselves decide how wide they search for a possible train path. However, freight traffic is usually flexible, and the most important aspect is that the transport can be carried out, not that a specific departure time is met. To reduce these rejections, and thereby reduce the time and resources needed to handle the changes, several RUs ask for better communication during the ad hoc process.

V. Cancelling Freight Trains

Cancellations are common: about 35% of freight trains in our data were cancelled. According to both the interviewees and the operational data pictured in Fig. 3, the most common causes by far for changes in the ad hoc process are different kinds of production adaptations at the railway undertakings (RUs). This happens for about 24% of freight trains, or 2/3 of all cancellations. The second biggest cause was adaptation to trackwork possessions, affecting about 7% of freight trains. The third most common cause of cancellations is erroneous planning, at about 3%. Other causes lead to cancellations for about 1% of trains.

According to the interviewees, these cancellations are necessary to handle changing demand from customers. While some RUs rarely cancel trains on their own initiative, other RUs describe that they apply for extra trains in the annual timetable, which are to varying degrees cancelled later. The purpose of having extra trains, that are then cancelled, is to ensure train paths for possible future customers or to handle uncertainty in customer demands. The latter is needed as some of the RUs' customers cannot determine exactly how many trains they will need, which days they need them or where they need the trains when they apply for the annual timetable. This is most frequent in the forestry sector, where it is difficult to know far in advance where the timber will be picked up. The RUs

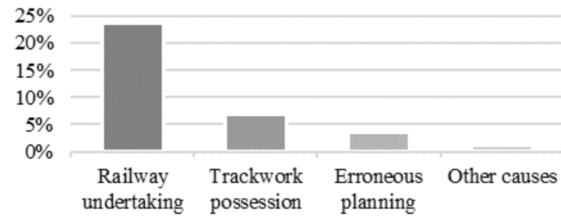

Fig. 3: Probability of train path being cancelled, by reported cause.

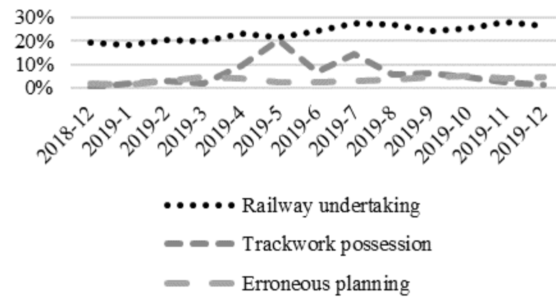

Fig. 4: Probability of cancellation of train by month, due to the top 3 reported causes. Cancellations due to railway undertaking and erroneous planning increase gradually over the year. Trackwork possessions peak in the summer and taper off towards the winter.

handle this by applying for a few trains to different locations each day, although only one of the train paths will be used. This behaviour is known to the IM. Unfortunately, no data is available about when cancellations are made.

The operational data illustrated in Fig. 4 shows that for two of the causes, RUs and erroneous planning, probability of cancellation increases steadily throughout the year. This is in line with the explanation that freight trains are difficult to plan well far in advance. For adaptations to trackwork possessions, Fig. 4 instead shows a clear peak in the summer months. This is the time of the year when most big trackworks take place, so it is not surprising. However, it reinforces the notion, expressed in interviews, that the timetables are often not adapted to trackwork possessions when they are created – such adaptations often happen later.

VI. Freight Train Delays

A. Beginning behind schedule

Approximately 40% of the freight trains in the case study had an initial delay when they entered the network, as illustrated in Fig. 5, top. These delays can arise in two ways: *before* or *after* the train reports that it is ready to depart. Fig. 5, bottom, illustrates that about 20% of trains experience delays *before* they are ready to depart. As Fig. 5 illustrates – as the plotted lines are roughly straight on a normal probability plot, where the Z-values indicate the number of standard deviations away from the mean – the probability of both types of delays are essentially normally distributed over both dates and train IDs, except that it is

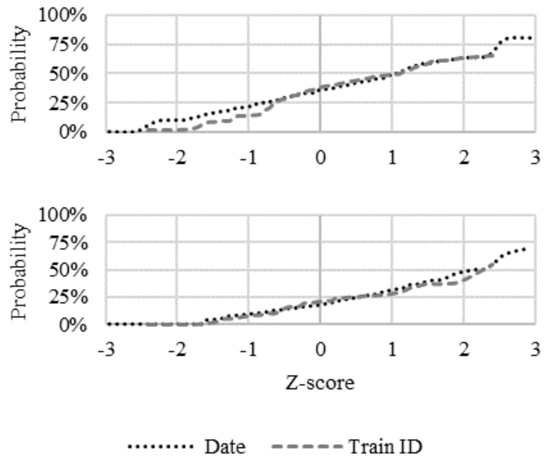

Fig. 5: Normal plot of the probability of late departures (top) and reports (bottom) across days and train IDs. Departure delays are composed of two parts: first the train needs to report that it is ready to depart, which can sometimes be delayed, then a dispatcher needs to give the train permission to depart, which can also be delayed. It turns out that the permission is almost always (>99%) delayed, so this is not plotted.

truncated at 0%. The standard deviation for the probability of departure delay is about 15%, and about 11% for late reports. Delays *after* the train reports that it is ready, are ubiquitous and affect more than 99% of trains. This only rarely leads to a delayed departure, however, because most trains report being ready *ahead of schedule*, a topic we will return to below.

The range of departure delays can be very large. Delays *before* trains are ready follow a flat distribution, with very long tails in both directions, up to a maximum of about 15 hours. The distribution of delays *after* the train is ready is more compact, usually up to about 20 minutes, but sometimes reaching up to 10 hours. In total, there exist some extreme departure delays of up to 22 hours during the year. Such extremes are rare, however, and less than 1% of freight trains depart more than 3 hours behind schedule.

Table 2 shows that *Late from depot* alone contributes a quarter of all delay hours, which is more than three times more than any other individual attribution code. It is difficult to determine the causes of delays occurring outside of Sweden. Trains being late from depots or yards within Sweden have widely varying causes, not all of which originate there. The most common cause, mentioned by all RUs, is that the loading is not done in time. This in turn can be due to late deliveries of goods from customers, issues with the loading at the yard or that the goods arrive in other trains that are delayed. Other major issues include locomotives or drivers arriving late, due to previously arriving trains' delays, or issues with rolling stock.

B. Run- and dwell time delays

Once they depart, freight trains can be affected by run- and dwell time delays. These are cases when the train takes longer to run between two stations than scheduled, or stops for longer than scheduled at a station, respectively. For context, our data contains some 575,000 runtimes and

TABLE 2:
TOP FIVE CAUSES OF FREIGHT TRAIN DELAYS IN THE SAMPLE, ACCORDING TO OPERATIONAL DELAY ATTRIBUTION DATA.

Rank order (of 33)	Delay attribution code	Delay hours during 2019	Share of total
1	Late from depot	1 516	25%
2	Delayed by other train	438	7%
3	Late from abroad	425	7%
4	Weather	364	6%
5	Deviating train composition	353	6%

56,000 dwell times, so there are about 11 runtimes for every dwell time. Runtimes are here defined as the time it takes for a train to run between two adjacent operating points, some of these points are stations where some trains stop, others are older stations or technical points, where this no longer happens. Dwell times are instead only counted (or included in the data) when the trains have a scheduled stop at a station (or other operating point). In fact, dwell times are only registered at points where trains have scheduled stops. Unscheduled stops at other points would instead lead to runtime delays. Put simply, the trains pass by many operating points without stopping. This naturally means that runtime delays are a bigger issue for freight trains.

Fig. 6 explores the influence of a train's departure status on the occurrence of both types of delays. With delays, we here mean instances when an activity in practice takes longer than scheduled – such as a particular runtime taking six minutes instead of four minutes, regardless of whether the train is behind schedule or not. In this way freight trains – which as we have seen very rarely depart on schedule – are different from passenger trains, where we instead would want a train that arrives early at a station to stay there longer, until the scheduled departure time, without considering that a delay. In this sense, we consider that freight trains operating ahead of schedule may still experience run- and dwell time delays, even if they do not have arrival delays. This is similar to a North American definition of freight train delays (e.g. [41]), where a delay means an extended travel time, without reference to specific departure or arrival times. For runtime delays (Fig. 6, top), we find that the probability of delay is highest for trains that are *behind schedule*, with little difference between those that are *ahead of schedule* or *on time*. The size and standard deviation of runtime delays also varies, being the lowest for trains that are *on time*, and higher when they are either *ahead of* or *behind schedule*. Being *on time* is thus preferable from the perspective of runtime delays.

Fig. 6 (bottom) also shows that dwell time delays are the smallest and least variable when trains are *on time*, but that the probability of delays is even lower if trains are *behind schedule*, and higher if they are *ahead of schedule.* This suggests that stops are sometimes skipped or reduced when trains are *behind schedule*, and that they sometimes are prolonged for trains that are *ahead of schedule*. Both practices are what one would expect and desire. Comparing

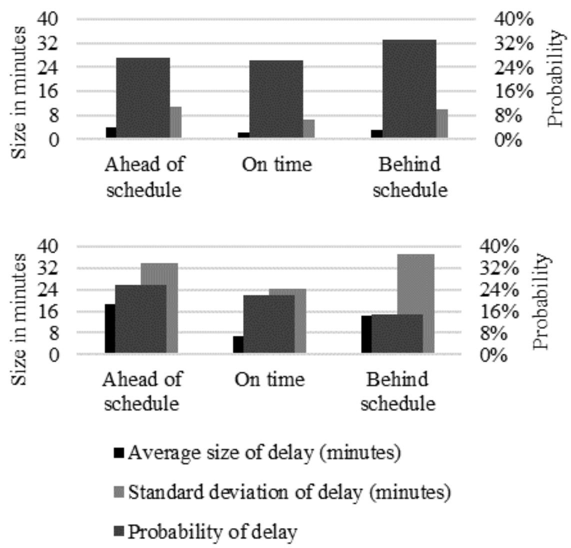

Fig. 6 Run- (top) and dwell time (bottom) delays by status of train.

run- and dwell time delays, the former are more frequent, both in sheer numbers (as freight trains make relatively few stops), and in terms of probability. When they happen, however, dwell time delays tend to last longer.

The different causes for of run- and dwell time delays are illustrated in Fig. 7. It shows us that *Railway Undertaking* is the most probable known cause for both types of delays, occurring in 2% of runtimes and 3% of dwell times. When this happens, the mean delay size is 10 minutes for runtime delays and 32 minutes for dwell time delays. For both delay types, *Accident/External* is slightly less probable, but leads to the biggest delays (averaging 14 and 36 minutes, respectively) and standard deviations. However, most delays by far are due to an *unknown cause*. There is a 27% probability of such a delay occurring on any given runtime section, and a 16% probability for dwell times. The average size for these runtime delays is 3 minutes, and 13 minutes for dwell time delays. The greater average size of dwell time delays somewhat compensates for the fact that there are many more run- than dwell times, and the slightly higher probability of run- compared to dwell time delays.

The variation in run- (top) and dwell time delays (bottom) across days, train IDs and line segments is illustrated in Fig. 8. All three fit rather well to normal distributions with average probabilities of about 30% for runtimes and around 20% for dwell time delays. As we saw in the previous paragraph, the average probabilities should be 29% and 22%, respectively, and the discrepancies here are mainly because the averages across each of these dimensions are unweighted, while a subset of stations (mainly, but also dates and train IDs) has many more observations than others.

The standard deviations which govern the slopes in Fig. 8 vary significantly across the three variables. For both run- and dwell time delays, it is the smallest across dates, at 3% and 4%, respectively, suggesting that there is a small

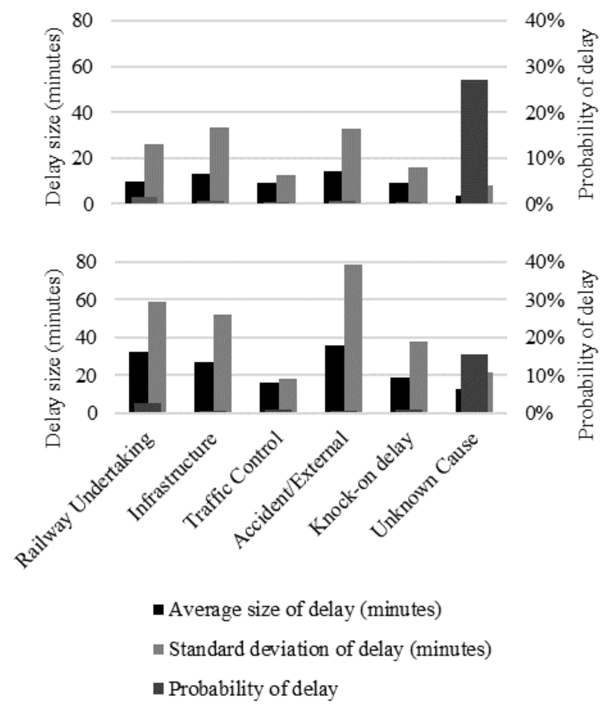

Fig. 7 Causes of run- (top) and dwell (bottom) time delays, and their associated sizes, standard deviations, and probabilities.

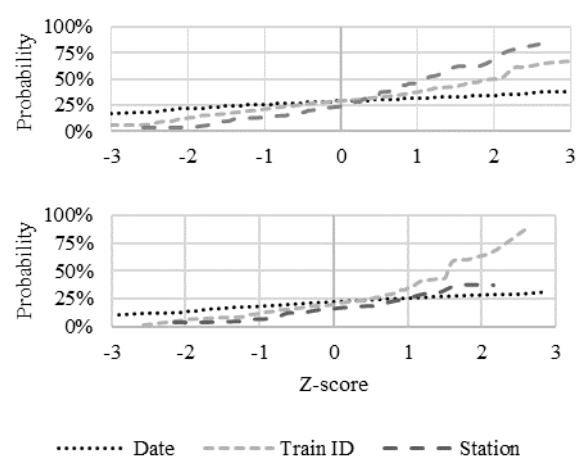

Fig. 8 Normal plot of the distribution of run- (top) and dwell (bottom) time delay probabilities across days, train IDs and line segments.

normal variation across time, but that this is not the dominant dimension. The variations across train IDs have high standard deviations of 9% for runtime delays, and as high as 14% for dwell time delays, for which it is the dominant factor. The geographical variation is larger (and dominant) for runtime delays, with a standard deviation of 17%, and still substantial for dwell time delays (a standard deviation of 9%). These probabilities are then modified slightly by whether the train departed on time or not (Fig. 8), by the (usually unknown) cause of delay (see Fig. 7).

Putting these dimensions together, there can be a large range of delay outcomes, both in terms of probabilities and sizes, but the baseline risk is quite high. A variation across days is difficult to compensate for in a timetable, but the variation across both train IDs and geography can in principle be adjusted for in scheduled run- and dwell times.

In fact, an average train travelling the 54 runtime sections along the studied line, making an average of almost 5 stops, would be expected to encounter 16 runtime delays and one dwell time delay, for an average of about 104 delay minutes. This is just an average case, and as we have seen there are substantial standard deviations along several dimensions here, so there is a large range around these numbers. Still, it highlights the need to recover time in various ways, a topic we shall now turn to.

VII. Freight trains running ahead of schedule

Trains often operate ahead of schedule due to a combination of causes, and this can spread to associated trains. For instance, a train might depart early and thus does not need to stop for a scheduled interaction, making the train run further ahead of schedule. When the train arrives at its destination early, the rolling stock is available earlier than planned for the next train in the traffic circulation, resulting in another early departure.

In general, the RUs consider early trains an indication that the planning could have been done in a better way. These deviations are not appreciated by the RUs, even though trains being early mean that the goods will *not* be delayed to the customers and associated trains will not be delayed. One way to eliminate the issues with early trains is to prohibit trains departing early. However, this is not recommended by either RUs or operational staff at the IM. Some brief trials have been carried out where no trains were allowed to depart early [42], but the results have been large delays both for freight and passenger trains. The following is a short account of different ways in which trains run ahead of schedule.

A. Beginning ahead of schedule

About 60% of all freight trains on the line depart ahead of schedule, and this is acknowledged by most of the interviewees. Slightly more do so from Malmö (65%), going north, than from Hallsberg (49%), going south. Fig. 9, top, shows that the probability of departing early is normally distributed over both dates and train IDs. Fig. 9, bottom, shows that almost 80% of trains report being ready to depart ahead of schedule, with a smaller discrepancy between the two yards (80% and 76%, respectively). From reporting to departing, about 20% of trains thus stop being ahead of time, having to wait for permission, slightly more from Hallsberg (27%) than Malmö (15%). There is a short single-track section just south of Hallsberg, so that more trains wait to depart from there is not surprising.

The interviewees mention several causes for why trains are ready to depart early. Two explanations are that the staff at the marshalling yards start the loading early, or that the loading takes less time than scheduled; for example, if the train has less goods than usual. Other common causes are that the train path has a later departure time than what was

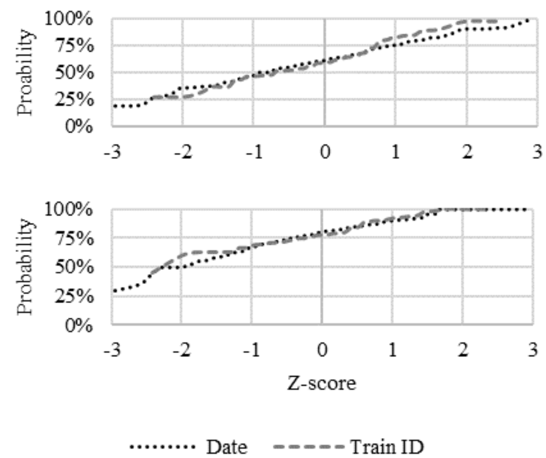

Fig. 9 Normal plot of the probability distribution of departing (top) and reporting (bottom) ahead of time days and train IDs.

applied for, or that that the time between two associated trains is longer than requested. Running ahead of schedule can thus spread from one train to another when goods or rolling stock arrive early to transhipments or turnarounds.

For dispatchers, early trains can be an asset as their scheduled train paths can be used by late trains. The extra time can also be used at meetings/overtakes to give other late trains a possibility to take back some time, e.g. by letting the early train take the cost of waiting at the meeting station. As a result, trains are often allowed to depart early when possible. However, the dispatchers cannot, for the most part, check the entire train path before releasing a train, and sometimes, early trains impede other trains, especially during disruptions.

B. Running faster than scheduled

Freight trains sometimes run faster than scheduled, once they are in the network. Fig. 10, top, shows that 40% of freight train run times are shorter than scheduled. The Fig. also shows how there is some variation around this mean, essentially following normal distributions across dates, train IDs, and line segments. The variation is largest across the latter of these, with a standard deviation of 15%, closely followed by the variation amongst train IDs, at 11%. The standard deviation is smallest across dates, only 4%.

The interviewees mention three main causes for trains running faster than scheduled: 1) When the formation of the trains varies from day to day, they can have different speed limits on different running days. In the application, the lowest speed is used to avoid delays, which means that the train can run faster other days. 2) Sometimes the runtime calculation is not correct, due to errors in or insufficient adaptation of the runtime templates to freight trains. 3) Many trains also run faster because they have excessive runtime supplements. Some trains also have time supplements for trackwork possessions all year, although the works only last short periods. Most of the RU interviewees think that the runtimes and time supplements

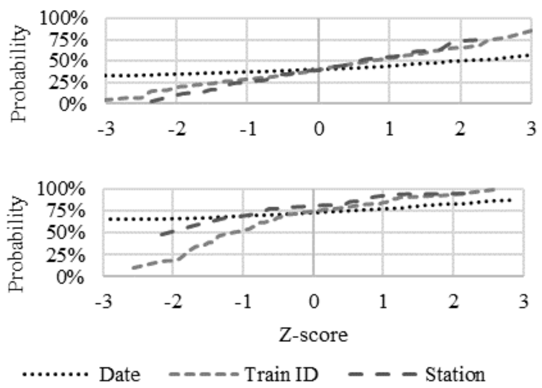

Fig. 10 Normal plot of the probability of negative run- (top) and dwell (bottom) time delays across days, train IDs and line segments or stations.

have increased in recent years. In several cases, the runtimes have passed a critical limit, forcing the RUs to take costly measures that, in hindsight, would not have been necessary. The IM also knows of the issue with excessive runtime supplements, but that it does not apply to all trains.

Trains can also make up time at stops, either completing them faster than scheduled, or skipping them entirely. Fig. 10, bottom, shows that this happens for as much as 75% of stops, and that the distribution of negative dwell time delays across dates, train IDs and stations all roughly fit to a normal distribution. The variation is biggest across train IDs (with as standard deviation of 17%), followed by stations (11%) and dates (4%).

The interviewees mention the following causes for reduced dwell times: 1) Technical stops can be avoided because one or both trains deviate from their train paths, so that a scheduled interaction between them is no longer necessary. 2) That the interacting train does not run at all. For instance, if a train running every day has a scheduled interaction with a train only running on Mondays, it will get the same train path every day, even though the interaction only occurs on Mondays. 3) Some stops are scheduled as breaks for the drivers, according to regulations, but the drivers can choose to shorten them or not to use them at all. 4) Commercial stops (such as for loading and unloading) can be completed faster than scheduled.

VIII. Summary of Deviation Parameters

The parameters described in sections VI-VII above are summarized in Table 3. These can be helpful for modelling purposes, and for better understanding how much and on which dimensions the different types of deviations vary. The *probability* of deviations when reporting ready at the yard, departing from the yard, running between stations, and stopping at stations, all roughly follow normal distributions with the means and standard deviations gathered in Table 3. That the means differ somewhat depending on whether one counts across dates, trains, or stations is because of differences in weights, and the differences are not very meaningful. The difference in

Table 3:
Summary of means and standard deviations across dates, trains and stations for different types of deviations.

Variation across:	Date		Train		Station	
Deviation type:	Mean	St.dev.	Mean	St.dev.	Mean	St.dev.
Ready late at yard	20%	12%	20%	11%	N/A	N/A
Late departure	36%	14%	35%	16%	N/A	N/A
Extended runtime	29%	3%	30%	9%	29%	17%
Extended dwell time	22%	4%	24%	14%	17%	9%
Ready early at yard	79%	12%	79%	11%	N/A	N/A
Early departure	61%	14%	62%	17%	N/A	N/A
Shorter runtime	41%	4%	39%	11%	40%	15%
Shorter dwell time	73%	4%	70%	17%	80%	11%

standard deviation is more important. For instance, for runtime deviations (both shorter and longer than scheduled) the geographical variation is the largest, while there is still substantial variation amongst trains, and only little variation between days. For departures from yards, however, the variation across dates and trains is of almost the same size. Notably, Table 3 does not contain information on the sizes of deviations, only the probability that they will occur. It is well known from earlier literature and experience that delay sizes are *not* normally distributed, and this paper does not attempt to model these size distributions.

IX. Analysis and Areas of Improvement

There are connections between the four types of deviations. One important set of connections relates to the lack of train path variants that cause early trains, illustrated in Fig. 11. The lack of timetable variants (such as to account for interactions occurring on some days but not others) could be caused by the need for many trackwork possessions. It might also occur when a daily train has the same train path every day, although the interactions with other trains differ on different running days.

We see that there are many factors contributing to the lack of train path variants. For example, if RUs apply for more train paths than they intend to use, this leads to a high number of applications and an increased workload at the IM. This means that the timetable planners do not have time to make multiple variants within the same train number, to adapt to varying traffic conditions. Because there are not enough variants to account for the varying conditions, many of the trains will in practice run early, compared to the schedule. RUs then attempt to optimise their timetables, to better utilise their resources, which leads to many applications in the ad hoc process. A lack of resources and system support at the IM makes it difficult for them to find new train paths, resulting in a high rate of rejections in the ad hoc process. This, in turn, means that RUs will have little trust in this process, and they compensate by instead applying for more train paths than they plan to use, completing the cycle and creating a feedback loop.

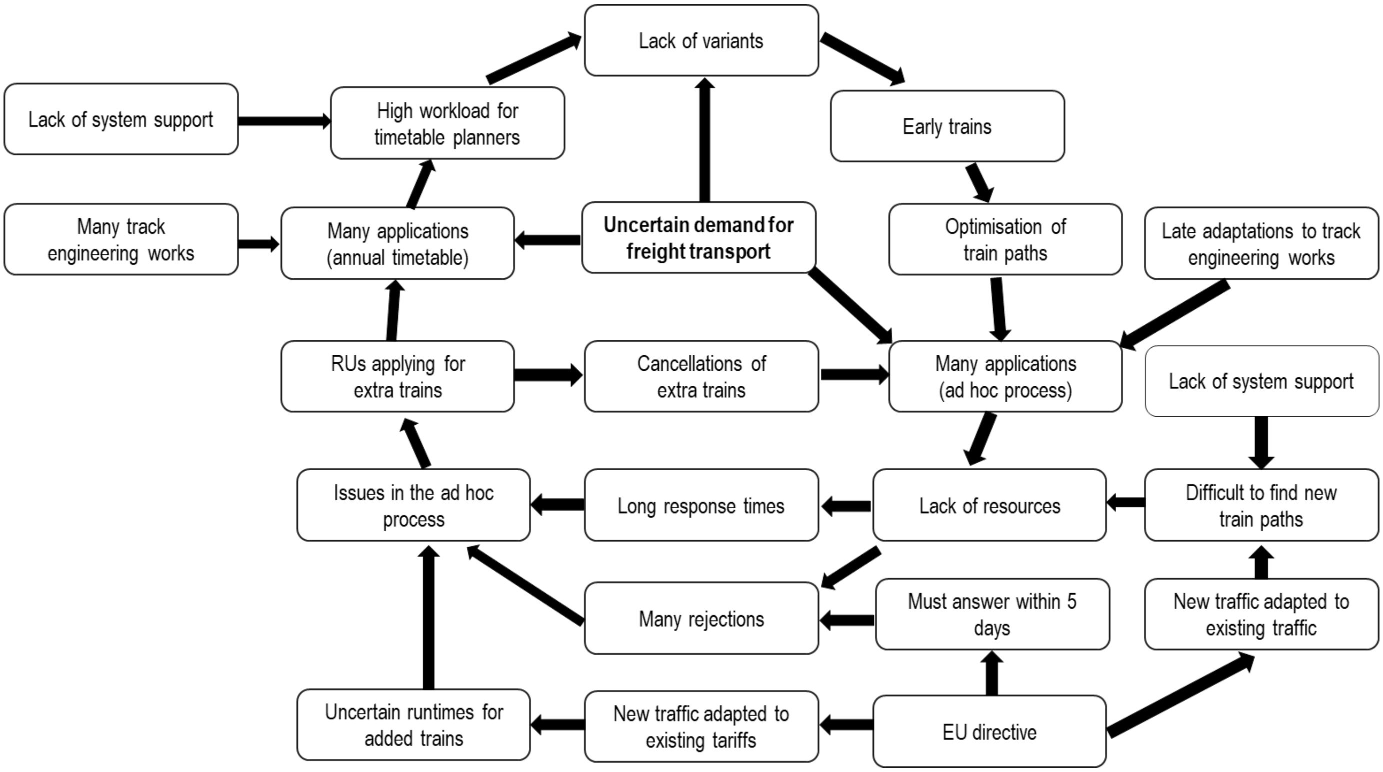

Fig. 11: An illustration of how different causes relate to one another.

The benefit of there being a loop, illustrated in Fig. 11, is that it can be interrupted at many different stages. For instance, if more of the required capacity for freight trains was instead allocated on a shorter time horizon, within the ad hoc process (such as is suggested in [36]), it is easy to see how the deviations would quickly decrease. This could be done through *reserving capacity* in the annual process, instead of trying (in vain) to *schedule* all freight trains in detail on an annual basis. The reserved capacity can then be applied for by RUs when necessary. As a result, the complexity of the scheduling could decrease. This could make it easier to plan the required trackwork possessions, with less need of train path variants and with greater certainty that the trains that are scheduled will run.

In addition, with fewer trains in the annual timetable, it would be easier to adapt the ad hoc applications for freight trains to existing traffic, and it would be easier to create new train paths for them. This could alleviate the lack of resources, lead to shorter response times and less rejections, thus greatly reducing the issues in the ad hoc process. With more confidence in this process, there would be less need for RUs to apply for extra trains that they will not use. This in turn could reduce the number of applications and the workload for the planners. The allocated train paths will then be better suited to the conditions when the train will run, such that there is no longer a lack of variants.

With more appropriate variants, there is less reason for trains to run ahead of or faster than scheduled, and train paths will be better optimised for the circumstances. As they are created for more specific times, this would result in a reduced need to reschedule trains to possessions, as these will already be known. With this need gone, one of the main reasons for ad hoc requests would be eliminated, and there are instead more resources left to schedule new freight trains, enabling shorter response times and lower rejection rates, in a virtuous circle. This is one example of how a change in one part of the process could cause a ripple effect and lead to a reduction of all four types of deviations.

Other interventions would be to improve the system support for timetabling at the IM (this is in progress, but delayed), to employ many more timetable planners, and/or to find a more predictable way to schedule trackwork possessions (perhaps such as in [43]), such that less adaptations of train paths need to be made for that reason. All of these are likely to be beneficial but shifting a greater part of the scheduling of freight trains into the ad hoc process rather than the annual timetabling process would be a relatively easy policy change and does not require any of the other interventions to yield benefits. The underlying unpredictability of freight transport demand itself is unlikely to change any time soon.

X. Conclusions

In this paper, we have studied how and why freight trains deviate from timetables, with a particular focus on the process of capacity allocation.

Once the annual timetable has been set, we have seen that adding trains is difficult but rather common, some 20%

of freight trains were added in this way. Cancellations are still more frequent, affecting about 35% of scheduled freight trains, mostly upon request by RUs.

The most common cause for freight train delays is that they are late from the yard or depot. 40% of freight trains in our case study begin in this way. Once a train is in the network, the cause of most delays is unknown, but 30% of runtimes and 20% of dwell times are delayed. These are averages, the probabilities vary across train IDs, geography, and day to day, essentially following normal distributions.

More commonly, freight trains run ahead of schedule. 80% ready to depart early, and 60% do so. Once in the network, 40% of runtimes and 75% of dwell times are shorter than scheduled. Again, these are averages, with large variations across train IDs and geography, and to a less extent between days. The reason that trains are ahead of schedule is essentially that timetables are not sufficiently adapted to the actual operational conditions.

Analysing the root causes for the deviations, we identify links and feedback loops between different causes. One type of deviation can lead to another one, which causes another, and so on. This suggests that improvements on one dimension could have effects on others too. For instance, adjusting how capacity for freight trains is allocated could lead to fewer deviations and a higher punctuality.

The study suggests that the underlying unpredictability of freight transport demand is inevitable, and that the process of allocating capacity should be further adapted to the specific requirements of freight transportation. One way to do this is to *reserve* more capacity for freight trains in the annual process, which can instead be allocated in the ad hoc process (or another short time horizon). In line with previous research on incremental planning and similar LEAN approaches, this could lead to more appropriate timetables for freight trains, with less deviations and a reduced amount of (wasted) planning effort that currently goes into first creating and then continuously readjusting the annual timetable. Another recommendation is to use IT support that is better able to handle multiple variants of timetables and more rapid, possibly automated, processing of requests for new or modified train paths. This would also help make the scheduled train paths more relevant during actual operations. In turn, this might make railways more attractive for freight transports, helping to induce a modal shift away from road towards rail instead.

XI. Acknowledgements

The authors thank the Swedish Transport Administration for funding and fully collaborating in this study. Additional funding was received from the Swedish Civil Contingencies Agency (MSB) through the CRITFLOW project. The authors also wish to acknowledge the reviewers' constructive comments that helped develop this paper. We are grateful to Kenneth Håkansson for his guidance and support, to Elin Björnsson and Roar Hermo for providing the necessary data. Our thanks also go to Lena Winslott Hiselius, Martin Aronsson, Ruben Kuipers, Daria Ivina, Sara Gestrelius, Emma Solinen, Michelle Ochsner and Kah Yong Tiong for constructive comments on earlier versions of this manuscript.

XII. Declarations of interest

None.

XIII. References

[1] C. Macharis and Y. M. Bontekoning, "Opportunities for OR in intermodal freight transport research: A review," *Eur. J. Oper. Res.*, vol. 153, no. 2, pp. 400–416, 2004, doi: 10.1016/S0377-2217(03)00161-9.

[2] T. A. Mathisen and T. E. S. Hanssen, "The academic literature on intermodal freight transport," *Transp. Res. Procedia*, vol. 3, no. 2352, pp. 611–620, 2014, doi: 10.1016/j.trpro.2014.10.040.

[3] J. Törnquist, "Computer-based decision support for railway traffic scheduling and dispatching: A review of models and algorithms," in *5th workshop on algorithmic methods and models for optimization of railways (ATMOS'05)*, L. G. Kroon and R. H. Möhring, Eds. Dagstuhl: Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2006.

[4] G. Caimi, L. Kroon, and C. Liebchen, "Models for railway timetable optimization: Applicability and applications in practice," *J. Rail Transp. Plan. Manag.*, vol. 6, no. 4, pp. 285–312, 2016.

[5] R. M. P. Goverde *et al.*, "A three-level framework for performance-based railway timetabling," *Transp. Res. Part C Emerg. Technol.*, vol. 67, pp. 62–83, 2016, doi: 10.1016/j.trc.2016.02.004.

[6] C.-W. Palmqvist, "Delays and Timetabling for Passenger Trains," Lund University, 2019.

[7] J.-E. Nilsson and M. Andersson, "Knapphet på järnväg: sammanställning av två promemorior," Linköping, 2011.

[8] Trafikanalys, "Tåglägen, gods och trängsel på spåren," Stockholm, 2016.

[9] European Union, *Directive 2012/34/EU of the European Parliament and of the Council of 21 November 2012 establishing a single European railway area, OJ L 343, 14.12.2012*. European Union, 2019.

[10] J. Rohde and M. Wagner, "Master planning," in *Supply chain management and advanced planning*, H. Stadtler and C. Kilger, Eds. Berlin: Springer-Verlag, 2008, pp. 161–180.

[11] H. Stadtler and B. Fleischmann, "Hierarchical planning and the supply chain planning matrix," in *Advanced planning in supply chains*, H. Stadtler, B. Fleischmann, M. Grunow, H. Meyr, and C. Sürie, Eds. Berlin: Springer-Verlag, 2012, pp. 21–34.

[12] M. Forsgren, M. Aronsson, and S. Gestrelius, "Towards shorter lead times in railway timetabling in Sweden," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 2013, pp. 1053–1058.

[13] H. Pouryousef, P. Lautala, and T. White, "Railroad capacity tools and methodologies in the U.S. and Europe," *J. Mod. Transp.*, vol. 23, no. 1, pp. 30–42, 2015, doi: 10.1007/s40534-015-0069-z.

[14] G. S. Roscoe, C. T. Dick, D. Ph, and P. E, "Incremental Free Path Heuristic Algorithm for Deadlock- Free Train Dispatching on Long Complex North American Rail Corridors," no. 1989, 2019.

[15] J. Fischhaber and A. Imable, "Tågplanens förändring," Lund University, 2016.

[16] J. Törnquist Krasemann, "Analys av planerad och utförd trafik," 2015.

[17] P. Hellström, "Problems in the integration of timetabling and train traffic control," Uppsala, 2014.

[18] A. H. Halse, V. Østli, and M. Killi, "Å måle det upresise: Årsaker til og konsekvenser av togforsinkelser," Oslo, Norway, 2015.

[19] N. A. Krüger and I. Vierth, "Precautionary and operational costs of freight train delays: a case study of a Swedish grocery

company," *Real-Time Syst.*, vol. 51, no. 2, pp. 1–10, 2015, doi: 10.1007/s12544-015-0155-7.
[20] Trafikverket, "Årsstatistik för punktlighet 2019," Borlänge, 2020. [Online]. Available: https://www.trafikverket.se/om-oss/var-verksamhet/Rapporter/Manatlig-trafikrapport/Transport-pa-jarnvag-i-ratt-tid/Statistik-for-punktlighet/arsstatistik/arsstatistik-for-punktlighet-2019/.
[21] B.-L. Nelldal, "Större trafikavbrott vid Sveriges järnvägar 2000-2013 och dess effekter på transportkunderna," Stockholm, 2014.
[22] Myndigheten för samhällskydd och beredskap, "Kostnader för störningar i infrastrukturen : Metodik och fallstudier på väg och järnväg," Karlstad, 2015.
[23] Trafikanalys, "Bantrafik 2019," Stockholm, 2020.
[24] A. Økland and A. Ekambaram, "Beregning av punklighet og pålitelighet for godstransport på jernbane - PUSAM teorigrunnlag," Trondheim, Norway, 2010.
[25] C.-W. Palmqvist, N. O. E. Olsson, and L. Hiselius, "Some influencing factors for passenger train punctuality in Sweden," *Int. J. Progn. Heal. Manag.*, 2017.
[26] F. Khoshniyat and A. Peterson, "Improving train service reliability by applying an effective timetable robustness strategy," *J. Intell. Transp. Syst. Technol. Planning, Oper.*, vol. 21, no. 6, pp. 525–543, 2017, doi: 10.1080/15472450.2017.1326114.
[27] Y.-C. Lai, M. H. Dingler, C.-E. Hsu, and P.-C. Chiang, "Optimizing train network routing with heterogeneous traffic," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2159, no. 1, pp. 69–76, 2010.
[28] A. Landex, "Methods to estimate railway capacity and passenger delays," Technical University of Denmark, 2008.
[29] M. J. C. M. Vromans, R. Dekker, and L. G. Kroon, "Reliability and heterogeneity of railway services," Rotterdam, ERS-2003-090-LIS, 2006. doi: 10.1016/j.ejor.2004.10.010.
[30] R. Hedström and O. Eriksson, "Godstågs avvikande hastighet - Analys av förekomst, orsaker och konsekvenser," Linköping, 2018.
[31] M. Wahlborg, A. Peterson, C. Schmidt, L. Jalili, and L. Gruosso, "D3.1 - Final pre-study for an improved methodology for timetabling planning including state-of-the-art and future work plan," 2018.
[32] M. Joborn, "Oplanerade stopp och potential for målpunktstyrande system," Stockholm, 2014.
[33] Trafikverket, "Network statement 2019," Borlänge, 2017.
[34] M. Aronsson, M. Forsgren, and S. Gestrelius, "The Road to Incremental Allocation & Incremental Planning," Stockholm, 2012. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:ri:diva-24071.
[35] S. Gestrelius, M. Bohlin, and M. Aronsson, "On the uniqueness of operation days and delivery commitment generation for train timetables," 2015, [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:ri:diva-24429.
[36] M. Aronsson, "Reservkapacitet i tågplaneprocessen: Förstudie," Stockholm, 2019. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:ri:diva-38524.
[37] RailNetEurope and Forum Train Europe, "Description of the Redesigned Timetabling Process (TTR)," 2017.
[38] R. K. Yin, *Case study research and application – Design and methods*, 6th ed. London: SAGE Publications Ltd, 2018.
[39] M. Veiseth, P. M. Hegglund, I. Wien, N. O. E. Olsson, and Ø. Stokland, "Development of a puncuality improvement method," *TMQ J.*, vol. 23, no. 3, pp. 268–283, 2011, doi: 10.1108/17542731111124334.
[40] B. Nyström, "A methodology for measuring the quality of deviation reporting: Applied to railway delay attribution," *Int. J. Qual. Reliab. Manag.*, vol. 25, no. 7, pp. 656–673, 2008, doi: https://doi.org/10.1108/02656710810890863.
[41] M. F. Gorman, "Statistical estimation of railroad congestion delay," *Transp. Res. Part E*, vol. 45, pp. 446–456, 2009, doi: 10.1016/j.tre.2008.08.004.
[42] Banverket, "Testvecka - släpp inte ut tidiga tåg 2-6 mars," 2009.
[43] T. Lidén, *Concurrent planning of railway maintenance windows and train services*. 2018.