Towards a robust traffic timetable for the Swedish Southern Mainline

A. Peterson
Department of Science and Technology, Linköping University, Sweden

Abstract

We identify the on-time performance as a key to evaluate a railway timetable's robustness to disturbances and evaluate the on-time performance for two single services on the Swedish Southern Mainline for the autumn period 2011. We analyse the punctuality by studying how the performance develops en route. Typically the time spent in the stations is underestimated, which partly is compensated for by time margins along the line, giving rise to a sawtooth formed delay muster with an increasing trend. The standard deviation in the delay reports seems to be a good indicator for the precision in the traffic. In this material it is almost linearly increasing with a good minute per hour scheduled running time. Two attempts are made to change the timetable for a better performance. By using socio-economic values a customer-oriented description of the current timetable is calculated as a trade-off between high punctuality and short travel time. This timetable does not affect any other traffic. Minor effects to other traffic would be the result if the existing margins are re-distributed to better match the demand en route in some sense. As long as no more time supplement is added, we will, however, not improve on the punctuality to the final destination. For future research is left, how the decrease in precision can be accounted for already in the timetable construction.
Keywords: railway timetable, punctuality, on-time performance, robustness.

1 Introduction

A tendency in the Swedish railway network for the last twenty years is a growing demand for capacity and train slots, which no longer can be fully met. During 2011 in total 188 million railway journeys were produced in Sweden corresponding to 11.4 billion passenger-kilometres [1]. Solely for the last five years, this means an increase of the passenger transport by more than 10% [1].

WIT Transactions on The Built Environment, Vol 127,
www.witpress.com, ISSN 1743-3509 (on-line)
doi:10.2495/CR120401

During the process of constructing the national traffic timetable for 2011, the capacity limit was reached at several crucial links in the mainline network.

High capacity utilization means a timetable which is sensitive to disturbances. UIC [2] has identified 'stability' as one of four parameters underpinning capacity. The other three are the number of trains, the average speed and the heterogeneity in the traffic. Qualitatively spoken, a dense, slow and widely diversified traffic makes it hard to achieve a traffic timetable, which is robust to disturbances.

There are different ways of defining and measuring robustness. One illustrative example is given by Kroon *et al.* [3], who describe the robustness of a timetable as its capability to "*deal as well as possible with relatively small disturbances in the real-time operations*" and identify three imposed effects, namely that: initial disturbances can be absorbed to some extent so that they do not lead to delays, there are few knock-on delays from one train to another, and delays disappear quickly, possibly with light dispatching measures.

A straightforward way to evaluate a timetable's robustness to the disturbances, that de facto occur (in practice, have occurred) is to study the on-time performance. Traditionally this is measured per service in terms of punctual arrivals to the terminal destination, with some tolerance level. It is reasonable to assume that the punctuality statistics for the services' terminal destinations alone does not contain enough information to fully evaluate the robustness in a given traffic timetable.

This paper comprises an in-depth study of the en route on-time performance for two northbound long-distance trains on the Swedish Southern Mainline. The purpose is to understand and visualize how the on-time performance develops en route. Whished for is information that can be used to construct a more robust timetable, both in a customer-oriented and an operational point of view.

Clearly, the robustness in a traffic timetable is a result of relations between all activities in the timetable. An increase of the time distance between two consecutive train movements at one point in general also means a decrease somewhere else. Therefore, a study of single services is of limited value for improving the robustness in the complete timetable. The level of ambition in this paper is focused on measures that can be taken for single services, without disturbing other traffic. We show how the observed on-time performance could be communicated to the travellers from a socio-economic point of view. We also analyse the effect of re-distributing the existing margin time according to some principles.

The reminder of this paper is organized as follows. In Section 2, we describe the two studied services and their respective timetables. Section 3 contains the on-time performance study and in Section 4 we analyse how it can be improved without disturbing other traffic. Section 5 concludes the paper.

2 Preliminaries

The Swedish Southern Mainline is a more than 500 Km double tracked railway line connecting Stockholm with Malmö and Copenhagen. The traffic along the line is dense and widely diversified. It includes long-distance services of both intercity and high-speed types. Several regional services and three commuter train systems are using parts of the line, as do cargo trains. Andersson *et al.* [4] have studied the traffic performance for October 2010 on this particular line.

This study concerns two northbound X2000 services, run with the fastest type of train in Sweden with a maximum speed of 200 km/h. One of the services (No. 500) is an extra fast connection with only two intermediate stops, whereas the other (No. 530) is a standard service with eight intermediate stops. The data material for the autumn 2011 covers all 85 working days from 15th August (first day after summer reduction) to 9th December (last working day before a major timetable shift).

Overall this was a disturbance free period, and no cancelled or rerouted trains were reported. Merely three days are protocoled, on which the trains for both services have arrived to the terminal destination more than 20 minutes delayed, indicating some general problems. Throughout the studied period infrastructural maintenance work has led to single track traffic on parts of an 18 Km long stretch north of Nässjö (roughly at scheduled running time minute 120). The exact position has varied over the period, but we assume that all studied trains are affected in a comparable way.

Cargo traffic along this line is in general very unpunctual, which means that overtakings of cargo trains most likely did not take place where they were scheduled to. Some scheduled overtakings might not have taken place at all, due to a large time difference, but on the other hand, there were unscheduled overtakings of cargo trains, being far ahead of or behind their respective schedules. In general, the cargo traffic is denser on the southern half of the line.

Next follows a basic description of the two services and their respective timetable, for a complete schedule, we refer to [5].

2.1 Service No. 500

For service No. 500 the first checkpoint is the arrival time at Malmö central station, which is scheduled at 05.48 (24 hour time format). The train departures at 05.51, has scheduled stop for passenger exchange in Lund and Hässleholm and arrives at the terminal destination, Stockholm central station, at 10.01. Total travel time is 253 minutes.

The service is scheduled to overtake two other passenger services during their respective stop for passenger exchange in Hässleholm (scheduled running time minute 50) and Linköping (minute 161), respectively. These overtakings are critical in the sense that they cannot be moved to other stations without a significant time loss. Furthermore, two cargo trains are scheduled to be overtaken on a regular basis and some additional on parts of the studied period, certain weekdays or just single days.

2.2 Service No. 530

The study of service No. 530 starts at 09.10 with the scheduled arrival at Malmö central station. The service is scheduled to departure Malmö at 09.17, has stop for passenger exchange in Lund, Hässleholm, Alvesta, Nässjö, Mjölby, Linköping, Norrköping and Södertälje Syd, and should arrive at the terminal destination, Stockholm central station, at 13.39. Total travel time is 269 minutes.

The service is scheduled to overtake one other passenger service, namely the intercity service No. 200, at its 24 minutes technical stop (no passenger exchange) in Vikingstad. The overtaking is non-critical and can, if advantageous, easily be moved to other stations. As for No. 500, also for No. 530 two cargo trains are scheduled to be overtaken on a regular basis and some additional on parts of the study period, certain weekdays or just single days.

3 On-time performance

Statistics for the on-time performance in Sweden is normally presented in terms punctual arrivals to the terminal destination. Here, we will also explore how the performance develops en route and analyze the standard deviation in the performance statistics.

3.1 Punctual arrivals to the terminal destination

Punctuality is one way of describing on-time performance. The Swedish Transport Administration [6] considers a passenger train as *punctual* if it arrives to its terminal destination with at most five minutes delay. During the period August–December 2011 the punctuality for all Swedish long-distance services of high-speed type was 83.4% (ibid.).

For the studied period, the punctuality for service No. 500 was 31.0%. The average delay when arriving Stockholm was 16.5 minutes. An arrival delay exceeding 20 minutes was observed on 15 of the 85 studied days and the average delay for these days was 56.2 minutes. For one day observations are missing Aiming for a robust timetable, these outliers are of no interest. If a train is more than 20 minutes delayed, other traffic surely will be prioritized and the timetable has lost is meaning as a tool for train dispatching. Throughout the rest of this paper, therefore, we only consider those 69 days, on which the train has reached Stockholm with at most 20 minutes delay. For these days, the average delay is 7.9 minutes and the median delay is 9 minutes.

The punctuality for service No. 530 when reaching the terminal destination was 58%, and the average delay was 8.1 minutes. On 10 of the 85 studied days the arrival delay exceeded 20 minutes, and the average delay for these 10 days was 36.9 minutes. As for service No. 500, from now on we will only consider those days, on which the train has reached Stockholm with at most 20 minutes delay. For these 75 days, the average delay is 4.3 minutes and the median delay is 2 minutes.

3.2 On-time performance en route

Beside the arrival time to the final destination, it is also interesting to see how the on-time performance develops en route. The data material contains departure (passing) times for 78 stations (checkpoints). In addition, the arrival time is collected at stations where there is a scheduled stop for passenger exchange. In total this means 82 checkpoints for No. 500, and 88 checkpoints for No. 530. All data is collected with minute precision, but for convenience the time presented here has been computed as a simple rolling average consisting of the passing time for three consecutive checkpoints.

The on-time performances en route for service No. 500 is shown in Figure 1.

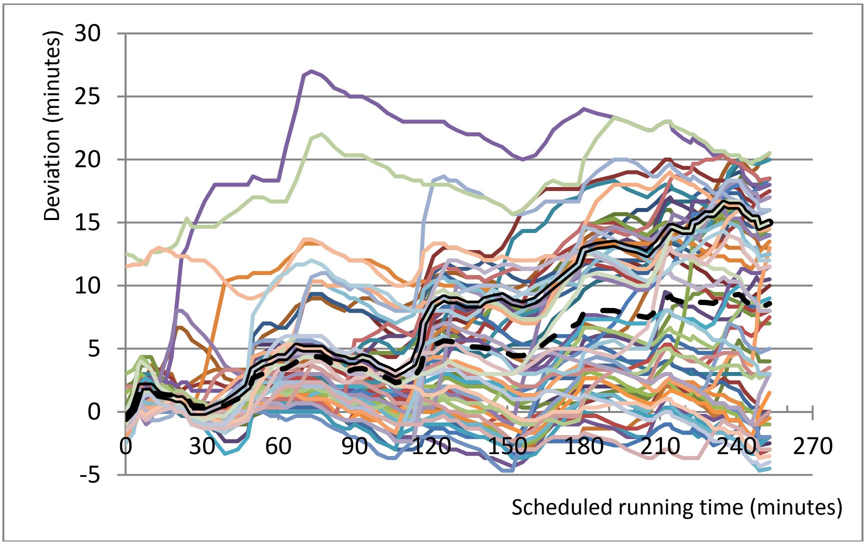

Figure 1: En route deviation from the timetable for service No. 500, including average value (dashed line) and the 75 percentile (double line).

As is mentioned before, service No. 500 is scheduled to overtake other trains in Hässleholm (scheduled running time minute 50) and Linköping (minute 161). From Figure 1 we can trace the discrete choice made by the dispatchers in cases where No. 500 is delayed, leading to a clustering (white areas) after these points.

The on-time performance en route for service No. 530 is depicted in Figure 2, below. This service shows one clustering tendency similar to that of No. 500. Clustering occurs for trains being more than 6 minutes delayed when departing Mjölby (minute 153), which are leaving just after a commuter train and thereby lose more time. The commuter train is scheduled to start from Mjölby 4 minutes after service No. 530.

Comparing Figures 1 and 2, both services have sawtooth formed performance curves, especially clear for trains being nearly on-time. Delays typically occur at distinct points, mainly stations, and are recovered gradually until next distinct delay. This muster primarily deflects the strategy to spread the inserted margins along the line. Generally, the stopping time at stations seems to be underestimated.

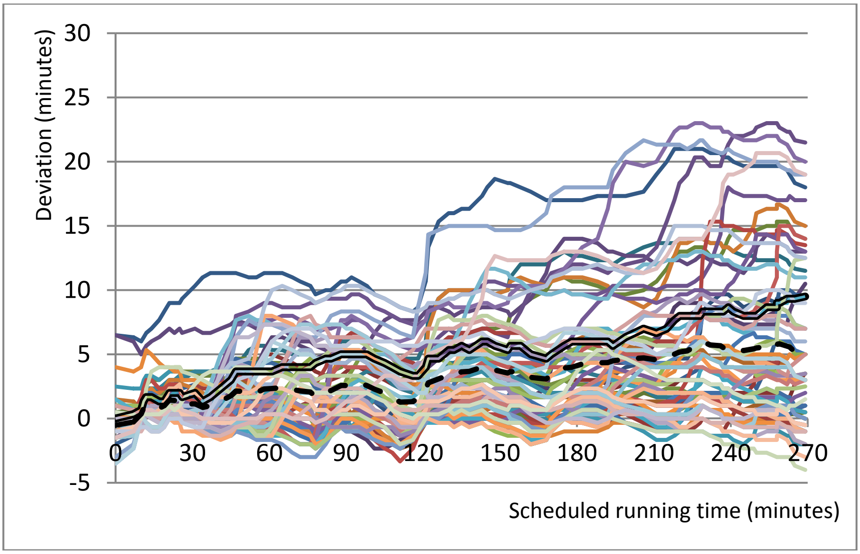

Figure 2: En route deviation from the timetable for service No. 530, including average value (thick dashed line) and the 75 percentile (double line).

The clustering effects in total are quite small and last for 30–60 minutes. They do not seem to affect the arrival time to the final destination. It should be reminded that also other traffic might be unpunctual, which evens out parts of the effect.

3.3 Precision in the train operations

The standard deviation for passage times at the checkpoints is an expression for the precision in the operational traffic. A large value indicates that there is hard to predict the exact passage time.

In Figure 3, below, the standard deviation in the on-time performance material for the services No. 500 and No. 530 has been computed. For comparison, the mean curves from Figures 1 and 2 also are re-plotted.

According to Figure 3, the development of the standard deviation, unlike that of the mean deviation, for both services is almost non-decreasing and quite well approximated by its linear trend line. Apparently the given margins can decrease the (average) delay, but are not enough to improve the precision. The trend line for No. 500 starts at 2.4 minutes and increases by 2.1%, whereas that for No. 530 starts at 1.2 minutes and increases by 1.9%. Roughly spoken, every train seems to lose some 2% precision in the on-time performance per time unit, equivalent to 1.2 minutes per hour scheduled running time.

An interesting question is whether this observation also holds for other services and other railway lines. A preliminary study, not presented here, confirms the observation for a regional service ('Öresundståg'), which serves the south most third of the Southern Mainline (Malmö–Alvesta). A pure local

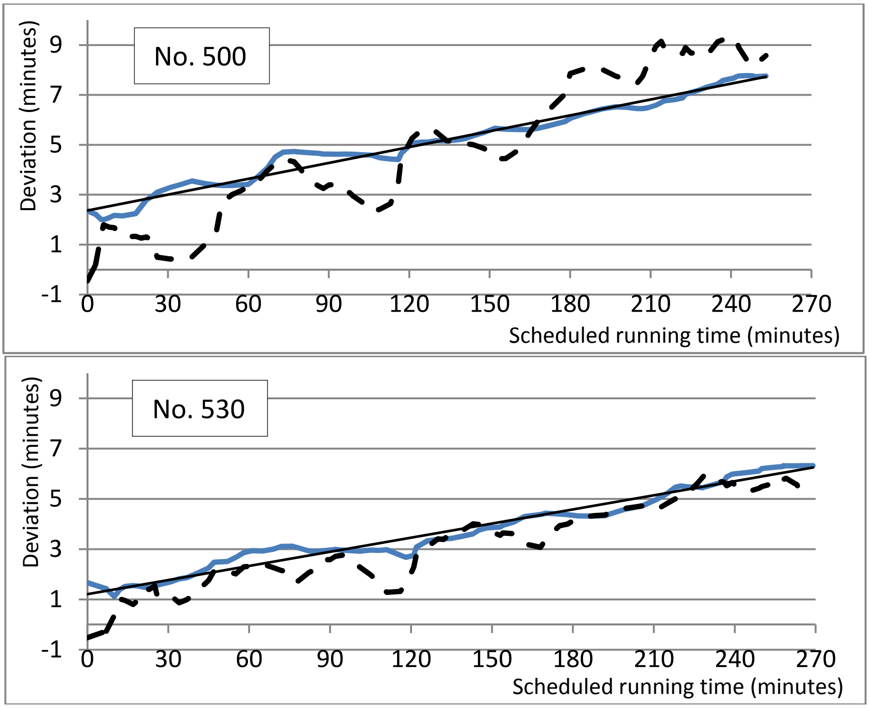

Figure 3: Standard deviation at the checkpoint passages (solid curve) and its trend line (straight line), compared to the average deviation from the timetable (dashed curve) for the services No. 500 and No. 530.

service ('Pågatåg') with very frequent stops, however, shows another behaviour, where precision is gained at the stops and has no increasing trend.

4 Improved on-time performance for a single service

Timetable construction is in its nature a combinatorial puzzle involving a whole railway network with a complete set of train movements to be scheduled. This study is limited in the way, that only the performance for single services is considered.

4.1 A customer-oriented socio-economic timetable

Given the current on-time performance, as it is presented in Section 3, what timetable should be communicated to the travellers? Adding time to the announced arrival times would make the customer reaching the destination less delayed. A longer journey time, on the other hand, would make the service less attractive.

There are different ways of estimating socio-economic values for travel time and delay time. According to key-figures, which are generally used for evaluation of socio-economic effects in the public transport in Sweden [7], delay time should be given the weight 1.87 relative to travel time. Interpreting being ahead of schedule as unused travel time with weight 1, and being behind schedule as delay with weight 1.87, we can ad hoc fit a timetable curve to the actual on-time performance, presented in Section 3.2. Postponing an announced departure time would change the production plan for the remaining journey, and hence affect the on-time performance on subsequent stations. Therefore only the adjustment of the arrival time for stations were passengers are leaving the train is considered. The results for the services No. 500 and No. 530 are presented in Table 1 and 2, respectively.

Table 1: Socio-economic timetable for service No. 500.

		Present timetable	New timteble
Departure	Malmö	05:51:00	
Arrival	Lund	06:01:00	06:01:40
Arrival	Hässleholm	06:36:00	06:36:20
Arrival	Stockholm	10:01:00	10:09:30

Table 2: Socio-economic timetable for service No. 530.

		Present timetable	New timteble
Departure	Malmö	09:17:00	
Arrival	Lund	09:27:00	09:27:40
Arrival	Hässleholm	09:55:00	09:56:20
Arrival	Alvesta	10:32:00	10:32:40
Arrival	Nässjö	11:06:00	11:06:20
Arrival	Mjölby	11:43:00	11:45:20
Arrival	Linköping	11:58:00	11:59:40
Arrival	Norrköping	12:22:00	12:24:40
Arrival	Södertälje Syd	13:19:00	13:22:00
Arrival	Stockholm	13:39:00	13:42:00

At the stops for service No. 500 (Table 1) in Lund and Hässleholm there is not much to be changed. For the 205 minutes non-stop journey from Hässleholm to Stockholm, however, one should apparently add some extra 8–9 minutes.

For service No. 530 (Table 2) the calculation means a postponing of the arrival time at the intermediate stations with 1–3 minutes. Since the announced stopping time for these stations typically is 2 minutes, a practical interpretation of the result in Table 2 would be to set arrival time equal to departure time for every intermediate station. To this some extra three minutes should be added to the arrival time at the terminal destination.

The calculation here of course is dependent of the choice of relative weights. Other estimations of the socio-economic values would have given a different result. The major contribution in Table 1 and 2 is not the particular revision of the timetable, but the idea to separate the timetable, which is communicated to the travelers, from that, which is used for the operative production and to use a socio-economic value to calculate the difference.

4.2 Re-distribution of margins

By subtracting the shortest travel time (including scheduled stopping time for passenger exchange) from the schedule, we get an expression for the inserted margins. Analogously, if this shortest travel time is subtracted from the observed average travel time, we get an expression for the consumed margins.

For service No. 500 the total scheduled margin time summarizes to 978 seconds, compared to a total travel time of 15 180 seconds (253 minutes), whereas the average consumed margins in total are 1 644 seconds. For service No. 530 the total scheduled margin time is 904 seconds, compared to a total travel time of 16 140 seconds (269 minutes), whereas the average consumed margins summarize to 1 187 seconds. In Figure 4, below, the accumulated values for inserted and consumed time margins are compared.

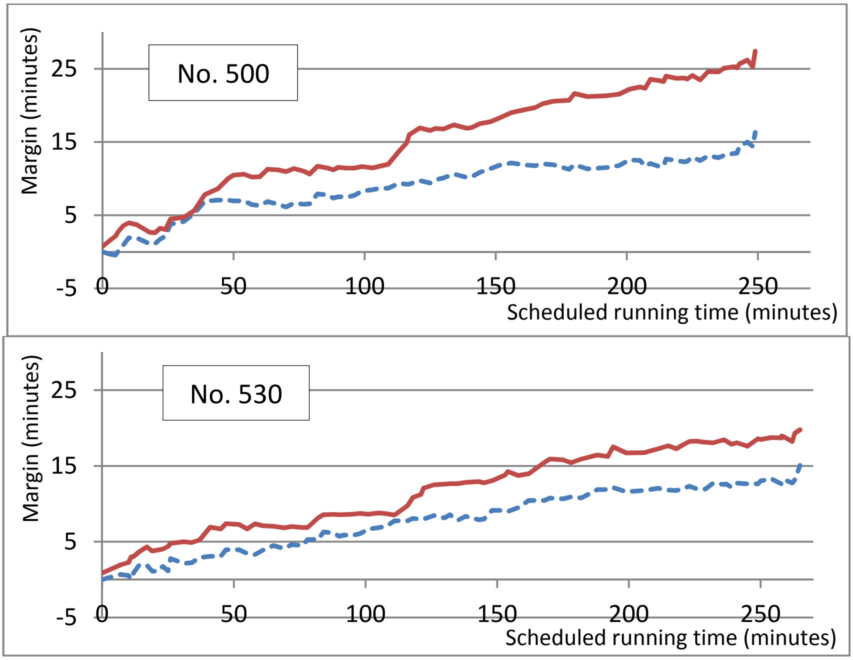

Figure 4: Average consumed (solid curve) versus scheduled (dashed curve) margins for the services No. 500 and No. 530, accumulated values.

The dotted curves in Figure 4 are entirely under the dashed ones, meaning that less margin time is scheduled than consumed. To meet the average demand, in total a margin of 666 seconds and 283 seconds should be added to the schedules for services 500 and 530, respectively. Such a change would, however, also affect the running time and interaction with other traffic in a non-controlled manner. Instead we will discuss three ways of re-distribute the current margin, assuming that such changes are small enough not to disturb other traffic.

If the scheduled total margin time is re-distributed as to reproduce the consumed margin in, for example, a least square meaning, we will insert the available margin time following the average demand as long as is possible, leaving no margin at all for the last part. Doing this in Figure 4 would be to let the dotted curves follow the increase of the solid curves as long as possible and being a straight line for the remaining part. For No. 500 the last margin time will be placed at Fliseryd (scheduled running time minute 121), and for No. 530 the last margin will be placed at Linköping (minute 170).

A second attempt to re-distribution is to insert the available margin in proportion to the consumed margin. Doing this in Figure 4 would be to make the dotted curves mimic the shape of the solid curves but ending in the same point. The change would be quite small, and basically concentrate more margin time at the delay increase around minute 110, which is induced by the one-track stretch at Nässjö.

A third attempt for a re-distribution is to insert the available margin time evenly distributed over time. Doing so would give straight lines from the first to the last point of the two dotted curves in Figure 4. From this third attempt we observe that overall more margin per time unit is added to the schedule in the beginning than in the end of this particular line. This observation, however, also seems to hold for the consumption.

5 Conclusions and future research

Part of the aim with this study is to understand and visualize how the punctuality develops en route. A sawtooth formed performance curve can be identified, where the delay occurs at distinct points, typically stations with passenger exchange, and recovers along the line. The discrete choices, which are made by the train dispatchers to solve conflicts in delayed situations, can be traced, but the phenomenon is not dominating and evens out after 30–60 minutes.

Changes in the timetable in general affect several services. This study is limited to single services, which strongly bounds the possibilities to adjust the timetable for a better on-time performance. Arrival times communicated to the travelers are, however, independent from the timetable which is used in the operative traffic production. By weighting delay relative to announced travel time, we can make a simple trade-off between short travel time and high on-time performance. For the faster service No. 500 this means that some 8–9 minutes should be added to the arrival time at the final destination, and for the standard service No. 530 it is probably enough to set arrival time equal to departure time at intermediate stations and add some three minutes at the final destination. Other socio-economic values might give other results.

Small effects to the production timetable would be the result of re-distributing the existing margins according to any alternative principle, such as trying to mimic the average margin consumption muster or simply to assign the margins in proportion to the shortest travel time for each part of the route. An overall comparison of scheduled and consumed margin time shows that on average more margin time should be added to the schedule. Another observation is that, in relation to the running time, more margin time is consumed in the southern than in the northern part of the Southern Mainline.

Whereas the average on-time performance basically expresses where margins are scheduled and consumed, the standard deviation in the on-time reports gives a measure for the precision in the train operations. For both studied services this standard deviation is rather stably increasing and well approximated by a linear function. Per hour running time each service loses a good minute of precision in this sense. This conclusion is interesting. This standard deviation seems to be a good indicator for the robustness in the schedule.

A question for future research is how to account for the observed decrease in precision already in the construction of the timetable. One approach would be to accept the increase and incorporate it in the plan, meaning that the minimum required time slot to other trains must increase with the running time. Another approach is to insert more margin time, for example at the stations, such that trains depart more precise than they arrive. A consequence is, however, that even modest delayed trains must be kept still, waiting for the scheduled departure time. The latter approach is probably easier to implement for local services with frequent stops, each of which contains a certain margin. Future techniques for informing the driver in real-time about speed recommendations other than the maximum speed limit could also be useful for increasing the precision. Taking measures to improve precious performance, however, most likely decreases line capacity or it implies a higher homogeneity in the traffic.

Finally, it should be reminded that, this study is limited to two northbound long-distance services on the Southern Mainline. For higher confidence in the conclusions, more services and other lines should also be considered.

References

[1] Person- och godstransport på järnväg, 2011 kvartal 4, report, Trafikanalys: Stockholm, Sweden, 2012.

[2] UIC, *Capacity,* Leaflet 406, 1st ed., 2004.

[3] Kroon, L., Huisman, D. and Maróti, G., Optimisation models for railway timetabling, *Railway Timetable & Traffic*, eds. Hansen, I.A. and Pachl, J., Eurailpress: Hamburg, Germany, pp. 135–154, 2008.

[4] Andersson, E., Peterson A. and Krasemann, J.T., Robustness in Swedish railway traffic timetables, *The 4th International Seminar on Railway Operations Modelling and Analysis*, Rome, Italy, 2011.

[5] 'Tidtabellsbok': sammanställning av planenliga körplaner, Tågplan T11, Trafikverket, 2010.

[6] Transport på järnväg i rätt tid, Swedish Transport Administration, Online. http://www.trafikverket.se/Om-Trafikverket/Trafikverket/Manatlig-trafikrapport/Transport-pa-jarnvag-i-ratt-tid/
[7] Swedish Road Administration, Effektsamband för vägtransportsystemet: kollektivtrafik – effektkatalog, report, Vägverket: Borlänge, Sweden, 2008.