



Causal Analysis of Railway Running Delays

Cerreto, Fabrizio; Nielsen, Otto Anker; Harrod, Steven; Nielsen, Bo Friis

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Cerreto, F., Nielsen, O. A., Harrod, S., & Nielsen, B. F. (2016). *Causal Analysis of Railway Running Delays*. Paper presented at 11th World Congress on Railway Research , Milan, Italy.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Causal Analysis of Railway Running Delays

Fabrizio Cerreto^{1*}, Otto Anker Nielsen¹, Steven Harrod¹, Bo Friis Nielsen².

¹ Department of Transport, Technical University of Denmark, Lyngby, Denmark

² Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark

*Corresponding author: facer@transport.dtu.dk

Abstract

Operating delays and network propagation are inherent characteristics of railway operations. These are traditionally reduced by provision of time supplements or “slack” in railway timetables and operating plans. Supplement allocation policies must trade off reliability in the service commitments against service transit times and railway asset productivity. Methods to investigate the quality of supplement time allocation are necessary to reduce the behavioral response and the waste of resources.

This is a preliminary study that investigates train delay data from the year 2014 supplied by Rail Net Denmark (the Danish infrastructure manager). The statistical analysis of the data identifies the minimum running times and the scheduled running time supplements and investigates the evolution of train delays along given train paths.

An improved allocation of time supplements would result in smaller overall aggregate timetable supplement, reduced transport travel times, and higher productive utilization of train rolling stock. The study results will lead eventually to both better allocation of time supplements in timetable structures, and identification of areas that should be a high priority for correction.

1 Introduction

The railway industry commonly benchmarks itself through key performance indicators such as punctuality and reliability. These compact measurements express the quality of the service, meant as the ability to respect the schedule promised to the passengers. The running time supplement is one of the timetabling tools used to improve punctuality. This paper gives an overview of the running time supplement design and use in operation. It also presents a statistical approach to analyze historical data of train timekeeping in Denmark, in order to investigate the quality of the timetable supplement allocation. The purpose is to present different strategies for the design of timetable supplements and to assess their impact on punctuality.

With the objective of evaluating the effectiveness of the slack currently scheduled in train paths, this paper proposes statistical methods to quantify the running time supplement and compare it with the delay evolution through the paths. It is possible to identify areas where the running times supplement is not used and therefore wasted, and sections of the train paths where delays are not recovered, suggesting a lack of running time supplement.

1.1 Punctuality, primary delays and secondary delays

Punctuality and delays are well known general concepts, but their definition and computation method vary among countries and railway companies. *Punctuality* refers to the number of trains that are not delayed,

compared to the total number of trains operated (Olsson & Haugland, 2004). It can be attributed to individual stations or trains over a period of time, or it can measure railway networks entirely or partly. Differences are found in the selection of the punctuality measurement points and of the trains to be included in the measure. Accordingly, also the punctuality targets are different in every country and may be train category specific (Schittenhelm, 2011). For example, punctuality is measured along the entire train path in Denmark, while only selected stations are counted in the Netherlands, Switzerland and Germany. Other countries measure punctuality only at the final destinations, like Italy and Norway. It is common to differentiate the punctuality target between passenger and freight trains. In several countries, passenger trains are further divided into long distance and regional/suburban trains.

Delays are positive deviations between the realized times and scheduled times of activities. In the literature, different classifications of delays are available. Most of the classifications distinguish between delays that are due directly to the variability of process times and delays that are originated by the subsequent conflicts in the actual operation (Goverde & Hansen, 2013). The *primary delays* are unexpected extensions of the planned times of the individual processes scheduled. For instance, equipment failures and large passenger flows generate primary delays. The *secondary delays*, on the other hand, are delays generated by operation conflicts, which are due to primary delays themselves. When a train is delayed, it needs to use infrastructure elements at different times than planned. A conflict arises when two or more trains request to use the same element at the same time: they will be queued by dispatching decisions, since only one train at time can use one element. The delay that generates from the queuing is called secondary delay (Cerreto, 2016). Dispatching decisions are crucial for the management of the delay propagation: Olsson and Haugland (2004) found that the dispatchers tend to use defined priority rules on single tracked lines or in cases of large delays. Personal judgment prevails, on the other hand, on double tracked lines or with small delays.

1.2 *Timetable supplement*

Scheduled times are usually longer than the minimum time required by processes. The difference between the scheduled times and the expected minimum realization times is referred to with different names by authors: *slack time*, *timetable allowance*, or *time supplement*. The timetable supplement is a tool that planners include in the timetables to compensate for natural variations of process times. It reduces the probability of generating primary delays, and it is expected to increase punctuality. On the other side, the supplement increases the travelling time and operating costs, resulting in a reduction of attractiveness and efficiency. To be effective and efficient, the timetable supplement should be properly dimensioned and distributed. Some strategies to allocate the supplement times are described below.

1.3 *Allocation strategies for the timetable supplement*

The allocation of the time supplement is a tradeoff between attractive travel times and timekeeping. General guidelines, built on empirical studies, are provided by the International Union of the Railways (UIC, 2000). The guidelines provide a fixed supplement to include in the train paths, proportional to the path length and increasing with the maximum speed, but they give no indication about the optimal distribution of the supplement along the paths. In addition, the recommendations are not mandatory and only suggest a minimum amount of supplement. Every railway planner has its own strategy to allocate the slack in the timetable and most western European countries use larger values than recommended. For example, the Danish railway Infrastructure Manager, RailNet Denmark, uses a flat distribution of the supplement on the regional and long distance trains, which, in some cases, doubles the UIC-recommended values. Condensation and compensation, instead, is the Swiss strategy for timetable

supplement allocation. The network is divided in zones according to the capacity utilization. The capacity bottlenecks areas are named *condensation zones*, where the supplement time is minimized to reduce the capacity utilization. In contrast, large supplement times are scheduled in the areas that are not capacity bottlenecks, called *compensation zones*, to recover possible delays accumulated in the previous condensation zones (Schittenhelm, 2011).

The national strategies for the supplement time allocation typically reflect the way the punctuality is measured: for instance, Denmark measures punctuality at all the stations and spreads the supplement along the train paths, except in the Copenhagen suburban railway network. Switzerland measures punctuality at larger stations and concentrates the supplement before those stations. Norway measures punctuality at the final destination and schedules large amounts of supplements in the last segments of the paths.

1.4 Effects of the time supplement

A properly designed time supplement should lead to better regularity of the scheduled process, improving railway punctuality. The relation between supplement time increase and punctuality improvement, though, is not straightforward. Carey (1998) formulated a behavioral response model to describe an observed phenomenon that reduced the benefit of supplement times. The main finding was that if more time is allowed to a process, the process self-adapts to the new schedule and takes a longer time on average. Train drivers tend to act slower in the departure procedures and to drive slower, passengers tend to take longer to board and alight, dispatchers tend to use the extra elasticity given by supplement times for train prioritization and delay management. In this sense, the supplement time could be thought as the capacity buffer between consecutive productive processes that absorbs the inherent variabilities in the production. The risk is to hide systematic failures in the process, which should be tackled individually to increase the reliability. The famous case of the Sunset Limited train in the USA is reported by Larson (1998): the train schedule included such a large slack time that it had been hiding wrong dispatching strategies for years, and was consistently attaining poor punctuality. Adding even larger supplement times did not improve the train punctuality, while the increased travel time reduced the attractiveness to passengers.

Carey's theoretical formulation (1998) finds a balanced supplement time allocation optimizing the total cost, which consists of the cost of the scheduled trips and the cost of the expected delays. The cost of the scheduled trips is proportional to the trip length, so it is minimized with short running times and, therefore, minimum running time supplements. The cost of the expected delays decreases non-linearly enlarging the supplement times. A reduction in the expected delay is mirrored by a relevant reduction in costs for fuel, equipment utilization, and overtime wages, as also mentioned by Johnston (2008).

2 Case study

New methods to design and allocate the running time supplements are subject of several studies with different methodologies. Our current research focuses on the statistical study of historical data to assess whether the timetable supplement in existing timetables fits the actual need and if it is properly used.

In the following subsection we present methods to investigate the actual use of the time supplement in train paths and compare it to the scheduled timetables through the statistical analysis of historical data from the daily operation.

2.1 Minimum running time

As described in the previous sections, the scheduled process times consist of the minimum process time and a slack time, or time supplement, to absorb inherent variations of the process time. Therefore, the planners need to compute the minimum running time between two stations. Different tools support this operation, each of them with a different approximation. Acceleration and deceleration models can provide approximated running time estimation, especially on simple plain lines. Micro-simulation of train motion allows a more accurate computation. It and can easily be combined with detailed infrastructure models to take into account slopes and the train's tractive effort and braking power (Cerreto, 2015). Real tests on the lines can be performed running trains on free tracks, but this type of tests is expensive and hard to realize. Each estimation method has its own uncertainties that should be evaluated.

We used historical data from RailNet Denmark from 2014, third quarter, to investigate the realized running times in the past. The actual minimum running times were identified on the railway line Copenhagen – Roskilde, the most congested line in Denmark. The investigation covers only the express trains (“Lyntog”) that stopped at the bigger stations. The scheme below outlines the 30 km long line and the stopping locations.

Figure 1 - Railway line Copenhagen - Roskilde. Express trains only stop at the major stations.

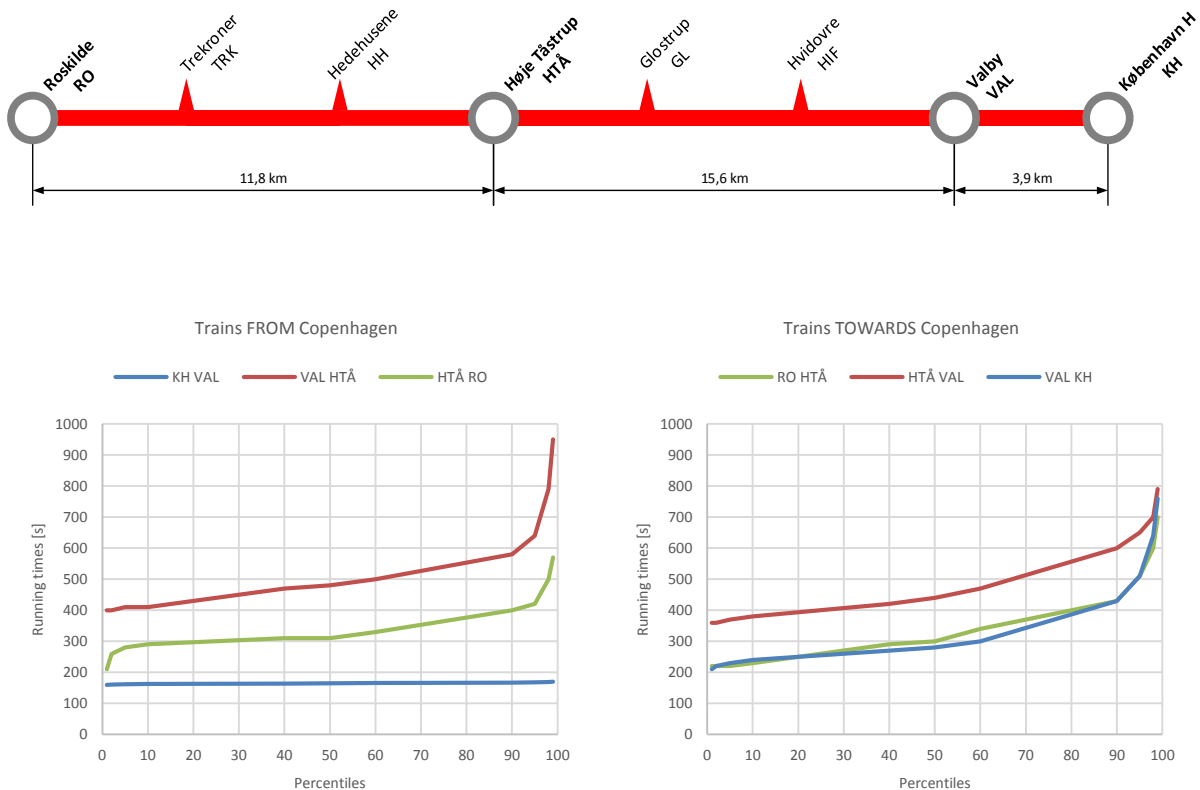


Figure 2 – Actual running time percentiles of express train on the railway line Copenhagen – Roskilde, divided by segment and direction.

The charts above represent percentiles of the actual running time distributions, divided by segments between stops, and by direction. The scheduled supplement time was filtered by referencing the minimum running time at the second percentile of the distributions. The second percentiles filtered well also running times that were too short, possibly due to the accuracy of the recordings or to random errors.

Differences in the distributions of the two directions are worth further investigation. The spread of running times by segment is considerably wider for trains from Copenhagen. The segment closest to Copenhagen changes significantly in stability between the two directions, being almost constant for trains from Copenhagen. The future investigation could highlight the existence of a behavioral response to supplement time allocated at the departure, as Copenhagen is often the origin of long distance trains.

The 2014 schedule varied considerably over the day, even for trains from the same category and scheduled with the same stopping patterns and rolling stock. The changes made it not possible to identify a unique running time supplement for each leg. Further research will investigate the variability of allocated supplement over the day. The supplement time will be estimated for individual trains through longitudinal statistics over the whole year.

2.2 Delay, delay variation and supplement times

Alongside the minimum running times, we compared the train delays at different stations to evaluate the delay development.

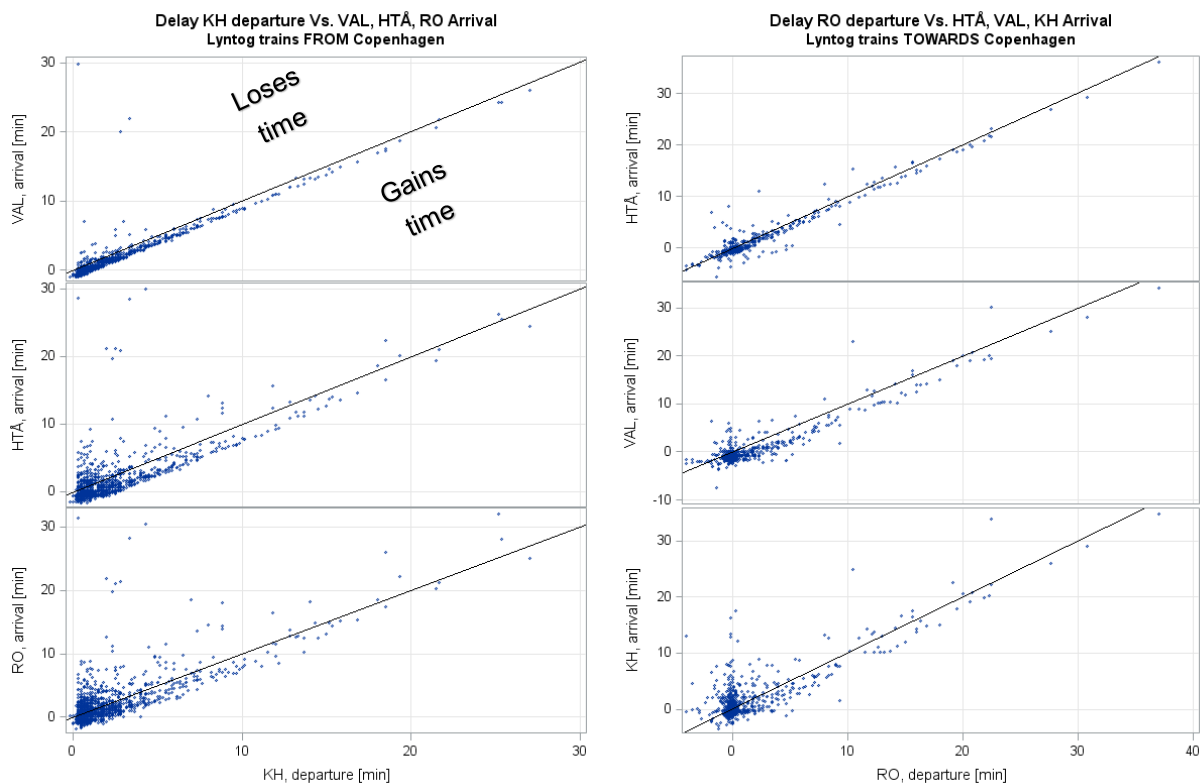


Figure 3– Departure delays at the beginning of the line compared to arrival delays at the following stations. Reference lines drawn at equal delays.

For both the directions, the delay departing from the first station was compared to the arrival delay at last three stations. A reference line is set to $x=y$ in the plots, where x is the departure delay at the first station and y is the arrival delay at the following stations. Points below this line represent trains catching up their delay, while points above the reference line mean that the trains increase their delay along the way. For both the directions, the charts follow the train path top-down. The majority of the points lay near the reference line, indicating natural variations in the delays that normally occur over 30 km of line.

Differences in the cloud density are visible between the two directions: while trains from Copenhagen are tight, trains bound for Copenhagen spread wider over the charts. The same phenomenon is visible in the comparison of delays from trains travelling in the same direction. The vertical distance between the individual points and the reference line represents the train's recovery or loss. The trains leaving Copenhagen show a clear recovery pattern at Valby, shaping a straight line parallel to the reference line. This shape fades out at Høje Tåstrup and Roskilde, with more dispersed points showing more variability. An explanation is found in the distance covered by trains, as mentioned by Olsson & Haugland (2004): the section of the line that we considered is the final segment of many long distance trains bound for Copenhagen, and the initial one for trains from Copenhagen. For these reason, trains to Copenhagen are subject to higher variability in delays. The realized recovery on one section could be modelled, thus, as an aleatory variable. The charts show that the realized recovery on consecutive line sections does not sum up linearly, but as aleatory distributions. A model is worth deeper investigation and theoretical formulation.

Early departures from Copenhagen are forbidden, as visible in the scatter plot. On the other side, some express trains that do not stop at Roskilde are allowed to travel early at this station. The right-hand scatter plot shows that the earliness of several trains at Roskilde translates into late arrivals in Copenhagen. An extension of Olsson and Haugland findings on dispatching decision (2004) could suggest the prevalence of dispatchers' personal judgment also for early trains and should be further investigated.

The association of the higher running times registered for trains from Copenhagen, and the variation in delay recovery between Høje Tåstrup and Roskilde, suggests the existence of scheduled supplement times that are not used to recover delays. This excess should be quantified to optimize the resources utilization in future timetables. On the other hand, early trains to Copenhagen, travelling out of their designated slot, could relate to an excess of supplement time in the section before Roskilde. An optimal distribution of the supplement time should prevent excessive earliness, reducing dispatching issues at the bigger nodes, and resulting in better punctuality.

3 Conclusion

This paper reports the preliminary results of a research on train delays under development at the Technical University of Denmark, within the research project IPTOP (Integrated Public Transport Optimization and Planning).

Today's access to large scale data makes it possible today to apply multivariate statistics to the recordings of railway operation, based on automated train detection systems.

Previous studies identified several influencing factors in punctuality. Nevertheless, new methods to identify excessive and insufficient timetable supplements are necessary. This paper shows that the actual supplement time can be detected in a train path by means of historical data. Further, the possibility to spot delay and recovery patterns is presented, and the impact of dispatching strategies will be developed in future research.

Recurring delay patterns may be found dependent on the infrastructure layout, the rolling stock performance and reliability, the time of the day and of the year, and the stationing on lines and at stations. Delay causes tracking is regulated under the UIC leaflet 450-2 (2009), which sets a standard codification, thus the structure of this analysis is applicable in many nations. The availability of detailed information on delay causes will also offer the possibility to deepen the previous studies on punctuality influencing factors. Delay causes recording is now required for international trains by the International Union of Railways and it is also being adopted for national trains among the railway infrastructure managers, giving access to data unavailable before. Primary and secondary delays should be explicitly recorded, in this way, making it possible to develop algorithms to link primary and secondary delays, and to further clarify how trains may auto-correlate their delays.

4 Bibliography

- Carey, M. (1998). Optimizing scheduled times, allowing for behavioural response. *Transportation Research Part B: Methodological*, 32(5), 329-342. doi:10.1016/S0191-2615(97)00039-8
- Cerreto, F. (2015). Micro-simulation based analysis of railway lines robustness. *6th International Conference on Railway Operations Modelling and Analysis*. Tokyo, Japan: IAROR. Retrieved 2015, from http://orbit.dtu.dk/fedora/objects/orbit:140777/datastreams/file_112408001/content
- Cerreto, F. (2016). A Cubic Function Model for Railway Line Delay. *6th European Transport Research Conference*. Warsaw, Poland.
- Goverde, R. M., & Hansen, I. A. (2013). Performance indicators for railway timetables. *IEEE International Conference on Intelligent Rail Transportation Proceedings*. IEEE. doi:10.1109/ICIRT.2013.6696312
- Johnston, B. (2008, September). Delays get attention in Congress. *Trains*, 68(9), p. 30.
- Larson, J. L. (1998, March). "The battle of the Sunset Limited" How Amtrak took SP to court over the performance of trains 1 and 2. *Trains*, 58(3), pp. 64-69.
- Olsson, N. O., & Haugland, H. (2004). Influencing factors on train punctuality—results from some Norwegian studies. *Transport Policy*, 11(4), 387-397. doi:10.1016/j.tranpol.2004.07.001
- Schittenhelm, B. H. (2011). Planning With Timetable Supplements in Railway Timetables. *Annual Transport Conference at Aalborg University*. Aalborg, DK: Trafikdage. Retrieved from http://www.trafikdage.dk/papers_2011/63_BerndSchittenhelm.pdf
- UIC. (2000). *UIC leaflet 451-1 Timetable recovery margins to guarantee timekeeping - Recovery margins, 4th Edition*. Union Internationale des Chemins de fer.