

Delays for passenger trains on a regional railway line in southern Sweden

C.W. Palmqvist¹, N.O.E. Olsson², L. Hiselius¹

¹*Lund University, Department of Traffic and Roads, Sweden*

²*NTNU Norwegian University of Science and Technology, Norway*

Abstract

The aim of this paper is to study how selected variables influence delays in train traffic. Data has been collected on train movements, timetables, weather and capacity utilization on a highly utilized single track railway line in southern Sweden during 2014. Based on this dataset, we have analysed how different factors affect delays in passenger traffic. We measure delays in a novel way, as deviations from the scheduled duration for each line section and station stop, not as deviations from a published or operational timetable, and this allows us to identify when and where the delays first occur. Average delays were much larger at station stops. The most significant factor affecting delays was the scheduled duration time at station stops, and the existence of margins on line sections. If trains arrive to a line section or station stop slightly delayed they speed up the activity, otherwise they are typically delayed. The influence of weather was less significant and somewhat contradictory: snow and cold temperatures increase delays on line sections but reduce them at station stops, while precipitation made no difference. Capacity utilization seems to have a negative correlation with delays, but we have too little variation in the levels to be confident. All studied variables, except for precipitation, have impacts that are statistically significant to a very high degree of confidence, using both t-tests and regression analysis. The results of this study have important practical implications for timetable construction, for instance we estimate that a reallocation of scheduled time at stations could reduce delays by as much as 80%.

Keywords: Delays, punctuality, travel time variation, railway, railroad, timetable, trains, allowances, margins, station stop.

1 Introduction

Punctuality is a key performance indicator and important success factor for the railway systems, and a failure in this key factor affects the competitiveness of railway transports relative other transport modes [1].

A number of factors can influence the precision of the railway traffic [2]. Harris [3] found that distance covered and train length, were statistically significant in determining punctuality.

Running time margins are generally added to the nominal running time when setting up a timetable, as described by Goverde [4]. Nominal running times are based on a maximal speed profile in normal conditions. To make a timetable realistic, margins are added to allow for variations in rolling stock performance, individual driving patterns for drivers, variations in weather conditions and other factors that may influence the running time.

Running time margins are typically expressed as a percentage of the nominal running time. International guidelines for running time margins are given in [5]. According to Pacht [6], common running time margins in Europe are in the range of three to seven percent, and six to eight percent in North America. According to Goverde [4], the Dutch Railways use a margin of 7% of the nominal running time. Palmqvist [7] found that the actual running time margins for a selected line in Sweden were about 10%. Punctuality, disturbances and the optimal distribution of supplements are discussed at length in Vromans [8]. Kroon, Dekker and Vromans [9] developed a measure to assess at which part of the line time supplements were added. Parbo et al. [10] discusses the use of percentiles of a sample of actual times instead of defining theoretical running times, and then adds a percentage as margin.

Ceder and Hassold [11] found that one of the main causes for delays in New Zealand was heavy passenger load, which increases dwell times. Bender et al. [12] states that time required for passenger boarding and alighting at stations is a critical element of overall train service performance. Harris, Mjøsund and Haugland [13] studied delays at stations in the Oslo area. They claim that these delays are often small in nature, poorly recorded, and not well understood.

Weather conditions can influence the railway traffic. Xia et al. [14] estimate the effects of weather conditions on railway operator performance of passenger train services. They found that wind gust, snow, precipitation, temperature and leaves contributed to infrastructure disruptions. Wei et al. [15] studied the cascade dynamics of delay propagation during severe weather. Ludvigsen, Klaboe [16] and UIC [17] present how weather conditions affect train operator and railway infrastructure performance.

Railway capacity is recognized as an important factor regarding train punctuality [18], [19].

A growing data volume and data availability has enabled researches to study delays and delay causes on a relatively detailed level. Markovic et al. [20] proposed machine learning models for analyzing the relation between train arrival delays and selected characteristics of the railway system. Gorman [21] used econometric methods on US freight rail data to predict congestion delays. Total

train running time was predicted based on free running time predictors (horsepower per ton, track topography and slow orders) and congestion-related factors (meets, passes, overtakes, number of trains, total train hours, train spacing variability, and train departure headway). Wallander and Mäkitalo [22] used a data-mining approach for analyzing rail transport delays with the aim of developing a more robust timetable structures, and provide tools for rail network planning.

This paper has a similar objective as Wallander and Mäkitalo [22] and shows an example of how this type of analysis could be carried out. We relate delays to deviations from the scheduled running times between stations, and to scheduled station times in order to distinguish between the places of occurrence of the delay, but also the actual use of the infrastructure.

There are well-established models based on physics for calculating nominal train running times, and such calculations are implemented in time tabling support tools. Time at stations has traditionally been set more empirically, or based on tradition. Our aim is to map deviations from these scheduled reference times, for each section along a selected line, and to investigate influencing factors. By focusing on how long an activity should take and how long it actually took, this approach also facilitates discussions related to the most efficient use of the infrastructure. We perform the analyses for both line sections and station stops.

The main aim of this paper is consequently to contribute to the methods by which deviations can be understood through this novel approach on defining deviations. The paper furthermore contributes to the much needed knowledgebase on deviations and punctuality in the railway system in order to reach a robust system, yet keeping an efficient use of the infrastructure. The analysis is based on data from a single railway track in Sweden and various variables possibly influencing the occurrence and size of the delays.

2 Method

2.1 Key terminology and calculations

Punctuality is measured as percentage of late trains [1] while delays are measured by a time unit, such as minutes. Punctuality and delays provide an overview of the performance of a railway system, and can be summarized to easily communicated numbers, such as punctuality of all trains on a line for a period, average delay or aggregated delay hours. However, from an analytical point of view, a lot of information is hidden in such aggregation.

When analyzing railway traffic, nuanced presentations with higher resolution are desired. Rietveld, Bruinsma and Vuuren [23] measure unreliability as deviations from a timetable. Noland and Polak [24] discuss travel time variability using a distribution of arrival times, without referencing a timetable. Nicholson et al. [25] instead propose three measures for the term resilience.

The analyses in this paper are based on differences between the actual and scheduled durations of activities, be they movements along line sections or station stops. If a stop at a station is scheduled to take two minutes but actually takes three,

we would call that a delay of one minute. In the rare case that the stop only took one minute instead of two, that would register as a *negative* delay of one minute (-1). This paper thus differs from what is typically meant by “delay” in that, beyond the scheduled duration and any margins in it, we are not concerned with the timetable per se. If a train is scheduled to arrive at a station at 12:00 and depart at 12:02, thus with a scheduled duration of two minutes, but it actually arrives at 12:01 and departs at 12:03, with an actual duration of two minutes, we would not register that as a delay. Instead, we would have registered the delay, with relation to the arrival and departure times, where and when the delay first occurred.

The paper furthermore differs from what is typically meant by “travel time variability” in that we use the timetable as the baseline how long activities “should” take, instead of the average duration as measured over some time span. If a stop at a station is scheduled to take two minutes, but actually takes three minutes, we would consider that a delay of one minute even if the average dwell time is three minutes. From the perspective of travel time variability, however, there would be no deviation.

In the paper, we use the term *average delay*. What we mean is an average *conditional* on the value of an explanatory variable. We also discuss *relative risk*, which indicates how much a factor increases or decreases the average delay in *relative* terms. The relative risk of delays is calculated as the average delay conditional on some value of an explanatory value, divided by the average delay across all station stops or line sections, as the case may be. This makes it easier to discern the effect of factors where the values or differences are small in absolute terms.

The analysis of this paper also makes use of the *delay contribution*. Here we multiply the average delay with the number of observations, and compare this to the total amount of delays. This allows us to identify which levels lead to the most delays, overall.

To ease comparison between the differing lengths of line sections, between the speeds of trains, and so on, we have chosen to use ratios instead of absolute time units. Throughout we have chosen to use the scheduled duration as the denominator. If the scheduled duration is 100 seconds and the total margin for that activity is 10 seconds, we would register that as an allowance of 0.10, or 10%.

For durations we normalize with the average duration for the respective activity. Across our sample those averages are approximately 100 seconds for station stops and 197 seconds for line sections. A scheduled stop of two minutes is thus translated to a duration of 1.20 while a scheduled movement along a line section of two minutes would translate to 0.61. We have then rounded these ratios to one or two decimal points, to limit the number of distinct values we work with, and to ensure that there are enough observations in each bin for their average delays to be reasonably stable.

2.2 Data

The railway line analyzed in this study is 113 km long, electrified, single track, with dense and heterogeneous regional traffic. This makes it a fairly typical case in a Scandinavian context. It is located in the far south of Sweden, and goes almost straight to the east from Helsingborg on the western coast to Kristianstad near the eastern shore, connecting the West Coast Line and the Southern Main Line.

The paper is based on four main datasets: train movements on the railway line for the year of 2015, detailed timetables for these trains, weather observations from around the region, and the calculated capacity utilization along the railway. These datasets have been linked together through a series of Access databases, resulting in data for over 198 000 stops at 13 stations and 363 000 movements along 20 line sections. Details about the component datasets follow below.

The core of our data is that of registered train movements. We do not have data from every signal, but at certain points, typically where trains can meet. Arrival and departure of the trains are logged for purposes of traffic control, punctuality measurement etc. From these logs we can calculate the actual duration of the activity, whether it was a stop at a station or a movement between them.

Timetables for trains in Sweden are created and stored in the tool TrainPlan. The geographical resolution is almost the same as above. We have received exports from this program, in which we can see every planned activity and its allotted time. This covers both stops and movements. From this we extract the scheduled duration and any margins that may have been allotted.

We have meteorological observations from measuring stations in connection to the studied railway line, over the variables snow depth, precipitation and temperature. The two former variables are measured once per day, the latter every hour. Snow depth is presented as meters, with a precision of two decimal points. Temperature is in degrees Celsius, precipitation in millimeters.

The infrastructure manager has adopted an interpretation of the UIC [18], [19] guidelines for how to calculate the capacity utilization on line sections, and provided us with the results. However, this framework does not include a method to calculate the utilization of capacity at stations. In order to look at this, we define our own measure. First we calculated, for every day and hour, how many trains passed through the station. Then we took the maximum value and set it as the maximum capacity for every given station. Finally, we compared the utilization over every day and hour with the maximum for that station to get two ratios, one for the daily utilization and one for the peak utilization.

A summary of studied variables is found in Table 1.

Table 1: Description of studied variables.

Variable	Description
Earlier delay	The train's delay at the beginning of the activity, as a fraction of the scheduled duration of the activity. Early arrivals are counted as negative delays.
Scheduled duration	The amount of time scheduled for an activity, be it line section or station stop, before margins are applied. Normalized by the average scheduled duration for the respective activity and presented as a fraction.
Margins	The amount of margins in the timetable for the current activity. Normalized by the scheduled duration and presented as a percentage.
Temperature	Degrees Celsius, measured hourly.
Precipitation	Rain- or snowfall, in [mm], measured daily.
Snow	Depth of snow in [m], measured daily.
Day of the week	Monday (1) through Sunday (7)
Month of the year	January (1) through December (12)
Capacity utilization	Percentage of maximum theoretical capacity on line section or at station. Calculated and analyzed both for an entire day and for peak load.

2.3 Analysis

In order to analyze the impact of these variables on delays we make use of three basic steps.

The first is to determine if the studied variables had a statistically significant impact on the delays. We used Welch's t-test to do this, which is a more general version of Student's t-test, which permits the studied samples to have different sizes and variances. Here the data is segmented based on the value of the explanatory variable. For example, we look at temperatures below zero and temperatures that are above zero, and whether the average delays for these samples are significantly different from each other.

A second step is to perform linear regressions for delays at station stops and line sections respectively, in order to give an overview of the trend of each studied variable. These are not intended to be the focus of the paper, or to provide good predictive models. This regression analysis assumes that, instead of there being two categories within each variable, there is a linear relationship between the studied variable and delays occurring.

The third step of the analysis includes three different plots for each variable across line sections and station stops, each emphasizing different aspects of the data. Each explanatory variable is plotted against the average delay, the relative risk and the delay *contribution* as defined in section 2.1 for line sections and station stops respectively. These plots together give a more complete and nuanced picture than either the t-tests or the regressions, and visual inspection can reveal if the studied variables vary along with delays in any recognizable pattern, or if there is no pattern at all. Due to limitation of space this paper only presents two diagrams. The result is thus mainly presented in text.

3 Results with discussion

3.1 Line sections and station stops

The presentation of the results begins with an overview over the average delays and probabilities of being delayed on line sections and at station stops. As Table 2 shows, the average delay for a train on a line section is very close to zero, while significant delays are to be expected at the scheduled station stops. The values in Table 2 are normalized with regard to the scheduled duration.

Table 2: Average delays and probabilities of delay

	Summary statistics		Probability of finishing activity ...			
	Average delay	Standard deviation	Early	On time	Delayed	Delayed by more than 100%
Line section	0,00	0,01	24%	58%	18%	0%
Station stop	0,44	1,21	19%	34%	47%	41%

Table 2 also illustrates that for movements on line sections, the probability of being exactly on time is higher than at stations, and the probability of gaining time is higher than that of losing it. At stations almost half of all stops take longer than scheduled. The delays that do occur also tend to be larger at station stops than on line sections.

3.2 T-test and regression analysis

To test the significance of the studied variables we first use Welch's t-test. The results of these are summarized in Table 3. Most of the variables were found to have a highly significant impact on delays.

Table 3: Results of t-tests and regression analysis

Input		Welch's t-test						Regression		
Variable	Source	Value	Average	St. Dev.	Value	Average	St. Dev.	p-value	Coefficient	p-value
Earlier delay of train	Line section	>0	0,00	0,01	<=0	0,00	0,01	0,00	0,00	0,00
	Station stop		0,14	1,11		0,78	1,22	0,00	-0,05	0,00
Scheduled duration relative to average	Line section	>1	0,00	0,01	<=1	0,00	0,01	0,27	0,00	0,00
	Station stop		0,08	0,73		0,54	1,30	0,00	-0,22	0,00
Margins in timetable	Line section	>0	0,00	0,01	<=0	0,00	0,01	0,00	-0,01	0,00
	Station stop		-0,43	0,92		0,44	1,21	0,00	0,00	0,00
Temperature	Line section	<0	0,00	0,01	>=0	0,00	0,01	0,00	0,00	0,00
	Station stop		0,35	1,20		0,44	1,21	0,00	0,00	0,01
Precipitation	Line section	>0	0,00	0,01	0	0,00	0,01	0,05	0,00	0,00
	Station stop		0,44	1,16		0,44	1,24	1,00	0,00	0,39
Snow, depth	Line section	>0	0,00	0,01	0	0,00	0,01	0,00	0,02	0,00
	Station stop		0,29	1,14		0,45	1,21	0,00	-1,16	0,00
Day of week	Line section	6-7	0,00	0,01	1-5	0,00	0,01	0,00	0,00	0,00
	Station stop		0,42	1,17		0,44	1,22	0,01	-0,01	0,00
Month of year	Line section								0,00	0,00
	Station stop								0,02	0,00
Capacity utilization calculated daily	Line section								0,01	0,00
	Station stop								0,06	0,00
Capacity utilization calculated at peak	Line section								-0,01	0,00
	Station stop								-0,03	0,10

We then perform regression analyses on the line section and station stop data respectively. Both of these yield results similar to the t-test, and the key results are also summarized in Table 3.

A number of coefficients differ in sign between the regression for line sections and station stops: the three weather variables increase delays on line sections while reducing them at stations, while month by month the delays decrease on line sections and increase at stations.

3.3 About the explanatory variables

In the following section we discuss the results of visual analysis of more than 60 plots over the explanatory variables impacts on delays. This analysis allows us to see patterns that are not visible in Table 3. We do not have the space to show more than a couple of these diagrams, but briefly discuss the others in text.

3.3.1 Timetable factors

Visual analysis indicates that small earlier deviations from the timetable affect the occurrence delays on line sections. One possible interpretation is that if trains arrive a little before the scheduled arrival time the drivers seem to slow down, if they are a little bit late they try to catch up as best they can. This effect disappears as the deviations become larger, both in positive and negative values.

Station stops were shortest when trains arrived a few minutes after the scheduled arrival time. Trains that arrived either on time or early to the station were on average delayed at the stations so that they depart with a delay of around half a minute, regardless of how early they were. Trains that arrived more than a few minutes late were on average delayed even more at the station.

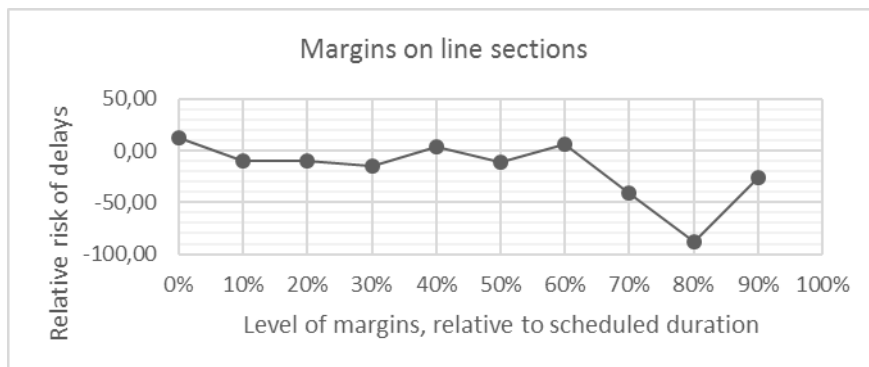


Figure 1: Relative risk of delays at different margin levels, on line sections

On line sections margins have greater impact on delays than the scheduled duration. While longer line sections on average had slightly lower delays than shorter ones, the effect is not statistically significant, as we see in Table 3. Most delays on line sections in our sample are caused when trains have no margins,

although negative margins do occur and contribute to a small degree. The relationship is illustrated in Figure 1. The most effective level for delay reduction is around 10%, while the greatest delay reduction is achieved with margins around 80%. On line sections we find this to be the single most important variable for explaining delays.

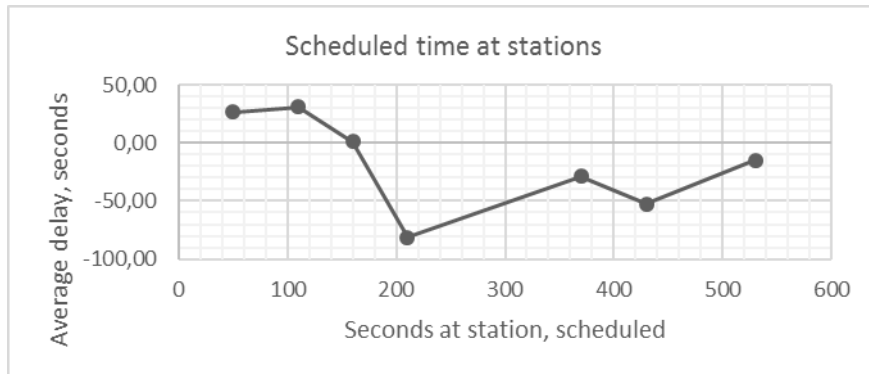


Figure 2: Average delays at stations, scheduled duration

Margins are less well defined at station stops compared to margins on line sections. In the sample we study only a few hundred trains have clearly labelled margins, and those are on the order of hours per station stop, clearly intended for other purposes. Instead, the scheduled duration of the stop is more important, and is manually set by the timetable-creator. Figure 2 illustrates that the average delay is significant if the scheduled duration is lower than 160 seconds. The greatest reduction in delays occur at a duration of 210 seconds, thereafter additional station time increases the average delays. This variable, scheduled time at station stops, explains more of the delays than any other, both at stations and in total.

Knowing this, we might seek to reallocate the time at stations so as to reduce the delays. To do this we chose a strategy where the total scheduled duration across all stations stops is not allowed to increase, while trying to minimize the total delay time. As a result, we set 21% of the station stops to 210 seconds and the rest at 50 seconds. Using the estimated effect on delays from Figure 2, we found that this reallocation could reduce delays by as much as 80% without increasing the scheduled travel times.

3.3.2 Weather variables

None of the weather variables we study are found to have a large impact on delays, the coefficients and differences in average delays are relatively small in Table 3, even if the impact is statistically significant for both snow and temperature.

Visual inspection reveals that an increasing snow depth increases the delays on line sections and *reduces* delays at stations, while precipitation does not seem

to have a significant effect on delays at all. Temperatures below zero contribute to raising the average delays on line sections, while milder temperatures pull the average towards trains arriving early. At stations we find quite the opposite: the risk of delays *decreases* around the freezing point, while more normal temperatures are associated with delays.

3.3.3 Time: days of the week, months of the year

Visual inspection reveals a noticeable and statistically significant difference between the average delays across the seven days of the week, with delays being the smallest at the weekend and largest on Mondays. This is almost entirely explained by how many trains run each day, as the correlation between the daily averages and the number of trains running the on same days is over 92% on line sections and 70% at stations.

The risk of being delayed also varies from month to month. On line sections we do not discern any pattern or seasonal effect, while at stations the risk increased slowly as the months went by. The risk was around 25% lower than average during the early part of the year, and about as much higher towards the end.

3.3.4 Capacity utilization

Capacity utilization is the factor that we are the least confident about, both for line sections and station stops, and especially using the peak measure.

The railway line we study is not very long and only has three different levels of capacity utilization for line sections, calculated by the infrastructure manager. The regression analysis reported in Table 3 showed a small positive correlation, but our visual inspection instead shows a *negative* correlation between capacity utilization and delays. We hypothesize that this mismatch is due to a too small number of distinct capacity utilization levels in our sample, and do not draw any strong conclusions about this variable.

Capacity utilization is not defined and calculated in the same way for stations as for line sections, so we define and use a measure of our own, described in the method chapter. When we use the daily measure for stations, the risk is somewhat elevated when the capacity utilization is 70%. Using the peak measure, which we are less confident about methodologically, we instead see a reduction at the same level. Again, we do not think any strong conclusions should be drawn from this.

4 Conclusions

In this paper we have used a novel approach of looking at where delays occur, a method which makes it easy to distinguish between delays occurring on line sections and those occurring at station stops. By focusing on how long an activity should take and how long it actually took, this approach also facilitates discussions about efficient use of the infrastructure.

The results indicate that variables influencing delays do so to different degrees, sometimes even with different signs, on line sections and at station stops. We also showed that, at least on this railway line, delays at station stops were by far the biggest problem.

We found that the timetable was the most important factor for explaining delays. Trains with slight delays make better time on average, both on line sections and at station stops. Those that instead are early or with larger delays typically require more time, increasing delays. When margins of around 10% exist on line sections the risk of delays falls, while larger margins are less efficient. On station stops a duration of 210 seconds is optimal for reducing delays, and we propose that these should be mixed with stops of 50 seconds to obtain a good balance between minimizing delays and minimizing the expected travel time. We estimate that by reallocating existing station time in this manner, it is possible to reduce delays by as much as 80%.

Weather had some contradictory impacts. On line sections delays increase as temperatures fall, or if there is snow. At stations the same conditions actually improve performance, and in our sample the effect at stations was more important. Precipitation had no significant impact on delays. The average delays are slightly smaller on weekends, especially on line sections, and this is mostly because there are fewer trains running then than during the week. The number of trains also explains part of the variation of delays between months when we look at stations, but not on line sections. We do not detect a clear pattern regarding capacity utilization and its effect on delays.

The results of this study have important practical implications for railway traffic planning. Some factors, such as weather, cannot be influenced, but their impact on traffic is still important to know when establishing plans. Other factors, such as the distribution of margins, and length of station stops are factors that planners can influence.

From a research perspective we have shown that it is both possible and useful to gather and analyze large volumes of data about train movements in this manner: analyzing even small deviations from the timetable, identifying when and where the deviations occur, considering the differences between line sections and station stops, and combining several datasets and explanatory variables.

We shall continue our research expanding the network size, time period and number of variables studied, and increasing the depth of analysis.

References

- [1] Nyström, B., Aspects of Improving Punctuality, Luleå University of Technology Doctoral thesis, 2008.
- [2] Olsson, N.O.E., Haugland, H., Influencing factors on train punctuality – results from some Norwegian studies. *Transport Policy*, 11(4), pp. 387–397. 2004.
- [3] Harris, N.G., *Planning Passenger Railways: A Handbook*, 1992.
- [4] Goverde, R.M.P. Punctuality of Railway Operations and Timetable Stability Analysis. The Netherlands TRAIL Research School Thesis Series no. T2005/10, 2005.
- [5] International Union of Railways, UIC Code 451 Timetable recovery margins to guarantee timekeeping – Recovery margins, 2000.
- [6] Pachel, J., *Railway Operation and Control*, 2002.

- [7] Palmqvist, C.W., Running time supplements for high speed trains. Lund University Master thesis 257, 2014.
- [8] Vromans, M.J.C.M., Reliability of Railway Systems. ERIM PhD series Research in Management 62, 2005.
- [9] Kroon, L., Dekker, R., Vromans, M., Cyclic railway timetabling: A stochastic optimization approach. *Algorithmic Methods for Railway Optimization* 16, 41–66, 2007.
- [10] Parbo, J., Nielsen, O.A., Prato, C.G., Passenger Perspectives in Railway Timetabling: A Literature Review, *Transport Reviews*, 2016.
- [11] Ceder, A., Hassold, S., Applied analysis for improving rail-network operations. *Journal of Rail Transport Planning & Management* 5, 50–63, 2015.
- [12] Bender, M., Büttner, S., Krumke, S.O., Online delay management on a single train line: beyond competitive analysis. *Public Transp.* 5, 243–266, 2013.
- [13] Harris, N.G., Mjøsund, C.S., Haugland, H. Improving railway performance in Norway, *Journal of Rail Transport Planning & Management* 3, 4, 172–180, 2013.
- [14] Xia, Y., Van Ommeren, J.N., Rietveld, P., Verhagen, W., Railway infrastructure disturbances and train operator performance: The role of weather. *Transportation Research Part D* 18, 97–102, 2013.
- [15] Wei, D., Liu, H., Qin, Y., Modeling cascade dynamics of railway networks under inclement weather. *Transportation Research Part E* 80, 95–122, 2015.
- [16] Ludvigsen, J. Klæboe, R., Extreme weather impacts on freight railways in Europe, *Nat. Hazards* 70, 707–787, 2014.
- [17] International Union of Railways, Winter and Railways, UIC/SIAFI Report, 2011.
- [18] International Union of Railways, UIC CODE 406 Capacity, 2004.
- [19] International Union of Railways, UIC CODE 406 Capacity, 2013.
- [20] Markovic, N., Milinkovic, S., Tikhonov, K.S., Schonfeld, P., Analyzing passenger train arrival delays with support vector regression. *Transportation Research Part C* 56, 251–262, 2015.
- [21] Gorman, M.F., Statistical estimation of railroad congestion delay. *Transportation Research Part E* 45, 446–456, 2009.
- [22] Wallander, J., Mäkitalo, M., Data mining in rail transport delay chain analysis, *Int. J. Shipping and Transport Logistics* 3, Vol. 4, 269–285, 2012.
- [23] Rietveld, P., Bruinsma, F. R. and van Vuuren D.J., Coping with unreliability in public transportation chains: A case study for Netherlands, *Transportation Research Part A* 35, 539-559, 2001.
- [24] Noland, R. B., Polak, J. W., Travel time variability: a review of theoretical and empirical issues, *Transport Reviews* 22, Vol. 1, 39-45, 2002.
- [25] Nicholson, G., Kirkwood, D., Roberts, C., Schmid, F., Benchmarking and evaluation of railway operations performance, *Journal of Rail Transport Planning & Management* 4, Vol. 5, 274-293, 2015.