Some Influencing Factors For Passenger Train Punctuality In Sweden

Palmqvist, Carl-William; Olsson, Nils; Hiselius, Lena

Some influencing factors for passenger train punctuality in Sweden

Carl-William Palmqvist[1], Nils O.E. Olsson[1,2], and Lena Winslott Hiselius[1]

[1] Lund University Faculty of Engineering, Box 118, Lund, 221 00, Sweden
[2] Norwegian University of Science and Technology, Høgskoleringen 1, 7491, Trondheim, Norway

carl-william.palmqvist@tft.lth.se

ABSTRACT

Punctuality is regarded as an important measure of the performance of a railway system, and is the one most commonly used and discussed measure both in the industry and among travelers. In many countries, the punctuality of trains, and thus the performance of the railway system, is deemed as lacking. The aim of this article is to study and quantify how several weather-, timetable, operational and infrastructure-related variables influence punctuality in passenger train traffic. This can contribute to better understanding of the performance of railway systems, and help identify possible improvements.

The study is based on a dataset containing detailed timetables and records of all 32.4 million train movements for all trains in Sweden during the year of 2015, over 1.1 million departures. Supporting this is a comprehensive register of over 80 000 infrastructure elements, and almost 87 million weather observations.

We consider the size and allocation of margins, the existence of negative margins, two measures for traffic volume, the journey time and distance, how often different vehicle individuals are used, the number of line and station interactions between trains, the amount of precipitation, the temperature, wind speed, snow depth, and eight types of infrastructure elements. We show how these variables affect punctuality, and estimate how much of the variation in punctuality can be explained by them.

The findings can be used to design timetables, change operational parameters and modify infrastructure design so that punctuality improves. They can also help identify areas which should be prioritized in planning, maintenance and research.

1. INTRODUCTION

Punctuality is an important factor for the attractiveness and efficiency of the railway sector. In Sweden, a target has been set and agreed upon by the industry, that by 2020 the punctuality across all trains should be 95%, measured as arriving at the destination with a delay of at most five minutes. Since 2012 the punctuality has been steady at around 90 % for passenger trains and slightly below 80% for freight trains (Transport Analysis, 2016). Large and rapid improvements are thus required, if the target of 95% is to be reached on time.

The purpose of this study is to identify and quantify the impact of several weather, timetable, operational and infrastructure variables on the punctuality of passenger trains in Sweden. This is a much broader scope than is typically seen in the literature, using an extensive dataset. Most previous research is focused on a single type of influencing variables, such as weather or timetable properties, using limited time periods and geographies to illustrate the effects. While this is often necessary, there is a gap in the research that attempts to synthesize this knowledge in a more holistic approach. This paper is intended to bridge that gap.

1.1. Previous Research

1.1.1. Weather

The influence of weather and climate change on train delays and punctuality has received considerable attention in the literature recently. Brazil et al. (2017) found that precipitation delayed trains on a metropolitan rail line in Dublin, combining a dataset of over 6 000 train departures with hourly observations of several weather variables. Zakeri and Olsson (2017) investigated the impact of weather on punctuality of local trains in the Oslo area, and found strong correlations between punctuality and temperatures below -7°C and snowfall of at least 15 cm. Xia, Van Ommeren, Rietveld and Verhagen (2013) estimate how wind, temperature and precipitation cause disturbances in the railway, mainly by damaging the infrastructure. Qin, Ma and Jiang (2017) model how rain, snow, and different temperature thresholds affect delays on a regional railroad in Sweden. Ludvigsen and Klæboe (2014) describe how long cold spells, heavy snowfall and strong winds have severely delayed freight trains across five European countries. Xu,

Corman and Peng (2016) analyze the disruptions in the Chinese high-speed railway, and find that almost 90% of these are due to bad weather. Nagy and Csiszar (2015) highlight the effects of weather conditions on the punctuality of Hungarian passenger trains. Ferranti et al. (2016) study how heat causes failures in the railroad infrastructure in England, particularly in the signaling systems. They also discuss the concept of failure harvesting, in which components that fail early in the season are replaced by newer and more resilient components, reducing the vulnerability and number of failures later in the season.

Falling leaves often impact punctuality in the autumn, as described by Xia et al. (2013), and Brazil et al. (2017) among others. The leaves create a mulch on the rails which lowers the adhesion between the rails and wheels, increasing both braking and acceleration times, and thus causing delays. Because falling leaves are not typically measured or recorded like other weather-related variables, both sets of authors use dummy variables for each month to try to capture this effect.

Tahvili (2016) describes extensively and in detail how snow, cold temperatures and strong winds cause problems for the railroad, and how the Norwegian rail sector is undertaking winterization measures to reduce these problems in the future. Lehtonen (2015) described the conditions of four snow-rich winters in the Helsinki-area, and how they caused significant issues for the railroads there. Jaroszweski, Hooper, Baker, Chapman and Quinn (2015) describe how a storm in the UK caused very severe delays for both road and rail transport, severing the main link between England and Scotland. Doll et al. (2014) present a case study for adapting rail roads in Austria to the different climate of 2050, and conclude that the damages due to weather will increase. Kellermann, Bubeck, Kundela, Dosio and Thieken (2016) simulated how the changing climate will affect the frequency of critical meteorological conditions in the Austrian railroads, concluding that while snowfall and extremely cold temperatures will become less frequent, intense rainfall and heat waves will become more common. Ford et al. (2015) simulate how extreme weather events like heat waves and floods will become more common in the UK as global warming continues, leading to increasing disruptions to the railroad system. Oslakovic, ter Maat, Hartmann and Dewulf (2013) also study how weather conditions cause failures in infrastructure elements in the Randstad region of the Netherlands, and how these conditions are likely to become more frequent as the climate changes.

1.1.2. Timetable

As Parbo, Nielsen and Prato (2016) show, timetable characteristics are important influencing factors for delays and robustness in railway traffic. Kim, Kang and Bae (2013) present and categorize different types of train delays, concluding that the important causes in South Korea are short headways, short scheduled run times, delays of preceding trains, and excessive passenger loads. Cerreto, Nielsen, Harrod and Nielsen (2016) present a preliminary study on the quality of time supplement allocation in timetables, and on how trains recover or increase delays that occur during a journey. Our own previous studies also indicate that properties of the timetable have significant impact on delays and punctuality (Palmqvist, Olsson & Hiselius, 2017a). This often stems from several strategic decisions that a planner must consider when designing a timetable. On a high level, these strategies address several issues. This includes the balance between precision and slack (Olsson et al., 2015), and the balance between using headways to assign buffers between trains or using time supplements to assign margins within train journeys (Nelldal, 2009). Other issues are the degree to which a cyclic timetable is desired, the heterogeneity of traffic and the degree to which homogenization measures are to be employed (Nelldal, Lindfeldt & Lindfeldt, 2009). In addition, questions arise about geographical accessibility and the design of the network to be utilized, the balance between stops at end points or intermediate stations, and other issues.

An overview of the state of the art in timetable research is provided by Hansen (2009). The author concludes that the key issue for high quality timetables is a precise estimation of blocking times, considering the signals, platforms, train processing, and using realistic run and dwell times. This is often not the case in practice. Similarly, queuing and simulation models inadequately reflect speed variations and the behavior of railway staff. Planners have tools to make timetables robust against delays, for example by adding time supplements, lowering heterogeneity in the timetable by having uniform stopping patterns, finding optimal speed and reducing interdependencies between trains (Parbo et al., 2016).

Carey (1999) discusses several different heuristic measures of timetable reliability, with special consideration to knock-on delays. These include probabilities of delays, calculated in several ways, and different headway based measures. The latter are found to be easier to use and calculate, because they do not require nearly as much data.

Scheepmaker and Goverde (2015) demonstrate that it is more energy-efficient to distribute time supplements evenly along a train route. Vromans (2005) introduces the measure WAD, or the Weighted Average Distance, to describe how supplements are distributed along the journey, and attempted to optimize this using both analytical and numerical methods for some hypothetical and real cases, concluding that a slight shift towards the beginning is best. This is further discussed in Andersson (2014). Using simulation-based methods, Vromans (2005) and Vekas, van der Vlerk and Haneveld (2012) found that a uniform distribution was sub-optimal for delay recovery, given some assumptions of the delay distributions.

1.1.3. Operational

Influencing factors on train punctuality in Norway are presented by Olsson and Haugland (2004). In short, the authors found that in congested areas the management of boarding and alighting passengers is the key factor, while on single track lines the management of train crossings is the key success factor. Gorman (2009) used statistical analysis to study which factors contributed the most to delays for freight trains in the US. He found that the number of meets, passes and overtakes consistently had the highest impacts, suggesting that congestion was the primary cause for delays.

Wiggenraad (2001) studied seven Dutch train stations in detail. He found that dwell times are longer than scheduled, that the dwell times at peak and off-peak were the same, and that passengers concentrated around platform access points. This suggests an improvement potential of shorter real dwell times if travelers could be distributed more evenly along the platforms. Along the same lines, Nie and Hansen (2005) studied trains in the station area of The Hague. They found that trains operate at lower than design speeds, and that dwell times at platforms are systematically extended because of other trains blocking their routes, and because of the behavior of train personnel.

1.1.4. Infrastructure

Veiseth, Olsson and Saetermo (2007) links infrastructure data with delay and punctuality data to study the infrastructure's influence on rail punctuality. They report that some 30 % of delay hours in Norway are caused by infrastructure failures, and suggest that the quality of punctuality data can be improved by connecting it with infrastructure and operational databases. Thaduri, Galar and Kumar (2015) discuss how the many systems and sub-systems in the railway can be studied using big data analytics, and gives an overview of the main databases used in the Swedish railways. Norrbin, Lin and Parida (2016) discuss the concept of robustness for railway infrastructure, and present a roadmap for studying and improving it. Stenström, Parida, Lundberg and Kumar (2015) develop a composite indicator for benchmarking and monitoring of rail infrastructure, considering four factors: failure frequency, train delays, logistic time and repair time. Nikolic et al. (2016) discuss the poor quality of the Serbian railway infrastructure, and adapts the measure of Overall Railway Infrastructure Effectiveness, which is another composite indicator developed in Sweden, for use in their national network.

2. METHOD

This section describes in turn the datasets used, the variables analyzed, and the method of analysis.

2.1. Datasets

This study is based on a database containing three main datasets. The core set contains all train movements in Sweden, over 32.4 million, derived from the track blocking and signaling systems for the timetable year of 2015, which we use to determine the punctuality of trains. The second set contains detailed exports of all train timetables in Sweden during the timetable year of 2015, this covers almost 46,000 distinct timetable versions and over 1.1 million departures. The third dataset contains all historical meteorological observations of snow depth, temperature, wind strength, precipitation in Sweden, which we use to estimate the weather conditions in which the trains operated. These datasets were linked together, and several filters were applied so that it focuses on only passenger trains, and excludes incomplete observations. The remaining data covers over 883 000 completed passenger train journeys across Sweden during one year. Freight and service trains are not included in the analysis, only passenger trains, because the preconditions between the different types are considered too different, as is the handling of both timetable planners and traffic control.

2.2. Analyzed Variables

We analyze how 36 variables across four categories affect punctuality. The breakdown by category is as follows: six variables related to weather, seven variables related to the timetable, seven variables related to operations, and 16 variables relating to eight types of infrastructure elements. These are described in the following sections.

2.2.1. Punctuality

We define punctuality in the following way: trains arriving at their scheduled stops with a delay not exceeding five minutes are considered punctual at that stop and are given a value of 1, otherwise a value of 0 is given. In this manner, cancellations are counted as non-punctual. For each train, we calculate the average of these values to arrive at a punctuality measure. A train that has four stops and arrives at three of them punctually, but is not punctual at the fourth, thus receives the punctuality value of 0.75. In any aggregate of trains, the punctuality is calculated as the average of these values and presented as a fraction. This additional step, of taking the average across all scheduled stops, gives a more holistic picture. In our case, it also improves the overall punctuality of passenger trains by 2.43 percentage points, to 92.17%, compared to when punctuality is only measured at the destination. We only consider the punctuality of trains, not the size, frequency or distributions for delays or disturbances.

Manually reported causes for delays are not utilized in this study, for a number of reasons. The reported error causes are already relatively well known and publicized in both Sweden and Norway (see for instance Swedish Transport

Administration, 2017, and Veiseth et al., 2007). Delays must be relatively large and obvious for causes to be reported, in our material 55% of delays are small enough that they would not be categorized. Manual attribution of delays is also prone to errors and often quite inconsistent (Nyström, 2008), with an estimated reliability around 80 % (Nilsson, Björklund, Pyddoke & Vierth 2015). Research is ongoing on combining the attributions with data of the kind that we use in this study, to cross validate among the different sources.

In practice calculating punctuality is more challenging than it may appear, as a substantial number of observations are missing in the data. These are points where the train has not been canceled but there is no record of the train arriving or departing, and there are records of it arriving at surrounding stations. One specific example is for airport trains which depart from Arlanda North and stop at Arlanda South one minute later, before heading towards Stockholm C. Very often the record for Arlanda South is missing, despite there being records of the train leaving Arlanda North, only 570 m to the north along the same track, and then subsequently arriving in Stockholm. There are many other examples, especially for regional trains. We have dealt with this by going through a loop of (1) using the average delay at the two adjacent stations, when both of these records exist, and otherwise (2) using the delay at the previous control point if there is an observation there, or if not then (3) using the delay at the next control point, if there is one there, and if the observation was missing there as well then starting again at (1). In this way, small gaps in the record are filled immediately, and larger gaps are filled in step by step, as the loop is repeated. By iterating this process six times, we reduced the share of missing observations from 7.5 to 0.1 %. After this we calculate the punctuality, in the way described above.

2.2.2. Weather data

Temperature is measured, on average, about 18.7 times per day and station. The average of these was taken to get a daily temperature value for each station. Wind strength was measured at fewer stations, with an average of 23 observations per day. To convert to a daily wind value for each station, we took the maximum value for each day, because we are mainly interested in stronger winds. Snow depth and precipitation is measured daily, but with data missing on average 9 % and 0.5 % of the days, respectively. An overview of this data is given in Table 1.

Table 1. Overview of weather variables

Variable	Unit	No. stations	Obs. freq.	No. obs.
Temperature	°C	247	Hourly	1 680 k
Wind speed	m/s	166	Hourly	1 400 k
Snow depth	cm	276	Daily	91 k
Precipitation	mm	585	Daily	212 k

Trains often travel long distances, through varying weather conditions. To account for this, and because the locations of meteorological observations are typically different from those of the train stations, we created an algorithm that matches each of the train stations to the nearest meteorological station. The matching was done separately for each weather variable, because not all meteorological stations observe the same variables. And because some stations lack observations on some days, the algorithm was set to match the two station sets for each day, to ensure that an observation could always be given.

As each train passes several train stations, which can have different values of the weather variables, there are several ways in which to convert these different values to one single variable. With wind, we were interested in the highest speeds and chose to take the maximum. With temperature, we tried both the average, the minimum and the maximum. We ended up choosing the minimum temperature for cold weather, and the maximum temperature for hot weather, arguing that we are most interested in the extremes, but found that the choice made little difference in the analysis. For precipitation and snow depth, we considered the average, maximum and sum of the measured variables. In the end, we found that the sum best explains the effect of precipitation on punctuality, despite the complicating factor of introducing the distance and number of measuring stations into the variable, we find that the added explanatory value more than makes up for the reduced independence of the variable. For snow depth, we found the average to work very well.

One of the variables we use to explain punctuality is the difference in temperature. It basically represents the difference in temperature that a train is exposed to. Because of how we have chosen to handle the temperature data, this temperature difference should be interpreted as being across the geography, not across time.

We do not have access to any information on falling leaves, and because the weather and change of seasons varies quite considerably within Sweden, we do not think that falling leaves can be captured as well by monthly dummies in our dataset, as may have been the case in countries like the Netherlands or Ireland. Accordingly, without access to data or good proxies, we exclude this phenomenon from our analysis.

2.2.3. Timetable variables

The seven timetable variables are summarized in Table 2.

In this paper, we look at the size of margins in two slightly different ways: as a percentage of the scheduled runtime without margins, and as seconds per kilometer.

To measure the distribution of margins within a timetable, we use the measure of Weighted Average Distance (WAD) described in Vromans (2005). This is used to describe how the various time supplements in a timetable are balanced,

Table 2. Overview of timetable variables

Variable	Unit
Total margins	%
Supplement density	s/km
WAD of margins	%
Negative margins	binary
Avg. speed with stops	km/h
Travel time (without margins)	h
Distance/no. of scheduled stops	km

Table 3. Overview of operational variables

Variable	Unit	Number of observations
Station interactions	count	2 091 619
Line interactions	count	20 945
Trains to same station & hour	count	4 355 911
Movements per vehicle	count	970
Movements per day	count	364
Days run per train number	count	5 726
Travel distance	km	883 715

being more towards the beginning or end of the journey, or in between. It is expressed as a decimal and normally takes values between 0 and 1, with lower values expressing a shift towards the beginning and higher values a shift towards the end of the journey.

In some timetables, there are negative margins: cases where the scheduled time has been manually set to be shorter than the technical minimum. We use a dummy variable with the value of 1 if there are any instances of this in a timetable, and 0 if there are not.

The travel time without margins, or the scheduled duration of the journey, is included to help differentiate between the distance covered and the time in the system.

The average speed is calculated to include stopping times, and is derived by dividing the distance by the duration of the journey (as defined in the timetables).

Another timetable characteristic is the average distance between stops, which is calculated by dividing the distance with the number of scheduled stops.

This paper does not consider headway or buffer times, or other measures of margins between trains, only margins assigned within each train path.

2.2.4. Operational variables

An overview of the seven studied operational variables is given in Table 3.

The distance covered by trains has been known to influence punctuality since at least Harris (1992), who showed that distance covered was statistically significant in determining punctuality. We include travel distance as a parameter, measured in kilometers.

In this paper, as in a previous one (Palmqvist et al., 2017a) we consider and count interactions between trains. We define an interaction as an instance when two trains are at the same place at the same time, which can happen at a station, or on a line section. When on a line section, the trains need to be traveling in the same direction to be counted as an interaction. This happens relatively frequently on double tracks, but is also possible on some single tracks, if there are multiple blocks between two stations. The number of interactions are

counted, and the count is used as a variable to explain punctuality variations.

As a measure of traffic intensity and station size, we count the number of trains arriving at the same station at the same hour as the train. Another measure of traffic intensity is the number of movements per day during the studied year, which we count across the whole network. Some vehicles are used more frequently than others, and to study the effect of this on punctuality we count the number of movements per vehicle individual during the timetable year. A movement is defined as crossing one line section, between two control points.

Some train numbers are also run more frequently than others: some run almost every day, others only a handful of times or even once during a year. Those that run more often might be expected to perform better, because of increased routine and increased incentives. To study this, we count the number of days run per train number.

While passenger volumes on both trains and stations are believed to be an important factor for delays, this type of data is often confidential and difficult to access, so the effects of these must be left for future studies.

2.2.5. Infrastructure elements

From the Swedish rail asset management database BIS, see Thaduri et al. (2015) for a description, we have high level information on eight types of infrastructure elements: their type and location. An overview of these is given in Table 4.

We match all but 1 500 of 82 700 elements to the train stations and railway links. We can distinguish between elements in station areas and those on links, but choose not to do so in this analysis to keep the number of variables from growing too large, and because the results are largely the same for most variables.

To add a dimension to this analysis, we also consider the density of elements, not only their number. We do this by dividing the distance traveled by a train by the count of elements of a given type passed by the same train, to arrive at an average distance between the elements.

Table 4. Overview of infrastructure elements

Infrastructure element	Approximate number of elements
Signal	31 000
Switch	15 500
Bridge	4 000
Tunnel	200
Level crossing	11 000
Cutting	10 000
Embankment	2 000
Fence	9 000

Whereas the number of infrastructure elements correlates strongly with the distance traveled, as trains covering longer distances pass by more infrastructure elements, the density measures instead depend on where the trains travel and how dense the infrastructure is there, they are not dependent on the distance covered.

We do not study the age or condition of infrastructure elements, or reported faults on them, because we do not have access to that data, only their number. Track works and temporary speed restrictions are not covered, for the same reason.

2.3. Data Analysis

The relationship between punctuality and the studied influencing factors is analyzed using a regression, t-tests and visual analysis of plots. Correlation coefficients are also presented.

A linear regression is performed over all the studied variables found in Table 7, with punctuality as the dependent variable and the other 36 variables as independent. The regression is performed in R. This is primarily done to see how much of the variation in punctuality is explained by the studied variables.

Thereafter the variables are studied individually, with regards to their influence on punctuality. This is done by first setting up several threshold values, then by performing t-tests to see if the punctuality above a threshold is significantly higher or lower than the average punctuality across all the trains in our dataset. See below for an example. Welch's two-sided t-test is used to allow for the fact that the samples are of unequal size and variances. The threshold is used to distinguish between observed trains with higher or lower values of the studied variable. When the p-value is found to be lower than 0.01, the punctuality for trains in the subset, surpassing the threshold, is compared to the punctuality across all trains in our dataset, and the difference is noted.

For instance, in our dataset there are 41 614 (out of 883 678) trains which have travelled for at least 400 km during their journey. They had a punctuality of 79.62 % which is 12.55
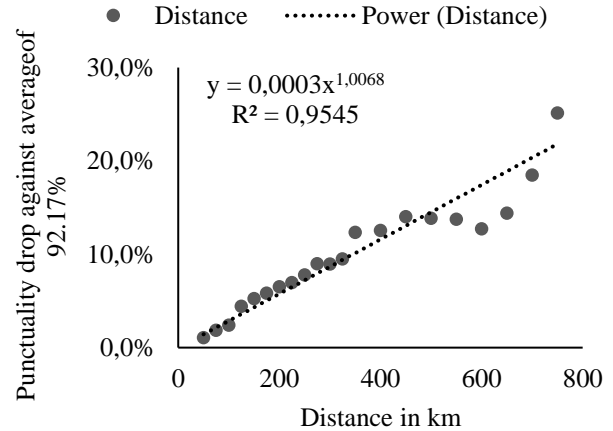

Figure 1. Punctuality and the distance traveled in km

%-points lower than the average of 92.17 %, plotted in Figure 1, and this difference is found to be statistically significant using a two-sided Welch's t-test, with a p-value approaching 0. From these t-tests we construct one plot per studied variable, as in Figure 1, with the threshold values on the x-axis and the differences in punctuality, compared to the average across all trains in our sample, on the y-axis. To continue the example above, we plot an x-value of 400 and a y-value of 12.55 %, and then repeat the same procedure again for the x-value of 450, determine the y-value and check whether this difference in punctuality is statistically significant with a p-value lower than 0.01.

This procedure is repeated at intervals for each studied variable, and the results plotted in scatter diagrams, to which trend lines are fitted. This is done in Microsoft Excel. The results are robust with regards to the range of the studied variable and the number of points per plot.

For each plot, we in turn try linear, exponential, logarithmic, second-degree polynomial and power functions, and choose the function which provides the best fit, as determined by the R^2-value. In some cases, where the difference in R^2 was less than 0.02, we instead opted for a linear trend line function, for simplicity.

3. RESULTS

The following section first describes the results of the linear regression containing all studied variables, and how each type of variable contributes to the overall picture, then a table summarizing the results for each variable separately. The results for each studied variable are then described and discussed category by category, to better illustrate the results.

3.1. Linear regressions

A linear regression across all 37 studied variables listed in Table 7 was performed in R. Due to the large number of variables considered, we omit the estimated coefficients for

Table 5. Summary results of the linear regression

Statistic	Value
Residual standard error:	0.2142 on 883641 degrees of freedom
Multiple R-squared:	0.04302
Adjusted R-squared:	0.04298
F-statistic:	1103 on 36 and 883641 DF
p-value:	< 2.2e-16

Table 6. Regressions by type of variable

Type of variables	No. variables	Adjusted R^2-value	Relative performance
Weather	6	0.02068	0.48
Timetable	7	0.02538	0.59
Operations	7	0.02943	0.68
Infrastructure	8	0.02355	0.55
Infrastructure density	8	0.00643	0.15
All of the above combined	37	0.04298	1.00

space concerns, and only present the summary statistics of the regression, which are presented in Table 5, below.

The results of the regression show that all but 5 variables (#2, #17, #33, #36 and #37 in Table 7) have significant impact with p-values < 0.00001. They also show that less than 5% of the variation is explained by this model, as well as a large residual standard error, showing that it has a low predictive accuracy even if the effect of the factors is shown to be significant. Changing the specification to consider polynomial functions only improves these numbers marginally.

We also carried out a series of regressions to study the impact of groups of variables, and their relative importance in explaining punctuality variations. Some summary results of these are presented in Table 6. These results show that the variables we have categorized as operational have the largest impact on punctuality, whereas the infrastructure density variables we used only contributed marginally. There is also some overlap between the types of variables, in the punctuality differences they explain. The rightmost column is calculated as the Adjusted R^2-value for that row, divided by the corresponding value in the bottom-most row, containing the combined model of all 37 variables.

3.2. Summary table

Table 7 summarizes the results for each variable, based on the method described in section 2.3. The #-column contains an identifier. The Variable-column contains brief, descriptive names of the variables. The Trend line function-column contains the trend line functions from the plots, described in section 2.3, where x is the threshold value of the variable, and F(x) is the decline in punctuality of trains where the threshold is superseded, compared to the average punctuality across all trains of 92.17 %. Punctuality is measured at all scheduled stops, and cancelations are treated as non-punctual. The R^2-column describes how well the trend line function fits the points plotted in each diagram. The Range plotted-column

lists the smallest and largest threshold values of the studied variable included in the plots, that are used to derive the trend line functions. The Pts./plot-column describes how many different thresholds were plotted to make up the diagrams, to which the trend lines were fitted. The Highest corr. to-column presents the correlation coefficient to the other studied variable which is the highest, as well as the identifier for that variable.

3.3. Weather

The result show that punctuality falls exponentially as the temperature drops below 0 °C. At -5 °C the punctuality is about 7.5 %-points lower than average, and at -30 °C it is about 50 %-points lower. The same pattern is found at high temperatures. At 23 °C punctuality is about 5 %-points lower than average, and by 27 °C it has fallen by 26 %-points. In the face of increasing temperatures and more frequent heat waves, this suggests that more ought to be done to increase the railway systems' resilience to high temperatures, similar to the findings of Ferranti et al. (2016), Ford et al. (2015) and others. Xia et al. (2013) used a series of dummy variables for temperature, and their plot of the effect on punctuality looks similar to ours.

The variation in temperature across the geography that the train passes through is also significant. We find a logarithmic relationship between punctuality and the difference between minimum and maximum temperature across the journey. Even a difference of 5 °C lowers punctuality by about 9 %-points. This variation is highly correlated (a correlation coefficient of 0.47) with the distance traveled, which is to be expected, as a train that travels longer passes through a larger geography and potentially larger temperature variations. That the temperature gradient affects punctuality is to be expected, some of the mechanisms are described in Tahvili (2016), but the effect we find is larger than in Xia et al. (2013).

Table 7. Summary of results from the analysis

#	Variable (x)	Trend line function	R^2	Range plotted	Pts. /plot	Highest corr. to
1	Punctuality, at all stops	Target variable, y-value	N/A	N/A	N/A	-0.16 to #21
	Weather					
2	Temperature, below 0°C	$0.035e^{-0.088x}$	0.97	3 - 32	18	-0.46 to #4
3	Temperature, above 0°C	$1E-04e^{0.2838x}$	0.96	15 - 27	12	0.13 to #2
4	Temperature, difference °C	$0.05041\ln(x)+0.0082$	0.90	1 - 26	25	0.50 to #26
5	Wind speed, max in m/s	$0.0001x^{2.1182}$	0.93	5 - 25	16	0.22 to #7
6	Snow depth, average in cm	$0.113\ln(x)-0.0134$	0.92	1 - 14	14	-0.37 to #2
7	Precipitation, sum of mm	$0.0002x + 0.0172$	0.97	1 - 1000	31	0.29 to #24
	Timetable					
8	Total margins, %	$1.3461x^2-0.632x+0.0526$	0.94	0.1 – 0.5	9	0.91 to #9
9	Supplement density, s/km	$0.0003x^2-0.0055x+0.01$	0.75	2 - 20	10	0.91 to #8
10	WAD of margins, %	$0.5178x^2-0.556x+0.1375$	0.69	0.35 – 0.80	11	-0.09 to #14
11	Negative margins, binary	$0.0277x-0.0116$	1.00	0 - 1	2	-0.35 to #8
12	Avg. speed with stops, km/h	$0.0003e^{0.0457x}$	0.81	40 - 120	10	-0.43 to #9
13	Travel time without margins, h	$0.0163x+0.0145$	0.95	0.5 - 10	20	0.89 to #21
14	Distance/Scheduled stops, km	$0.0013x-0.0199$	0.98	10 - 125	12	0.49 to #26
	Operations					
15	Station interactions, count	$0.0104x-0.0208$	0.95	2 - 20	10	0.73 to #24
16	Line interactions, count	$0.022x+0.0236$	0.97	1 - 5	5	0.08 to #24
17	Trains to same station & hour	$-0.0014x+0.0059$	0.92	3 - 30	11	0.53 to #18
18	Movements per vehicle	$-1E-12x^2+4E-07x-0.025$	0.98	100k–300k	9	0.53 to #17
19	Trains per day	$1E-11x^2-2E-06x+0.0819$	0.94	70k - 110k	5	0.15 to #17
20	Days run per train number	$-4E-05x+0.0012$	0.99	100 – 350	6	-0.20 to #32
21	Travel distance, km	$0.0003x^{1.0068}$	0.96	50 – 750	21	0.89 to #13
	Infrastructure					
22	Signal, count	$5E-05x-0.0055$	0.93	1 - 4500	10	0.95 to #24
23	Distance/Signal, km	$0.0081x^2-0.0059x+0.013$	0.88	0.1 - 5	16	0.79 to #25
24	Switch, count	$-6E-08x^2+0.0002x-0.027$	0.88	1 - 2000	14	0.95 to # 22
25	Distance/Switch, km	$0.0033x^2-0.012x+0.0154$	0.98	0.1 - 10	15	0.82 to #27
26	Bridge, count	$0.0007x-0.0017$	0.93	1 - 300	19	0.93 to # 22
27	Distance/Bridge, km	$0.0029\ln(x)+0.0085$	0.55	0.1 - 60	18	0.82 to #25
28	Tunnel, count	$0.0001x^2-0.0006x+0.009$	0.94	1 - 40	10	0.75 to #22
29	Distance/Tunnel, km	$0.0.0156x^{0.1501}$	0.77	0.1 - 160	25	0.44 to #13
30	Level crossing, count	$0.0005x+0.0057$	0.98	1 - 160	12	0.65 to #24
31	Distance/Level crossing, km	$0.0016x^2-0.0096x+0.004$	1.00	0.1 - 10	12	0.34 to #23
32	Cutting, count	$0.0004x+0.0007$	0.96	1 - 200	14	0.83 to #36
33	Distance/Cutting, km	$-8E-06x^2-0.0013x+0.004$	0.92	0.5 - 100	16	0.19 to #35
34	Embankment, count	$0.0053x^{0.5479}$	0.87	1 - 160	12	0.48 to #32
35	Distance/Embankment, km	$1E-06x^2+0.0001x+0.004$	0.94	0.1 - 160	25	0.25 to #30
36	Fence, count	$7E-07x^2+3E-05x+0.002$	0.97	1 - 400	21	0.92 to #26
37	Distance/Fence, km	$-4E-05x^2+0.002x+0.009$	0.76	0.1 - 60	19	0.48 to #25

The result suggests that a power curve best describes the influence of wind speed on punctuality. When wind speeds exceed 10 m/s punctuality is almost 2 % lower than average, and about 9 %-points lower when they exceed 23 m/s. This is a larger effect than found by Xia et al. (2013), who estimated that wind speeds of 23-26 m/s reduced punctuality by about 3.3 % in the Netherlands.

A linear function approximates the relationship between precipitation, measured as the sum of precipitation across the train stations passed by the train, and punctuality. A quarter of all trains accumulate at least 30 mm of precipitation, associated with a punctuality drop of 1.8 %-points compared to the average, and the drop increases about 2 %-points per 100 mm. Xia et al. (2013) used a slightly different measure, but found mostly linear effects of a similar magnitude.

The effect of snow depth on punctuality is best described by a logarithmic function fitted to the average snow depth, recorded at stations across the journey. While less than 6 % of the observations in our dataset have average snow depths larger than 1 cm, the magnitude of the effect, when it is present, is quite large: at an average of 5 cm the drop in punctuality is about 17.5 %-points. These effects are substantially larger than the estimates by Xia et al. (2013). The logarithmic function may suggest an increased preparedness and ability to deal with snow in the regions where large snow depths are often found, which decreases the harmful influence of snow.

3.4. Timetable

The results regarding timetable variables suggest that increasing margins benefits punctuality up to a point, after which it begins to decline. That point occurs at around 12 s/km, or 25-30 % of the minimum run time, at which point punctuality is around 2 %-points higher than the average of 92.17 %. These two different ways of measuring punctuality are very highly correlated with one another: the correlation coefficient is 0.91. This is well in line with earlier studies on margins (Palmqvist et al., 2017a).

Similarly, increasing the weighted average distance (WAD) of margins raises punctuality up to a point. The highest punctuality, almost 1 %-point higher than average, can be seen when the WAD is around 0.60. This confirms earlier findings by the authors, and shows that the effect exists even when punctuality is measured at intermediate stops, rather than just at the end destination.

The presence of negative margins in a timetable has a negative impact on punctuality, lowering it by on average 2.8 %-points compared to when there are no negative margins present. This figure is slightly lower than what we have found in earlier research (Palmqvist et al., 2017a), but on the same order of magnitude.

Overall, increasing average speeds of trains is linked with decreasing punctuality. Between average speeds of 60 and 120 km/h, including stopping times, an exponential curve provides the best fit. There is a very notable decrease in punctuality around 120 km/h, as the airport trains have an average speed of 118 km/h and a punctuality which is 5 %-points better than average, whereas the high-speed trains average 128 km/h (including stops) and 12 %-points lower punctuality than average. This makes for a clear break in the plot, and suggests that the speed is perhaps more of a proxy variable than the real issue.

Plotting punctuality against the scheduled duration of journeys, without margins, the result is a linear decrease in punctuality of about 1.6 %-points per hour. The duration of the journey is very highly correlated to the distance expressed in kilometers, with a correlation coefficient of around 0.87. The average distance between stops also appears to affect punctuality in a mostly linear fashion, by about 1.3 %-point for every 10 km. These findings are largely in line with our expectations, and with the fact that long distance trains often perform significantly worse in terms of punctuality.

3.5. Operations

The results in Table 7 furthermore suggest that punctuality drops linearly with the number of interactions at stations and on line sections. For interactions on line sections, which are rare in our data, the drop is about 2.2 %-points per interaction. For interactions at stations, which are much more common, the decrease in punctuality is approximately 1 %-point per interaction. These findings largely confirm what we have found in earlier research (Palmqvist et al., 2017a), and with the research on congestion by Gorman (2009).

The number of trains that arrive at a station during an hour is linked to punctuality in a linear manner. Punctuality increases slightly as the stations are handle more trains. At volumes of at least 20 trains per station and hour, the punctuality is about 2.5 %-points higher than average. This is an interesting finding, as increasingly congested stations are often suggested as a problem for punctuality (Palmqvist et al., 2017b).

We find that vehicles that are used more often have a higher punctuality than those that are used less frequently, with a quadratic relationship between the two variables. At 125 000 movements punctuality is about 0.5 %-point better than average, and at 300 000 it is 5 %-points higher. This is somewhat surprising, but suggests that operators prefer to use the more reliable vehicles for more intensive routes, and does a good job of keeping them in a working condition. One possibly confounding factor is that airport train vehicles seem, in our data, to be utilized much more heavily than other types of passenger trains, and the punctuality for these trains is very good.

How the number of trains per day affects punctuality is best described using a quadratic function, but the effect is relatively small. The largest number of trains operating in a

day we consider is 110 000, which is associated with a punctuality drop of 1.2 %-points.

Train numbers that are run more frequently are slightly more punctual than those that run less frequently. We find a linear relationship, with those running almost every day being about 1.2 % more punctual than average. This is in line with the suggestion earlier, that punctuality is improved when vehicles are operated more frequently, or that more reliable vehicles are used to run the most important routes. Unfortunately, the risk of the variable working as a proxy for the kind of passenger train is also present, as airport trains have the most days run and the highest punctuality, followed by commuter trains, and so on for regional, long distance and high-speed trains.

The single best indicator for punctuality is the distance traveled by a train. A linear function best fits our observational data, with a decline of about 3 %-points for every 100 km. The correlation coefficient with punctuality is -0.20, which is the highest in our findings.

3.6. Infrastructure

Finally, plotting the number of switches, tunnels and fences against punctuality, we find that they fit best to quadratic functions. Bridges, signals, level crossings, and cuttings show linear relationships to punctuality. Embankments fit best to a power function. Signals have the largest effects in terms of magnitude, being associated with punctuality drops of around 22 %-points at the most.

The quadratic relationships we find between punctuality and the number of switches, suggests a potential of gaining disproportionately large punctuality benefits by limiting their number. Particularly in large stations, where the current numbers are large and even small gains in punctuality are highly valuable.

When considering the distance between elements, the results are broadly similar across the different types of infrastructure, and the overall picture is that punctuality is better where the infrastructure is dense. An exception is for cuttings, where larger distances between them is associated with higher punctuality. Level crossings are another exception, where punctuality first improves as the distance between them rises to around 3 km, before it declines rapidly with increasing distances.

4. CONCLUSIONS, DISCUSSION AND FUTURE RESEARCH

In this paper, we have quantified how temperature, precipitation, snow depth and wind speed affect punctuality. Especially high and low temperatures can have large impacts on punctuality, and the effects are exponential. This is an important finding, which indicates that much more attention should be given to increasing the resilience of the railways with regards to heat, as the climate changes and temperatures rise. Overall, our findings with regards to weather are in line

with what others have found, though the magnitude of the effect is larger than in the Netherlands, for instance.

For timetabling, the highest punctuality is obtained when margins are around 25% of the minimum run time, or 12 s/km, with a slight shift towards the end of the journey and no negative margins. Punctuality also improves slightly when the train numbers are run more frequently and the vehicles less frequently. Traffic volume does not appear to be a major concern: variations due to a higher number of trains per day are small, and punctuality is slightly higher at more busy stations and times. However, the number of interactions between trains should still be minimized, as they are shown to lower punctuality. Many of these variables can be affected directly by planners and managers at train operating companies and infrastructure managers, such that punctuality improves. Even simple measures, such as reducing the number of trains with negative margins, have large impacts on punctuality.

Similar impacts are found with different types of infrastructure elements. Most infrastructure elements are highly correlated both with each other, and with the distance traveled, and for that reason it may be appropriate to construct a sort of infrastructure complexity index. The overall picture, however, is that a simple infrastructure with less components performs better. This, too, is something that planners can affect over time, as existing infrastructure can be modified, and new infrastructure can be designed in a manner that supports good punctuality.

The method of conducting a series of t-tests and plotting the results was successful in illustrating the relationship between punctuality and the studied variables individually. It illuminates and quantifies impacts that are elusive when using other methods. However, more work needs to be done to disentangle the effects of many of these variables from each other. The number of infrastructure elements depends to a high degree on the distance traveled, which was already known to be correlated with punctuality. Dense infrastructure and stops are, on the other hand, associated with local trains traveling shorter distances at lower speeds with more margins and higher punctuality. The linear regression of all studied variables together in this paper was intended to handle the sometimes-significant covariation of different variables.

We believe that one reason for the mediocre R^2-value in this regression, compared to earlier ones, is that we study all passenger trains in the national network for one year whereas other studies have looked at a smaller subset of trains, on one or two selected lines or regions, during which the conditions have been adverse, or using a measure that is more tailored to only find the faults that are under consideration. What we have tried to do is much broader, and it is no surprise that the degree to which we can explain the variation is smaller, simply because the variation studied is much larger. We also expect that data on passenger volumes, more detailed dwell times, as well as headway times in both the timetable and in

realized traffic would help explain more of the variation in punctuality.

It is our hope that this research can help both researchers and practitioners hone in on what can be done to improve punctuality. From improved heat-resilience of components, to better allocation of margins, more standardized train routes, and a simpler infrastructure with fewer components, there are many things that can be improved. The findings in this paper can also be used to assess the possible impacts of different measures, and to help prioritize between them.

REFERENCES

Andersson, E. V., Peterson, A. & Thörnquist Krasemann, J., (2013). Quantifying railway timetable robustness in critical points. *Journal of Rail Transport Planning & Management*, issue 3, pp. 95-110.

Andersson, E. V., Peterson, A. & Thörnquist Krasemann, J., (2015). Reduced railway traffic delays using a MILP approach to increase robustness in Critical Points. *Journal of Rail Transport Planning & Management*, issue 5, pp. 110-127.

Andersson, E., (2014). Assessment of Robustness in Railway Traffic Timetables. Licentiate thesis. Linköping University, Norrköping.

Brazil, W., White, A., Nogal, M., Caulfield, B., O'Connor, A. & Morton, C. (2017). Weather and rail delays: Analysis of metropolitan rail in Dublin. *Journal of Transport Geography*, vol. 59, pp. 69-76.

Carey, M., (1998). Optimizing scheduled times, allowing for behavioral response. *Transportation Research Part B*, vol. 32, no. 5, pp. 329-342.

Carey, M., (1999). Ex ante heuristic measures of schedule reliability. *Transportation Research Part B*, issue 33, pp. 473-494.

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 (WCRR 2016)*, Milan, Italy.

Cicerone, S., D'Angelo, G., Di Stefano, G., Frigioni, D. & Navarra, A. (2009). Recoverable robust timetabling for single delay: Complexity and polynomial algorithms for special cases. *Journal of Combinatorial Optimization*, issue 18, pp. 229-257.

Dewilde, T., Sels, P., Cattrysse, D. & Vansteenwegen, P., (2013). Robust railway station planning: An interaction between routing, timetabling and platforming. *Journal of Rail Transport Planning & Management*, issue 3, pp. 68-77.

Doll, C., Trinks, C., Sedlacek, N., Pelikan, V., Comes, T. & Schultmann, F. (2014). Adapting rail and road networks to weather extremes: case studies for southern Germany and Austria. *Nat Hazards,* vol. 72, pp. 63-85.

Ferranti, E., Chapman, L., Lowe, C., McCulloch, S., Jaroszweski, D. & Quinn, A. (2016). Heat-Related Failures on Southeast England's Railway Network:

Insights and Implications for Heat Risk Management. *Weather, Climate and Society*, vol. 8, issue 2, pp. 177-191.

Ford, A., Jenkins, K., Dawson, R., Pregnolato, M., Barr, S. & Hall, J. (2015). Simulating Impacts of Extreme Weather Events on Urban Transport Infrastructure in the UK. *International Symposium for Next Generation Infrastructure Conference (ISNGI 2014)*, pp. 233-238, Vienna, Austria.

Gorman, M. F. (2009). Statistical estimation of railroad congestion delay. *Transportation Research Part E*, pp. 446-456.

Haldeman, L. (2003). *Automatische Analyse von IST-Fahrplanen.* Zürich.

Hansen, I. A. (2009). Railway Network Timetabling and Dynamic Traffic Management. *2nd International Conference on Recent Advances in Railway Engineering (ICRARE-2009)*, pp. 135-144, Tehran, Iran.

Harris, N. (1992). *Planning Passenger Railways: A Handbook*. London: Transport Publishing Company Limited.

International Union of Railways (2000). *UIC Code 451-1 Timetable Recovery margins to guarantee timekeeping - Recovery margins.* Paris: International Union of Railways.

Jaroszweski, D., Hooper, E., Baker, C., Chapman, L. & Quinn, A. (2015). The impacts of the 28 June 2012 storms on UK road and rail transport. *Meteorological Applications*, vol. 22, pp. 470-476.

Kellermann, P., Bubeck, P., Kundela, G., Dosio, A. & Thieken, A. H. (2016). Frequency Analysis of Critical Meteorological Conditions in a Changing Climate – Assessing Future Implications for Railway Transportation in Austria. *Climate*, vol. 4, issue 25.

Kim, H., Kang, J., & Bae, Y. (2013). Types of train delay of high-speed rail: indicators and criteria for classification. *Journal of the Korean Operations Research and Management Science Society*, vol. 38, issue 3, pp.37-50.

Lehtonen, I. (2015). Four consecutive snow-rich winters in Southern Finland: 2009/2010-2012/2013. *Weather*, vol. 70, issue 1, pp 3-8.

Ludvigsen, J. & Klæboe, R. (2014). Extreme weather impacts on freight railways in Europe. *Nat Hazards*, vol. 70., pp-767-787.

Nagy, E. & Csiszar, C. (2015). Analysis of delay causes in railway passenger transportation. *Periodica Polytechnica Transportation Engineering,* vol. 42. pp. 73-80.

Nelldal, B.-L. (2009). *Kapacitetsanalys av järnvägsnätet i Sverige. Delrapport 3: Förslag till åtgärder för att öka kapaciteten på kort sikt*. Stockholm: KTH Royal Institute of Technology.

Nelldal, B.-L., Lindfeldt, A. & Lindfeldt, O. (2009). *Kapacitetsanalys av järnvägsnätet i Sverige. Delrapport 1: Hur många tåg kan man köra? En analys av teoretisk*

och praktisk kapacitet. Stockholm: KTH Royal Institute of Technology.

Nie, L. & Hansen, I. A. (2005). System analysis of train operations and track occupancy at railway stations. *European Journal of Transport and Infrastructure Research*, vol. 1, issue 5, pp. 31-54.

Nikolic, I., Arsovski, S., Eric, M., Vujcic, S., Manojlovic, G. & Jovanovic, J. (2016). Overall railway infrastructure effectiveness as a quality factor for Serbian railways. *Tehnicki vjesnik*, vol. 2, pp. 547-554.

Nilsson, J.-E., Björklund, G., Pyddoke, R. & Vierth, I. (2015) Regress – en god idé i järnvägssektorn? VTI Rapport 850. Linköping: VTI.

Norrbin, P., Lin, J. & Parida, A. (2016). Infrastructure Robustness for Railway Systems. *International Journal of Performability Engineering*, vol. 13, issue 3, pp 249-264.

Olsson, N. & Haugland, H. (2004). Influencing factors on train punctuality - results from some Norwegian studies. *Transport Policy*, issue 11, pp. 387-397.

Olsson, N., Halse, A. H., Hegglund, P. M., Killi, M., van der Kooij, R., Landmark, A., Seim, A., Sørensen, A. Ø., Økland, A. & Østli, A. (2015). *Punktlighet i jernbanen - hvert sekund teller.* Oslo: SINTEF akademisk vorlag.

Oslakovic, I. S., ter Maat, H. W., Hartmann, A. & Dewulf, G. (2013). Risk Assessment of Climate Change Impacts on Railway Infrastructure. *Engineering Project Organization Conference 2013,* July, 9-11, Devil's Thumb Ranch, Colorado, USA.

Pachl, J. (2002). *Railway Operation and Control.* Mountlake Terrace: VTD Rail Publishing.

Palmqvist, C.-W., Olsson, N. O. E. & Hiselius, L. (2017a). An Empirical Study of Timetable Strategies and Their Effects on Punctuality. Paper presented at the *7th International Conference on Railway Operations Modelling and Analysis (RailLille 2017)*, April, 4-7, Lille, France.

Palmqvist, C.-W., Olsson, N. O. E. & Winslott Hiselius, L. (2017b). Punctuality problems from the perspective of timetable planners in Sweden. Paper presented at *EURO Working Group on Transportation Meeting 2017 (EWGT 2017),* September, 4-6, Budapest, Hungary.

Parbo, J., Nielsen, O. & Prato, C. (2016). Passenger Perspectives in Railway Timetabling: A Literature Review. *Transport Reviews*, vol. 36, issue 4, pp. 500-526.

Qin, Y., Ma, X. & Jiang, S. (2017). A Stochastic Process Approach for Modeling Arrival Delay in Train Operations. *Transportation Research Board 96th Annual Meeting*, January, 8-12, Washington, D. C., USA.

Rietveld, P., Bruinsma, F. & van Vuuren, D. (2001). Coping with unreliability in public transport chains: A case study for Netherlands. *Transportation Research Part A*, vol. 35, pp. 539-559.

Scheepmaker, G. M. & Goverde, R. M. (2015). The interplay between energy-efficient train control and scheduled running time supplements. *Journal of Rail Transport Planning & Management*, issue 5, pp. 225-239.

Schöbel, A. & Kratz, A. (2009). A bicriteria approach for robust timetabling. In Ahuja, R. K., Möhring, R. H. & Zaroliagis, C. D. (Eds.) *Robust and online large-scale optimization*, pp. 119-144. Berlin: Springer Berlin Heidelberg.

Stenström, C., Parida, A., Lundberg, J., & Kumar, U. (2015). Development of an integrity index for benchmarking and monitoring rail infrastructure: application of composite indicators. *International Journal of Tail Transportation*, vol. 3, issue 2, pp 61-80.

Swedish Transport Administration. (2017). Tillsammans för tåg i tid - Resultatrapport 2017. Borlänge: Trafikverket.

Tahvili, N. (2016). *Winterization of Railways – issues and effects.* Master thesis, Norwegian University of Science and Technology, Trondheim.

Thaduri, A., Galar, D. & Kumar, U. (2015). Railway assets: A potential domain for big data analytics. *Procedia Computer Science*, vol. 53, pp 457-467.

Transport Analysis (2016). *Punktlighet på järnväg 2015.* Stockholm: Transport Analysis.

Veiseth, M., Olsson, N. O. E. & Saetermo, I. A. F. (2007). Infrastructure's influence on rail punctuality. *WIT Transactions on The Built Environment*, vol. 96, pp 481-490.

Vekas, P., van der Vlerk, M. & Haneveld, W. (2012). Optimizing existing railway timetables by means of stochastic programming. *Stochastic Programming E-Print Series*, vol. 8.

Wiggenraad, I. P. B. (2001). *Alighting and boarding times of passengers at Dutch railway stations.* Delft, TRAIL Research School.

Vromans, 2005. *Reliability of railway systems.* Doctoral thesis, Erasmus Research Institute of Management, Erasmus University Rotterdam, Rotterdam.

Xia, Y., Van Ommeren, J. N. , Rietveld, P. & Verhagen, W. (2013). Railway infrastructure disturbances and train operator performance: the role of weather. *Transportation Research Part D*, vol. 18, pp. 97-102.

Xu, P., Corman, F. & Peng, Q. (2016). Analyzing Railway Disruptions and Their Impact on Delayed Traffic in Chinese High-Speed Railway. *IFAC-PapersOnLine*, volume 49, issue 3, pp 84-89.

Yuan, J. & Hansen, I. A. (2007). Optimizing capacity utilization of stations by estimating knock-on train delays. *Transportation Research Part B*, vol. 41, pp. 202-217.

Yuan, J. & Hansen, I. A. (2008). Closed form expression of optimal buffer times between scheduled trains at railway bottlenecks. *11th International IEEE Conference on Intelligent Transportation Systems*, October, 12-15, Beijing, China.

Zakeri, G. & Olsson, N. O. E. (2017). Investigation of punctuality of local trains: The case of Oslo area. Paper presented at *EURO Working Group on Transportation Meeting 2017 (EWGT 2017),* September, 4-6, Budapest, Hungary.