A data-driven approach for generation of tactical planning rules regarding buffer time in initial railway timetables

A case study on the differentiation of buffer times in the railway timetable of Nederlandse Spoorwegen

S. E. (Susan) Bouman

Delft University of Technology

A data-driven approach for generation of tactical planning rules regarding buffer time in initial railway timetables

A case study on the differentiation of buffer times in the railway timetable of Nederlandse Spoorwegen

by

S. E. (Susan) Bouman

in partial fulfilment of the requirements for the degree of
Master of Science in Transport, Infrastructure & Logistics
at Delft University of Technology,
to be defended publicly on Friday July 1, 2022 at 16:00.

June 17, 2022

Student:	Susan Esmee Bouman	4558340
Project duration:	November 2021 - June 2022	
Supervisors:	Prof. Dr. R. M. P. (Rob) Goverde	TU Delft, CEG, Transport & Planning
	Dr. N. (Nikola) Bešinović	TU Delft, CEG, Transport & Planning
	Dr. S. (Stefano) Fazi	TU Delft, TPM, Transport & Logistics
	Ir. P. (Patrick) Looij	Nederlandse Spoorwegen, Netwerkontwikkeling & Ontwerp

An electronic version of this thesis is available at http://repository.tudelft.nl/

Preface

The thesis in front of you marks the end of my master thesis of the master programme Transport, Infrastructure and Logistics and with it the six years I spent studying at Delft University of Technology. The project was conducted in cooperation with Nederlandse Spoorwegen. During this period, I've learned so much about research, programming, the railway industry and above all myself.

First, I would like to express my gratitude to my graduation committee of the Delft University of Technology for their supervision. I would like to thank my chair Rob Goverde for first introducing me to the thesis topic, overseeing the entire process, and asking challenging questions. I would like to thank my daily supervisor Nikola Bešinović for his enthusiasm and sincerity in our meetings, keeping me on the right track and always challenging me to sharpen the overarching goal of the research. My other daily supervisor, Stefano Fazi, I want to thank for his valuable other perspective on the topic. His feedback ensured that there was a clear an consistent story line and that this thesis remained understandable for readers less involved in the process.

I would also like to thank my external supervisor from NS, Patrick Looij, for introducing me to various colleagues with extended knowledge on the topic, the detailed discussion we had on using the data to its fullest potential. It was a great pleasure to work with such a supportive and dedicated supervisor. Additionally, I am grateful for all colleagues at IPO for making me feel part of the team, sharing their expert knowledge, and always showing interest in the progress of my research.

A special words of thanks goes to my parents for always supporting and encouraging me throughout my study, but especially during the last eight months. A big thank you goes to my boyfriend Bob for encouraging me to keep progressing and for putting up with my complaints when yet another error emerged in my code. Lastly, I would like to thank my friends, especially Zara-Vé and Jolijn, for always lending me an ear when needed or distracting me when necessary. You have truly made my study period a pleasant balance between studying, having fun and self-development.

Many thanks to all, enjoy reading!

Susan Esmee Bouman
Delft, June 17, 2022

Summary

Introduction
The Dutch railway network is one of the busiest in Europe, with as many as 147 train per day per route kilometer in 2019. Under regular conditions, the timetable is adhered to apart from minor variations. But train operations are still subject to uncertainties that can disturb train services, cause delay to multiple trains, and propagate through the network. Delay can be dealt with in different stages of the railway transport management. One option is to mitigate delay already in the timetable design. To provide a high quality of services, buffer times are added to the minimum difference between the times two successive trains of either direction enter a section.

It is still common practice to design buffer times based on a deterministic value in combination with personal experience, decreasing operational capacity and requiring large amount of manual checking by planners. Existing approaches in literature to effectively allocate buffer time in timetables lack flexibility and require an initial timetable, but agree on the idea that effectively scheduling buffer times should be performed considering the actual size of the delay. This raises the question how more suitable buffer time planning rules can be designed that can be applied to the tactical timetable design stage when no initial timetable is present. Therefore, the following research question was formulated:

How can buffer time planning rules focused on reducing delay propagation be extracted from realisation data?

Methodology
The research question was addressed by applying a combination of methods. First, a literature review of the various theories and methods on buffer times in railway timetable design was performed. Second, grey literature and internal documents from ProRail and NS was consulted to gather insight into the practical background regarding buffer times. It was used to address the current planning norms, timetable design process and design practices. Third, the main research method was a statistical analysis of observed mean secondary delay and hindrance percentage in critical conflict points, retrieved from historical railway traffic realisation data. Two stepwise regression analyses were performed aiming to predict the two variables with the use of timetable characteristics related to headway situations of two succeeding trains. The timetable characteristics included to help predict the two target variables are (i) direction; (ii) number preceding conflicts; (iii) scheduled buffer; (iv) scheduled headway; (v) technical minimum headway; (vi) timetable point; (vii) type headway situation; (viii) type timetable point; (ix) event; (x) carrier; (xi) driving characteristic; (xii) previous and next timetable point; (xiii) rolling stock type; (xiv) rolling stock amount carriages; (xv) scheduled running time supplement; (xvi) train series (pattern); and (xvii) type of previous and next timetable point. The results of the regression analysis are used to determine the amount of scheduled buffer time that would ensure a certain amount of hindrance percentage given a specific headway situation

Results
The methodology has been applied to a case study of the Dutch railway network, between Haarlem, Leiden Centraal and Schiphol Airport. The general impression from the cae study is that the delay propagation at different critical conflicts is quite heterogeneous. The regression models were both able to explain 90.7% of this heterogeneity, meaning that structural delay propagation patterns can largely be contributed to timetable design and infrastructure characteristics. In particular, factors related to the route and train series of the trains contributing to the critical conflict appeared to be influential in explaining and predicting mean secondary delay and the probability of hindrance. Still, critical conflicts with the same mean secondary delay and hindrance percentage can have different secondary delay distributions, caused by among others (i) number of realisations of the conflict; (ii) number dispatching options; (iii) proactive attitude of traffic controllers; (iv) delay distributions of the first and second train at the previous timetable point; (v) the variety in scheduled buffer; (vi) the variety in rolling stock; and (vii) time of day.

The planning rules emerging from the regression models can be integrated in timetable design processes as (i) the size of the wake added in the simulation; or (ii) an indication of when slightly hindered headways are acceptable. To reduce the workload for planners further, it would be favourable that the needed scheduled buffer for a specific

conflict is automatically shown in the planning application. To this end, all relevant factors explaining these aspects of delay propagation ought to be integrated in planning applications. The standardisation could help streamline data-driven timetable design.

Conclusion

The general conclusion of this paper is that an estimation or prediction of the expected delay propagation between two trains in a given situation could be done quite accurately (adjusted R-squared of both models equals 0.907) is essential to determining suitable buffer time planning rules, because it provides insight into where buffer time can be placed in order to be most effective. This research has operationalised delay propagation as the mean secondary delay and probability of hindrance a succeeding train encounters caused by a delay of a preceding train. Using the prediction model, the planning rules can be determined by setting a desired value for these metrics. This value could possibly be determined by making a trade-off between delay propagation and capacity based on a chosen objective (e.g. passenger punctuality, infrastructure consumption). Interesting to note is that the mean secondary delay is not significantly impacted by the scheduled buffer, but the probability of hindrance is. This indicates that the delay distributions of both trains have a bigger impact on the size of the secondary delay, while the size of the scheduled buffer determines how often hindrance occurs. With regard to planning buffer time focused on reducing delay propagation in the Netherlands, improvements can be made concerning design practices and data processing. In particular, it is favourable that the flexible needed scheduled buffer for a specific conflict is automatically shown in the planning application, to reduce the workload for planners. This standardisation could help streamline the data-driven timetable design.

Contents

List of abbreviations

CRM Critical Robustness Measure

FRISO Flexible Rail Infra Simulation Environment (Dutch: Flexibele Rail Infra Simulatie Omgeving)

ILP Integer Linear Programming

MIP Mixed-Integer Programming

PESP Periodic Event Scheduling Problem

PoH Percentage of Headways equal to or less than the minimum value

RCP Robustness in Critical Points

ROBERTO Running time and Headway time Calculation Tool (Dutch: Rij en Opvolgtijd Berekenings Tool)

SAHR Sum of Arrival Headway Reciprocals

SSHR Sum of Shortest Headway Reciprocals

TTP Train Timetabling Problem

VIF Variance Inflation Factor

WAD Weighted Average Distance

List of definitions

adjusted R-squared value the R-squared value, while accounting for the number of predictor variables included in the regression model, leading to a more parsimonious model

block section an area protected by signals within which only one movement at a time may take place (ProRail, 2015)

blocking time the time interval in which a section of track is allocated to the exclusive use of one trains and therefore blocked to all other vehicles (Pachl, 2014)

buffer time an extra time that is added to the minimum line headway to avoid the transmission of small delays from a train to a succeeding train (Pachl, 2014)

buffer time conflict overlap in extended blocking times (Scheepmaker and Goverde, 2021)

buffer time effectiveness the share of total buffer time that is actually used to reduce delays

coasting a driving regime where no traction nor braking is applied

critical block section the block section where the blocking time stairways of two successive trains would touch each other without any tolerance, or in case of timetable conflicts, the block section where the overlap in (extended) blocking times of two successive trains is the biggest

critical conflict the conflict between two train series patterns on the critical timetable point

crossover a timetable point for changing tracks on open track via at least two switches (Dutch: overloop; ProRail (2015))

delay jump the difference in delay between consecutive scheduled events of a train (Dutch: vertragingssprong; Goverde and Meng (2011))

Donna a micro-scopic planning tool used by all railway carriers in the Netherlands for short term planning

DONS a micro- and macroscopic timetable planning tool used by NS and ProRail for long term studies

dwell time supplement

halt a timetable point on open track intended and designed for stopping train and also for boarding and disembarking passengers and/or for accepting and delivering goods, where trains cannot overtake (Dutch: halte, ProRail (2015))

hindrance an operating situation in which a train will encounter a signal showing a restrictive aspect and will therefore have to modify its speed, due to a conflict with a preceding train (Pachl, 2014). Note that timetable conflicts do not necessarily lead to hindrance, because in the realisation phase a second train can also be delayed before it arrives at the location of the conflict, therefore eliminating the conflict.

junction a timetable point that is exclusively equipped for the convergence of 3 or more open tracks (Dutch: aansluiting; ProRail (2015))

knock-on delay a delay caused by other trains due to either short headway times or late transfer connections (Pachl, 2014)

R-squared value represents the proportion of the variance of the target variable that is explained by the predictor variables in the regression model

route conflict overlap in technical minimum blocking times

running time supplement a time allowance added ti the shortest possible running time between scheduled stop, to facilitate a recovery or reduction of a delay within the timetable of one train (Dutch: rijtijdspeling, Pachl (2014))

secondary delay a delay caused by preceding trains, see knock-on delay

section an area delimited by section boundaries that is reported available or occupied as a whole. So the smallest unit of rail infrastructure that can be used for one purpose at a time

station a timetable point intended and designed for stopping, terminating, overtaking or crossing trains and provided with at least one switch and also for boarding and disembarking of passengers and/or for accepting and delivering goods (Dutch: station; ProRail (2015))

stop indicates a stop of a train on a halt or a station to board and disembark passenger and/or to accept and deliver goods (Dutch: haltering; ProRail (2015))

time allowance collective term for slack time in the timetable, i.e. buffer time, running time supplement and dwell time supplement

timetable conflict overlap in (extended) blocking times in the timetable. Collective term for route conflict and buffer time conflict

timetable conflict a (buffer time) conflict observed in the timetable

timetable point primary area that forms a continuous, delimited part of the railway network and that fulfills a function in defining the timetable. It moreover functions as a measuring point to determine any deviation from the timetable. (Dutch: dienstregelpunt, ProRail (2015))

wake represents buffer time in a microscopic simulation. The block after the signal is released 60 seconds later than would be possible according to the blocking time model (Dutch: kielzog)

1
Introduction

This thesis explores the ability to create buffer time planning rules based on realisation data suitable for usage in an initial timetable. In this chapter, the main problem and contributions addressed in this research are described. Section 1.1 starts with an introduction of the problem. Next, Section 1.2 states the scope and goal of the research as well as the research questions. Section 1.3 follows with the scientific and societal contributions of this research. Finally, in Section 1.4 an overview of the document structure is presented.

1.1. Problem description

This section first discusses the general problem and motivation of this research. Thereafter the problem statement that follows from the literature research and practical background is presented.

1.1.1. Problem context

The Dutch railway network is one of the busiest in Europe, with as many as 147 trains per day per route kilometer in 2019 (Independent Regulators' Group - Rail, 2021), and its network usage intensity will become even higher with the frequency increase of train services under the High-Frequency Rail Transport Programme (Ministerie van Verkeer en Waterstaat, 2010). The programme aims to realise frequencies that enable timetable-free travel for passengers, with up to 12 intercity trains and 6 sprinter trains every hour on high-utilised trajectories in the country, as opposed to a maximum of 8 and 4 trains of each respective category per hour before the start of the program. This increase in frequency requires more efficient use of the current infrastructure capacity as well as reliable train operations. Timetabling is performed to serve exactly these functions (Pachl, 2014), and becomes even more important and complicated with the higher network usage intensity.

Under regular conditions, the timetable is adhered to apart from minor variations in the train services due to differences in driving behaviour and stochasticity in process times. Despite advanced communication, monitoring, and control facilities, train operations are still subject to uncertainties that can disturb train services, cause delays to multiple trains, and propagate through the network. Delays have a negative impact on the performance of the railway system, such as increased travel time, missed transfers and extended working hours of staff. Thereby, they lead to higher operational costs and a lower level of service.

Delay can be dealt with in different stages of the railway transport management. In real-time, dispatching decisions can focus either on reducing the impact of delays, aiming for overall passenger punctuality by accepting some cancellations, or on minimising cancellation of delayed trains. In the Netherlands, the first strategy is applied, resulting in an average cancellation ratio and a high passenger punctuality compared to other train operators in Europe (Independent Regulators' Group - Rail, 2021; Nederlandse Spoorwegen, 2020).

Another option is to mitigate delay already in the timetable design. To provide a high quality of services, time allowances are added to the shortest possible running time between scheduled stops (i.e. running time supplement) and minimum line headway, which corresponds to the minimum difference between the times two successive trains of either direction enter a section (Pachl, 2014) (i.e. buffer time). The former aims to facilitate a recovery or reduction of a delay within the timetable of one train, whilst the latter enables prevention and reduction of delay propagation between trains. Both types of time allowance accommodate variation in driving behaviour and small deviations in process times (Goverde, 2014). This thesis' focus is on contributing to the timetable design part with a particular emphasis on buffer times.

1.1.2. Problem statement and knowledge gaps

It is still common practice to design buffer times based on a deterministic value in combination with personal experience and tacit knowledge. Although easy to apply, this method (i) often decreases operational capacity in heavily utilised networks; and (ii) requires a large amount of manual checking in case of smaller scheduled buffer times. Both aspects affect the extent to which buffer times are effective in reducing delay propagation. This is supported by findings from Carey and Kwieciński (1994), Yuan and Hansen (2007), Yuan and Hansen (2008), and Zieger et al. (2018), showing that the realised headway and the type of buffer time distribution have a significant influence on the expected knock-on delays and capacity. Therefore, there is a need for buffer time allocation techniques that achieve a high effectiveness of buffer times.

When creating an initial timetable from scratch, the literature predominantly considers minimum buffer time, or minimum headway time including an explicit minimum buffer time as a further simplification, as a deterministically given input value (Cacchiani and Toth, 2012; Caprara et al., 2002; Scheepmaker and Goverde, 2021; Serafini and Ukovich, 1989). Most works consider this to be a hard constraint, while in practice a smaller buffer is only rejected when (i) an unplanned stop would occur, or (ii) an arrival delay bigger than 30 seconds would occur (ProRail, 2020). Thus, the optimisation models in the literature lack flexibility regarding the trade-off between headway time and running time. One way of introducing flexibility to these type of models is to make the minimum buffer time input variable. Therefore, an approach that can differentiate buffer time is required.

With regard to allocation of running time supplements and buffer times to reduce delay propagation and to improve timetable robustness, several approaches have been introduced in literature (Hauck and Kliewer, 2019; Huang et al., 2019; Jin et al., 2019; Jovanović et al., 2017; Khoshniyat and Peterson, 2017; Kroon et al., 2008; Yang et al., 2020). The main limitation of these approaches is that they are only able to adjust an existing timetable by reallocating the current running time supplements and buffer times. Thus, a predefined train order is assumed, the total amount of scheduled buffer remains unchanged and an initial timetable is required. This raises the question how more suitable buffer time planning rules can be designed that can be applied to the tactical timetable design stage when no initial timetable and train order is present.

The rising availability of data has enabled analysis of empirical delay distribution, delay propagation and headway variation (Corman and Kecman, 2018; Dietzenbacher, 2021; Zieger et al., 2018). The findings have been applied to predict train delays and reallocate buffer times, but do not explicitly specify how buffer times should be allocated to alleviate delay propagation most effectively. Furthermore, some approaches have explicitly aimed to improve the effectiveness of running time supplements and dwell time supplements to recover from delays based on historical data (Huang et al., 2019; Yang et al., 2020). Thus, the proposed methods mainly focus on the reduction of primary delays and are unable to generalise the findings, such that it can be applied to create initial timetables where no realisation data is available yet. The latter is crucial to cope with delay propagation effects already while constructing an initial timetable.

Combining the scientific gaps, identified in Chapter 2, and the practical gaps, identified in Chapter 3, this research addresses the following knowledge gaps:

- It is unknown to what extent specific timetable characteristics (e.g. scheduled buffer time, scheduled running time supplement, train characteristics, headway situation, location) influence delay propagation.
- The effective allocation of buffer times to alleviate delay propagation based on historical realisation data has not been widely addressed in the tactical planning stage of railway timetables where an initial timetable is constructed.
- Planning rules for buffer time based on historical realisation data have not been widely explored in literature.

1.2. Research design

This section describes the design of the research. First the scope of the research is defined. Next, the goal and research questions are stated. Finally, the methods used to answer the research questions are explained.

1.2.1. Research scope

The scope of this research is defined as follows:

- *Focus on buffer times*
  running time supplements allocation in an initial timetable in relation to improving timetable robustness has been addressed in Khoshniyat and Peterson (2017) and Scheepmaker and Goverde (2015). In these studies, running time supplements have proven not only to be beneficial to reduce the delay of one train, but also to

facilitate energy efficient driving in case of no delays. Moreover, Huang et al. (2019) and Yang et al. (2020) have presented methods to enhance allocation of running time supplements and dwell time supplements based on realisation date. Analytical approaches that aim to enhance timetable robustness based on headway have been developed by Carey and Kwieciński (1994) and Vromans et al. (2006), but they do not explicitly consider buffer time. There is less knowledge about buffer time allocation in an initial timetable, whilst its distribution has proven to be of great influence to the built-up of knock-on delays (Zieger et al., 2018). Therefore, this research only focuses on the differentiation of buffer times.

- *Focus on reduction of delay propagation*
  As mentioned in Section 1.1, buffer time is added to the minimum line headway to provide a high quality of services by enabling prevention and reduction of delay propagation between trains. Currently, in practice, the emphasis in scheduling with buffer times is on assuming a certain value and only changing it when tighter headways are desired in order to gain extra capacity. The focus in this research will be reversed; to investigate when a greater buffer time is desired in order to reduce delay propagation.

- *Focus on planning rules for an initial timetable based on realisation data*
  The railway traffic system can be divided into the strategic, tactical, and operational phase (Ibarra-Rojas et al., 2015). Respectively they (i) define desired connections and required performances, (ii) define train services and timetables, and (iii) monitor train operations and manage disturbances and disruptions. In this research, realisation data from the operational stage is used to improve the tactical rules and procedures with the goal to improve the quality of the timetable with regard to delay propagation.

1.2.2. Research objectives and research questions

To fill the knowledge gaps identified in Section 1.1, the aim of this research is twofold. On the one hand, the goal is to identify timetable characteristics that influence the delay propagation. On the other hand, the aim is to provide a data-driven approach for determination of buffer time planning rules suitable for usage in an initial timetable. These planning rules are not necessarily generic, but can depend on the characteristics that have been identified by achieving the first objective.

To achieve the objectives, this research will answer the following main research question:

How can buffer time planning rules focused on reducing delay propagation be extracted from realisation data?

In order to provide an answer to the main research question, the following sub-questions are composed:

1. What can be learned from previous quantitative, data-driven approaches for scheduling time allowances in railway timetable design?

2. What is the current state of the practice regarding planning norms for buffer times and regarding data-driven timetable design approaches in the Netherlands?

3. What data-driven model(s) can be used to determine buffer time planning rules in the tactical timetable design phase?

4. How are buffer time planning rules determined from the data-driven model(s) and how do they compare to the current planning norms for buffer times?

5. Which insights do the data-driven model(s) and buffer time planning rules bring that may help practitioners effectively allocate buffer times in the tactical timetable design phase?

1.2.3. Research approach

The research questions are answered with a combination of desk research, quantitative data analysis and case study. Below is an explanation of each of the methods.

Desk research

The first part of this research is a desk research, consisting of a review of literature and grey literature. It starts with a literature review on buffer times in railway timetable design. The review discusses (i) timetable design concepts; (ii) timetable design methods; (iii) data-driven delay distribution, propagation and recovery models; and (iv) data–driven buffer time scheduling. Furthermore, it highlights the research gaps addressed in this thesis and it answers the first sub-question. The discussed literature consists of scientific papers and conference papers found in Scopus based on a keyword search within the article title, abstract and keywords. Additional articles were found by applying the forward and backward snowballing principle (Wohlin, 2014).

In addition grey literature and internal documents from ProRail and Nederlandse Spoorwegen are consulted to gather insight into the practical background regarding buffer times. It is used to address the current planning norms, timetable design process and design practices. The literature review and practical background are presented in Chapter 2 and Chapter 3 respectively.

Data analysis
The second method is a quantitative data analysis. The analysed data concerns conflict data from the timetable, complemented with historical railway traffic realisation data. The conflict data is used to identify what headway situations are planned tighter than the norms and the realisation data is used to assess how these critical headway situations turned out in the operation.

Using this information, it is determined how delay has propagated from a preceding train to a succeeding train on critical conflict locations. Then, regression models are established to describe and predict the delay propagation. Next, it is investigated which model fits the data the best to extract buffer time planning rules from. This method is in line with current buffer time allocation approaches and research gaps (Huang et al., 2019; Wen et al., 2019; Yang et al., 2020). More details on the data analysis are provided in Chapter 4 and Chapter 5.

The data analysis is performed in Jupyter notebook and coded in Python with the use of Pandas (McKinney, 2010), Matplotlib (Hunter, 2007), Seaborn (Waskom, 2021), Scipy and Scikit-learn (Pedregosa et al., 2011; Virtanen et al., 2020), and Statsmodels (Seabold and Perktold, 2010).

Case study
According to Gerring (2004), a case study can be best defined as "an intensive study of a single unit for the purpose of understanding a larger class of (similar) units". This research's case study focuses on a subsystem of the Dutch railway network to show how the methodology can be applied to infer buffer time planning rules from railway traffic realisation data and how its effects can be assessed. The case study is elaborated on in Chapter 5.

1.3. Scientific and societal contributions

The main scientific contribution of this research consists of a new data-driven approach to allocate buffer times in an initial timetable. The approach is generic and can be applied to a larger scope and to railway networks in other countries. Additionally insights are generated into what timetable characteristics contribute to delay propagation in critical conflict points.

The main societal contribution of this research consists of recommendations for designing buffer time planning rules in a data-driven fashion. This information could support the commissioner, Nederlandse Spoorwegen, in improving their timetable to better avoid delay propagation and in identifying new timetable structures due to more specific buffer time planning rules. Furthermore, this work adds in gaining a better understanding of the interactions with buffer time. A last contribution is the potential this research demonstrates for timetable design based on traffic realisation data, by showing that the necessary data are available and that relevant conclusions can be drawn from the analysis of headways.

1.4. Structure of the thesis

The remainder of the report is structured as follows: Chapter 2 starts with a literature review on timetable design concepts and methods, data-driven delay models and buffer time scheduling and ends with an overview of the scientific gaps this research fills. Subsequently, Chapter 3 presents the current norms and working practices in the Netherlands regarding timetable design and buffer time scheduling. Chapter 4 follows to explain the approach of the research in more detail, including the framework used to determine buffer time planning rules. The case study and its results are presented in Chapter 5. Finally, a discussion of the results is provided in Chapter 6 and conclusions are presented in Chapter 7. An overview of the report structure is depicted in Figure 1.1.

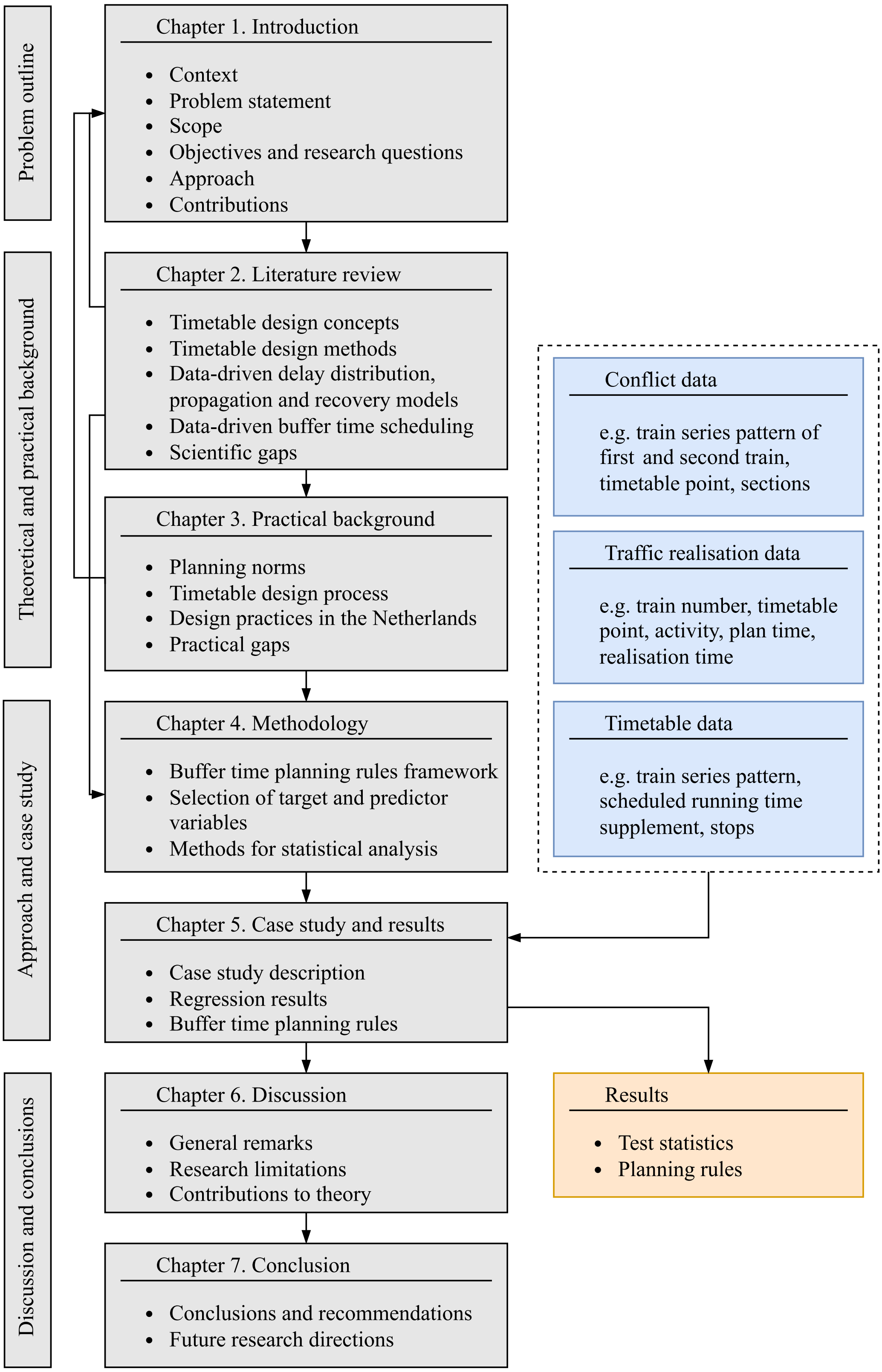


Figure 1.1: Research flow diagram

2
Literature review

This chapter provides a literature review of the various theories and methods on buffer times in railway timetable design in scientific literature. First, Section 2.1 starts with an introduction on timetable design concepts. Section 2.2 presents nominal, robust and multi-level timetable design methods. Section 2.3 discusses previous data-driven methods on buffer time scheduling and delays. Subsequently, existing research gaps are highlighted in Section 2.4. Finally, Section 2.5 concludes the chapter and answers sub-question 1.

2.1. Timetable design concepts

This section introduces the basic principles of timetable design. First, the level of detail of the design is discussed and later time allowances and periodicity in the timetable are addressed.

2.1.1. Level of detail

When designing a timetable, multiple levels of detail can be applied for representing infrastructure and determining headways. Literature distinguishes the macro and micro level, briefly defined below. For a more elaborate distinction between macro- and microscopic infrastructure modelling, see Radtke (2014).

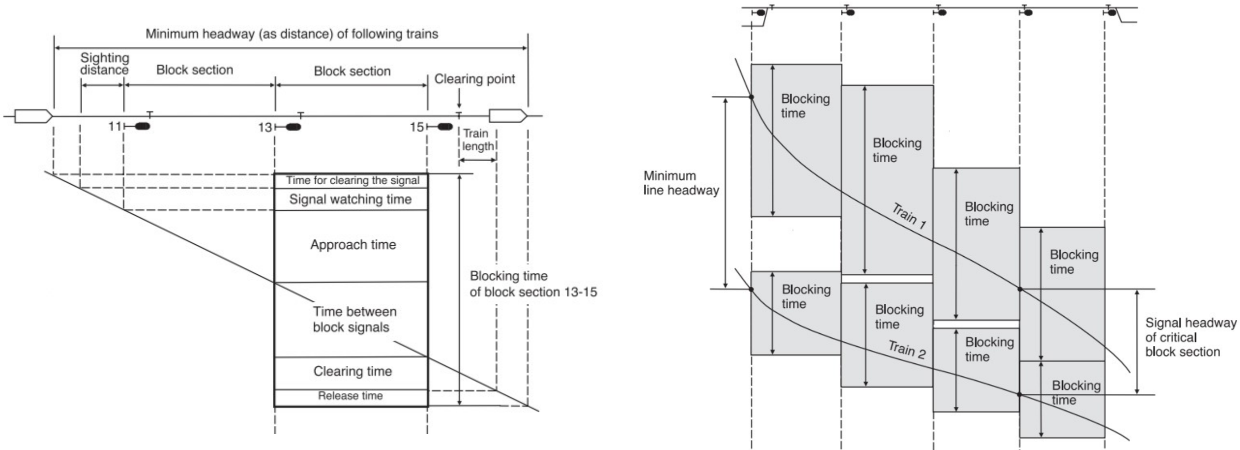
Macroscopic timetabling

Macroscopic modelling of infrastructure contains a network with homogeneous sections (i.e. constant values for speed restrictions and track gradients). Nodes represent timetable points, such as stations or junctions, and links correspond with tracks. These type of models are mainly useful for long term planning and traffic assignment. The running time can be based on tables of running times for certain train configurations and for different combinations of station stops and passing through non-stop. These tables are based on microscopic modelling of infrastructure, but are generalised to fit macroscopic timetabling. With regard to headway determination, in macroscopic modelling headways are often based on a generic planning norm. The norms differ for different events of the first and second train. Events considered by ProRail (2020) are arrival, passing, short stop and departure.

Microscopic timetabling

Microscopic models represent the infrastructure in a highly detailed manner. A typical microscopic model contains all tracks, signals, gradients, sections, block sections and stopping positions on lines, junctions and stations. This makes it a suitable approach for runtime calculations and simulations. The running time between an arrival and departure can be calculated based on how the train accelerates, runs and brakes. For this detailed infrastructure and train data are required, so microscopic modelling is necessary. A microscopic model that enables minimum headway determination is the blocking time model, explained in detail by Pachl (2014).

This theory is based on the principle that the railway infrastructure is divided into block sections, each with a signal that indicates whether a train can enter the section. The blocking time refers to the time a block section is devoted to a specific train and therefore not accessible to other trains, visualised in Figure 2.1a. When computing the blocking time for multiple blocks for a train, the shape of a staircase emerges. The blocking time staircases of consecutive trains are stacked as shown in Figure 2.1b to determine the minimum headway when no buffer time is considered. The minimum headway is influenced most by the speed profile of the train and the critical block section (i.e. the block section where the blocking time stairways of two successive trains would touch each other without any tolerance). In the Netherlands the tool ROBERTO was developed to enable generation of large sets of minimum headway times (Middelkoop, 2010), and recently this functionality was added to the timetable design applications DONS and Donna, making ROBERTO obsolete.



(a) Components of the blocking time

(b) Blocking time stairway to determine minimal headway

Figure 2.1: Blocking time model, adopted from Pachl (2014)

2.1.2. Time allowances

To provide a high quality of services, time allowances are added to the shortest possible running time between scheduled stops (i.e. running time supplement), minimum dwell time at stations (i.e. dwell time supplement) and minimum line headway (i.e. buffer time). The former two types of slack aim to facilitate a recovery or reduction of a delay within the timetable of one train, whilst the latter enables prevention and reduction of delay propagation between trains. Moreover, all three time allowances accommodate variation in driving behaviour and small deviations in process times (Goverde, 2014). Buffer time can be visualised by adding space between the stacked blocking time stairways in Figure 2.1b, as depicted in Figure 2.2. Then, the buffer time is the smallest slot between the blocking time stairways of two trains (Pachl, 2014). Running time supplement is added to minimum running time of a train. Figure 2.2 shows an example where the running time supplement is evenly spread over the train path, but there are also situations where the running time supplement is concentrated at the end of the run or at large intermediate stations.

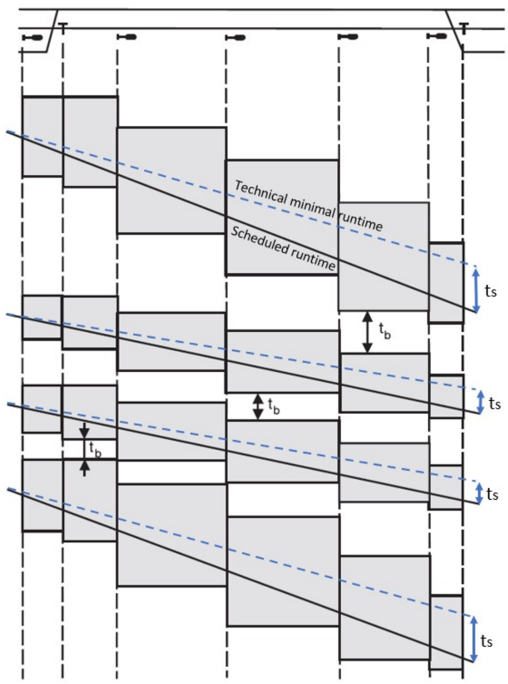
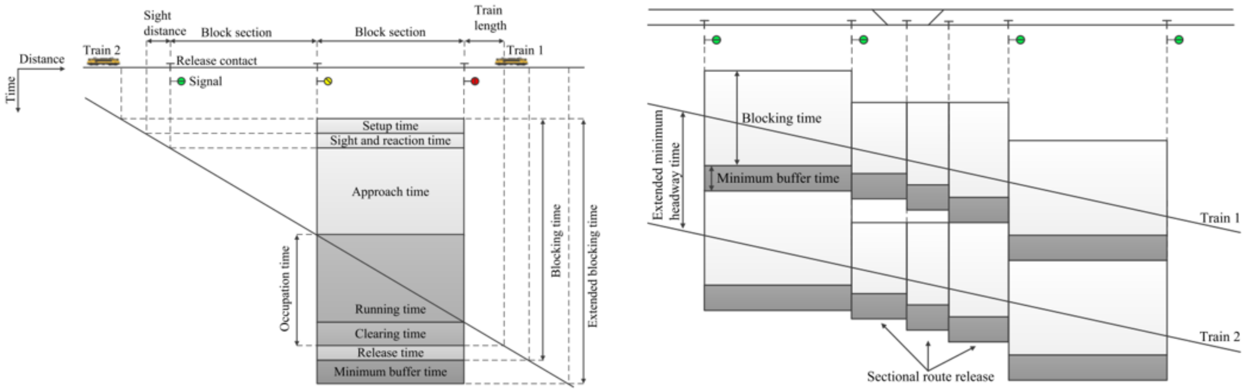


Figure 2.2: Simplified microscopic distance-time graph showing the running time supplement (t_s) and buffer time (t_b), adjusted from Pachl (2014)



(a) Components of the extended blocking time

(b) Extended blocking time stairway

Figure 2.3: Extended blocking time model, adopted from Scheepmaker and Goverde (2021)

In practice, a minimum buffer time is often considered. Therefore, Scheepmaker and Goverde (2021) propose the notion of extended blocking time, which includes a minimum buffer time, as shown in Figure 2.3a. Stacking the extended blocking times of two trains results in the extended blocking time stairway depicted in Figure 2.3b. This concept is also used in practice to gain insight into the robustness of the timetable.

This enables identification of two types of timetable conflicts; route conflicts occur when blocking times of two trains overlap, and buffer time conflicts occur when the minimum buffer time overlaps with the blocking time of another train.

2.1.3. Periodicity

A periodic, or cyclic, timetable has a pattern that repeats itself. Usually, a cycle time of one hour is scheduled. The advantages are convenience for travellers, convenience for the planning, and compact in presentation. Disadvantages are the lack of flexibility and increased costs. A distinction can be made between non-symmetrical, symmetrical and integrated timetables, where the latter is focused on aggregating all trains at a connecting station at the same time.

2.2. Timetable design methods

This section reviews literature on timetable design methods, introducing optimisation methods and analytical approaches for buffer time determination.

2.2.1. Timetable optimisation methods

Given a proposed line plan with desired frequencies and stops, timetabling problems entail determining timetables for a set of lines, satisfying track capacity constraints with the aim of optimising an objective function (Cacchiani and Toth, 2012; Lusby et al., 2011). A distinction can be made between nominal and robust versions. The former has a focus on determining efficient timetables, whilst the latter concerns creating schedules that avoid, in case of disruptions in the railway network, delay propagation as much as possible.

Nominal methods

Timetabling problems are in literature usually referred to as the Periodic Event Scheduling Problem (PESP) (Serafini and Ukovich, 1989) or Train Timetabling Problem (TTP) (Caprara et al., 2002). The PESP model is a mathematical model that periodically schedules events (i.e. arrivals or departures from a given station) and represents a feasibility problem. Extensions have been made regarding objective functions (e.g. minimising travel time, capacity consumption and maximising passenger satisfaction) and solving methods. The TTP aims at finding a feasible, efficient and stable non-cyclic timetable for a single one-way corridor. Some works have used TTP as a basis to include real-world constraints and apply the model to a railway network. For an overview of PESP- and TTP-based research, see Cacchiani and Toth (2012).

A drawback of nominal methods is that they are purely focused on feasibility of the timetable. Thereby, they do not consider the distribution of time allowances, but rather assume norms for running time and headway time as an input. This restricts optimally coping with stochastic variations of process times and disturbances.

Robust methods

In order to create timetables that can absorb delays as much as possible and that can avoid delay propagation, attention has been given to the robust version of the timetabling problem. Many works have been dedicated to defining and measuring robustness, but no unanimous formulation has been defined. In the light of this research, the definition of Goverde and Hansen (2013) is adopted: *"Robustness comprises the ability of a timetable to withstand design errors, parameter variations, and changing operational conditions"*.

Salido et al. (2008) acknowledge that robustness of a given timetable can be increased by adding time allowances, decreasing the number of trains and decreasing heterogeneity in types of trains. To this end, buffer time and running time supplement are reallocated. In timetabling, multiple approaches to achieving robustness have been introduced, such as (i) stochastic programming (Jin et al., 2019; Kroon et al., 2008); (ii) light robustness (Fischetti and Monaci, 2009); (iii) recoverable robustness (Cicerone et al., 2009; Liebchen et al., 2009); (iv) delay management (Liebchen et al., 2010); and (v) bi-objective approaches (Schöbel and Kratz, 2009) An overview of robustness-centred timetable design methods can be found in Cacchiani and Toth (2012) and Lusby et al. (2018).

The main limitation of existing robust timetabling methods is that they require an initial nominal timetable, which is adjusted to reallocate the current time allowances. Thereby the studies exclude evaluating the total size of the time allowances and are incapable of changing the order of the trains. Furthermore, in the cases where no iterative approach is used, the quality of the initial timetable is a high determining factor of the final timetable's quality.

Multi-level timetabling

The PESP- and TTP-based models, both in its nominal and robust version, are predominantly modelled at a macroscopic level, thereby they neglect timetable feasibility at the microscopic level. Therefore, approaches have been introduced to better integrate macroscopic models with more detailed models in a multi-stage approach (Bešinović et al., 2016; Goverde et al., 2016).

These types of models usually contain a macroscopic timetabling part for constructing the timetable, and a microscopic simulation part for evaluating the performance of the timetable. Bešinović et al. (2016) and Goverde et al. (2016) propose a timetabling approach that incorporates robustness in a micro-macro framework and enables the iterative construction of a microscopically conflict-free and stable timetable that is optimised at a macroscopic level. This type of multi-level scheduling is nowadays the standard approach for most timetabling approaches in literature, as it enables optimising the schedule whilst ensuring a feasible timetable.

2.2.2. Statistical methods for buffer time scheduling

With regard to distributing time allowances in the timetable, Palmqvist et al. (2017a) identify five strategies from literature: (i) A uniform allocation of running time supplement percentage (Scheepmaker and Goverde, 2015); (ii) Shifting running time supplement towards the beginning or end of the line (Vromans, 2005); (iii) Place dwell time supplement at or near strategic locations (Vromans, 2005); (iv) Based on where disturbances happen most frequently (Huang et al., 2019; Yuan and Hansen, 2008); and (v) Place time allowances at critical points (Abid et al., 2017; Andersson et al., 2013). Critical points here are defined as locations where two trains follow, cross, or overtake, and are most sensitive to delays.

To appraise the effect of time allowances on delay propagation, several ex-ante metrics have been introduced. Vromans (2005) defined WAD as the Weighted Average Distance of running time supplements from the starting point of the train line. WAD describes how running time supplements are distributed along the journey and attempts to optimise this process, using both analytical and numerical methods. Carey (1999) and Kroon et al. (2008) introduced a metric known as the Percentage of Headways equal to or less than the minimum value (PoH), to investigate how many tight headways exist in the timetable. The strategy of placing margins at critical points was developed by Andersson et al. (2013) (Robustness in Critical Points (RCP)) and enhanced by Abid et al. (2017) (Critical Robustness Measure (CRM)).

Concerning buffer time, always using the same deterministic value as an indicator for the size of the time allowance is the most simple form of distribution. Carey and Kwieciński (1994) suggest to base the scheduled headway on its relation with knock-on delays, delays transferred from a train to its successor. They perform a stochastic simulation to derive an approximate relationship between scheduled headways and knock-on delays. Their approach can be used to adjust train timetables to better avoid and reduce knock-on delays. Vromans et al. (2006) argue that the uniform spread of trains over time contributes to larger headways, less delay propagation, and thus more robust timetables. Therefore, buffer time should be scheduled such that trains are spread evenly over time. To this end, the

Sum of Shortest Headway Reciprocals (SSHR) and the Sum of Arrival Headway Reciprocals (SAHR) are introduced to measure the robustness of headways.

These approaches do not explicitly analyse buffer time, but examine headways as a whole. Nowadays more knowledge on (technical) minimum headway generation has been gathered, enabling more detailed approaches to reduce delay propagation. Thus, methods to identify the suitable amount of buffer time per situation are desired in order to achieve more precise specification of headways as a whole.

Jensen et al. (2017) provide a first step in achieving these situation specific buffer times. They define the minimum buffer time (i.e. critical buffer time) as *"the slack deemed necessary by planners to reduce the risk of delay propagation"*. Computation of the minimum buffer times is done by simulating the delay propagation of a timetable without any buffer time with the use of a sample of initial delays. Then, the estimated delay propagation is an indication of where there should have been added more buffer time in order to resolve the delay propagation.

2.3. Data-driven approaches

Historical realisation data has proven to be beneficial for monitoring operational quality, improving the schedule, diagnosing operational problems, and evaluating countermeasures in transportation science. In this section, literature related to data-driven knowledge on delays and buffer times is reviewed. For a more elaborate overview of big data applications in railway transportation and train dispatching in particular, see Ghofrani et al. (2018) and Wen et al. (2019) respectively.

2.3.1. Delay distribution, propagation and recovery

Delays have been proven to be most relevant when arriving or departing from stations (Yuan and Hansen, 2007). Therefore, the probability distributions of arrival and departure delays have been studied in the literature. Wen et al. (2017b) investigated the main statistics of primary delays, including (i) delay causes, (ii) delay frequencies, (iii) delay occurrences in time and space, (iv) affected number of trains, and (v) delay recovery patterns. Their study shows that the frequency of primary delay and the affected number of trains distribution could be fit to respectively a log-normal distribution and a non-linear regression model. Moreover, they conclude there is a high probability of delay occurrence whenever there is a high capacity utilisation. Minbashi et al. (2021) address a similar topic, focusing on individual shunting departure deviations. They aim to find a probability distribution that fits a large spectrum of departure deviations best, but they find that in particular early departures are difficult to fit.

The fitted probability distributions help better represent delays, but do only consider primary delays. To incorporate secondary delays, more sophisticated prediction models of delays and delay propagation are needed. Such a delay prediction approach is introduced by Corman and Kecman (2018). In their study they present a stochastic prediction of train delays in real-time using Bayesian network. They characterise the effect that the prediction horizon and incoming information about running trains may have on the probability of the future train delays. Thus, when two trains use the same part of the infrastructure within short time, a delay of the first train can be used to predict the delay of the second. Şahin (2017) takes into account the stochasticity of train operations as well by prediction of train delays and recovery with the use of a Markov chain model. Given the initial state of a train, its future states can be predicted with probabilistic terms. This approach furthermore provides a novel approach for assessing the effectiveness of time allowances to maintain punctuality and robustness.

Alongside prediction models for secondary delays, research has focused on the measurement of this aspect of delay propagation. Weeda and Wiggenraad (2006) introduce a conceptual model to approximate secondary delay from realisation data, based on the delay of two conflicting trains and the scheduled buffer. They find that the estimation approach presents values similar to the delay jump, but depending on the situation could show both and overestimation as an underestimation. A more accurate and effective metric is introduced by Goverde and Meng (2011), measuring the time loss associated with a route conflict compared to a reference running time on a block section of unhindered operations. This requires detailed scheduling and realisation data of signal passages.

W. H. Lee et al. (2016) introduced a supervised decision tree method that estimates the key factors contributing to knock-on delays. It is found that unscheduled waiting time for meeting or overtaking causes the majority of knock-on delays. More specifically, Zieger et al. (2018) analysed the effects of different buffer time distributions on the formation of knock-on delays. It is shown that the choice of distribution has a significant impact on performance metrics. For example, to achieve the same level of service, the necessary buffer time can differ up to 1 minute buffer time between trains among various types of buffer time distributions.

Furthermore, Huang et al. (2019) provide a methodology to determine how and to what extent running time supplements and dwell time supplements affect delay recovery. They assess the utilisation rate of these two time allowances based on the fraction of the recovered time (i.e. the delay jump) and the total scheduled time allowances. It is found that the utilisation rate is much higher in sections than in stations, indicating that running time supplements are more effective in recovering from delay that dwell time supplements. Similarly, Yang et al. (2020) focus on the effect of running time supplements and dwell time supplements on delay recovery. They identify three situations: (i) the delay of a train is smaller than the dwell time supplement at a specific station, or (ii) the delay of a train is bigger than the dwell time supplement at a specific station, but smaller than the dwell time supplement plus the running time supplement to the next station, thus does not affect the station arrival time, or (iii) the delay is bigger than the sum of the dwell time supplement and the running time supplements, thus results in a follow-up delay to the same train at the next station.

From these situations can be derived that the effectiveness of time allowances to facilitate delay recovery is highly dependent on the size of the delay. Thus, allocation of time allowances should be performed considering the actual impact of the delay. Combining the findings from these studies enables construction of cause-based predictive models for delay recovery, such as regression models (Guo et al., 2015; Huang et al., 2019; Wen et al., 2017a).

Existing studies on delay propagation concern delay prediction and delay recovery. In general, models are able to fit and predict primary delays, but only based on specific lines. Models that can fit multiple lines or a complex network are currently lacking. Furthermore, studies on delay prediction focus primarily on primary delays, whilst secondary and knock-on delays can provide more information about the effective distribution of buffer times. Delay recovery studies do take into account secondary and knock-on delays, enabling assessment of the effectiveness of buffer times to reduce delays.

2.3.2. Timetable design based on realisation data

Statistical methods

A conceptual framework is introduced by Weeda and Wiggenraad (2006) to acquire a joint design standard for running times and headway times. With the use of historical realisation data, an empirical relationship between headways and knock-on delays is derived. The first thing that can be concluded is that the need for supplements is greatest around stops in order to avoid delay propagation, rather than equally distributed along the way. Secondly, the buffer time proves to reduce knock-on delays, but the benefits decrease with higher values of buffer time, in line with findings by Yuan and Hansen (2007) and Zieger et al. (2018). It is however, in contrast with the optimal approach enabling energy-efficient timetabling, where a uniform allocation is preferred (Scheepmaker and Goverde, 2015). Lastly, they find a relationship that enables buffer time and running time supplement to be interchanged when designing the timetable.

In line with the work of Carey and Kwieciński (1994), a data-driven method of determining minimum headways for conflicting train movements is proposed by Cerreto and Jonasson (2019). By means of a statistical analysis of the relationship between planned headways and the delay jump, the section specific minimum feasible headway between conflicting train movements in a railway system is estimated with a linear regression.

Dietzenbacher (2021) explored the spread and trend of headways within realisation data of train operations. This information can be used to derive conclusions about buffer times. The study recommends a comparatively higher buffer for plan situations with a higher standard deviation of headway spread or with few headways below the minimum headway from the planning stage. These situations were (i) departure of the first train and arrival of the second train; and (ii) arrival of both trains.

Optimisation methods

One of the most applied techniques to study allocation of time allowances are optimisation methods. Ho et al. (2012) were one of the first to integrate timetable optimisation with historical data to cope with disturbances. They present a methodology that establishes a set of operational rules for real-time decision making.

Hauck and Kliewer (2019) enhance existing solution approaches for the train timetabling problem as discussed in Section 2.2 by considering historical delay information. With the use of a Mixed-Integer Programming (MIP) model, the difference between the realised running time and the planned running time is minimised. The optimisation is initiated with a given feasible timetable and adjusted in terms of arrival times, departure times and buffer times to adapt the timetable to avoid systematic delays.

Huang et al. (2019) address allocation of time allowances in high-speed railway train operations. Based on historical data, a ridge regression model of delay recovery is generated regarding primary delay severity, running time supplements and dwell times. Subsequently, an Integer Linear Programming (ILP) model is presented that maximises the delay recovery by reallocating running time supplements and dwell times in an initial timetable.

Yang et al. (2020) base their study on the insight that allocation of time allowances should consider the demand of train delay recovery in order to make full use of the available time allowances and not waste capacity. Therefore, a timetable design model with weighted average delay expectation as the objective function was constructed and solved by a mathematical analysis method. The solution is based on a given feasible timetable, whereof the time allowances have been redistributed.

Time allowances are not only effective to reduce delays and enhance punctuality, but can moreover be inserted to improve energy efficiency. To reduce energy consumption, running time supplements enable speed reduction by applying coasting. And in case of delays, the amount of buffer time determines whether the signal aspects become restrictive, changing the envisioned (energy-optimal) speed profile.

In order to accomplish energy-efficient train operations, Scheepmaker and Goverde (2015) create an energy-efficient train control model that determines the energy-efficient driving strategy by calculating the optimal cruising speed and coasting point given a specific timetable. Additionally, they find that energy savings are higher when the running time supplements are distributed evenly, rather than concentrating it near the main stations. Thus, the extent to which the train can apply the driving strategy depend on the allocation of the running time supplements in the timetable. Therefore, energy consumption should already be considered in the timetabling stage.

Goverde et al. (2016) address energy consumption as objective in the timetable construction, but only consider it after a feasible, stable and robust timetable structure was determined. Scheepmaker and Goverde (2021) optimises the objective of minimising energy consumption simultaneously with minimising the total running time, capacity utilisation and maximising the robustness of the timetable by reallocating running time supplements. The optimisation is performed with a minimum buffer time of 0 seconds, 30 seconds and 60 seconds. The variation in buffer time is found to impact the total buffer time scheduled, the extended cycle time, frequency and capacity consumption, but does not show differences in total running time or energy consumption. It turns out that a negative relationship exists between the distance between stops and the optimal relative running time supplement.

A delay recovery strategy to reduce the delay and increase the energy-efficiency simultaneously was proposed by Li et al. (2020). They classify delays in historical data of metro systems according to the relationship between the delay time and the following headway, and present optimised delay recovery strategies for various initial delay distributions.

The main limitation of these data-driven optimisation methods is that the majority requires an initial timetable, which is adjusted to reallocate the current time allowances. Thereby, the studies exclude evaluating the total size of the time allowances and are incapable of changing the order of the trains. The latter is only enabled in the study of Khoshniyat and Peterson (2017), and they are able to generate larger headways, while having only small reductions in capacity, heterogeneity and speed. Furthermore, Hauck and Kliewer (2019), Huang et al. (2019), and Yang et al. (2020) mainly focus on prevention of primary delays by adjusting running time supplements and dwell times. Hence, the studies do not optimise the schedules to prevent delay propagation with the use of buffer times.

2.4. Scientific gaps

To generate situation specific time allowances, in particular buffer times, (i) optimisation approaches, applying mathematical models; (ii) simulation approaches, iteratively enhancing a current solution; (iii) statistical approaches, applying statistics to infer relationships between variables; and (iv) data-driven approaches, indicating recurring bottlenecks and tightly planned headways with the use of historical data, have been introduced. The works have been discussed in the sections above and are summarised in Table 2.1.

Reference	Goal	Time allowance	Approach	Initial timetable required
Burggraeve and Vansteenwegen (2017)	Passenger robustness	BT, RTS, DTS	Optimisation, Simulation	X
D'Acierno et al. (2018)	Energy consumption	BT, DTS	Statistics	
Hauck and Kliewer (2019)	Punctuality	BT, RTS, DTS	Optimisation, data-driven	X
Ho et al. (2012)	Total operation time	RTS, DTS	Optimisation, data-driven	X
Huang et al. (2019)	Delay recovery	RTS, DTS	Optimisation, data-driven	X
Jensen et al. (2017)	Capacity consumption	BT	Simulation, statistics	
Jin et al. (2019)	Average delay time	RTS	Optimisation	X
Jovanović et al. (2017)	Capacity and delay propagation	BT	Optimisation	X
Khoshniyat and Peterson (2017)	Timetable reliability	BT	Optimisation	X

Kroon et al. (2008)	Average weighted delay	BT, RTS	Optimisation, simulation	X
Y. Lee et al. (2017)	Average delay	BT. RTS	Simulation	X
Li et al. (2020)	Cumulative delay time and energy consumption	RTS, DTS	Optimisation, data-driven	
Lu et al. (2017)	Efficiency and robustness	BT	Optimisation	
Meng et al. (2019)	Sum of train delays	BT, RTS	Optimisation	X
Restel et al. (2021)	Delay propagation and energy consumption	RTS	Simulation, data-driven	X
Scheepmaker and Goverde (2021)	Running time, infrastructure occupation, timetable robustness and energy efficiency	BT, RTS	Optimisation	
Weeda and Wiggenraad (2006)	Punctuality	BT, RTS, DTS	Statistics, data-driven	
Yang et al. (2020)	Delay expectation at stations	RTS, DTS	Optimisation, data-driven	X
Yuan and Hansen (2008)	Weighted sum of knock-on delays	BT	Optimisation	
This research	Delay propagation	BT	Statistics, data-driven	

Legend: BT = Buffer Time, RTS = Running Time Supplement, DTS = Dwell Time Supplement

Table 2.1: Situation specific time allowances in previous studies

So far, in literature, quite some attention has been given to planning with time allowances in railway timetable design. Some analytical approaches aim to enhance the robustness of the timetable by deriving relationships regarding planned headways (Carey and Kwieciński, 1994; Vromans et al., 2006). As these approaches do not explicitly consider buffer time, their applicability to reduce delay propagation is limited. Cerreto and Jonasson (2019) extend the knowledge by introducing a method for minimum headway determination from historical data. Although, their aim is to estimate the minimum feasible headway between conflicting train movements, their approach can also be used to derive actual available buffer time in the realised operations by subtracting the minimum feasible headway from the headway in the realisation.

When creating an initial timetable from scratch, minimum buffer times, or minimum headways as a further simplification, are predominantly considered as a given input (Cacchiani and Toth, 2012; Caprara et al., 2002; Scheepmaker and Goverde, 2021; Serafini and Ukovich, 1989). Only in later timetable design stages, buffer times have been adjusted in order to optimise a wide range of objectives (e.g. energy consumption, delay time, reliability, capacity consumption, delay propagation). The majority of studies that explicitly consider allocation of buffer times require an initial timetable, which is adjusted to reallocate the current buffer time and running time supplement (Burggraeve and Vansteenwegen, 2017; Hauck and Kliewer, 2019; Ho et al., 2012; Huang et al., 2019; Jin et al., 2019; Jovanović et al., 2017; Khoshniyat and Peterson, 2017; Kroon et al., 2008; Y. Lee et al., 2017; Meng et al., 2019; Restel et al., 2021; Yang et al., 2020). Thereby, the studies exclude evaluating the total size of the time allowances and are incapable of changing the order of the trains.

Overall, some studies have aimed to improve the effectiveness of running time supplements and dwell time supplements to reduce delays based on historical data (Huang et al., 2019; Yang et al., 2020). Yet, the proposed methods mainly focus on primary delays and are unable to generalise the findings, such that it can be applied to create initial timetables where no realisation data is available. The latter is crucial to cope with delay propagation effects already while constructing an initial timetable.

Considered all together, this leads to the identification of the following scientific gaps:

- It is unknown to what extent specific timetable characteristics (e.g. scheduled buffer time, scheduled running time supplement, train characteristics, headway situation, location) influence delay propagation.
- The effective allocation of buffer times to improve delay recovery based on historical realisation data has not been widely addressed in the tactical planning stage of railway timetables where an initial timetable is constructed.
- Planning rules for buffer time based on historical realisation data have not been widely explored in literature.

Concluding, this research differs from existing works regarding generating situation specific buffer times by providing an approach that (i) does not require an initial timetable; (ii) predicts delay propagation based on factors known in the timetable design; and (iii) specifies planning rules that can be applied to new situations.

2.5. Chapter conclusion

This chapter started with a general overview of railway timetable design concepts and methods. Indicators assessing the robustness of a timetable were introduced and approaches aiming to enhance the effectiveness of time allowances were reviewed. Three scientific gaps were identified from the reviewed literature. In Chapter 1 the scientific gaps are consolidated with the practical gaps, as identified in Chapter 3. The remainder of this report addresses the consolidated knowledge gaps. Using the information gathered in this chapter, the first sub-question is answered.

Sub-question 1: What can be learned from previous quantitative, data-driven approaches for scheduling time allowances in railway timetable design?

Research on scheduling time allowances is strongly related to robustness, aiming to achieve a high effectiveness and utilisation rate of time allowances such that the timetable can withstand design errors, parameter variations, and changing operational conditions. Robustness approaches that consider buffer times aim for example to (i) reduce the total delay; (ii) reduce delay propagation; (iii) diminish the needed real-time dispatching action; and (iv) enhance punctuality.

Metrics that facilitate comparisons between different timetables in terms of robustness have been introduced. In most cases the metric focuses on one type of time allowance, making it straightforward to propose timetable improvements. It appears that metrics addressing buffer time predominantly focus on ex-ante situations and headways as a whole, thus not on buffer times specifically. Two studies were found where the evaluation is done ex-post in a data-driven fashion, based on how effective running time supplements and dwell time supplements have been in reducing delays. It appears hard to define ex-post metrics of buffer time effectiveness, since the focus of buffer time is on avoiding and decreasing secondary delays rather than reducing existing delays. The amount of secondary delay that has been avoided by insertion of buffer time is challenging to determine. However, measuring propagated delay can provide an indication of the effectiveness of buffer time.

The reviewed articles suggest that the effectiveness of time allowances to facilitate delay recovery is highly dependent on the size of the delay. Regarding the data-driven approaches in the reviewed articles, the focus rarely lies on delay propagation and buffer times.

3
Practical background

This chapter provides background information on the timetable design process in the Netherlands. To start, Section 3.1 states the current norms for running times, headways and dwell times in the Netherlands. Next, Section 3.2 describes the current timetable design process and the tools utilised. The way the planning norms and design process interact is explained in Section 3.3. Subsequently, Section 3.4 highlights the existing research gaps. The chapter ends with a conclusion and an answer to the second sub-question in Section 3.5.

3.1. Current planning norms in the Netherlands

In the Netherlands, the railway timetable planning process is directed by ProRail, the Dutch rail infrastructure manager. Planning norms are created and updated annually, and applied to satisfy a safe and executable plan. Ideally, location specific planning norms are followed, with a technical minimum time and a time allowance for a specific process (ProRail, 2020). When technical minimum times of a process are unknown, standardised norms that combine the technical minimum time and time allowance can be used as well.

ProRail (2020) specifies the minimum scheduled running time as the technical minimum running time as calculated by Donna plus 7% running time supplement for passenger trains with a crew (i.e. no automatically closing doors). The planning norm for minimum scheduled headways is similarly composed of a technical minimum headway time and a minimum buffer time. The technical minimum headway time is based on the situation and location with the use of the blocking time model as explained in Section 2.1. The necessary minimum buffer time is tested with a deterministic, microscopic simulation during the *Medium Term (MLT)* phase, which is explained in the next section. The buffer time needed for robust execution of the timetable is simulated by adding 60 seconds wake behind every train movement (i.e. the section is released 60 seconds later than would be possible according to the blocking time model). Hence, the wake represents the buffer time between the trains. Then, if this results in hindered operations it could still be acceptable, except when (i) an unplanned stop occurs, or (ii) an arrival delay bigger than 30 seconds occurs at certain stations. Planning with minor hindrance requires simulation and consideration of buffer time and running time supplement (Nederlandse Spoorwegen, 2019; ProRail, 2020). Dwell time planning norms are currently specified based on the rolling stock used and differ as well between trains which are operated with automatically closing doors and trains with a crew. Location specific dwell times are applied for hub stations, and take into account connections from different platforms.

3.2. Timetable design process in the Netherlands

Timetable design is a complex procedure where multiple parties and dependencies are present. Moreover, it is a long-term process as it already starts up to 20 years in advance with infrastructure requirements and capacity studies (Planting, 2016). The design process is a collaboration between the infrastructure manager and various carriers for passengers and freight.

In the Netherlands, the timetable design process is composed of multiple steps. From long term planning to real-time operations the phases are *Long Term (LT), Medium Term (MLT), Preliminary Design (VO), Basic Hour (BU, in literature also referred to as BUP), Basic Day (BD), Basic Day update (BDu), Specific Day (SD)*, and *Traffic Control (VL, in literature also referred to as VKL)*. The most recent timetable in each of these planning phases is referred to as the name of the abbreviation of the phase followed by the word timetable. At the end of the operations, traffic realisation data is collected to compute and interpret the performance. In this research, this is indicated as the *Performance Analysis (PA)* phase. There exists an official feedback loop from realisation data to the various planning stages, where mistakes and inconsistencies of the timetable are reported. In recurrent meetings that incorporate lessons from the realisation data in the BU, BD, BDu and SD phases, a vast majority of cases is evaluated manually based on expert judgement, increasing the risk of subjectivity. An overview of the steps categorised in the strategic, tactical and operational phase is shown in Figure 3.1. More information on the various phases of timetable design in the Netherlands

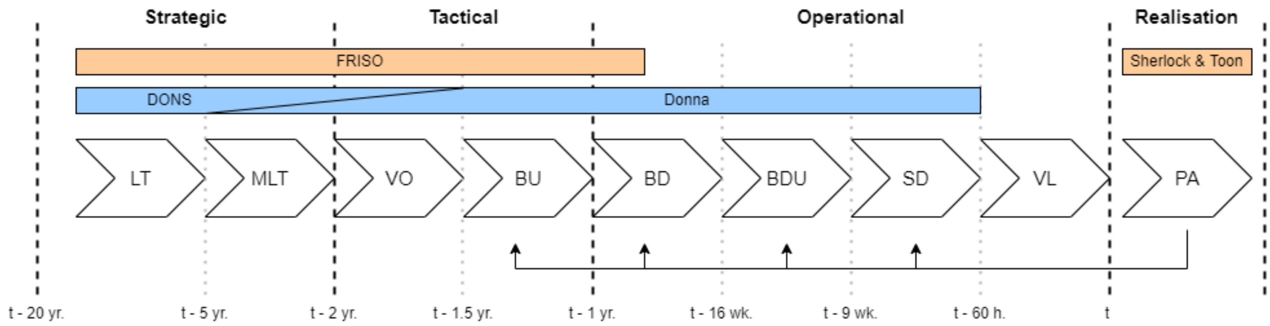


Figure 3.1: Overview of timetable design phases and tools (analysis depicted in orange; planning depicted in blue), adjusted from Planting (2016)

can be found in Planting (2016).

During the timetable design, many different tools are used. For this research, the most important instruments are software regarding planning and analysis. Concerning the former, Nederlandse Spoorwegen incorporates design tools based on planning norms and mathematical programming, such as DONS for long term, micro- and macroscopic planning and Donna for short term, microscopic planning. The latter comprises the microscopic simulation software FRISO and the registration and visualisation of realisation data in Sherlock and Toon respectively. An overview of the phase in which the different tools are used in shown in Figure 3.1. More information on the various tools utilised in the timetable design process can be found in Planting (2016).

3.3. Design practices in the Netherlands

Every year, Nederlandse Spoorwegen creates a cyclic timetable, where frequencies higher than once per hour are divided into regular intervals (e.g. a train line than runs 4 times per hour is scheduled exactly 15 minutes apart). The planning norms offer some guidance with respect to the necessary size of the time allowances. Since 2020, the Dutch timetable is composed with a precision of 6 seconds.

With regard to running time supplement for passenger trains, it is common practice to evenly spread it over the train path between critical points (Nederlandse Spoorwegen, 2019). Critical points are determined by the design and consist of important (hub) stations, overtaking locations and crossing points. The advantages of this practice are that (i) symmetry is guaranteed; (ii) the impact of the running time supplement on the buffer time is limited; (iii) the method of the distribution is easily traceable and reproducible for planners and train crew; and (iv) the distribution meets the desire from energy-efficient driving to have an equal amount of driving time available for Sprinters at as many intermediate stops as possible (Nederlandse Spoorwegen, 2018).

As mentioned in Section 3.1, the necessary minimum buffer time is tested in microscopic simulations with the use of 60 seconds buffer, where hindrance might be acceptable in some cases. In the Netherlands, the tracks in the Randstad are so busy that the rule of completely unhindered headways plus an additional 60 seconds buffer are met only sporadically. In Donna, therefore, the planners work with microscopic headway times plus 30 seconds buffer. Additionally, by using slightly hindered headways, 'hidden' capacity becomes available to use. Here slightly hindered headways refer to headways where there is a small overlap in blocking times, so that a conflict is detected. Slightly hindered headways are as a rule of thumb accepted when the gain in headway time is bigger than the loss in running time. A disadvantage of this method is that each slightly hindered headway has to be evaluated on its own, a very labour intensive process. But they accommodate flexibility in the planning process, being one of the main reasons why microscopic optimisation models in the planning are used less often; they only consider fully unhindered headways and thus are not flexible. Thus, while microscopic optimisation models with a 60 seconds buffer are too strict (i.e. some constructions that are possible in practice are rejected or not identified in the microscopic optimisation model), models without a 60 seconds buffer are too flexible (i.e. not enough time for delay recovery is in place). This flexibility is added by using 30 seconds buffer in Donna.

As mentioned in Section 3.2, a feedback loop from the realisation stage is increasingly employed to enhance the timetable design. For example, when, based on practical measurements, the running time appears different than calculated by Donna, ProRail may rectify the technical minimum runtime after consultation of the affected carrier (ProRail, 2020). Additionally, if certain headway situations appear not to work in practice, they are reactively adapted in the Basic Day update phase and in the timetable of the following year. For dwell times, Nederlandse Spoorwegen

(2019) allows deviation from the planning norms for predetermined stations or train series by taking the 50th percentile observed in the realisation. This approach is not used for running times, since for heavy rail spread in the realisation of running times is caused primarily by the planning (e.g. increasing running time or dwell time to avoid conflicts) (Nederlandse Spoorwegen, 2018). It is undesired to incorporate the structural hindrance from the previous timetable into the next timetable. Currently, the focus of the feedback is mainly on individual situations for a specific year, but the aim is to also incorporate insights from realisation data in planning rules, running time calculations and conflict detection.

3.4. Practical gaps

Recently, planning norms have been made more location- and situation specific with the integration of the blocking time model in DONS and Donna. This allows the breakdown of a planned time into technical minimum time and a time allowance for a specific process. The determination of the size of the running time supplement is dependent on the runtime, whilst for buffer time always a value of 60 seconds is taken as input. The buffer time value can vary slightly depending on the extent to which hindrance occurs in a deterministic, microscopic simulation. Slightly hindered headways are as a rule of thumb accepted when the gain in headway time is bigger than the loss in running time. Thus, the focus here is to reactively satisfy feasibility, rather than proactively reduce delay propagation.

Additionally, feedback from the realisation to the planning phase is done structurally with the use of meetings where the interesting observations are shared. However, currently only specific cases are discussed. There is no feedback loop that considers the large and heterogeneous set of realisation data to improve the timetable design.

What stands out is that, although optimisation models seem very attractive in literature, they are not used that often in practice. This is due to their inflexibility regarding planning. For example, a slightly hindered headway situation that has been evaluated and accepted with the use of microscopic simulation, may be rejected or not even considered in an optimisation model, because it violates the inputted minimum buffer time. Thus, there is a practical need for understanding how flexible buffers can be created, so that they can be used as input for optimisation models.

Considered all together, this leads to the identification of the following practical gaps:

- The differentiation of buffer times has not been addressed in practice with regard to the improvement of delay recovery.
- Planning rules for buffer time based on historical realisation data to improve delay recovery have not been identified and applied in practice.

3.5. Chapter conclusion

This chapter started with a general overview of the current Dutch norms for running times, headways and dwell times. Next, the process of timetable design and working practices were discussed. Two practical gaps were identified from the reviewed grey literature. In Chapter 1 the practical gaps are consolidated with the scientific gaps, as identified in Chapter 2. The remainder of this report addresses the consolidated knowledge gaps. Using the information gathered in this chapter, the second sub-question is answered.

Sub-question 2: What is the current state of the practice regarding planning norms for buffer times and regarding data-driven timetable design approaches in the Netherlands?

Planning rules for buffer times are imposed by ProRail to ensure safe railway operations. Currently, 60 seconds buffer time between two trains is set as a norm, but in some cases less buffer time is accepted. In Donna, even, a buffer time of merely 30 seconds is displayed to facilitate flexibility. While slightly hindered headways are beneficial for capacity reasons, its consequence on delay propagation is less often considered explicitly. Headway situations are mainly investigated individually and manually, and are both actively and reactively adapted based on information from realisation data. Although this improves the timetable, it doesn't show structural needed changes to headway situations. Thus, a more structural way of analysing the impact of buffer times on delay propagation is lacking. Overall, realisation data is consulted to identify places of improvement in the timetable, but much could be gained by formalising and intensifying the feedback process.

4
Methodology

This chapter describes the methodology used to derive buffer time planning rules from. First, Section 4.1 introduces the framework used to determine buffer time planning rules from realisation data. In particular, it addresses what steps are being taken and how the different steps are related to each other. Subsequently, Section 4.2, Section 4.3, Section 4.4, Section 4.5 and Section 4.6 address the contents of the conceptual framework by respectively presenting (i) the target and predictor variables used in the analysis; (ii) the data collection; (iii) the data pre-processing; (iv) the statistical analysis, consisting of a regression analysis; and (v) the output. The chapter finishes with a conclusion and an answer to the third sub-question in Section 4.7.

4.1. Conceptual framework for buffer time rules determination

Figure 4.1 summarises the methodology for the determination of buffer time planning rules from realisation data. The framework is explained below and has been divided into four parts: input, processing, statistical analysis and output, each of which is discussed in detail in Section 4.3 through Section 4.6.

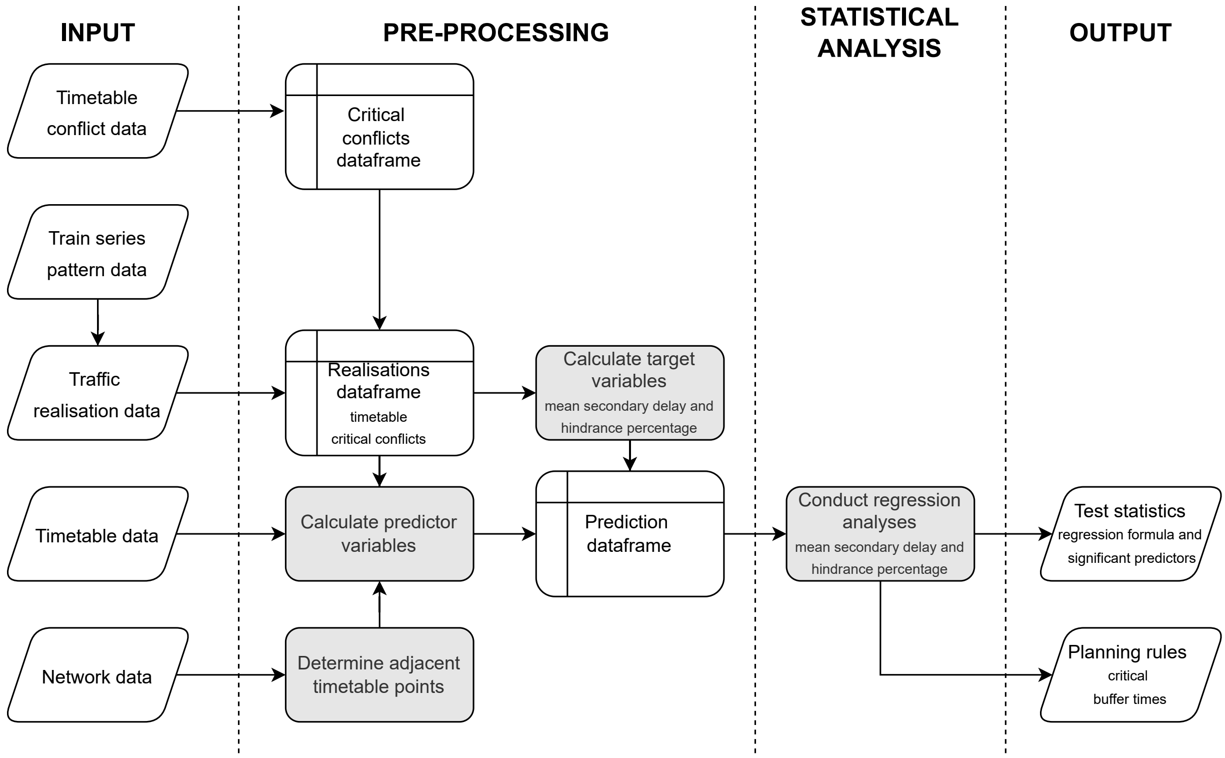


Figure 4.1: Framework for buffer time rules determination

Various types of input data are collected for this analysis, consisting of *Timetable conflict data, Train series pattern data, Traffic realisation data, Timetable data*, and *Network data*. Details of the data collection are presented in Appendix C. The *Timetable conflict data* is used to identify the route conflicts and buffer time conflicts of the timetable (i.e. situations where the start of the blocking time of a train on a section is planned within the (extended) blocking time of a preceding train). From the *timetable conflict data*, critical conflicts are identified, indicating a timetable conflict between two specific train series patterns located in a critical point. The choice to only consider critical conflicts is based on the aim of this research to determine precise critical buffer times for tight headway situations. Then, of these critical conflicts the realisations are found by checking the *traffic realisation data*. With the realisations of the critical conflicts, the target and predictor variables can be computed. Subsequently, hindrance distribution and a prediction dataframe with one row for each conflict are created. Using the prediction dataframe, two regression analyses were performed aiming to predict the two target variables introduced in Section 4.2. From these two prediction models, test statistics and buffer time planning rules are extracted. Regression was used as a prediction model due to its transparency and easy interpretation, making it straightforward to determine planning rules from.

4.2. Selection of target variables and predictor variables

Prior to the data analysis, it is decided what variables are investigated to determine buffer time planning rules from. In Chapter 2 several metrics for the effectiveness of time allowances have been discussed. It is found that metrics addressing buffer time predominantly focus on ex-ante situations and headways as a whole, thus not on buffer times specifically. It appears hard to define ex-post metrics of buffer time effectiveness, since the purpose of buffer time is to avoid and decrease secondary delays rather than reducing existing delays. However, measuring the propagated delay can provide an indication of the effectiveness of buffer time. To this end, this study investigates the aspects of delay propagation to gain insights into the effects of buffer time.

4.2.1. Target variables

The target variables in this study are related to the aspects of delay propagation. The literature review has shown few methods to assess the extent to which delay propagates and the size of secondary delays. Another interesting feature of delay propagation is the probability that delay propagates given a certain situation.

Weeda and Wiggenraad (2006) study the mechanism of secondary delays and buffer time. As Gibson et al. (2002) state secondary delay itself is not straightforward to measure. Therefore, Weeda and Wiggenraad (2006) approximate the secondary delay based on the delays of two succeeding train series patterns and the scheduled buffer on a conflict location. To do so, a conceptual model is presented in Figure 4.2 that shows for approximately 700 fictional realisations the delays of both trains of a fictional headway situation with a buffer of 60 seconds. Each dot represents a fictional realisation of a headway situation between two trains, before the interaction at the conflict point. As long as the delay of train 1 is smaller than the buffer time, train 2 can proceed its normal operation. The diagonal line represents the instant when train 1 just leaves the conflict section when train 2 enters the conflict section. Observations to the right of the line indicate hindrance of train 2 caused by train 1, because train 1 has not left the conflict section yet. Various categories can be identified: (Ia) train 1 is running on time or has a delay smaller than the buffer time and train 2 is on time or late and is not perturbed; (Ib) train 2 is running early and is held up, but it does not matter since it only makes train 2 run on time again; (Ic) train 1 is late, but train 2 is even later and does not suffer an additional delay. Although train 2 should not run late, there is no hindrance; (II) train 2 reaches the junction before it is released by train 1. Hindrance occurs: train 1 delays train 2, or in case of an order change train 2 delay train 1; and (III) train 1 is running that late that the train dispatcher gave priority to train 2.

The conceptual model enables an estimation of the secondary delay (i.e. the increase in delay the hindered train has as opposed to when it would not be hindered) train 2 suffers because of the delay of train 1. The secondary delay is defined as the horizontal distance between an observation in category II and the diagonal left border line of category II (which represents the divide between hindered observations and undisturbed observations) and can be calculated with Equation 4.1, where r_2 is the secondary delay to train 2, d_i is the initial delay to train i, and b_{12} is the scheduled buffer time between both trains. For observations outside category II (i.e. unhindered observations), the secondary delay equals zero.

$$r_2 = d_1 - max(0, d_2) - b_{12} \tag{4.1}$$

This research utilises the insights from this framework in defining two target variables. Firstly, the observed percentage of secondary delayed trains is utilised as an estimation of the probability of hindrance in a given conflict situation. Thus, this study defines hindrance as an operating situation in which a train will encounter a signal showing a restrictive aspect and will therefore have to modify its speed, due to a conflict with a preceding train, in line with the definition of disturbed train operation by Pachl (2014). Secondly, the expected amount of hindrance that occurs in a specific situation is denoted by the mean secondary delay observed in that situation. The hindrance percentage and the mean secondary delay serve as target variables and are both estimated with its own regression model.

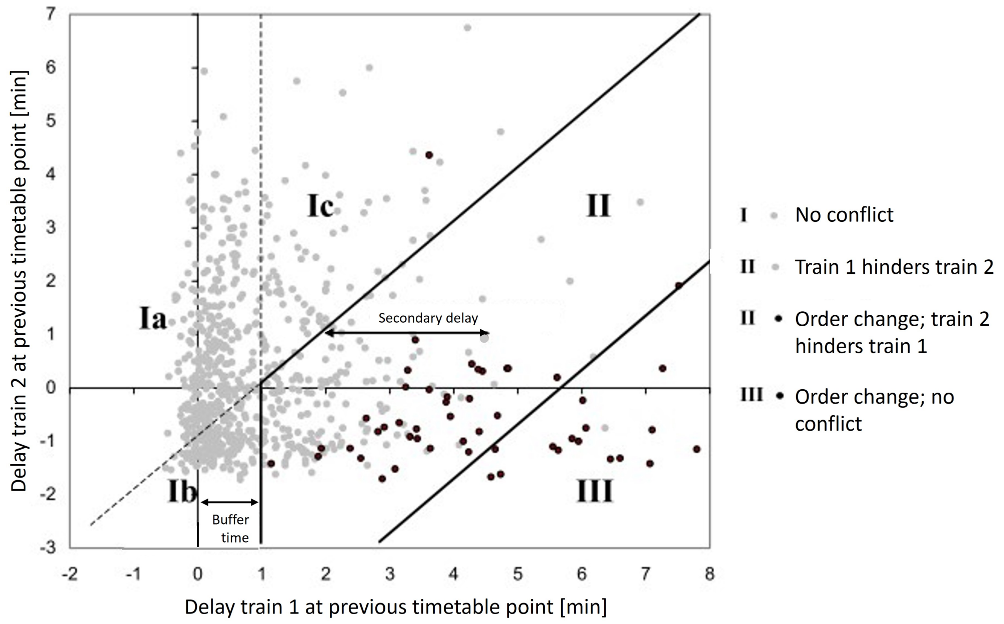


Figure 4.2: Determining secondary delay from initial delays of train 1 and train 2. Own creature based on Weeda (2005)

Limitations and assumptions

The conceptual model does not represent all aspects of reality. The most important model reductions and assumptions of the conceptual model are stated below (Weeda, 2005):

- This simplified calculation of secondary assumes that train 2 stops at the yellow sign and, as soon as the sign turns green continues with the original speed, depicted in Figure 4.3. On average, this results in a approximation of secondary delay similar to the true secondary delay. But investigating per conflict, this could lead to both an overestimation as an underestimation of the secondary delay, because in reality two opposing effects can occur: (i) the second train has a lower speed or has even stopped when the block section is released, but (ii) the second train is closer to the entrance of the block section. When the first effect dominates, the true secondary delay will be bigger than the approximation, while domination of the second effect will reduce the headway time. Which effect dominates and what impact it has on the secondary delay depends on numerous factors:
  - The configuration of block lengths and speed. With high speeds or heavy trains, the first effect can be dominant, because the start time loss is huge. On stations, where the maximum speed is lower, being closer to the entrance of the block section could reduce the headway time.
  - Driver behaviour at a yellow sign varies between braking as late as possible, hoping that the following sign improves, or braking early such that the train still has some speed when passing the next signal is permitted.
  - The visibility of the signal that follows the restrictive signal.
- The delay distributions of both trains should be independent. Therefore the delay of both trains is measured at a location where they cannot have influenced each other. For crossings of trains that originate from different tracks, this could be the previous timetable point, but in the case of a succession the trains frequently drive on the same infrastructure for a long time. In this research, for successions the delay of both trains is also measured at the previous timetable point, because a measure point further away may not catch the most recent delay. Fr starting trains without a previous timetable point the most recent known delay is used. Although this means the delay distributions of both trains are not fully independent, the occurrence of secondary delay at the previous timetable point is limited, since the conflict is smaller compared to the critical conflict.

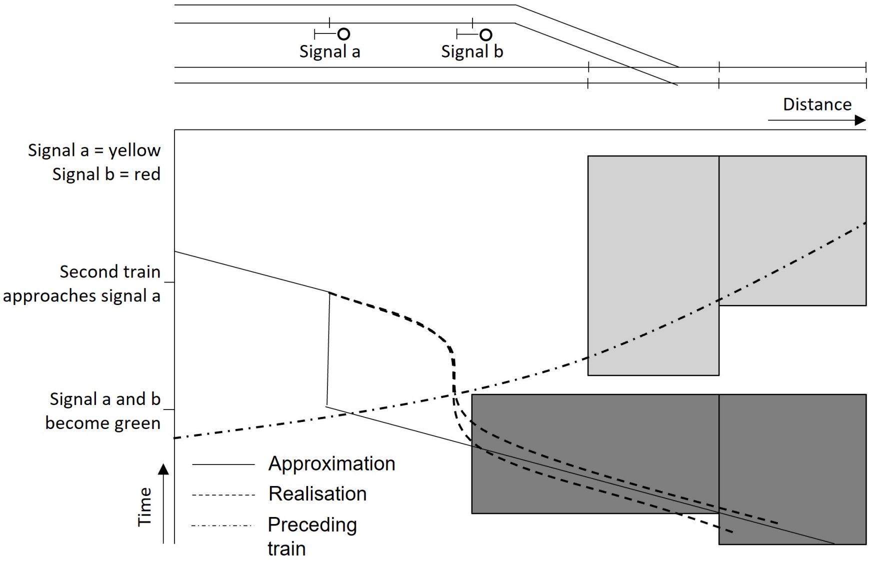


Figure 4.3: Visualisation of uncertainty on true hindrance. Adopted from Weeda (2005)

4.2.2. Predictor variables

The target variables are impacted by a variety of variables that function as predictor variables in the statistical analysis. An overview of the added predictor variables and its computation are provided in Table 4.1. Among the predictor variables only factors known in the construction of a timetable are considered, because the buffer time planning rules should be applicable in an early stage of timetable development. Moreover, the reason to include certain predictor variables is substantiated by a source, or in case of own hypotheses explained below. Not all source-based predictor variables indicate a relationship to the target variables used in this research, but might be related to delay, delay propagation or punctuality. It is expected that the same factors may influence the secondary delay, and thus the variables are considered in this research as well.

The numeric predictor variables are accommodated with a hypothesised sign of the correlation with the target variable. For the categorical variables a hypothesised sign does not make sense, since the sign of the relation is expected to differ per category and depends on the reference category used in the coding scheme, discussed in Section 4.5.

	Variable (source)	**Description**	**Expected sign**
For the headway situation	Direction (W. H. Lee et al., 2016)	Whether the two trains run in the same direction	
	Number Preceding Conflicts (Gorman, 2009)	Number of prior conflicts where the first train was involved in on the timetable point of the timetable conflict	+
	Scheduled Buffer (Palmqvist et al., 2017b; Vromans et al., 2006)	The buffer time that was scheduled in the timetable	-
	Scheduled Headway (Gorman, 2009)	The headway that was scheduled in the timetable	-
	Technical Minimum Headway	The minimum headway needed to ensure a conflict-free headway situation	+
	timetable point (W. H. Lee et al., 2016)	The timetable point the conflict is closest to	
	Type Headway Situation (W. H. Lee et al., 2016)	Whether the conflict relates to a succession or a crossing of the two trains	
	Type timetable point (W. H. Lee et al., 2016)	The type that the timetable point belongs to. Options include junction, bridge, station, crossover, shunting yard, and halt	
For both trains	Event (Dietzenbacher, 2021)	The event on the track	
	Carrier	The carrier of the train	
	Driving Characteristic (W. H. Lee et al., 2016)	The type of train	
	Previous timetable point; Next timetable point	The previous / next timetable point the train is scheduled on	
	Rolling Stock Type	The type of rolling stock the train has driven with	

Rolling Stock Amount Carriages	The number of carriages the train has driven with	+/-
Scheduled Running Time Supplement (Goverde and Meng, 2011; Palmqvist et al., 2017b)	The percentage of running time supplement that was scheduled in the timetable, determined from the previous stop to the timetable point of the headway situation	-
Train Series (W. H. Lee et al., 2016)	The train series the train belongs to	
Train Series Pattern (W. H. Lee et al., 2016)	The train series pattern the train belongs to	
Type Previous timetable point; Type Next timetable point	The type that the timetable point belongs to. Options include junction, bridge, station, crossover, shunting yard, and halt	

Table 4.1: Overview of candidate predictor variables for predicting mean secondary delay and hindrance percentage

The variables *Direction* and *Type headway situation* are expected to influence the secondary delay and hindrance percentage, because in case of the same direction and successions the trains may follow each other for a longer time. This means there is a longer period in which they are depended on one and another. Therefore, a higher probability of hindrance could be explainable.

The variable *number preceding conflicts* indicates how many prior conflicts on a timetable point have occurred that may directly influence the punctuality of the first train. It is hypothesised that more preceding conflicts influences the secondary delay and hindrance percentage negatively, due to there being less slack to compensate for delay of preceding trains.

Currently buffer time norms are specified in absolute values, whilst running time supplements are more often defined as a percentage. The technical minimum headway is included in the analysis to investigate whether a buffer time percentage would be more appropriate. When the technical minimum headway appears to impact the target variables, this could indicate a specification in terms of percentages could be more suitable. A positive relationship with the target variables is expected, since the buffer of 60 seconds gets smaller with increasing technical minimum headway, and thus might have less ability to avoid secondary delays.

A difference in secondary delay and hindrance percentage might exist between different carriers. For example, international trains are more often prioritised than national trains. This relates not only to the driving characteristic of the train, but also to the power and prestige the various carriers have.

The itinerary of the train is expected to impact the target variables as well as the type of timetable points of the itinerary. This is operationalised by the previous and next timetable point.

Lastly, the type of rolling stock and the amount of carriages driven with can impact the secondary delay and the hindrance percentage. Due to the different characteristics of various rolling stock, differences in punctuality and sensitivity to delay are expected.

4.3. Input

This section introduces the datasets that were used as an input for the framework. Consecutively, traffic realisation data, train series pattern data, timetable conflict data, timetable data and network data are discussed. The way in which the data is obtained is described in Appendix C.

4.3.1. Traffic realisation data

The first type of input data is railway *traffic realisation data*, containing information of trains passing timetable points in the Dutch railway network. The dataset has a long format, meaning that each row contains one observation (i.e. one train event) and each column corresponds with one variable. The predefined variables in this dataset are presented in Table C.1.

Variable	**Column name**	**Description**
Date	SL_VERKEERSDATUM	The date of the event
Train number	SL_TREINNR	The train number of the train
Train series	SL_TREINSERIE	The train series of the train number. Usually equal to the train number in hundreds
Driving characteristic	SL_RIJKARAKTERISTIEK	The type of train
Carrier	SL_VERVOERDER	The carrier of the train
Timetable Point	SL_DRGLPT	The timetable point of the event
Event	SL_ACT_SRT	The event on the track
Scheduled time	SL_BASIC_PLANTIJD	The most recent plan time of the train
Realised time	SL_BASIC_UITVTIJD	The most accurate available realised time of the train
Rolling stock	SL_MATSOORT_TYPE	The type of rolling stock the train has driven with
Carriages	SL_MATSOORT_BAKKEN	The number of carriages the train has driven with

Table 4.2: Overview of traffic realisation dataset

The exact realisation time of a train event is generally unknown, and is therefore approximated using the most accurate source available. Missing data may occur due to physical infrastructure, cancelled trains, renumbering of trains or software malfunctions. In the data pre-processing the the missing data will be addressed.

4.3.2. Train series pattern data

The second type of input data is *train series pattern data*, specifying to what series pattern each train number corresponds. Domestic passenger trains in the Netherlands usually run as part of a train series, a collection of services following the same route and stopping pattern. Such a train series has a format of xxx00, where xxx denotes the different series. To indicate the direction of the train, the train realisation data uses E for Even and O for Odd train numbers, whereas in the *timetable conflict data* the trains with different directions are distinct by an additional letter in front of the train series. The additional letter in the timetable is used to create multiple alternatives for a train series (e.g. a slightly different pattern may be used for peak hours and off-peak hours).

For example, the 2400 series represents the intercity trains between Dordrecht (Drd) and Lelystad Centrum (Lls). In the *traffic realisation dataset* all trains headed towards Dordrecht are identified as train series 2400 with direction O, while the *timetable conflict data* categorises these trains as part of either the B2400 or D2400 train series pattern, depending on the time of the hour they depart from Lelystad Centrum. Thus, in order to connect the *traffic realisation data* with the *timetable conflict data*, it must be known what train series pattern corresponds with what direction. Therefore, with the use of this dataset, the train series in the *traffic realisation data* are updated with a letter in front, based on the train number. This process is shown in Table 4.3, where on the left a part of the two initial datasets is presented and on the right the updated version of the *traffic realisation data* is depicted.

Realisation data				**Train series pattern data**			**Updated realisation data**			
Series	Train Number	Direction	Weekday	Series pattern	Train Number	Weekday	Series	Train Number	Direction	Weekday
2400	2430	E (Lls)	1 (Monday)	A2400	2430	1 (Monday)	A2400	2430	E (Lls)	1 (Monday)
2400	2477	O (Drd)	1 (Monday)	B2400	2477	1 (Monday)	B2400	2477	O (Drd)	1 (Monday)
...	...	...	...	...	...	...	...	...	...	...

Table 4.3: Conversion of train series pattern column

4.3.3. Timetable conflict data

The third type of input data is *timetable conflict data*, presenting both the route conflicts and buffer time conflicts of the timetable (i.e. situations where the start of the blocking time of a train on a section is planned within the (extended) blocking time of a preceding train), based on a minimum buffer of 60 seconds and 120 seconds. The choice to only consider headway situations with timetable conflicts is substantiated by the higher probability of secondary delay emerging in these situations.

The data was exported from DONS and shows all the route conflicts and buffer time conflicts of a BU timetable. Each row in the dataset represents an overlap in (extended) blocking time between two train series on a timetable point and on one or multiple sections. The dataset is used to identify the critical conflicts, their scheduled buffer times, and their realised events by coupling this dataset with the *traffic realisation dataset*. The used dataset is presented in Table 4.4 and consists of predefined variables extracted from microscopic conflict detection based on blocking time in DONS (first five in the table) (Nederlandse Spoorwegen, 2019) and manually constructed variables (last three in the table).

Variable	**Column name**	**Description**
Timetable Point	Gebied	The timetable point the conflict is closest to
Train Series Pattern 1	Trein1	The series pattern of the first train of the conflicting headway situation
Train Series Pattern 2	Trein2	The series pattern of the second train of the conflicting headway situation
Overlap	Overlap	The amount of seconds that the extended blocking times of both trains overlap
Sections	Secties	The sections on which the conflict occurs
Headway situation	Opvolgsituatie	The event pair of the trains contributing to the conflict
Direction	Zelfde richting	Whether the train contributing to the conflict are driving in the same direction
Type headway situation	Type opvolging	Whether the conflict relates to a succession or a crossing of the two trains

Table 4.4: Overview of timetable conflict dataset

4.3.4. Timetable data

The fourth type of input data is *timetable data*, specifying the planned events in terms of departure, passage and arrival times of trains at timetable points. The data was exported from DONS and shows all the planned events in a BU timetable of NS on timetable points and tracks. Each row in the dataset corresponds with an event of a train series pattern on a specific location. The dataset is used to identify the amount of scheduled running time supplement from the previous stop to the timetable point of the timetable conflict. The used predefined variables in this dataset are presented in Table 4.5.

Variable	**Column name**	**Description**
Train Series Pattern	Serie	The series pattern of the train
Timetable Point	Drp	The timetable point of the scheduled event
Track	Spoor	The track of the scheduled event
Stops	Stopt	Whether the train is scheduled to stop on the timetable point or not
Technical Minimum Running Time	RtKaal	The technical minimum running time of the train from the previous timetable point
Scheduled Running Time	RtPlan	The scheduled running time of the train from the previous timetable point
Scheduled Arrival Time BU	Aankomst	The scheduled arrival time in the BU timetable
Scheduled Departure Time BU	Vertrek	The scheduled departure time in the BU timetable

Table 4.5: Overview of timetable dataset

4.3.5. Network data

The fifth type of input data is *network data*, specifying how each timetable point is connected to its neighbour(s) in the network. The text file that was retrieved for this purpose is an infrastructure file from Infra Atlas, which normally serves as input for Donna. Infra Atlas is one of the applications that ProRail uses and maintains with information about the railway infrastructure in the Netherlands. Each row in the file corresponds with a connection to a neighbouring timetable point by a single track. By looping through the rows, an adjacency list was created that specifies the neighbour(s) of each timetable point, and a dictionary was created that specifies for each timetable point what type it belongs to. The adjacency list enables identification of the previous and next timetable point as well as the scheduled and realised times at these point.

4.4. Pre-processing

This section describes the pre-processing steps taken to acquire the final dataset used for the statistical analysis. In particular, it addresses the identification of critical conflicts and its realisations and the operationalisation of the predictor variables.

4.4.1. Critical conflicts

From the *timetable conflict data*, critical conflicts are identified, indicating a timetable conflict between two specific train series patterns located in a critical point. Although in literature a critical point is predominantly determined on block level, the *timetable conflict data* used in this research aggregates the timetable conflicts of multiple sections based on the closest timetable point. Therefore, this research uses the notion of critical timetable point rather than critical block section. The choice to only consider critical conflicts is based on the aim of this research to determine precise critical buffer times for tight headway situations. For instance, with a headway of ten minutes and a technical minimum of two minutes, it matters little what the minimum headway buffer would be.

4.4.2. Realisations of timetable critical conflicts

First the critical conflicts are filtered to show only (buffer time) conflicts where (i) the train series patterns of the conflict do not belong to the same pattern; (ii) the scheduled headway is smaller than a set threshold; and (iii) the timetable point of the conflict is within the chosen geographical scope.

Secondly, the *traffic realisation data* for the first and second train is coupled to the *timetable conflict data* by using Pandas' merge function. Based on the train series pattern, timetable point and event, the realisation data of the first train is inner merged (i.e. only the rows are kept of which the keys train series pattern and timetable point are in both datasets). This results in a dataset with all realisations of the first train of all critical conflicts within the selected scope. Next, the realisation data of the second train is added in the same manner, also taking into account the date.

Lastly, only the rows are kept where (i) the scheduled headway is at most 15 minutes (based on a frequency of two times per hour); and (ii) the scheduled headway is positive, such that the scheduled order of the trains in the most recent plan is equal to the order in the BU timetable.

4.4.3. Operationalisation of predictor variables

The candidate predictor variables have been identified in Section 4.2. This section presents the computation method of the various predictors. An overview of the candidate predictor variables and its computation are provided in Table 4.6. For some of the computed predictor variables additional details on the calculation are provided in Appendix D.

	Variable	**Computation method**
For the headway situation	Direction	Appended from *timetable conflict data*
	Number Preceding Conflicts	Iterative algorithm that determines the number of prior timetable conflicts on a timetable point in which train 1 has acted as train 2 plus the number of preceding conflicts of each of these prior conflicts. See Appendix D for detailed explanation
	Scheduled Buffer	Subtract technical minimum headway from scheduled headway
	Scheduled Headway	From traffic realisation data, subtract scheduled time of train 1 from scheduled time of train 2
	Technical Minimum Headway	Scheduled headway minus the planning norm for buffer time plus the duration of the buffer time conflict. See Appendix D for detailed explanation
	Timetable Point	Appended from *timetable conflict data*
	Type Headway Situation	Appended from *timetable conflict data*
	Type Timetable Point	Appended from dictionary created with the *network data* based on timetable point column
For both trains	Event	Appended from *traffic realisation data*
	Carrier	Appended from *traffic realisation data*
	Driving Characteristic	Appended from *traffic realisation data*
	Previous Timetable Point; Next Timetable Point	Minimum gap between scheduled times at adjacent timetable points, known from the adjacency created with the *network data*. See Appendix D for detailed explanation
	Rolling Stock Type	Appended from *traffic realisation data*
	Rolling Stock Amount Carriages	Appended from *traffic realisation data*
	Scheduled Running Time Supplement	Calculated from the previous stop to the timetable point of the timetable conflict, with the use of *timetable data*. See Appendix D for detailed explanation
	Train Series	Appended *from traffic realisation data*
	Train Series Pattern	Appended from *traffic realisation data*
	Type Previous Timetable Point; Type Next Timetable Point	Appended from dictionary created with the *network data* based on Previous / Next timetable point

Table 4.6: Operationalisation of candidate predictor variables

4.5. Statistical analysis

This section introduces the statistical models created and analyses performed to derive buffer time planning rules from.

4.5.1. Regression analysis

To predict a target variable based on a wide variety of predictor variables, while maintaining easy interpretability, a regression analysis can be performed. With a regression analysis it could be determined whether the mean secondary delay and hindrance percentage can be inferred from the values of the candidate predictor variables. The intensity of the effects the predictor variables have on the target variables are represented by parameters (i.e. betas) in a regression function. More specifically, the betas indicate the change in the prediction of the target variable for a one unit increase in the predictor variables (Allison, 1999). In the context of this research, a larger value for the target variables indicates a bigger effect on delay propagation. The parameters are estimated by ordinary-least squares, which minimises the sum of the squares of the deviation between the target variables' observed and predicted values (Berry and Feldman, 1985).

This study estimates the linear relationship between the target variable and the predictor variables. The basic formula for such a linear regression is defined in Equation 4.2, where Y represents the target variable (i.e. either secondary delay or hindrance percentage), β_0 refers to the estimated constant, β_1 ... β_n are estimated coefficients for the scores of the target variable, X_1 ... X_n are observed scores for the predictors variables, and ϵ represents the estimated residual, capturing the variance of the target variable not explained by the predictors in the regression model. The additive model assumes that the effect of a change in a predictor variable on the target variable is not affected by the level of other predictor variables. This implies that the effect can be described without stating the level of other predictors variables.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \epsilon \tag{4.2}$$

In order to perform a linear regression, the following assumptions must be met (Aljandali, 2017):

- **Linearity** - there exists a linear relationship between each predictor variable and the response variable.
- **Absence of multicollinearity** - none of the predictor variables are highly correlated with each other.
- **Homoscedasticity of residual errors** - the residuals have constant variance at every point in the linear model.
- **Independence of residual errors** - the observations are independent.
- **Multivariate normality** - the residuals of the model are normally distributed.

In case the linearity assumption is violated, one can either choose to transform the data to a linear form (e.g. by taking the logarithm or square root) or resort to nonlinear regression analysis (e.g. polynomial regression, ridge regression, add interaction effects). When multicollinearity is present, one can remove the predictor variable that causes multicollinearity from the model. Heteroscedasticity can be fixed by either (i) transforming the target variable, (ii) redefine the target variable, or (iii) use weighted regression. Regarding the last two assumptions, the Gauss-Markov theorem states that they do not have to be met as long as the errors are uncorrelated and have a mean expected value of zero (Theil, 1971).

4.5.2. Dummy coding

Since regression analysis is based on correlation between the target variable and predictors, all variables in the model are required to be numeric (Alkharusi, 2012). However, some of the predictor variables are categorical variables. By applying dummy coding to all non-numerical predictor variables, the categorical variables can be included in the regression analysis. Dummy coding transforms a categorical variable with k categories to k-1 dummy variables that each represent one category and have numerical values of zero and one. A one is assigned when a observation belongs to the category the dummy variable represents, otherwise a zero is assigned.

The category that is not represented by a dummy variable is called the reference level. The regression coefficients (β) of the included dummies indicate the difference between the category the dummy represents and the reference level. In this research, the reference category is the level that belongs to the conflict whose target variable is closest to the population mean.

4.5.3. Model estimation procedure

In case multiple candidate predictor variables have been identified, a regression model including all candidate predictors is often hard to interpret, not parsimonious and suffers from multicollinearity. To avoid this, only the most important candidate predictors are incorporated in the model (Galvão et al., 2008; Ghani and Ahmad, 2010). To this end, there are three common variable selection methods for multiple linear regression in literature: (i) backward elimination, (ii) forward selection, and (iii) stepwise.

This research adopts a variable selection strategy that compares best to the stepwise method, a combination of backward elimination and forward selection. A stepwise model estimation procedure, starts with no predictors in the model. Then, the effects of adding a predictor to the model is evaluated for each candidate predictor individually. If adding any predictor to the model appears significant (i.e. the p-value of the predictor is below a certain threshold), then the candidate predictor that yields the lowest p-value is added to the model. This process continues until adding a candidate predictor does no longer result in a significant p-value or when all candidate predictors have been added to or excluded from the model. Simultaneously, already included predictors have to maintain significance. When the p-value of an already included predictor rises above a certain threshold, the predictor is removed from the model.

The model estimation procedure in this research, shown in Figure 4.4, differs from the above-mentioned stepwise method in four ways. Firstly, the adding criterion becomes the candidate predictor that results in the highest increase in the adjusted R-squared value. The adjusted R-squared value represents the proportion of the variance of the target variable that is explained by the predictor variables. This statistic ranges from 0, meaning the model explains no variance, to 1, meaning the model explains the full variance of the target variable. Unlike the regular R-squared value, the adjusted R-squared value accounts for the number of predictor variables included in the regression model, and thus leads to a more parsimonious model.

Secondly, the elimination criterion of already included variables differs. This research keeps either the previous version of the model (i.e. without inclusion of the candidate predictor) or the version with the candidate predictor excluding the predictors that have become insignificant, depending on which model has the highest adjusted R-squared value.

Thirdly, an extra stopping condition is added. If the candidate predictor that causes the highest increase in the adjusted R-squared value is below a certain threshold, the model development is stopped and the last model is considered the model. This has one major advantage: only the most important predictors are incorporated into the model, resulting in a more parsimonious model. This results in a model that is interpretable and has a high accuracy at the same time. What is the most appropriate value for the threshold, is experimented in the next chapter by evaluating the adjusted R-squared value as well as the number of significant variables in the model for a variety of thresholds.

Finally, the estimation procedure eliminates multicollinearity. Multicollinearity implies that at least two predictors variables correlate both with each other and with the target variable (Akinwande et al., 2015). As a result, the estimated coefficients as well as the p-values might be unstable and incorrect, leading to potential false inferences about relationships (Midi et al., 2010). To avoid this phenomenon, the Variance Inflation Factor (VIF) are inspected before adding the candidate predictor to the model. The VIF measures the structural independence of each term from all other terms in the model and is calculated by Equation 4.3, where R_i^2 is the proportion of the variance explained by other predictor variables in the model. In this study, an arbitrary value of VIF > 10 is used as an indication of multicollinearity. A disadvantage of preventing multicollinearity is that it ignores the unique contribution of the excluded variable and thereby can result in a loss of the explanatory power (Graham, 2003).

$$VIF_i = \frac{1}{1 - R_i^2} \tag{4.3}$$

4.5.4. Limitations of the regression analysis

Wiegand (2010) find that regression model created with variable selection methods tend to find the correct model more often in situations with numerical candidate predictors as compared to binary predictors. They explain this by the amount of information both types of variables contain. Numerical predictors contain more variation than dichotomous predictors and thus more information. Predictors with less information may more closely overlap information with other (binary) predictors and be more prone to multicollinearity. In particular this can raise problems for selecting the categorical variables that have been coded into dummy variables.

4.6. Output

The output of the statistical analysis is threefold; (i) hindrance distributions; (ii) regression statistics, that state how well the models can predict and what predictor variables have the most impact; and (iii) buffer time planning rules, given an arbitrary value for acceptable amount of secondary delay and hindrance percentage.

The linear regression as presented in Equation 4.2 can be rewritten in the format of Equation 4.4, which can be used to determine the amount of scheduled buffer that would ensure a certain amount of hindrance percentage or mean secondary delay. As the acceptable amount of secondary delay is an arbitrary choice, a sensitivity analysis is performed to gain insights into the effects of various acceptable thresholds.

$$X_{scheduled\ buffer} = \frac{Y - \beta_0 - \beta_1 X_1 - \beta_2 X_2 - \ldots - \beta_{n-1} X_{n-1} + \epsilon}{\beta_{scheduled\ buffer}} \tag{4.4}$$

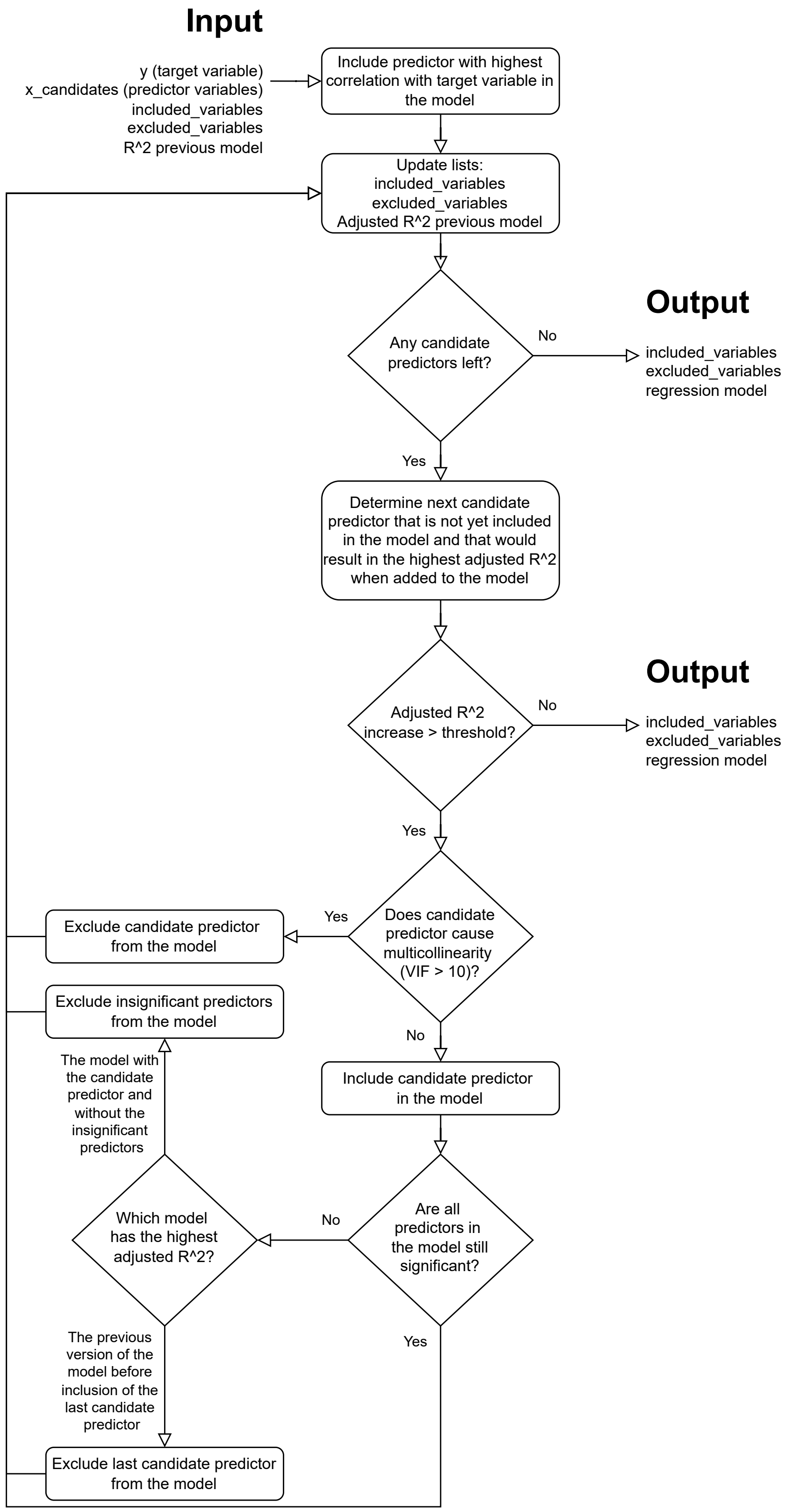


Figure 4.4: Conceptual model of multiple regression model

4.7. Chapter conclusion

This chapter started with the introduction of the framework used to determine buffer time planning rules from realisation data, followed by the selection of target and predictor variables for the statistical analysis and a description of the input data, pre-processing procedure, statistical analysis, and output. Using the information gathered in this chapter, the third sub-question is answered.

Sub-question 3: What data-driven model(s) can be used to determine buffer time planning rules in the tactical timetable design phase?

Buffer times are implemented to avoid or reduce delay propagation effects. Therefore, the development of the buffer time planning rules should consider the expected delay propagation, which can be represented by the mean secondary delay and probability of hindrance a succeeding train encounters caused by a delay of a preceding train. The observed mean secondary delay and the observed hindrance percentage occurring in a specific critical conflict situation can be appraised from realisation data. In this research the conceptual model of Weeda and Wiggenraad (2006) is utilised to determine these target variables. With the use of a prediction model on both target variables and setting an acceptable threshold for expected delay propagation, buffer time planning rules can be generated. The prediction model is based on the relation the target variables have with various timetable characteristics that function as predictor variables. Although literature identifies various other categories of variables to influence delay propagation, this research only considers timetable characteristics as predictor variables, because these are decision factors in the tactical planning stage of railway timetables where an initial timetable is constructed.

Due to the need for interpretability of the prediction model, a multiple linear regression model is suitable. To automate the data-driven regression model, a model estimation procedure resembling a stepwise variable selection method was introduced. The model estimation procedure in this research differs from commonly used stepwise variable selection methods in (i) adding condition; (ii) elimination condition; (iii) stopping condition; and (iv) a check on multicollinearity.

5
Case study and results

This chapter presents the case study and its results based on the conceptual framework for determination of buffer time planning rules. Section 5.1 introduces the case study's geographical and temporal scope. On top of the general pre-processing steps, some case specific pre-processing steps have been performed, listed in Section 5.2. Some of the results are presented for specific conflicts, to provide the reader with a more intuitive feel on how the results are to be interpreted. Which conflicts are highlighted in the results is explained in Section 5.3. Results were obtained by performing a exploratory data analysis in Section 5.4 and a regression analysis in Section 5.5. Next, Section 5.6 combines all the presented results to create buffer time planning rules. Section 5.7 summarises the chapter and provides the answer to sub-question four.

5.1. Case description and scope

The conceptual framework for buffer time rules determination was applied in a case study on a part of the Dutch railway network. A case study is performed, for it is too laborious for a master thesis research to evaluate planning standards for the entire rail network. A case study provides an idea of the state of affairs and the usefulness of such research. This section addresses the geographical and temporal demarcation.

5.1.1. Geographical scope

The study area consists of the corridor between Schiphol Airport and Leiden Centraal and the corridor between Haarlem and Leiden Centraal. The four main reasons for choosing this demarcation are (i) the combination of a busy and less busy corridor enables analysing the impact of capacity consumption on mean secondary delay and hindrance percentage; (ii) the (expected) shortage of capacity on the station Schiphol Airport and the corridor Schiphol Airport - Leiden Centraal (Planting, 2016); (iii) a variety of train types make use of the infrastructure and a variety of timetable point types are present in the scope; and (iv) relatively low amount of interaction with freight trains and trains of other carriers of which no data was available.

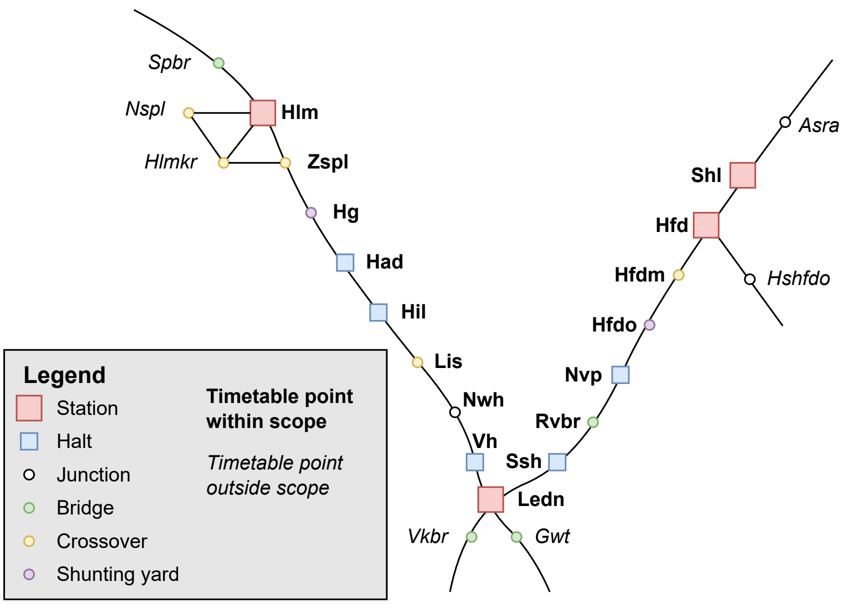


Figure 5.1: Schematic overview of geographical scope. Own creation based on ProRail (2022)

A schematic overview of the geographical scope is depicted in Figure 5.1, where the names of the timetable points are abbreviated. An overview of the name that belongs to an abbreviation in presented in Appendix E in Table E.1. Excluding the timetable points right outside the timetable point (depicted in italics in Figure 5.1), in total the study area contains 16 timetable points, of which four stations, five halts, one junction, one bridge, three crossovers and two shunting yards.

Line plan

The train series operating within the case study area are shown in Figure 5.2 and Table 5.1. Moreover, conflicts including a train not driving through the case study can occur at timetable points at the border of the geographical demarcation. Therefore, Table 5.2 includes an overview of all additional train series operating on a timetable point at the border of the geographical scope (e.g. Haarlem, Leiden Centraal & Schiphol Airport).

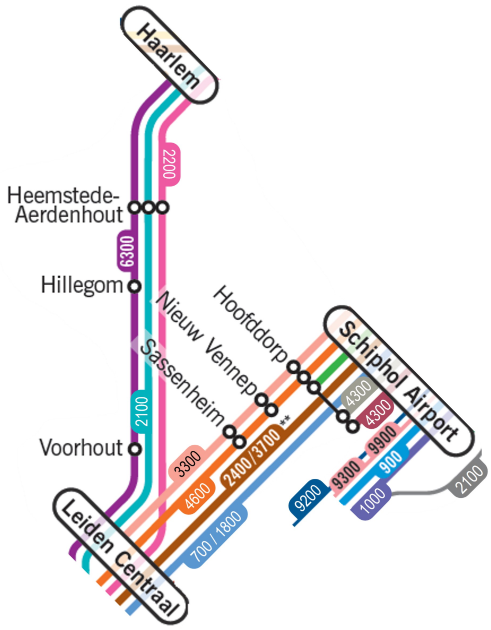


Figure 5.2: Line plan of the 2020 timetable within the case study area

Train Series	Trajectory	Trajectory in case study area	Driving charac-teristic	Frequency per direction	Start/end in case study area
700	Groningen - Den Haag Centraal	Schiphol Airport - Leiden Centraal	IC	1/h	
1800	Leeuwarden - Den Haag Centraal	Schiphol Airport - Leiden Centraal	IC	1/h	
2100	Amsterdam Centraal - Den Haag Centraal	Haarlem - Leiden Centraal	IC	2/h	
2200	Amsterdam Centraal - Vlissingen	Haarlem - Leiden	IC	2/h	
2400	Lelystad Centrum - Dordrecht	Schiphol Airport - Leiden Centraal	IC	2/h	
3300	Hoorn Kersenboogerd - Leiden Centraal	Schiphol Airport - Leiden Centraal	SPR	2/h	x
3700	Venlo - Dordrecht	Schiphol Airport - Leiden Centraal	IC	2/h, ** only in super off-peak hours	
4300	Almere Oostvaarders - Hoofddorp	Schiphol Airport - Hoofddorp	SPR	2/h	x
(1)4600	Zwolle / Amsterdam Centraal - Den Haag Centraal	Schiphol Airport - Leiden Centraal	SPR	2/h	
5700	Utrecht Centraal - Hoofddorp	Schiphol Airport - Hoofddorp	SPR	2/h	x
(1)5800	Amersfoort Vatforst / Amsterdam Centraal - Hoofddorp	Schiphol Airport - Hoofddorp	SPR	2/h	x
6300	Haarlem - Den Haag Centraal	Haarlem - Leiden Centraal	SPR	2/h	x

Table 5.1: Line plan of the 2020 timetable within the case study area

Train Series	Timetable Point in case study	Driving characteristic	Frequency per direction	Start/end in case study area
900	Schiphol Airport	IC Direct	2/h	
1000	Schiphol Airport	IC Direct	2/h	
14400	Haarlem	SPR	2/h, only during summer	x

15400	Haarlem	SPR	2/h, only during summer	x
(1)1600	Schiphol Airport	IC	2/h	x
3100	Schiphol Airport	IC	2/h	x
3400	Haarlem	IC	2/h	x
3500	Schiphol Airport	IC	2/h	x
4800	Haarlem	SPR	2/h	
5400	Haarlem	SPR	2/h	
8800	Leiden Centraal	IC	2/h	x
8900	Leiden Centraal	SPR	2/h	x
9100	Schiphol Airport, passage only	ES	1/h	
9200	Schiphol Airport	IC Direct	1/h	
9300 / 9900	Schiphol Airport	THA	1/h	

Table 5.2: Line plan of the 2020 timetable touching the case study area

5.1.2. Temporal scope

The study period is limited to December 15th 2019, until March 15th 2020, because it is the last period before the COVID-19 pandemic and the first timetable to be communicated to drivers in tenths of minutes (as opposed to minutes formerly). This timetable is considered as the benchmark for returning to the regular timetable now that COVID-19 restrictions are lifted. Realisation data from March 15th 2020 onwards is less representative, because there is less (secondary) delay due to (i) less passengers, thus shorter dwell times; (ii) less operating trains, thus less conflicts.

5.2. Data pre-processing

On top of the general pre-processing steps described in Section 4.4, some case specific pre-processing steps have been performed:

- Delays longer than ten minutes and early arrivals of more than four minutes are not taken into account, because in these cases rescheduling is often applied. In this study the focus is on secondary delays due to normal operation where only smaller initial delays occur, and thus no dispatching is considered. This approach is coherent with Jensen et al. (2017) and Kroon et al. (2008).
- Order changes in realisation are excluded from the analysis. The reason for this is that the scheduled first train can no longer affect the delay of the scheduled second train. Although the scheduled first train should not run this late, the effect on the scheduled train is limited because of the order change.
- This study does not distinguish between short stops (event = K_A and K_V) and regular stops (event = A and V), because in the realisation short stops can become regular stops according to choices made in later planning stages. When seeing them as different events, the merging of the dataset could exclude some realisations, as the events in the BU timetable are not the same as the one in the realisation.
- This study only considers passenger trains, so no shunting movements are analysed.

Both regressions are based on a subset of the dataset, complying with the four points mentioned above. As the mean secondary delay and hindrance percentage are aggregated values per conflict situation, the regression is performed on observations that are aggregated per conflict. To this end, for numerical predictor variables the mean of all realisations of that conflict was taken, whilst for categorical predictor variables the mode is taken.

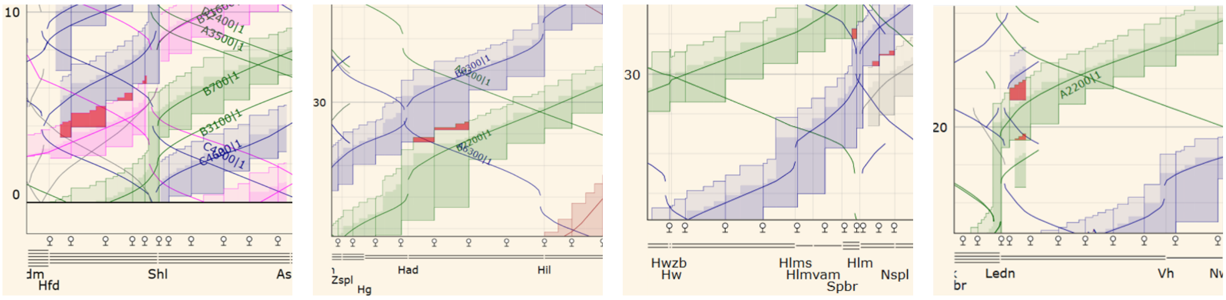
5.3. Example conflicts

In the remainder of the analysis, four conflicts will serve as an example to gain tangible results. These four conflicts have been selected, because they represent a large variety of situations. An overview of the main characteristics of the example conflicts is presented in Table 5.3 and train traffic diagrams are shown in Figure 5.3. More detailed figures of the scheduled routes that cause the conflicts are presented in Appendix E.

Train Series Pattern Train 1	Train Series Pattern Train 2	Timetable Point	Event pair	Driving Characteristics	Type Headway Situation	Number of realisations	Hindrance	Order Change
B700	A900	Shl	Arrival - Arrival	IC - IC Direct	Succession	1295	7 (0.54%)	44 (3.40%)

B2200	B6300	Had	Departure - Arrival	IC - SPR	Succession	893	342 (38.30%)	2 (0.22%)
B4800	D2100	Hlm	Departure - Arrival	SPR - IC	Succession	1181	217 (18.37%)	0 (0.00%)
A2200	C3300	Ledn	Departure - Departure	IC - SPR	Crossing	831	359 (43.20%)	32 (3.85%)

Table 5.3: Overview of example conflicts



(a) Schiphol (b) Heemstede-Aerdenhout (c) Haarlem (d) Leiden

Figure 5.3: Train traffic diagram with the critical conflict in red

5.4. Exploratory data analysis

This section presents the results of an exploratory data analysis. First, descriptive statistics and distributions of the target variables are discussed. Then, some relations of predictor variables with the target variables are explored to form hypotheses on which variables will have a significant impact on the target variables. Here it is also observed that secondary is not expected to adequately represent relationships with the predictor variables. Hence, it is explained that another target variable will be used in the remainder of the analysis.

5.4.1. Descriptive statistics

This section presents descriptive statistics of the target variables to provide some insights to which extent and where hindrance occurs before performing the main analysis. A brief overview of the main statistics by timetable point is presented in Table 5.4. It shows the number of (realisations of) timetable (critical) conflicts, hindrance percentage and mean secondary delay per timetable point, ordered according to their geographical location of the infrastructure. Per column it is indicated with darker colours where the higher values are located.

Although timetable conflicts take place on more than half of the timetable points within the scope, critical timetable conflicts occur predominantly on stop locations of trains (i.e. stations and halts). The highest number of critical conflicts can be found at Schiphol, indicating that this is a highly utilised timetable point. Interesting to see is that the hindrance percentage and mean secondary delay do not necessarily correlate with the number of realisations of critical conflicts on a timetable point. Hindrance percentage and mean secondary delay do align on the insights, as a higher hindrance percentage on a timetable point is paired with a higher mean secondary delay.

A more detailed overview of the descriptive statistics of the target and predictor variables can be found in Appendix F.

Timetable Point	**Abbreviation**	**Timetable conflicts [#]**	**Timetable critical conflicts [#]**	**Realisations of timetable critical conflicts [#]**	**Hindrance percentage [%]**	**Mean secondary delay [s]**
Haarlem	Hlm	57	51	16243	12,78	13,99
Haarlem Zuidelijke Splitsing	Zspl	0	0	0		
Haarlem Goederenstation	Hg	4	0	0		
Heemstede-Aerdenhout	Had	8	8	4551	31,84	24,98
Hillegom	Hil	4	0	0		
Lisse	Lis	0	0	0		
Noordwijkerhout	Nwh	0	0	0		
Voorhout	Vh	0	0	0		
Leiden	Ledn	54	54	21443	10	22,08

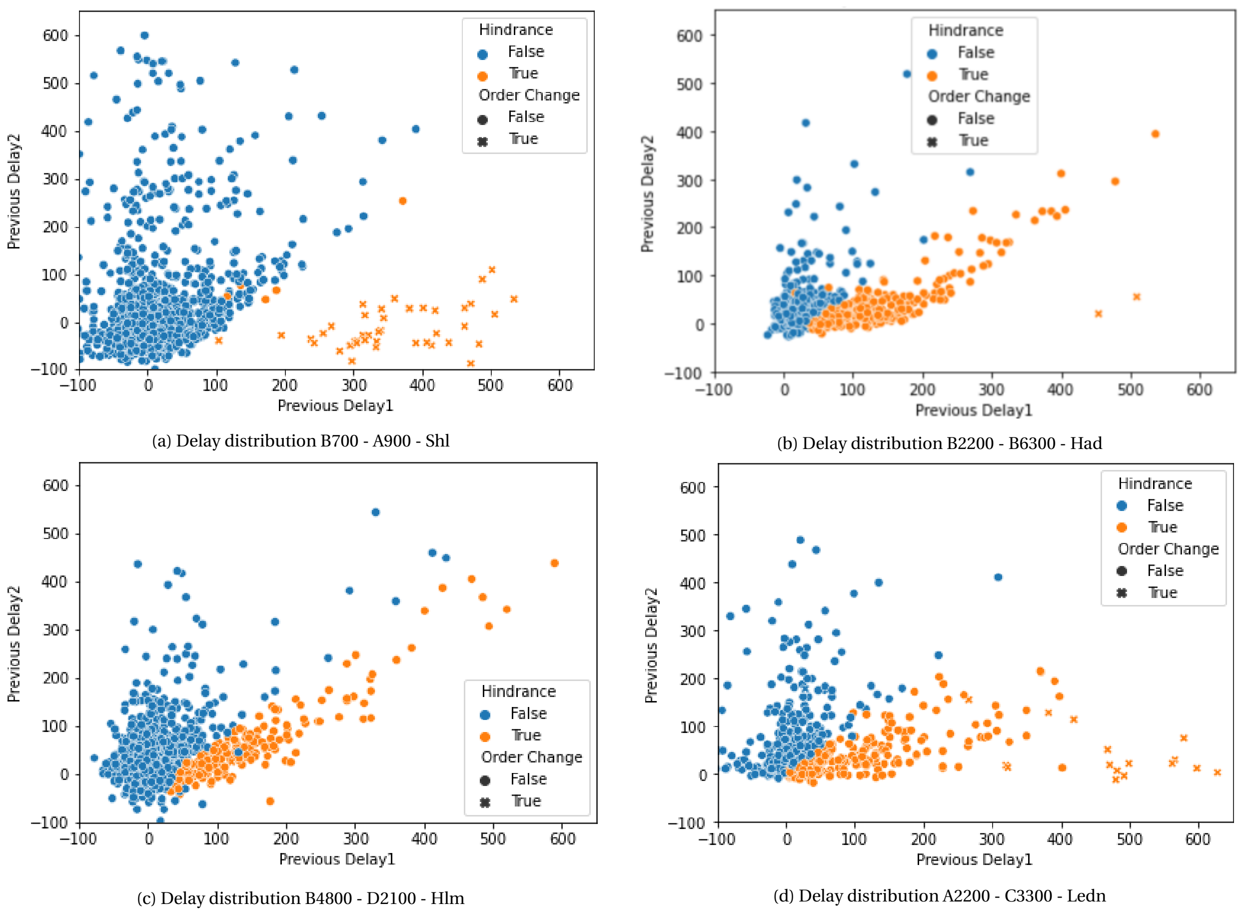
Sassenheim	Ssh	0	0	0		
Ringvaartbrug (brug o/d - bij Sassenheim)	Rvbr	0	0	0		
Nieuw Vennep	Nvp	2	0	0		
Hoofddorp Opstelterrein	Hfdo	0	0	0		
Hoofddorp Middenspoor	Hfdm	22	4	0		
Hoofddorp	Hfd	94	13	2214	31,17	40,16
Schiphol	Shl	134	127	36740	12,92	23,35

Table 5.4: Statistics per timetable point

5.4.2. Distributions of target variables

Figure 5.4 shows the realisations of the four example critical conflicts in the format of the conceptual model discussed in Chapter 4. The four critical conflicts differ in the number of realisations, hindrance percentage and order change percentage. It can be observed that the extent to which hindrance and order changes occur varies greatly among the critical conflicts that here serve as an example. This implies that hindrance, indeed, can be (partly) contributed to the timetable and network design. One striking observation is that the share of order changes at Schiphol appears considerably larger than at the other example conflicts. This could be caused by the choice that secondary delays at this timetable point are less accepted by traffic control due to the location specific infrastructure (e.g. tunnel).

The magnitude of the hindrance is less obvious from this figure, because the scheduled buffer can vary among the observations of a conflict, due to changes in BDu or SD timetable design phases (see Figure 5.5). But in general, observations further to the bottom right have a higher secondary delay, except when both trains have changed order, then the secondary delay as defined in this study is unknown. That is also the reason why observations with order changes are excluded from the remainder of the analysis.



(a) Delay distribution B700 - A900 - Shl

(b) Delay distribution B2200 - B6300 - Had

(c) Delay distribution B4800 - D2100 - Hlm

(d) Delay distribution A2200 - C3300 - Ledn

Figure 5.4: Delay distributions

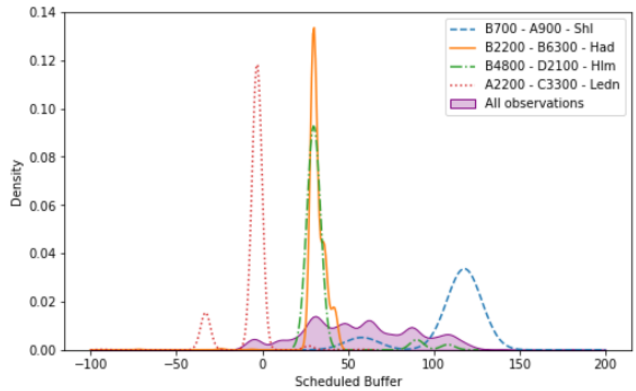
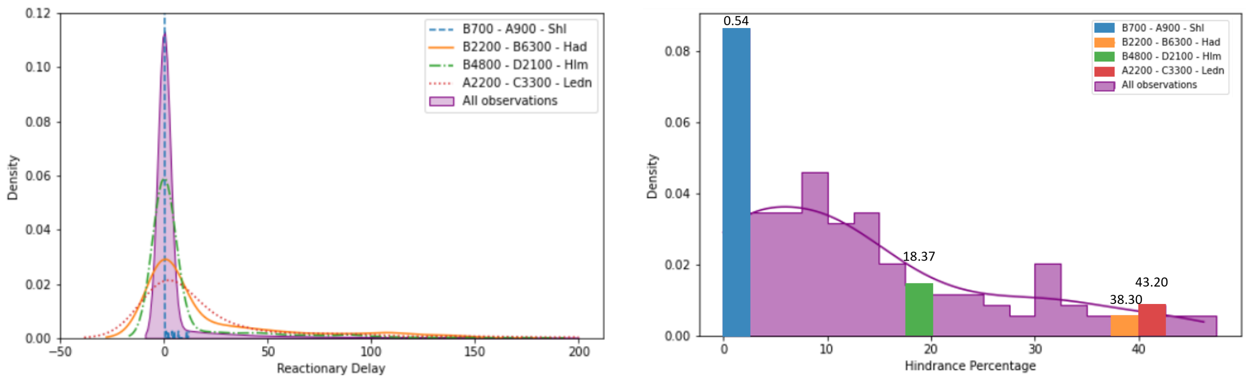


Figure 5.5: Distribution of scheduled buffer per conflict

The distributions of the target variables over all critical conflicts within the scope are presented in Figure 5.6, where the values of the four example critical conflicts are highlighted. The graph of the secondary delay follows a right-skewed curve, so the majority of the secondary delay has a small value. This is confirmed by the mean of 7.32 and the median of 0.00 over all observations. The conflicts with a wider distribution indicate both more variance in secondary delay and a higher hindrance percentage.

The graph of the aggregated conflicts shows that the hindrance percentage of the conflicts is mostly between 0% and 20%, and there are a few critical conflicts with up to 46.27% of realisations experiencing hindrance. The mean and median over all conflicts are respectively 13.32 and 9.58. The kernel density function has a drop at 0%, because there are almost no conflicts where hindrance does never occur in the realisation. For each of the conflicts that serve as an example it is indicated to what bin of the histogram they belong.



(a) Secondary delay distribution

(b) Hindrance percentage distribution

Figure 5.6: Distributions of target variables

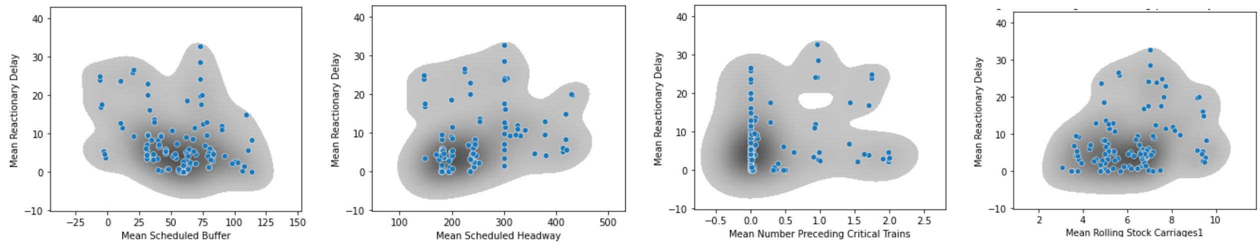
5.4.3. Relationships with target variables

To explore the relationships between the target variables and its candidate predictors, some figures are presented. For numerical predictors a scatter plot is drawn, so that it can be assessed whether a significant relationship is expected. A scatter plot plots all individual observations as dots, based on its values for a specified y-axis and x-axis. In case of a high amount of observations, the points in the plot can overlap, making it unclear where the most observations are positioned. To overcome this problem, a kernal density function is drawn behind the points, indicating where the observations are concentrated. Darker parts contain more observations than lighter parts.

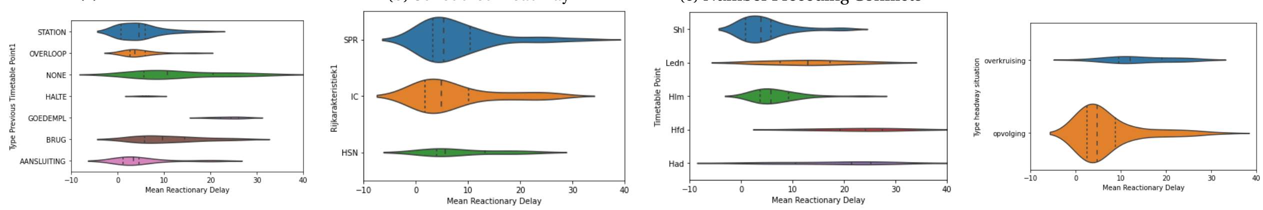
For categorical predictors scatter plots are an inconvenient way of presenting the relationship with the target variable, because all observations of one category would lie on the same line, making it unclear where the most observations are positioned. Since a kernal density function cannot be applied to categorical variables, another visualisation technique is chosen. The relationships of categorical variables with the predictor variables are depicted with the use of violin plots, which are a combination of a boxplot and a kernal density function. From the violin plots the percentage of observations in each category can be deducted by inspecting the size of each violin plot. Moreover,

the 25%, 50% and 75% interval of all observation within each category are shown with lines in the plot.
Figure 5.7 shows the relationships between mean secondary delay per conflict and some of the candidate predictor variables. As expected, the scheduled buffer seems to have a negative relationship with the mean secondary delay (Figure 5.7a). Longer preceding trains, tends to increase the mean secondary delay, which is explained by the longer occupation time of the section. Also, clear differences can be observed within the categories of the type previous timetable point of train 1 (Figure 5.7e) and type headway situation (Figure 5.7h). Therefore, these variables are expected to prove significant impact in the regression analysis.

Figure 5.8 shows the relationships between the hindrance percentage per conflict and some of the candidate predictor variables. Similar to Figure 5.7 this figure depicts more clear relationships than Figure F.1. Principally, the relationships observed are similar to Figure 5.7, with a few exceptions. For the driving characteristic of train 1 (Figure 5.8f), the HSN category seems to have some higher values compared to the other two categories. Additionally, timetable point Ledn (Figure 5.8g) shows a lower hindrance percentage compared to the other timetable points, which was not expected based on Figure 5.7g. Finally, the difference in type headway situation (Figure 5.8h) seems bigger for the target variables hindrance percentage than for mean secondary delay (Figure 5.7h).

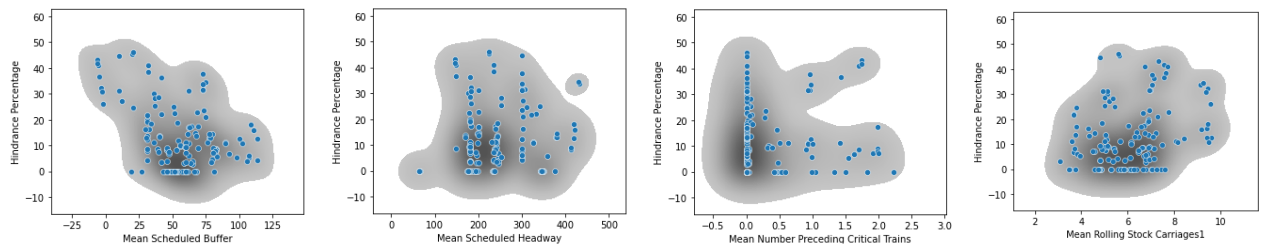


(a) Scheduled Buffer (b) Scheduled Headway (c) Number Preceding Conflicts (d) Rolling Stock Carriages Train 1

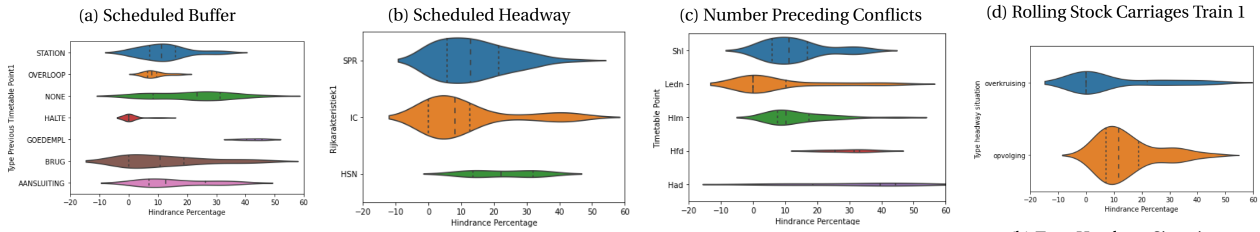


(e) Type Previous Timetable Point Train 1 (f) Driving Characteristic Train 1 (g) Timetable Point (h) Type Headway Situation

Figure 5.7: Relationships with mean secondary delay



(a) Scheduled Buffer (b) Scheduled Headway (c) Number Preceding Conflicts (d) Rolling Stock Carriages Train 1



(e) Type Previous Timetable Point Train 1 (f) Driving Characteristic Train 1 (g) Timetable Point (h) Type Headway Situation

Figure 5.8: Relationships with hindrance percentage

5.5. Regression results

This section presents the results of the regression analyses. It starts with an exploration of a suitable threshold for adjusted R-squared value increase to be used in the model estimation procedure. Next, it is checked whether the assumptions of linear regression analyses are violated, and if so, how this is handled. Subsequently, final regression models are presented and interpreted.

5.5.1. Explore threshold for adjusted R-squared increase

In the model estimation procedure two contrasting objectives are discovered. On the one hand, the adjusted R-squared value should be as high as possible, to explain as much variance of the target variable. Adding predictors to the model by definition increases the adjusted R-squared value of the model. On the other hand, for good interpretability of the results, as little significant predictor variables as possible is desired.

The threshold for adjusted R-squared value increase to pursue or end the model estimation procedure determines the trade-off of the model to add another predictor against its increase in adjusted R-squared value. To determine an appropriate value for the threshold, some experiments are done with varying values. Figure 5.9 presents the results of these experiments on adjusted R-squared value and the number of significant predictors in the final model. From the figure can be observed that the adjusted R-squared value for both target variables steadily decreases with larger values for the threshold. The number of significant variables in the model for both target variables takes a big drop first and then stabilises. The big drop in significant predictor variables together with a steady decrease of adjusted R-squared value indicates that the interpretability of the model is improved remarkably without compromising too much on model accuracy. For that reason, the remainder of the analysis will be performed with a threshold of 0.005, the point in the figure where both these aspects can be observed the best.

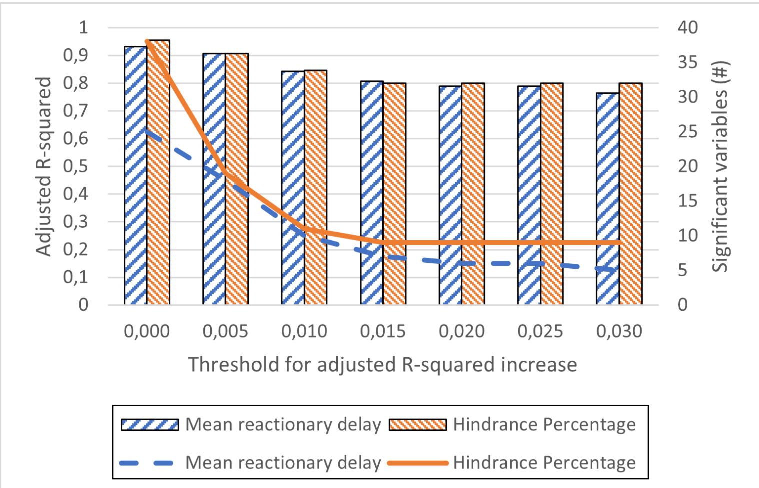


Figure 5.9: Effect of the threshold for adjusted R-squared increase on the regression output. Bar plot refers to the adjusted R-squared of the final model (left y-axis), line plot refers to the number of significant variables in the model (right y-axis)

5.5.2. Check on assumptions

The five assumptions of multivariate linear regression models have been discussed in Chapter 4. To recap, (i) the relationship between each predictor variable and the target variable should be linear (Linearity); (ii) none of the predictor variables can be highly correlated with each other (Absence of multicollinearity); (iii) the standardised residuals of the model should have constant variance at every point in the linear model (Homoscedasticity); (iv) the observations should be independent (Independence); and (v) the standardised residuals of the model should be normally distributed (Multivariate normality). The linearity assumption is checked during the preliminary data analysis in Section 5.4 and the multicollinearity assumption is ensured by the model estimation procedure as described in Chapter 4. The other three assumptions are evaluated below for each estimated linear model.

Homoscedasticity is checked with a plot of the model's standardised residuals versus the predicted values. In case the point in the scatter plot follow a pattern, then heteroscedasticity is present. Oftentimes, the pattern that indicates heteroscedasticity resembles a 'cone' shape, where the standardised residuals become more spread out as the predicted values get larger. The most common approaches to fix heteroscedasticity are (i) applying a nonlinear transformation to the response variable (Osborne, 2002); (ii) redefining the response variable; and (iii) using weighted regression.

Independence of observations is checked by performing a Durbin-Watson test, which is a formal statistical test that states whether the standardised residuals of the model exhibit autocorrelation. Generally, when the test statistic, denoted by d, is between 1.5 and 2.5 autocorrelation is unlikely to be present. The most common approaches to fix dependence of observations are (i) adding lags of the target and predictor variables to the model; (ii) checking that none of the variables in the model is overdifferenced; and (iii) adding seasonal variables to the model.

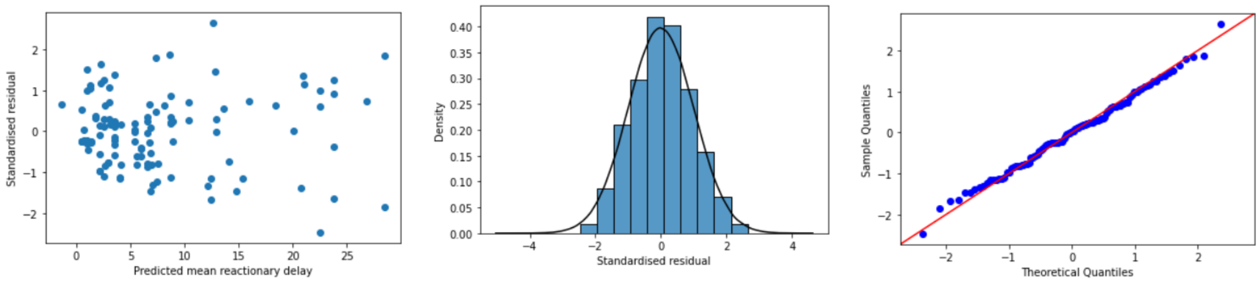
Multivariate normality can be checked in two ways; (i) visually using Q-Q plots; and (ii) using a formal statistical test like Shapiro-Wilk. Since, this statistical test is sensitive to large sample sizes, which are used in this research, this study uses Q-Q plots to assess the multivariate normality. The most common fixes when the normality assumption is violated are (i) verifying that there are no extreme outliers present in the data; (ii) if there are extreme outliers, resort to robust regression techniques or remove the outliers from the models; and (iii) applying a nonlinear transformation to the response variable (Osborne, 2002).

Regarding the last two assumptions, the Gauss-Markov theorem states that they do not have to be met as long as the errors are uncorrelated and have a mean expected value of zero (Theil, 1971).

Mean secondary delay

Before the choice was made to resort to mean secondary delay as a target variable, the assumptions of a regression model on secondary delay were also evaluated. Four out of five (with the exception of multicollinearity) were violated (see Appendix F for the detailed results of the tests). This is in line with the observations discussed in Section 5.4 and therefore presents yet another argument to analyse the mean secondary delay as target variable.

The figures used to check the assumptions are presented in Figure 5.10. Figure 5.10a shows that no heteroscedasticity is present, because the standardised residuals do not seem to become spread out as the predicted values get larger. Figure 5.10b and Figure 5.10c show that the standardised residuals mostly follow a normal distribution. Moreover, the mean of the standardised residuals equals -7.31e-13, which is very close to zero, and with a d-statistic of 2.30 the Durbin-Watson test concludes that the observations are dependent. Therefore, it is concluded that all regression assumptions are met.



(a) Plot to investigate homoscedasticity

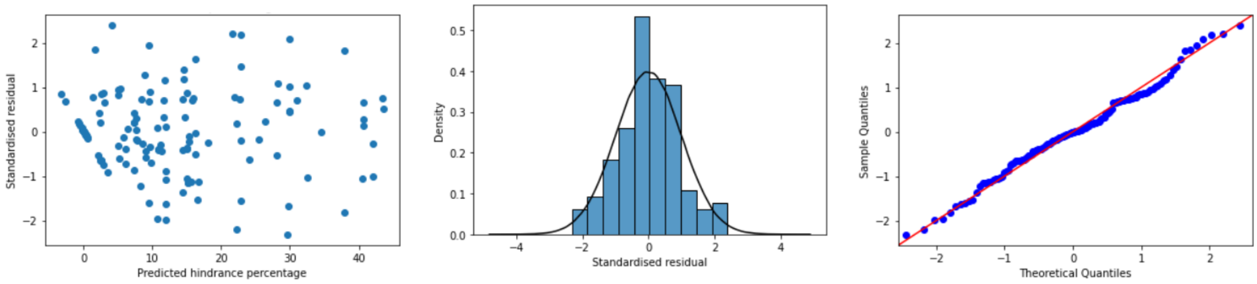
(b) Histogram showing the distribution of the standardised residuals

(c) Q-Q plot to investigate multivariate normality

Figure 5.10: Plots to check regression assumptions of the model on mean secondary delay (n = 112)

Hindrance Percentage

The figures used to check the assumptions are presented in Figure 5.11. Figure 5.11a shows that no heteroscedasticity is present, because the standardised residuals do not seem to become more spread out as the predicted values get larger. As for multivariate normality, Figure 5.11b and Figure 5.11c show that the standardised residuals mostly follow a normal distribution. Moreover, the mean of the standardised residuals equals 6.29e-14, which is very close to zero, and with a d-statistic of 2.32 the Durbin-Watson test concludes that the observations are independent. Therefore, it is concluded that all regression assumptions are met.



(a) Plot to investigate homoscedasticity

(b) Histogram showing the distribution of the standardised residuals

(c) Q-Q plot to investigate multivariate normality

Figure 5.11: Plots to check regression assumptions of the model on hindrance percentage (n = 139)

5.5.3. Regression statistics

In total two regression models for two target variables are created; mean secondary delay and hindrance percentage. The example conflict at Haarlem is taken as reference conflict, because the mean secondary delay and hindrance percentage of this conflict best resemble the statistics of the entire population as presented in Figure 5.6. This means that for the categorical variables, the level assigned to this conflict is taken as the reference level, see Table F.2 in Appendix F for an overview. As explained in Chapter 4, the coefficients of the dummy variables included in the model indicate the difference between the category the dummy represents and the reference level.

The unstandardised coefficients ought to be interpreted as the change in the value of the target variable when the predictor variable is increased with one unit, while controlling for all other predictors. Standardised predictors indicate the change in the standard deviation of the target value when the predictor variable increases with one standard deviation, while controlling for all other predictors.

Mean secondary delay

The first model aims to predict the mean secondary delay per conflict and is able to explain 90.7% of the variance within this variable (adjusted R-squared value = 0.907). Table 5.5 presents the coefficients of the included variables, which have a significant impact on the secondary delay. What stands out is that the scheduled buffer, nor the scheduled headway is found to have a significant effect on the mean secondary delay. Yet, interesting to see is that the technical minimum headway does have an impact on the secondary delay, suggesting that it might be effective to identify buffer time in percentages of the technical minimum headway rather than absolute values. This furthermore is in line with the way running time supplements are defined.

To conclude, it is stated that the model can sufficiently predict the mean secondary delay and therefore can be used to extract buffer time planning rules from. Although the scheduled buffer has no significant influence and realisations deviate from the mean, the model provides insights into how variables compare to the reference conflict in Haarlem and into the unit that buffer time could be scheduled with.

		Unstandardised coefficients		Standardised coefficients		
Variable	**Category**	**Beta**	**Standard error**	**Beta**	**t**	**Sig.**
Constant		-4.356	0.876		-4.970	0.000
Previous Timetable Point Train 2 (ref = Spbr)	Hfdm	16.854	1.361	2.250	12.379	0.000
	Vkbr	-15.127	1.250	-2.019	-12.098	0.000
Rolling Stock Type Train 1 (ref = SGMM3 SGMM3)	SNG3 SNG4	-5.128	1.400	-0.685	-3.664	0.000
	SNG4	2.587	0.897	0.345	2.885	0.005
Rolling Stock Type Train 2 (ref = VIRM6)	SNG3 SNG3	6.069	1.421	0.810	4.272	0.000
Scheduled Running Time Supplement Train 1		0.167	0.046	0.126	3.644	0.000
Technical Minimum Headway		0.036	0.004	0.318	9.266	0.000
Train Series Pattern Train 1 (ref = B4800)	B1800	-2.652	1.209	-0.354	-2.194	0.031
	D3700	7.985	1.682	1.066	4.749	0.000
	D5800	7.703	2.068	1.028	3.724	0.000
	D6300	6.617	2.047	0.883	3.233	0.002
Train Series Pattern Train 2 (ref = D2100)	B3700	11.724	2.430	1.565	4.825	0.000
	E3300	-6.582	1.805	-0.879	-3647	0.000
	E4800	5.652	1.889	0.769	2.992	0.004

	G4800	5.991	1.696	0.800	3.532	0.001
Train Series Train 1 (ref = 4800)	3400	4.876	1.257	0.651	3.881	0.000
Train Series Train 2 (ref = 2100)	2400	9.914	1.077	1.323	9.204	0.000
Type Next Timetable Point Train 2 (ref = Crossover)	Halt	19.801	0.886	2.643	22.349	0.000

Table 5.5: Regression results of mean secondary delay

To show how the regression output can be used to predict the mean secondary delay, the formula for the example conflict at Heemstede-Aerdenhout is presented in Equation 5.1. Note that this equation has been simplified by omitting components of $\beta_i X_i$ resulting in 0 (i.e. the train series of train 1 is 4800, thus the coefficient belonging to train series of train 1 when the category is 6300 is omitted). Figure 5.12 supports the conclusion that the model can sufficiently predict the mean secondary delay, by showing that for all four example conflicts the predicted value is close to the observed value and that both under- and over-estimations occur. A complete overview of the full regression formula in the format of Equation 4.2 is presented in Appendix F.

$$\textit{Predicted mean secondary delay} = -4.356 + 0.167 * 5.8 + 0.036 * 204 + 19.801 * 1 = 23.757 \tag{5.1}$$

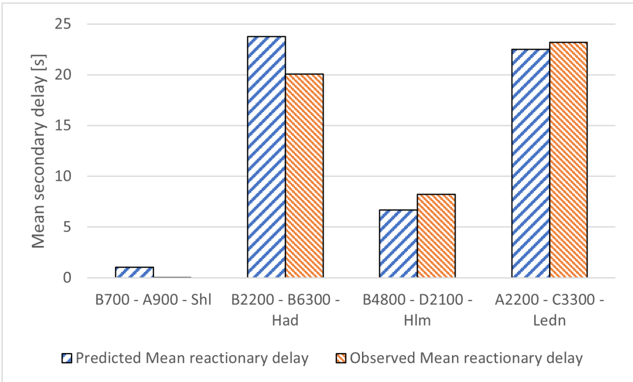


Figure 5.12: Prediction deviation mean secondary delay

Hindrance Percentage

The second model aims to predict the hindrance percentage per conflict and is able to explain 90.7% of the variance within this variable (adjusted R-squared value = 907). Table 5.6 presents the coefficients of the included variables, which have a significant impact on the secondary delay. For the significant categorical variables it is indicated what the reference level is.

When preceding trains are of the type IC Direct (abbreviation: HSN), the chance on hindrance increases with 13.822% as compared to a sprinter being the preceding train. Moreover, the second train being a starting train at the location of the conflict instead of coming from a bridge results in 23.708% more chance to be hindered. On the brighter side, crossings lead to 9.530% less chance on hindrance as compared to successions, which is explained by the fact the the time both trains are dependent on each other is smaller in case of crossings.

As expected, the scheduled buffer reduces the change on hindrance. Unexpectedly, the percentage of scheduled running time supplement the preceding train has, increases the hindrance percentage. Longer succeeding trains appear to experience less hindrance which might be because they gain a higher priority as more passengers are transported.

To conclude, it is stated that the model can sufficiently predict the hindrance percentage and therefore can be used to extract buffer time planning rules from.

		Unstandardised coefficients		Standardised coefficients		
Variable	**Category**	**Beta**	**Standard error**	**Beta**	**t**	**Sig.**
Constant		12.773	1.280		9.978	0.000
Driving Characteristic Train 1 (ref = SPR)	HSN	13.822	1.458	1.107	9.479	0.000

Next Timetable Point Train 1 (ref = Nspl)	Shl	25.769	1.828	2.063	14.099	0.000
	Vh	9.939	1.986	0.796	5.004	0.000
	Zspl	24.633	1.994	1.972	12.355	0.000
Previous Timetable Point Train 1 (ref = Spbr)	Hg	31.161	2.064	2.495	15.095	0.000
Previous Timetable Point Train 2 (ref = Spbr)	None	23.708	2.988	1.898	7.934	0.000
Rolling Stock Type Train 2 (ref = VIRM6)	SLT6	5.192	1.517	0.416	3.421	0.001
Scheduled Buffer		-0.119	0.015	-0.249	-7.647	0.000
Scheduled Running Time Supplement Train 1		0.337	0.081	0.139	4.144	0.000
Train Series Pattern Train 1 (ref = B4800)	A3300	-6.699	1.695	-0.536	-3.952	0.000
	B2400	12.643	2.841	1.012	4.450	0.000
	C3300	-6.963	2.313	-0.557	-3.010	0.003
	D3700	24.019	2.829	1.923	8.490	0.000
	D6300	11.241	3.197	0.900	3.516	0.001
Train Series Pattern Train 2 (ref = D2100)	B3700	17.067	4.030	1.366	4.235	0.000
	C1000	4.001	1.409	0.320	2.841	0.005
Train Series Train 1 (ref = 4800)	4600	8.352	1.799	0.669	4.644	0.000
Train Series Train 2 (ref = 2100)	2400	13.322	1.601	1.067	8.323	0.000
Type Headway Situation (ref = succession)	Crossing	-9.530	0.911	-0.763	-10.461	0.000

Table 5.6: Regression results of hindrance percentage

To show how the regression output can be used to predict the hindrance percentage, the formula for the example conflict at Heemstede-Aerdenhout is presented in Equation 5.2. Note that this equation has been simplified by omitting components of $\beta_i X_i$ resulting in 0 (i.e. the train series of train 1 is 4800, thus the coefficient belonging to train series of train 1 when the category is 4600 is omitted). Figure 5.13 supports the conclusion that the model can sufficiently predict the hindrance percentage, by showing that for all four example conflicts the predicted value is close to the observed value. A complete overview of the full regression formula in the format of Equation 4.2 is presented in Appendix F.

$$\textit{Predicted hindrance percentage} = 12.773 + 31.161 * 1 - 0.119 * 31.90 + 0.337 * 5.79 = 42.09 \tag{5.2}$$

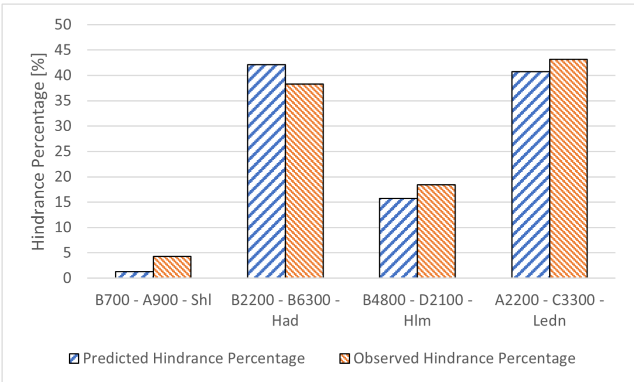


Figure 5.13: Prediction deviation hindrance percentage

5.5.4. Interpretation of regression analyses

The previous section has discussed two regression models, both being identified to be accurate enough to be used as prediction models. The models on mean secondary delay and hindrance percentage are both able to explain 90.7% of the variance on the target variable.

Limited overlap is observed in the significant variables of both regression models. This implies that hindrance itself and the extend to which hindrance occurs (i.e. hindrance percentage and mean secondary delay) are not likely to be influenced by the same factors. Only the following variables are present as significant predictor variables in both regression models: (i) Scheduled Running Time Supplement Train 1 (ii) Train Series Pattern Train 1 - D3700 & D6300; (iii) Train Series Pattern Train 2 - B3700; and (iv) Train Series Train 2 - 2400. As both target variables are affected by other factors, influencing both of them by changing the value of only one factor can only be achieved by changing the listed factors. But, since changing the name of a train series (pattern) is an arbitrary choice, the only real option to change is the scheduled running time supplement of train 1. Then, optimisations to both target variables separately can be done by changing their unique significant variables.

One striking observation is that the candidate predictor variables *timetable point*, *Number Preceding Conflicts*, *Same Direction*, *Event* and *Carrier* do not appear to have a significant impact on any of the target variables. The same is concluded for interaction effects between the scheduled buffer and (i) *timetable point*; (ii) *Type timetable point*; (iii) *Type Headway Situation*; (iv) *Driving Characteristic*; and (v) *Direction*. The latter implies that the effect of a unit of scheduled buffer is the same among all types of conflicts characterised by the listed predictor variables.

Interesting to see is that the scheduled buffer has a significant impact on hindrance percentage, whilst mean secondary delay is influenced by technical minimum headway. This indicates that it might be effective to identify buffer time in percentage of the technical minimum headway rather than absolute value.

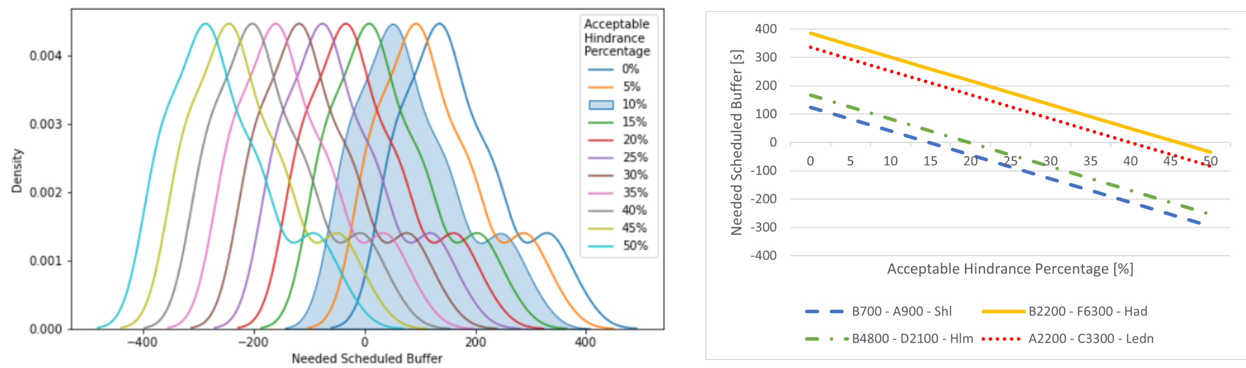
Using the standardised coefficients it can be determined what factors influence the target variables the most. The biggest influential factors on mean secondary delay are (i) Previous timetable point of train 2 being Hfdm (β = 2.250) or Vkbr (β = -2.019); and (ii) Type next timetable point of train 1 being a halt (β = 2.643). The biggest influential factors on hindrance percentage are (i) Next timetable point of train 1 being Shl (β = 2.063) or Zspl (β = 1.972); (ii) Previous timetable point of train 1 being Hg (β = 2.495); and (iii) Train 2 being a starting train (β = 1.898).

5.6. Determination of case-specific buffer time planning rules

The linear regression as presented in Equation 4.2 can be rewritten in the format of Equation 4.4, which can be used to determine the amount of buffer that would ensure a certain amount of hindrance percentage. As the scheduled buffer is not a significant predictor for the mean secondary delay, this method can not be applied to determine the amount of buffer that would ensure a certain amount of mean secondary delay.

Figure 5.14 shows that the amount of needed scheduled buffer to achieve a certain amount of hindrance percentage differs per conflict. Due to the linear nature of the regression model the distributions in Figure 5.14a are the same for all hindrance percentages, but are shifted on the x-axis and the lines in Figure 5.14b are linear. As the effect of one unit buffer does not differ among various conflicts the lines in Figure 5.14b are parallel.

Compared to the currently mean scheduled buffer per conflict (standard deviation = 26) the spread in the predicted values for the needed scheduled buffer per conflict is significantly higher (standard deviation = 93), leading to an increase of 257%. This increase in spread of buffer can also be observed in Table 5.7, where the range (i.e. difference in minimum and maximum value) of the currently mean scheduled buffer per conflict is 119 seconds, though for the predicted values for the needed scheduled the range equals 343 seconds. This corresponds with an increase of 188%.



(a) Distribution of needed scheduled buffer for critical conflicts in the geographical scope for different acceptable hindrance percentages

(b) Needed scheduled buffer for the example conflicts for different acceptable hindrance percentages

Figure 5.14: Plots to show variance in needed scheduled buffer

		Needed scheduled buffer [s] at hindrance percentage [%]										
	Currently scheduled buffer [s]	**0**	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**
Std	26	93	93	93	93	93	93	93	93	93	93	93
Mean	55	167	125	83	41	-1	-43	-85	-127	-169	-211	-254
Min	-6	44	2	-39	-81	-124	-166	-208	-250	-292	-335	-377

25%-quartile	40	107	74	42	9	-23	-55	-88	-121	-153	-186	-219
Median	56	146	104	62	20	-22	-64	-106	-148	-190	-232	-274
75%-quartile	73	186	153	120	88	55	22	-9	-42	-75	-107	-140
Max	113	387	345	302	260	218	176	134	91	49	7	-34
Sum	7716	23339	17473	11607	5741	-124	-5990	-11856	-17722	-23588	-29454	-35320

Table 5.7: Regression results of hindrance percentage

For example, using Equation 4.4, when one aims for a hindrance percentage of 15% at the example conflict of Heemstede-Aerdenhout, the amount of buffer that would ensure this can be calculated with the use of Equation 5.3, and results in 261 seconds. Note that this equation has been simplified by omitting components of $\beta_i X_i$ resulting in 0. The resulting needed buffer of 2611 seconds is much bigger than the buffer of 31.9 seconds that is currently scheduled at this conflict. This large deviation is explained by the fact that the hindrance percentage of the realisations is 38.29%, which is quite high. Thus to lower this, the scheduled buffer should be much bigger.

$$\textit{Needed scheduled buffer} = \frac{15 - 12.773 - 31.161 * 1 - 0.337 * 5.79}{-0.119} = 261 \tag{5.3}$$

5.6.1. Practical buffer time planning rules

Ideally, Equation 4.4 is integrated into the planning applications, so that the needed amount of buffer to achieve a certain hindrance percentage for a given conflict is identified automatically. To achieve this, still many steps have to be taken, which are described in Chapter 7. Until then, a manual approach is still required, and although Equation 4.4 resembles an accurate prediction of the needed scheduled buffer time for a given conflict, it is a time-consuming task to perform manually. To ease the direct utilisation of the planning rules, a more simplified, generalised and therefore practical approach is developed.

To this end, a base value for the needed scheduled buffer is presented for a range of hindrance percentages in Table 5.8a, based on (i) Equation 4.4; (ii) the constant of the regression formula; and (iii) the coefficient for scheduled buffer reported in Table 5.6. Essentially, the base value corresponds with a regression model where only the scheduled buffer is incorporated as predictor variable, which has an adjusted R-squared value of 0.161. Depending on the specific conflict, additional buffer time should be added to or deducted from this value. For each significant categorical predictor variable, it is stated in Table 5.8b what amount of scheduled buffer should be added to the base value when this category is present in the conflict, based on (i) Equation 4.4; (ii) the coefficient of the significant categorical predictor variable; and (iii) the coefficient for scheduled buffer as reported in Table 5.6. For each percentage of running time supplement scheduled to train 1, the predicted increase this gives to hindrance percentage can be cancelled out by a 2.83 seconds increase in scheduled buffer.

Still, to consider all variables listed in Table 5.8b can be an extensive job. Therefore, the column *Cumulative prediction accuracy* is added to represent the adjusted R-squared value of the prediction model when a variable, and all variables placed above that one are included in the model. The abverb *Cumulative* hence refers to the fact that the variables placed above the one looked up have to be in the model to achieve the prediction accuracy mentioned. The variables are ordered from highest to lowest increase in prediction accuracy, so that the most influential variables are included into the model first. This could help practitioners make a trade-off between model accuracy and workload in scheduling buffer time. Given the steep increase in cumulative prediction accuracy in Table 5.8b, Table 5.8a is likely to be used only in combination with the supplement.

For example, given a crossing conflict of train series patterns A2200 and B2400, a prediction accuracy of 73.5% is achieved when considering the *Next Timetable Point of Train 1 being Vh* and the *Train Series of Train 2 being 2400*. Whereas, also adding the variable *Type Headway Situation being a crossing* lifts the prediction accuracy to 88.2%. Hence the prediction accuracy can be read from the last line above the first excluded variable that is part of the conflict.

5.6.2. Reflection on buffer time planning rules

Chapter 3 discussed the current planning norms with regard to scheduling headway times. Essentially, the scheduled headways are composed of a technical minimum headway and a minimum buffer time of 60 seconds. As a rule of thumb, smaller buffer times are acceptable when the gain in headway time is bigger than the loss in running time, which is evaluated for each situation individually.

The planning rules presented in this chapter identify the minimum buffer time case-specifically, while evaluating the percentage of realisations that would be hindered. The planning rules provide more flexibility compared to the

Acceptable hindrance percentage[%]	Predicted scheduled buffer [s]
5	65
10	23
15	-18
20	-60
25	-102
30	-144
35	-186
40	-228
45	-270
50	-312

(a) Base values for predicted scheduled buffer

Variable	Category	Predicted scheduled buffer supplement [s]	Cumulative prediction accuracy [%]
Previous Timetable Point Train 1	Hg	262	26.9
Next Timetable Point Train 1	Shl	217	38.7
Driving Characteristic Train 1	HSN	116	47.7
Next Timetable Point Train 1	Vh	84	55.6
Next Timetable Point Train 1	Zspl	207	62.5
Train Series Pattern Train 1	D3700	202	67.1
Train Series Train 2	2400	112	71.8
Previous Timetable Point Train 2	None	199	73.5
Type Headway Situation	Crossing	-80	83.6
Train Series Train 1	4600	70	84.7
Train Series Pattern Train 2	B3700	94	85.7
Train Series Pattern Train 1	A3300	-56	86.6
Train Series Pattern Train 1	D6300	94	87.4
Train Series Pattern Train 1	B2400	106	88.2
Scheduled Running Time Supplement Train 1		2.83*	88.8
Rolling Stock Type Train 2	SLT6	44	89.5
Train Series Pattern Train 1	C3300	-59	90.2
Train Series Pattern Train 2	C1000	34	90.7

* For each percentage of scheduled running time supplement of the first train, an additional 2.83 seconds of buffer should be scheduled.

(b) Supplements per category for predicted scheduled buffer

Table 5.8: Practical buffer time planning rules

current norms, because the size of the minimum needed buffer depends on the preferred hindrance percentage.

Large differences in terms of predicted scheduled buffer are registered between the various conflicts in the case study. To obtain the same hindrance percentage the difference in predicted scheduled buffer between conflicts reaches up to 342 seconds. Thus, compared to the mean observed scheduled buffer per conflict (standard deviation = 26 seconds), the spread in the predicted values for the predicted scheduled buffer per conflict is significantly higher (standard deviation = 93), corresponding with an increase of 257%.

When the observed mean hindrance percentage of the realised conflicts (=13.32%) is taken as acceptable hindrance percentage for all conflicts in the prediction, the hindrance distribution across all conflicts becomes uniform and the sum of all predicted scheduled buffers equals the total observed scheduled buffer, namely 7716 seconds. For every percentage increase in hindrance percentage, the sum of predicted scheduled buffers drops with 1173 seconds. Additionally, the mean of all observed scheduled buffers equals the mean of predicted scheduled buffer across all conflicts. For every percentage increase in hindrance percentage, the mean of predicted scheduled buffers drops with 8 seconds. Yet, as shown in Table 5.9 the allocation of the scheduled buffers across the timetable points changes. In particular buffer times currently scheduled at Leiden are moved to Heemstede-Aerdenhout and Haarlem.

Timetable Point	Sum of observed scheduled buffer [s]	Share of observed scheduled buffer	Share of predicted scheduled buffer
Heemstede-Aerdenhout	211	2.74%	11.30 %
Hoofddorp	440	5.70 %	12.45%
Haarlem	1648	21.36 %	23.58 %
Leiden	1815	23.53%	8.30%
Schiphol	3600	46.67 %	44.37%

Table 5.9: Comparison of total scheduled buffer per timetable point (observed vs. predicted)

5.7. Chapter conclusion

This chapter presented a case study and its results based on the conceptual framework for determination of buffer time planning rules presented in Chapter 4. The case study's scope was introduced and case specific pre-processing steps were discussed. Next results in terms of exploratory data analysis and regression statistics were presented for all critical conflicts in the case study as well as for four example conflicts in more detail. Lastly, it was shown how to extract buffer time planning rules from the regression statistics. Using the information gathered in this chapter, the fourth sub-question is answered.

Sub-question 4: How are buffer time planning rules determined from the data-driven model(s) and how do they compare to the current planning norms for buffer times?

Two regression models were estimated, one on mean secondary delay and one on hindrance percentage. Both models appeared to have the same prediction accuracy of 90.7%. The model on mean secondary delay did not contain the scheduled buffer as a significant predictor variable and thus cannot be used to derive buffer time planning rules from. The model on hindrance percentage did contain the scheduled buffer as a significant predictor variable and thus can be used to derive buffer time planning rules from. The predicted needed scheduled buffer can be calculated with the use of Equation 5.4, where Y refers to the acceptable hindrance percentage chosen by the planner, β_0 corresponds with the constant; and $\beta_1 ... \beta_{n-1}$ and $X_1 ... X_{n-1}$ represent the coefficients and values of the other significant predictor values.

$$X_{scheduled\ buffer} = \frac{Y - \beta_0 - \beta_1 X_1 - \beta_2 X_2 - ... - \beta_{n-1} X_{n-1} + \epsilon}{\beta_{scheduled\ buffer}} \tag{5.4}$$

Compared to the current planning norms for buffer times, the planning rules presented in this chapter provide more flexibility, because the size of the needed scheduled buffer depends on the preferred hindrance percentage. The planning rules presented in this chapter however, do not indicate how the scheduled buffer compares to the mean secondary delay that can be expected. To gain the same mean hindrance percentage as the realisation data, the total amount and mean value of needed scheduled buffer remains unchanged. Yet, the variation in buffer sizes increases from a standard deviation of 26 seconds to a standard deviation of 93 seconds, meaning a 257% rise. Thus, the spread in scheduled buffer time is increased, and moreover the allocation of the scheduled buffer time changes geographically.

6
Discussion

This chapter further elaborates on the results from the regression analyses, discussing the quantitative but also qualitative aspects of secondary delay and hindrance percentage. Understanding these aspects helps create insights into how to prevent, anticipate and mitigate delay propagation effects.

6.1. General remarks

The general impression from the results in Chapter 5 is that the delay propagation at different critical conflicts is quite heterogeneous. The regression models were both able to explain 90.7% of this heterogeneity when using aggregated values for conflict situations, being mean secondary delay and hindrance percentage, meaning that structural delay propagation patterns can largely be contributed to timetable design and infrastructure characteristics. Still, critical conflicts with the same mean secondary delay and hindrance percentage can have different secondary delay distributions, caused by among others (i) number of realisations of the conflict; (ii) number dispatching options; (iii) proactive attitude of traffic controllers; (iv) delay distributions of the first and second train at the previous timetable point; (v) the variety in scheduled buffer; (vi) the variety in rolling stock; and (vii) time of day.

Integration of practical buffer time planning rules in the Dutch timetable design process

Currently, the necessary minimum buffer time is tested with a deterministic, microscopic simulation during the MLT phase. The buffer time needed for robust execution of the timetable is simulation by adding 60 seconds wake behind every train movement. Then, if this results in hindered operations it could still be acceptable, except when (i) an unplanned stop occurs, or (ii) an arrival delay bigger than 30 seconds occurs at certain stations. In reality, slightly hindered headways enables 'hidden' capacity to come available, and are as a rule of thumb accepted when the gain in headway time is bigger than the loss in running time.

The planning rules could be used to define the size of the wake added in the simulation, or can be used as an indication of when slightly hindered headways are acceptable. The former resonates best on the verge of the strategic and tactical planning stage, where microscopic simulations are used. The latter is more suitable on the interface between the tactical and operational design phase where conflicts are evaluated and approved.

6.2. Research limitations

Given the scope of this thesis and the followed methodology, there are some limitations to this research which are addressed here. As a result, there are still plenty of opportunities for future research which are discussed in the next chapter.

6.2.1. Methodological limitations

The analysis performed in this study can be made more accurate by addressing some simplifications and limitations of the methodology. In particular the accuracy of the analysis is limited by the following aspects:

- The chosen methodology only takes into account delay propagation between two trains, but does not contain information about the propagation of delay within a network. This makes it impossible to infer insights on vulnerabilities and bottlenecks to be used in the network design.

- In this research, a linear regression model was created, assuming a linear relationship with the target variables and buffer time. While literature agrees that buffer time limits secondary delay, it also shows the benefits decrease with increasing buffer time. More precisely, multiple studies demonstrated the presence of a negative exponential relationship between secondary delay and buffer time (Weeda, 2005; Yuan and Hansen, 2007; Zieger et al., 2018). Assuming a linear relation limits the accuracy of the model and may overestimate the effects of adding extra units of buffer time to reduce delay propagation.Moreover, due to the nature of the linear additive regression model, hindrance percentages below 0% and higher than 100% could be predicted, whilst this is not possible in reality.

- The target variable mean secondary delay is an aggregated value of multiple realisations of one critical conflict situation. So information on the spread of mean secondary delay within the conflict is lost, whilst this could be useful information in assessing the sensitivity of secondary delay with regard to the scheduled buffer time. A large spread could furthermore indicate that the specific values are caused by incidental causes, whereas a minimal spread in secondary delay may suggest more structural relationships with the timetable design.
- In this research, critical conflicts have been identified, indicating a timetable conflict between two specific train series patterns located in a critical point. It is assumed that each pair of train series patterns in the timetable conflict data has exactly one critical conflict. However, theoretically, multiple unrelated critical conflicts could exist between two train series patterns when they occur far enough from one another such that they do not influence each other. All independent critical conflicts could be captured by for example (i) having a maximum of one critical conflict per pre-specified area; or (ii) defining a minimum distance that the critical conflicts should be separated from each other; or (iii) accepting multiple critical conflicts if there is at least one timetable point in between where no timetable conflict is scheduled.
- The choice to use a case study limits the range of some predictor variables. For example, when the buffer time rules are developed based on location or train series it is impossible to integrate all possible categories into the case study. This research aimed to overcome this limitation by choosing a geographical scope that includes a large variety of type timetable points and driving characteristics. Moreover, quite specific variables where used, such as the train series (pattern) and location, making the approach less straightforward to generalise. This implies that the conclusions can only be drawn with regard to the investigated scope of the case study. To generate a more general model one could (i) append the model with a classification model that creates groups of the various categories (based on e.g. demand per timetable point, size of timetable point); or (ii) only include general variables in the model (exclude e.g. timetable point, train series (pattern)).
- Due to the limited scope of the case study, there is a risk of overfitting while creating the model with the used model estimation procedure. This could be resolved by splitting the dataset in a training dataset and test dataset. To this end, the model estimation procedure should be performed on the training set and expanded with an extra stop condition addressing the adjusted R-squared value improvement of the test set.

6.2.2. Data limitations

The analysis performed in this study can be made more accurate and scalable by using more sophisticated data. In particular the data of the following variables can be improved:

- **Secondary delay** - This research used an approximation method introduced by Weeda and Wiggenraad (2006). Although on average this estimation of secondary delay seems to resemble the true secondary delay, the estimation is particularly unsuitable for conflicts including starting trains and successions. This is because the method computes the target variable secondary delay at a conflict point based on the delay of two trains, before they have interacted with each other. In case of starting trains, no data of the previous pathing of the train was present in the dataset, therefore no previous delay was known. For this reason, in this study the latest known delay was taken. More detailed data could aid in determining the previous pathing of the rolling stock the train is operated with. In case of a succession, the trains could have been following each other for quite some time before the critical conflict, meaning that the location before the first interaction is ambiguous. In this study the timetable point before the conflict was used to determine the previous delay, but here the delays of both trains are not necessarily independent.

  Both problems can be solved by using another target variable, that is measured instead of approximated. One possible metric could be the time loss directly related to a route conflict (Daamen et al., 2009; Goverde and Meng, 2011). As time losses are determined locally at the signal of occurrence, more detailed data on signal passages and reference running times of unhindered operations have to be collected to enable this approach.
- **Number preceding conflicts** - This variable is computed based on a conflict tree of realisations. As the data collected in this study only entails realisations of train operated by NS passengers and NS international, conflicts including trains of other carriers are excluded. This results in an underestimation of the number of consecutive conflicts, and can therefore influence the significance of this variable. To accurately assess the impact

this variable has on the secondary delay and hindrance percentage, the paths of other carriers' trains should be known. Since this is known in the timetable design, one can opt to base this variable on the conflicts in the timetable, rather than the conflicts in the realisation. However, in case of cancelled trains, order changes, or big delays, the timetable does not provide an accurate representation of reality. Therefore, it is preferred to determine this variable based on realisation data of all carriers.

- **Type headway situation** - This variable was manually constructed for this analysis. To make the analysis more scalable, it is preferable to extract this data from one of the planning applications. To determine what type of headway situation constitutes the conflict, it should be known from what tracks the trains originate and to which track the train is headed and whether the directions of both trains are the same. A succession takes place when two train movements successively take place in the same direction, and both trains have the same entry or exit track. A conflict is called a crossing if two train movements to or from a timetable point travel on routes that share partially common infrastructure and do not have the same entry or exit track. A crossing can take place in the same or opposite direction. The data collected in this study only contained the track of the train movement at the timetable point, but not the entry or exit track.

- **Direction** - This variable was manually constructed for this analysis. To make the analysis more scalable, it is preferable to extract this data from one of the planning applications. Although, this variable did not appear to be a significant predictor, it could turn out to be when analysing a larger variety of headway situations. To determine whether the two trains are moving in the same direction, it should be assessed what adjacent timetable points of the location of the conflict are located at the same side of the timetable point. Then if either the previous or next timetable point of both trains are in the same direction, it can be concluded that the trains are heading in the same direction. This still requires some manual construction of a set of adjacent timetable points that are at the same side of the timetable point.

  Another option would be to analyse the time-distance diagrams of the timetable. When the lines of two conflicting trains are somewhat parallel, one can conclude that they run in the same direction.

- **Event pair of the conflict** - The event of both trains contributing to the conflict was manually constructed for this analysis. To make the analysis more scalable, it is preferable to extract this data from one of the planning applications. It could be assumed that the events of both trains that are scheduled the tightest are responsible for the conflict.

- **Technical minimum headway** - This variable is determined based on the scheduled times of both trains, the planning norm for buffer time and the overlap in extended blocking times in the timetable. This approach does not take into account the length of the train, since this is determined in a later planning stage. However, the length of the train does very much influence the section occupation time, and therefore the release time of the block. Thus, trains with different lengths have different blocking times and thus a different buffer.

6.3. Contributions to theory

The results from this thesis contribute to the theoretical understanding of delay propagation in railways.The contributions are summarised by reflecting on the identified knowledge gaps.

Research gap 1: It is unknown to what extent specific timetable characteristics (e.g. scheduled buffer time, scheduled running time supplement, train characteristics, headway situation, location) influence delay propagation.

This research has identified the impact timetable characteristics have on secondary delay and hindrance percentage. A general finding is that secondary delay is a multidimensional construct, as many factors were identified that affect the amount of secondary delay and the chance of hindrance a succeeding train experiences from a preceding train. It could be stated that secondary delay is not an intrinsic property related to just the conflict, but rather, it is an emergent property of the headway situation as a whole, including train and route characteristics. Consequently, various conflicts show a range of different behaviours, some of which are more prone to hindrance than others. This confirms the idea that buffer times should be tailored to the specific headway situation in order to reduce secondary delay.

In contrast with the existing body of research, the prediction model was based solely on known factors or decision variables in the timetable design. This eases the applicability to timetable design in practice.

Research gap 2: The effective allocation of buffer times to alleviate delay propagation based on historical realisation data has not been widely addressed in the tactical planning stage of railway timetables where an initial timetable is constructed.

The effective allocation of buffer times in the tactical planning stage has been addressed in several studies in the form of ex-ante robustness measures (Abid et al., 2017; Andersson et al., 2013; Carey and Kwieciński, 1994; Kroon et al., 2008; Şahin, 2017; Vromans et al., 2006). The effective allocation of buffer times based on historical data to be used in the tactical planning stage is less studied in literature. Weeda (2005) introduces a joint design standard for running times, dwell time and headway times based on historical data, however it can only be applied to headway situations of which realisations are known. To the best of the author's knowledge, this thesis marks the first time that buffer times were effectively allocated to a large and heterogeneous set of headway conflicts for situations where not all combinations of timetable characteristics are determined yet.

Research gap 3: Planning rules for buffer time based on historical realisation data have not been widely explored in literature.

Despite the unique characteristics of each conflict, it is shown that the regression models on mean secondary delay and hindrance percentage can provide clues about where to focus effort and resources to reduce these aspects of delay propagation. Information from the realisation data should be complemented with timetable data to enable the comparison between scheduled and reported operations. Doing so enables generation of planning rules that are more flexible than current norms for buffer times.

6.4. Chapter conclusion

This chapter provided a reflection of the research, highlighting the limitations as well as the contributions to railway delay propagation theory by reflecting on the identified knowledge gaps. Using the information presented in this chapter, the fifth sub-question is answered.

Sub-question 5: Which insights do the data-driven model(s) and buffer time planning rules bring that may help practitioners effectively allocate buffer times in the tactical timetable design phase?

The results showed that the structural delay propagation between two trains at different critical conflicts is quite heterogeneous. This confirms the vision that planning rules for buffer time should be generated based on the specific characteristics of the headway situation. With the use of regression models, the factors that influence structural delay propagation aspects can be exposed quite accurately. In particular, factors related to the route and train series of the trains contributing to the critical conflict appeared to be influential in explaining and predicting mean secondary delay and the probability of hindrance.

The planning rules emerging from the regression models can be integrated in the timetable design process of Nederlandse Spoorwegen as (i) the size of the wake added in the simulation; or (ii) an indication of when slightly hindered headways are acceptable. To reduce the workload for planners further, it would be favourable that the needed scheduled buffer for a specific conflict is automatically shown in the planning application. The standardisation could help streamline data-driven timetable design.

7 Conclusions

Based on the findings from the literature review, practical background and statistical analysis, this chapter presents the conclusions, practical recommendations and future research directions of this research. Section 7.1 presents the conclusions by recapitulating the answers to the sub-questions and answering the main research question. Next, Section 7.2 presents recommendations with regard to design practices and data processing at NS. Finally, future scientific research directions are identified in Section 7.3.

7.1. Conclusions on research questions

The answers to the five sub-questions defined in Chapter 1 were already provided in the last section of each chapter from Chapter 2 onwards. This section briefly restates the answers to sub-questions after which a general answer to the main research question is given.

Sub-question 1: What can be learned from previous quantitative, data-driven approaches for scheduling time allowances in railway timetable design?

Research on scheduling time allowances is strongly related to robustness, aiming to achieve a high effectiveness and utilisation rate of time allowances such that the timetable can withstand design errors, parameter variations, and changing operational conditions. Robustness approaches that consider buffer times aim for example to (i) reduce the total delay; (ii) reduce delay propagation; (iii) diminish the needed real-time dispatching action; and (iv) enhance punctuality.

Metrics that facilitate comparisons between different timetables in terms of robustness have been introduced. In most cases the metric focuses on one type of time allowance, making it straightforward to propose timetable improvements. It appears that metrics addressing buffer time predominantly focus on ex-ante situations and headways as a whole, thus not on buffer times specifically. Two studies were found where the evaluation is done ex-post in a data-driven fashion, based on how effective running time supplements and dwell time supplements have been in reducing delays. It appears hard to define ex-post metrics of buffer time effectiveness, since the focus of buffer time is on avoiding and decreasing secondary delays rather than reducing existing delays. The amount of secondary delay that has been avoided by insertion of buffer time is challenging to determine. However, measuring propagated delay can provide an indication of the effectiveness of buffer time.

The reviewed articles suggest that the effectiveness of time allowances to facilitate delay recovery is highly dependent on the size of the delay. Regarding the data-driven approaches in the reviewed articles, the focus rarely lies on delay propagation and buffer times.

Sub-question 2: What is the current state of the practice regarding planning norms for buffer times and regarding data-driven timetable design approaches in the Netherlands?

Planning rules for buffer times are imposed by ProRail to ensure safe railway operations. Currently, 60 seconds buffer time between two trains is set as a norm, but in some cases less buffer time is accepted. In Donna, even, a buffer time of merely 30 seconds is displayed to facilitate flexibility. While slightly hindered headways are beneficial for capacity reasons, its consequence on delay propagation is less often considered explicitly. Headway situations are mainly investigated individually and manually, and are both actively and reactively adapted based on information from realisation data. Although this improves the timetable, it doesn't show structural needed changes to headway situations. Thus, a more structural way of analysing the impact of buffer times on delay propagation is lacking. Overall, realisation data is consulted to identify places of improvement in the timetable, but much could be gained by formalising and intensifying the feedback process.

Sub-question 3: What data-driven model(s) can be used to determine buffer time planning rules in the tactical timetable design phase?

buffer times are implemented to avoid or reduce delay propagation effects. Therefore, the development of the buffer time planning rules should consider the expected delay propagation, which can be represented by the mean secondary delay and probability of hindrance a succeeding train encounters caused by a delay of a preceding train. The observed mean secondary delay and the observed hindrance percentage occurring in a specific critical conflict situation can be appraised from realisation data. In this research the conceptual model of Weeda and Wiggenraad (2006) is utilised to determine these target variables. With the use of a prediction model on both target variables and setting an acceptable threshold for expected delay propagation, buffer time planning rules can be generated. The prediction model is based on the relation the target variables have with various timetable characteristics that function as predictor variables. Although literature identifies various other categories of variables to influence delay propagation, this research only considers timetable characteristics as predictor variables, because these are decision factors in the tactical planning stage of railway timetables where an initial timetable is constructed.

Due to the need for interpretability of the prediction model, a multiple linear regression model is suitable. To automate the data-driven regression model, a model estimation procedure resembling a stepwise variable selection method was introduced. The model estimation procedure in this research differs from commonly used stepwise variable selection methods in (i) adding condition; (ii) elimination condition; (iii) stopping condition; and (iv) a check on multicollinearity.

Sub-question 4: How are buffer time planning rules determined from the data-driven model(s) and how do they compare to the current planning norms for buffer times?

Two regression models were estimated, one on mean secondary delay and one on hindrance percentage. Both models appeared to have the same prediction accuracy of 90.7%. The model on mean secondary delay did not contain the scheduled buffer as a significant predictor variable and thus cannot be used to derive buffer time planning rules from. The model on hindrance percentage did contain the scheduled buffer as a significant predictor variable and thus can be used to derive buffer time planning rules from. The predicted needed scheduled buffer can be calculated with the use of Equation 7.1, where Y refers to the acceptable hindrance percentage chosen by the planner, β_0 corresponds with the constant; and $\beta_1 ... \beta_{n-1}$ and $X_1 ... X_{n-1}$ represent the coefficients and values of the other significant predictor values.

$$X_{scheduled\ buffer} = \frac{Y - \beta_0 - \beta_1 X_1 - \beta_2 X_2 - ... - \beta_{n-1} X_{n-1} + \epsilon}{\beta_{scheduled\ buffer}} \tag{7.1}$$

Compared to the current planning norms for buffer times, the planning rules presented in this chapter provide more flexibility, because the size of the needed scheduled buffer depends on the preferred hindrance percentage. The planning rules presented in this chapter however, do not indicate how the scheduled buffer compares to the mean secondary delay that can be expected. To gain the same mean hindrance percentage as the realisation data, the total amount and mean value of needed scheduled buffer remains unchanged. Yet, the variation in buffer sizes increases from a standard deviation of 26 seconds to a standard deviation of 93 seconds, meaning a 257% rise. Thus, the spread in scheduled buffer time is increased, and moreover the allocation of the scheduled buffer time changes geographically.

Sub-question 5: Which insights do the data-driven model(s) and buffer time planning rules bring that may help practitioners effectively allocate buffer times in the tactical timetable design phase?

The results showed that the structural delay propagation between two trains at different critical conflicts is quite heterogeneous. This confirms the vision that planning rules for buffer time should be generated based on the specific characteristics of the headway situation. With the use of regression models, the factors that influence structural delay propagation aspects can be exposed quite accurately. In particular, factors related to the route and train series of the trains contributing to the critical conflict appeared to be influential in explaining and predicting mean secondary delay and the probability of hindrance.

The planning rules emerging from the regression models can be integrated in the timetable design process of Nederlandse Spoorwegen as (i) the size of the wake added in the simulation; or (ii) an indication of when slightly hindered headways are acceptable. To reduce the workload for planners further, it would be favourable that the needed scheduled buffer for a specific conflict is automatically shown in the planning application. The standardisation could help streamline data-driven timetable design.

Answer to main research question

How can buffer time planning rules focused on reducing delay propagation be extracted from realisation data?

As an answer to the main research question, the general conclusion from this thesis is that an estimation or prediction of the expected delay propagation between two trains in a given situation could be done quite accurately (adjusted R-squared value of both models equals 0.907) and is essential to determining suitable buffer time planning rules, because it provides insight into where buffer time can be placed in order to be most effective. This research has operationalised delay propagation as the mean secondary delay and probability of hindrance a succeeding train encounters caused by a delay of a preceding train. Using the prediction model, the planning rules can be determined by setting a desired value for these metrics. This value could possibly be determined by making a trade-off between delay propagation and capacity based on a chosen objective (e.g. passenger punctuality, infrastructure consumption). Interesting to note is that the mean secondary delay is not significantly impacted by the scheduled buffer, but the probability of hindrance is. This indicates that the delay distributions of both trains have a bigger impact on the size of the secondary delay, while the size of the scheduled buffer determines how often hindrance occurs.

Another explored approach in this study was to ex-post appraise the effectiveness of scheduled buffer times, and to aim to maximise the effectiveness. But it appeared hard to define ex-post metrics of buffer time effectiveness, since the focus of this time allowance is on avoiding and decreasing secondary delays rather than reducing existing delays. The amount of secondary delay that has been avoided by insertion of buffer time is challenging to determine.

7.2. Recommendations for practice

Building on the findings from this thesis research, recommendations are made with regard to design practices and data processing at NS. The recommendations are formulated as incremental steps that can be performed to mature the timetable design process with regard to planning buffer times focused on reducing delay propagation. Identifying intermediate steps to the most ideal usage of the outcomes of this data-driven research is in line with the way data-driven organisations are formed and evaluated according to literature on maturity models (Berndtsson and Svahn, 2020; Davenport, 2018). The provided order is a suggestion that currently seems suitable and feasible, but naturally extra steps can be added or superfluous steps may be skipped.

1. **Use case study conclusions** - Based on the regression model on mean secondary delay (i) Previous timetable point of train 2 being Hfdm or Vkbr; and (ii) Type next timetable point of train 1 being a halt were found to be the most influential factors. The biggest influential factors on hindrance percentage were found to be (i) Next timetable point of train 1 being Shl or Zspl; (ii) Previous timetable point of train 1 being Hg; and (iii) Train 2 being a starting train. As these factors have a large impact on the all-encompassing delay propagation, it would be worthwhile to explicitly consider the predicted scheduled buffer supplements for these categories when scheduling conflicts that contain one of these categories.

2. **Analyse expected delay propagation of scheduled situations** - Based on the found regression formulas in the case study, the expected mean secondary delay and expected hindrance percentage can be appraised for all critical conflicts scheduled in a concept timetable. This can provide an indication to planners what situations require more buffer time and what situations could suffice with less buffer. For example, an ordered list of all critical conflicts can be made based on the expected mean secondary delay or expected hindrance percentage. The planner can do another check on either (i) a predefined number of critical conflicts with the highest values; or (ii) the critical conflicts with values higher than an acceptable threshold.

3. **Setting an acceptable threshold** - Determine an acceptable threshold for mean secondary delay and hindrance percentage per critical conflict, so that the needed buffer to suffice this can be determined. This can differ per conflict, but it is advised to determine an overall maximum acceptable value. The value per conflict can be decided upon by making a trade-off with the available capacity at a certain conflict location. Another more straightforward approach might be to take 30 seconds as the maximum acceptable value for mean secondary delay, since this is currently the norm for acceptable arrival delays at certain stations due to hindered operations caused by small buffer times as discussed in Chapter 3.

4. **Integrate relevant factors in planning applications** - To gain better insights into all factors of the scheduled conflict it is advised to enhance the planning applications with conflict specific details. For instance, for a conflict situation it could be added what the scheduled buffer, direction of the trains and event pair of the conflict are. When this approach has become common practice, it is favourable that the needed scheduled buffer for a specific conflict is automatically shown in the planning application, to reduce the workload for planners. This standardisation could help streamline the data-driven timetable design.

5. **Automatically update regression models with more (new) realisation data** - Regularly updating the regression models results in more accurate parameters of the significant predictors, as the most recent realisation data is based on the most recent timetable (design process). Moreover, basing the analysing on data based on multiple timetables can provide a more holistic perspective on the effects of various ranges of scheduled buffer times.

7.3. Future research directions

With regard to the methodology, the limitations mentioned in Chapter 6 indicate that prediction methods other than linear regression may be more suitable to describe secondary delay. It could be worthwhile to extend the current models with (i) interaction effects (Aguinis and Gottfredson, 2010); (ii) polynomial effects (Ostertagová, 2012); or (iii) robust versions (Huber, 1981). The models themselves may also be extended by adding more predictor variables to reduce the unexplained variance in mean secondary delay and hindrance percentage.

Furthermore, it could be investigated how to determine the critical conflicts, such that multiple critical conflicts can be identified between two trains, in case multiple groups of succeeding conflicts exist which are unrelated to each other.

The last methodological research direction relates to the scope of the analysis. The analysis can be expanded by including freight trains, shunting movements, and trains of other carriers. For freight trains, in particular, other coefficients are expected, because they are often heavier and longer than passenger trains. This implies that they have a slower acceleration and longer braking and thus an unplanned stop has more impact. Although Pachl (2014) argue that a so-called buffer path after a certain number of trains might be more appropriate for such trains as they often run out of their schedule. Moreover, it might be useful to gain insights into what factors may cause order changes and to add factors that can explain temporal variation.

With respect to scientific research into planning rules for buffer times, various contributions can still be made. For starters, future research may include concrete simulations of different timetables with varying buffer times, to obtain a better estimate of the effect of tweaking these planning norms on robustness and capacity utilisation. An important question for applying such more specific planning rules in practice, is which timetable structures would be made possible or impossible, and whether the capacity on busy lines will be increased, stay equal, or be decreased.

Secondly, assessing the marginal utility and costs of adding an extra unit of buffer time could probably be explored in a simulation context. Based on this, the maximum performance of the headway situation could be determined based on the economical theory that marginal utility equals marginal costs. Marginal utility of an extra unit of buffer time has been explored by Jovanović et al. (2017).

Lastly, among practitioners implicit and tacit knowledge exists about when scheduling tighter buffer times than the planning rules does not result in (too much) secondary delay. This implicit knowledge could potentially be caught in a decision model, where the trade-offs of scheduling buffer times become concrete.

Bibliography

Abid, M. M., Khan, M. B., & Jiang, X. (2017). Ex-ante robustness measures for single track train timetables. *Tehnicki Vjesnik*, *24*(5), 1533–1542. https://doi.org/10.17559/TV-20150623101554

Aguinis, H., & Gottfredson, R. K. (2010). Best-practice recommendations for estimating interaction effects using moderated multiple regression. *Journal of Organizational Behaviour*, *31*, 776–786. https://doi.org/10.1002/job.686

Akinwande, M. O., Dikko, H. G., & Samson, A. (2015). Variance Inflation Factor: As a Condition for the Inclusion of Suppressor Variable(s) in Regression Analysis. *Open Journal of Statistics*, *05*(07), 754–767. https://doi.org/10.4236/ojs.2015.57075

Aljandali, A. (2017). *Multivariate Methods and Forecasting with IBM® SPSS® Statistics*. Springer. https://doi.org/10.1007/978-3-319-56481-4

Alkharusi, H. (2012). Categorical Variables in Regression Analysis: A Comparison of Dummy and Effect Coding. *International Journal of Education*, *4*(2), 202. https://doi.org/10.5296/ije.v4i2.1962

Allison, P. D. (1999). *Multiple Regression: A Primer*. Pine Forge Press.

Andersson, E. V., Peterson, A., & Törnquist Krasemann, J. (2013). Quantifying railway timetable robustness in critical points. *Journal of Rail Transport Planning and Management*, *3*(3), 95–110. https://doi.org/10.1016/j.jrtpm.2013.12.002

Berndtsson, M., & Svahn, T. (2020). Strategies for Scaling Analytics: A Nontechnical Perspective. *Business Intelligence Journal*, *25*(1), 53. http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-18504

Berry, W., & Feldman, S. (1985). *Multiple Regression in Practice*. Sage. https://doi.org/10.4135/9781412985208

Bešinović, N., Goverde, R. M. P., Quaglietta, E., & Roberti, R. (2016). An integrated micro-macro approach to robust railway timetabling. *Transportation Research Part B: Methodological*, *87*, 14–32. https://doi.org/10.1016/j.trb.2016.02.004

Burggraeve, S., & Vansteenwegen, P. (2017). Optimization of supplements and buffer times in passenger robust timetabling. *Journal of Rail Transport Planning and Management*, *7*(3), 171–186. https://doi.org/10.1016/j.jrtpm.2017.08.004

Cacchiani, V., & Toth, P. (2012). Nominal and robust train timetabling problems. *European Journal of Operational Research*, *219*(3), 727–737. https://doi.org/10.1016/j.ejor.2011.11.003

Caprara, A., Fischetti, M., & Toth, P. (2002). Modeling and solving the train timetabling problem. *Operations Research*, *50*(5). https://doi.org/10.1287/opre.50.5.851.362

Carey, M. (1999). Ex ante heuristic measures of schedule reliability. *Transportation Research Part B: Methodological*, *33*(7), 473–494. https://doi.org/10.1016/S0191-2615(99)00002-8

Carey, M., & Kwieciński, A. (1994). Stochastic approximation to the effects of headways on knock-on delays of trains. *Transportation Research Part B*, *28*(4), 251–267. https://doi.org/10.1016/0191-2615(94)90001-9

Cerreto, F., & Jonasson, M. H. (2019). Fact-checking of Timetabling Principles: a Case Study on the Relationship Between Planned Headways and Delays. *Proceedings of the 8th International Conference on Railway Operations Modelling and Analysis - RailNorrköping 2019*, 312–323.

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*, *18*(3), 229–257. https://doi.org/10.1007/s10878-009-9247-4

Corman, F., & Kecman, P. (2018). Stochastic prediction of train delays in real-time using Bayesian networks. *Transportation Research Part C: Emerging Technologies*, *95*, 599–615. https://doi.org/10.1016/j.trc.2018.08.003

Daamen, W., Goverde, R. M. P., & Hansen, I. A. (2009). Non-discriminatory automatic registration of knock-on train delays. *Networks and Spatial Economics*, *9*(1), 47–61. https://doi.org/10.1007/s11067-008-9087-2

D'Acierno, L., Botte, M., Gallo, M., & Montella, B. (2018). Defining Reserve Times for Metro Systems: An Analytical Approach. *Journal of Advanced Transportation*, *2018*. https://doi.org/10.1155/2018/5983250

Davenport, T. H. (2018). From analytics to artificial intelligence. *Journal of Business Analytics*, *1*(2), 73–80. https://doi.org/10.1080/2573234X.2018.1543535

Dietzenbacher, D. S. (2021). *Scrutinizing headway buffer norms for railway timetabling* (tech. rep.). Eindhoven University of Technology.

Fischetti, M., & Monaci, M. (2009). Light robustness. In R. K. Ahuja, R. H. Möhring, & C. D. Zaroliagis (Eds.), *Robust and online large-scale optimization, lecture notes in computer science* (pp. 61–84). Springer-Verlag. https://doi.org/10.1007/978-3-642-05465-5{_}3

Galvão, R. K. H., Araújo, M. C. U., Fragoso, W. D., Silva, E. C., José, G. E., Soares, S. F. C., & Paiva, H. M. (2008). A variable elimination method to improve the parsimony of MLR models using the successive projections algorithm. *Chemometrics and Intelligent Laboratory Systems*, *92*(1), 83–91. https://doi.org/10.1016/j.chemolab.2007.12.004

Gerring, J. (2004). What is a case study and what is it good for? *American Political Science Review*, *98*(2), 341–354. https://doi.org/10.1017/S0003055404001182

Ghani, I. M. M., & Ahmad, S. (2010). Stepwise multiple regression method to forecast fish landing. *Procedia - Social and Behavioral Sciences*, *8*, 549–554. https://doi.org/10.1016/J.SBSPRO.2010.12.076

Ghofrani, F., He, Q., Goverde, R. M. P., & Liu, X. (2018). Recent applications of big data analytics in railway transportation systems: A survey. *Transportation Research Part C: Emerging Technologies*, *90*(September 2017), 226–246. https://doi.org/10.1016/j.trc.2018.03.010

Gibson, S., Cooper, G., & Ball, B. (2002). Developments in transport policy: The evolution of capacity charges on the UK rail network. *Journal of Transport Economics and Policy*, *36*(2), 341–354.

Gorman, M. F. (2009). Statistical estimation of railroad congestion delay. *Transportation Research Part E: Logistics and Transportation Review*, *45*(3), 446–456. https://doi.org/10.1016/j.tre.2008.08.004

Goverde, R. M. P. (2014). Timetable Stability Analysis. In I. A. Hansen & J. Pachl (Eds.), *Railway timetabling & operations* (Second, pp. 118–134). DVV Media Group GmbH.

Goverde, R. M. P., Bešinović, N., Binder, A., Cacchiani, V., Quaglietta, E., Roberti, R., & Toth, P. (2016). A three-level framework for performance-based railway timetabling. *Transportation Research Part C: Emerging Technologies*, *67*, 62–83. https://doi.org/10.1016/j.trc.2016.02.004

Goverde, R. M. P., & Hansen, I. A. (2013). Performance indicators for railway timetables. *IEEE ICIRT 2013 - Proceedings: IEEE International Conference on Intelligent Rail Transportation*, 301–306. https://doi.org/10.1109/ICIRT.2013.6696312

Goverde, R. M. P., & Meng, L. (2011). Advanced monitoring and management information of railway operations. *Journal of Rail Transport Planning and Management*, *1*(2), 69–79. https://doi.org/10.1016/j.jrtpm.2012.05.001

Graham, M. H. (2003). Confronting multicollinearity in ecological multiple regression. *Ecology*, *84*(11), 2809–2815. https://doi.org/10.1890/02-3114

Guo, J., Meng, L., Kecman, P., & Corman, F. (2015). Modeling delay relations based on mining historical train monitoring data: a Chinese railway case. *Proceedings of the 6th International Conference on Railway Operations Modelling and Analysis*.

Hauck, F., & Kliewer, N. (2019). A Data-Driven Optimization Approach to Improve Railway Punctuality. In B. Fortz & M. Labbé (Eds.), *Operations research proceedings* (pp. 539–545). Springer, Cham. https://doi.org/10.1007/978-3-030-18500-8{_}67

Ho, T. W., Chen, T. C., & Chou, C. C. (2012). Integrating of optimization and data mining techniques for high-speed train timetable design considering disturbances. *2012 Proceedings of the 29th International Symposium of Automation and Robotics in Construction, ISARC 2012*. https://doi.org/10.4017/gt.2012.11.02.385.00

Huang, P., Wen, C., Peng, Q., Lessan, J., Fu, L., & Jiang, C. (2019). A data-driven time supplements allocation model for train operations on high-speed railways. *International Journal of Rail Transportation*, *7*(2), 140–157. https://doi.org/10.1080/23248378.2018.1520613

Huber, P. J. (1981). *Robust Statistics*. John Wiley & Sons, Ltd. https://doi.org/10.1002/0471725250

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment). *Computing in Science and Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Ibarra-Rojas, O. J., Delgado, F., Giesen, R., & Muñoz, J. C. (2015). Planning, operation, and control of bus transport systems: A literature review. *Transportation Research Part B: Methodological*, *77*, 38–75. https://doi.org/10.1016/j.trb.2015.03.002

Independent Regulators' Group - Rail. (2021). *Ninth Annual Market Monitoring Working Document* (tech. rep.).

Jensen, L. W., Landex, A., Nielsen, O. A., Kroon, L. G., & Schmidt, M. (2017). Strategic assessment of capacity consumption in railway networks: Framework and model. *Transportation Research Part C: Emerging Technologies*, *74*, 126–149. https://doi.org/10.1016/j.trc.2016.10.013

Jin, B., Feng, X., Wang, Q., Wang, X., & Liu, C. (2019). Improving Timetable Robustness through Optimal Distribution of Runtime Supplement. *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019*, 2803–2808. https://doi.org/10.1109/ITSC.2019.8917477

Jovanović, P., Kecman, P., Bojović, N., & Mandić, D. (2017). Optimal allocation of buffer times to increase train schedule robustness. *European Journal of Operational Research*, *256*(1), 44–54. https://doi.org/10.1016/j.ejor.2016.05.013

Khoshniyat, F., & Peterson, A. (2017). Improving train service reliability by applying an effective timetable robustness strategy. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, *21*(6), 525–543. https://doi.org/10.1080/15472450.2017.1326114

Kroon, L. G., Maróti, G., Helmrich, M. R., Vromans, M. J., & Dekker, R. (2008). Stochastic improvement of cyclic railway timetables. *Transportation Research Part B: Methodological*, *42*(6), 553–570. https://doi.org/10.1016/j.trb.2007.11.002

Lee, W. H., Yen, L. H., & Chou, C. M. (2016). A delay root cause discovery and timetable adjustment model for enhancing the punctuality of railway services. *Transportation Research Part C: Emerging Technologies*, *73*, 49–64. https://doi.org/10.1016/j.trc.2016.10.009

Lee, Y., Lu, L. S., Wu, M. L., & Lin, D. Y. (2017). Balance of efficiency and robustness in passenger railway timetables. *Transportation Research Part B: Methodological*, *97*, 142–156. https://doi.org/10.1016/j.trb.2016.12.004

Li, W., Peng, Q., Wen, C., Wang, P., Lessan, J., & Xu, X. (2020). Joint optimization of delay-recovery and energy-saving in a metro system: A case study from China. *Energy*, *202*. https://doi.org/10.1016/j.energy.2020.117699

Liebchen, C., Lübbecke, M., Möhring, R., & Stiller, S. (2009). The Concept of Recoverable Robustness, Linear Programming Recovery, and Railway Applications. In R. H. Ahuja, R. H. Möhring, & C. D. Zaroliagis (Eds.), *Robust and online large-scale optimization, lecture notes in computer science* (pp. 1–27). Springer-Verlag. https://doi.org/10.1007/978-3-642-05465-5{_}1

Liebchen, C., Schachtebeck, M., Schöbel, A., Stiller, S., & Prigge, A. (2010). Computing delay resistant railway timetables. *Computers & Operations Research*, *37*(5), 857–868. https://doi.org/10.1016/J.COR.2009.03.022

Lu, C., Tang, J., Zhou, L., Yue, Y., & Huang, Z. (2017). Improving recovery-to-optimality robustness through efficiency-balanced design of timetable structure. *Transportation Research Part C: Emerging Technologies*, *85*, 184–210. https://doi.org/10.1016/j.trc.2017.09.015

Lusby, R. M., Larsen, J., & Bull, S. (2018). A survey on robustness in railway planning. https://doi.org/10.1016/j.ejor.2017.07.044

Lusby, R. M., Larsen, J., Ehrgott, M., & Ryan, D. (2011). Railway track allocation: Models and methods. *OR Spectrum*, *33*(4), 843–883. https://doi.org/10.1007/s00291-009-0189-0

McKinney, W. (2010). Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*, 56–61. https://doi.org/10.25080/majora-92bf1922-00a

Meng, L., Muneeb Abid, M., Jiang, X., Khattak, A., & Babar Khan, M. (2019). Increasing Robustness by Reallocating the Margins in the Timetable. *Journal of Advanced Transportation*, *2019*. https://doi.org/10.1155/2019/1382394

Middelkoop, A. D. (2010). Headway generation with ROBERTO. *WIT Transactions on the Built Environment*, *114*, 431–439. https://doi.org/10.2495/CR100401

Midi, H., Sarkar, S. K., & Rana, S. (2010). Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics*, *13*(3), 253–267. https://doi.org/10.1080/09720502.2010.10700699

Minbashi, N., Palmqvist, C. W., Bohlin, M., & Kordnejad, B. (2021). Statistical analysis of departure deviations from shunting yards: Case study from Swedish railways. *Journal of Rail Transport Planning and Management*, *18*. https://doi.org/10.1016/j.jrtpm.2021.100248

Ministerie van Verkeer en Waterstaat. (2010). *Priority decision High-Frequency Rail Transport Programme* (tech. rep.). Ministerie van Verkeer en Waterstaat. https://silo.tips/download/priority-decision-high-frequency-rail-transport-programme

Nederlandse Spoorwegen. (2018). Werkwijze logistieke uitwerkingen v10.

Nederlandse Spoorwegen. (2019). Werkwijze dienstregelingstudies IPO.

Nederlandse Spoorwegen. (2020). International Benchmark 2014-2018. https://open.overheid.nl/repository/ronl-5cff90f8-dc43-464f-9e73-e4058bce62fc/1/pdf/bijlage-8-internationale-benchmark-ns.pdf

Osborne, J. (2002). Notes on the use of data transformations. *Practical Assessment, Research, and Evaluation*, *8*. https://doi.org/10.7275/4vng-5608

Ostertagová, E. (2012). Modelling using polynomial regression. *Procedia Engineering*, *48*, 500–506. https://doi.org/10.1016/j.proeng.2012.09.545

Pachl, J. (2014). Timetable Design Principles. In I. A. Hansen & J. Pachl (Eds.), *Railway timetabling & operations* (Second, pp. 13–46). DVV Media Group GmbH.

Palmqvist, C. W., Olsson, N. O. E., & Hiselius, L. W. (2017a). An Empirical Study of timetable Strategies and Their Effects on Punctuality. *7th International Conference on Railway Operations Modelling and Analysis*. https://doi.org/10.1109/HICSS.2015.400

Palmqvist, C. W., Olsson, N. O. E., & Hiselius, L. W. (2017b). Some influencing factors for passenger train punctuality in Sweden. *International Journal of Prognostics and Health Management*, *8*(Special Issue 7).

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourh, V., VanderPlas, J., Passos, A., Cournapeau, D., Brucher, M., Perror, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Planting, T. J. (2016). *Ontwerpmethoden van Dienstregelingen* (tech. rep.). Delft University of Technology.

ProRail. (2015). *Business Information Document - Begrippenlijst Spoorse Begrippen* (tech. rep.). https://nsdigitaal.sharepoint.com/teams/Logistiek/Basis%20documenten/Forms/Alle%20documenten.aspx?id=%2Fteams%2FLogistiek%2FBasis%20documenten%2FSpoorse%20Begrippen.pdf&parent=%2Fteams%2FLogistiek%2FBasis%20documenten

ProRail. (2020). *Startdocument Voorbereidingsfase dienstregeling 2022* (tech. rep.). https://nsdigitaal.sharepoint.com/teams/Logistiek/KritischeDocumenten/Plannormen%20dienstregeling%202022%20V1.0%202%20oktober%202020.pdf

ProRail. (2022). Basiskaart. https://maps.prorail.nl/portal/home/webmap/viewer.html?url=https%3A%2F%2Fmaps.prorail.nl%2Farcgis%2Frest%2Fservices%2FProRail_basiskaart%2FFeatureServer&source=sd

Radtke, A. (2014). Infrastructure modelling. In I. A. Hansen & J. Pachl (Eds.), *Railway timetabling & operations* (Second, pp. 47–63). DVV Media Group GmbH.

Restel, F., Wolniewicz, Ł., & Mikulčić, M. (2021). Method for Designing Robust and Energy Efficient Railway Schedules. *Energies*, *14*(24), 8248. https://doi.org/10.3390/en14248248

Şahin, I. (2017). Markov chain model for delay distribution in train schedules: Assessing the effectiveness of time allowances. *Journal of Rail Transport Planning and Management*, *7*(3), 101–113. https://doi.org/10.1016/j.jrtpm.2017.08.006

Salido, M. A., Barber, F., & Ingolotti, L. (2008). Robustness in railway transportation scheduling. *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)*, 2880–2885. https://doi.org/10.1109/WCICA.2008.4594481

Scheepmaker, G. M., & Goverde, R. M. P. (2015). The interplay between energy-efficient train control and scheduled running time supplements. *Journal of Rail Transport Planning and Management*, *5*(4), 225–239. https://doi.org/10.1016/j.jrtpm.2015.10.003

Scheepmaker, G. M., & Goverde, R. M. P. (2021). Multi-objective railway timetabling including energy-efficient train trajectory optimization. *European Journal of Transport and Infrastructure Research*, *21*(4), 1–42. https://doi.org/10.18757/ejtir.2021.21.4.5453

Schöbel, A., & Kratz, A. (2009). A bicriteria approach for robust timetabling. In R. K. Ahuja, R. H. Möhring, & C. D. Zaroliagis (Eds.), *Robust and online large-scale optimization, lecture notes in computer science* (pp. 119–144). Springer-Verlag. https://doi.org/10.1007/978-3-642-05465-5{_}5

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*, 92–98. http://conference.scipy.org/proceedings/scipy2010/pdfs/seabold.pdf

Serafini, P., & Ukovich, W. (1989). A Mathematical Model for Periodic Scheduling Problems. *SIAM Journal on Discrete Mathematics*, *2*(4), 550–581. https://doi.org/10.1137/0402049

SporenplanOnline. (2022). Emplacementstekeningen NS.

Theil, H. (1971). Principles of econometrics. https://archive.org/details/principlesofecon0000thei/page/n5/mode/2up

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R., Jones, E., Kern, R., Larson, E., ... Vázquez-Baeza, Y. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/S41592-019-0686-2

Vromans, M. J. (2005). *Reliability of Railway Systems* (Doctoral dissertation). Erasmus University Rotterdam. Rotterdam.

Vromans, M. J., Dekker, R., & Kroon, L. G. (2006). Reliability and heterogeneity of railway services. *European Journal of Operational Research*, *172*(2), 647–665. https://doi.org/10.1016/j.ejor.2004.10.010

Waskom, M. L. (2021). Seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. https://doi.org/10.21105/JOSS.03021

Weeda, V. A. (2005). *Plannen met speling, spelen met de planning* (tech. rep.). Delft University of Technology. Delft.

Weeda, V. A., & Wiggenraad, P. B. L. (2006). Joint design standard for running times, dwell times and headway times. *Computers in Railways X*, *1*, 611–620. https://doi.org/10.2495/CR060601

Wen, C., Huang, P., Li, Z., Lessan, J., Fu, L., Jiang, C., & Xu, X. (2019). Train dispatching management with data-driven approaches: A comprehensive review and appraisal. *IEEE Access*, *7*, 114547–114571. https://doi.org/10.1109/ACCESS.2019.2935106

Wen, C., Lessan, J., Fu, L., Huang, P., & Jiang, C. (2017a). Data-driven models for predicting delay recovery in high-speed rail. *2017 4th International Conference on Transportation Information and Safety, ICTIS 2017 - Proceedings*, 144–151. https://doi.org/10.1109/ICTIS.2017.8047758

Wen, C., Li, Z., Lessan, J., Fu, L., Huang, P., & Jiang, C. (2017b). Statistical investigation on train primary delay based on real records: evidence from Wuhan-Guangzhou HSR. *International Journal of Rail Transportation*, *5*(3), 170–189. https://doi.org/10.1080/23248378.2017.1307144

Wiegand, R. E. (2010). Performance of using multiple stepwise algorithms for variable selection. *Statistics in Medicine*, *29*(15), 1647–1659. https://doi.org/10.1002/sim.3943

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/2601248.2601268

Yang, X., Hou, Y., & Li, L. (2020). Buffer time allocation according to train delay expectation at stations. *International Journal of Rail Transportation*, *8*(3), 249–263. https://doi.org/10.1080/23248378.2019.1655805

Yuan, J., & Hansen, I. A. (2007). Optimizing capacity utilization of stations by estimating knock-on train delays. *Transportation Research Part B: Methodological*, *41*(2), 202–217. https://doi.org/10.1016/j.trb.2006.02.004

Yuan, J., & Hansen, I. A. (2008). Closed form expressions of optimal buffer times between scheduled trains at railway bottlenecks. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 675–680. https://doi.org/10.1109/ITSC.2008.4732539

Zieger, S., Weik, N., & Nießen, N. (2018). The influence of buffer time distributions in delay propagation modelling of railway networks. *Journal of Rail Transport Planning and Management*, *8*, 220–232. https://doi.org/10.1016/j.jrtpm.2018.09.001

A
Scientific paper

This appendix presents the scientific paper.

A data-driven approach for generation of tactical planning rules regarding buffer time in initial railway timetables

S. E. Bouman[a,b,*]

[a]MSc. Student Transport, Infrastructure & Logistics, Delft University of Technology, Delft, The Netherlands
[b]Nederlandse Spoorwegen, Netwerkontwikkeling & Ontwerp, Utrecht, The Netherlands

ARTICLE INFO

Keywords:
Railways
Timetable design
Buffer time
Data-driven
Multiple linear regression
Delay propagation
Secondary delay
Hindrance percentage

ABSTRACT

Despite advanced communication, monitoring, and control facilities, train operations are still subject to uncertainties that can disturb train services, cause delay to multiple trains, and propagate through the network. One option is to mitigate delay propagation in the timetable design by adding buffer time to the minimum difference between the time two successive train of either direction enter a section. It is still common practice to design buffer times based on a deterministic value, decreasing operational capacity and requiring large amount of manual checking by planners. Existing approaches to effectively allocate buffer time in timetables lack flexibility and require an initial timetable. In this paper, a data-driven approach for determination of buffer time planning rules suitable for usage in an initial timetable is presented. These planning rules are not necessarily generic, but rather depend on timetable characteristic. Two metrics that describe delay propagation, mean secondary delay and hindrance percentage, are extracted from literature and predicted in a regression analysis with the use of timetable characteristics related to headway situations of two succeeding trains. The results of the regression analysis on a case study of the Dutch railway network between Haarlem, Leiden Centraal and Schiphol Airport are used to determine the amount of scheduled buffer time that would ensure a certain amount of hindrance percentage given a specific headway situation. The results show that the mean secondary delay and hindrance percentage for various headway situations can both be predicted with an accuracy of 90.7% based on timetable characteristics and is quite heterogeneous. Mean secondary delay appeared not significantly impacted by the scheduled buffer time, contrary to hindrance percentage which is significantly influenced by the scheduled buffer time.

1. Introduction

The Dutch railway network is one of the busiest in Europe, with as many as 147 trains per day per route kilometer in 2019 (Independent Regulators' Group - Rail, 2021), and its network usage intensity will become even higher with the frequency increase of train services under the High-Frequency Rail Transport Programme (Ministerie van Verkeer en Waterstaat, 2010). This increase in frequency requires more efficient use of the current infrastructure as well as reliable train operations. Under regular conditions, the timetable is adhered to apart from minor variations in the train services due to differences in driving behaviour and stochasticity in process times. Despite advanced communication, monitoring, and control facilities, train operations are still subject to uncertainties that can disturb train services, cause delay to multiple trains, and propagate through the network.

Delay can be dealt with in different stages of the railway transport management. One option is to mitigate delay already in the timetable design. To provide a high quality of services, time allowances are added to the shortest possible running time between scheduled stops (i.e. running time supplement) and minimum line headway, which corresponds to the minimum difference between the times two successive trains of either direction enter a section (Pachl, 2014). The former aims to facilitate a recovery or reduction of a delay within the timetable of one trains, whilst the latter enables prevention and reduction of delay propagation between trains.

It is still common practice to design buffer times based on a deterministic value in combination with personal experience and tacit knowledge. Although easy to apply, this method (i) often decreases operational capacity in heavily utilised networks; and (ii) requires a large amount of manual checking in case of smaller scheduled buffer times. Both aspects affect the extent to which buffer times are effective in reducing delay propagation. This is supported by findings from Zieger et al. (2018); Carey and Kwieciński (1994); Yuan and Hansen (2007, 2008), which show that the realised

*Corresponding author: susan@boumanmail.nl

headway and the buffer time distribution have a significant influence on the expected knock-on delays and capacity. Therefore, there is a need for buffer time allocation techniques that achieve a high effectiveness of buffer times.

When creating an initial timetable from scratch, the literature predominantly considers minimum buffer time, or minimum headway time including an explicit minimum buffer time as a further simplification, as a deterministically given input value (Cacchiani and Toth, 2012; Caprara et al., 2002; Scheepmaker and Goverde, 2021; Serafini and Ukovich, 1989). Most works consider this to be a hard constraint, while in practice a smaller buffer is only rejected when (i) an unplanned stop would occur, or (ii) an arrival delay bigger than 30 seconds would occur (ProRail, 2020). Thus, the optimisation models in the literature lack flexibility regarding the trade-off between headway time and running time. One way of introducing flexibility to these type of models is to make the minimum buffer time input variable. Therefore, an approach that can differentiate buffer time is required.

With regard to allocation of running time supplements and buffer times to reduce delay propagation and to improve timetable robustness, several approaches have been introduced in literature (Hauck and Kliewer, 2019; Huang et al., 2019; Jin et al., 2019; Khoshniyat and Peterson, 2017; Kroon et al., 2008; Yang et al., 2020; Jovanović et al., 2017). The main limitation of these approaches is that they are only able to adjust an existing timetable by reallocating the current running time supplements and buffer times. Thus, a predefined train order is assumed, the total amount of scheduled buffer remains unchanged and an initial timetable is required. This raises the question how more suitable buffer time rules can be designed that can be applied to the tactical timetable design stage when no initial timetable and train order is present.

The rising availability of data has enabled analysis of empirical delay distribution, delay propagation and headway variation (Corman and Kecman, 2018; Zieger et al., 2018; Dietzenbacher, 2021). The findings have been applied to predict train delays and reallocate buffer times, but do not explicitly specify how buffer times should be allocated to alleviate delay propagation most effectively. Furthermore, some approaches have explicitly aimed to improve the effectiveness of running time supplements and dwell time supplements to recover from delays based on historical data (Huang et al., 2019; Yang et al., 2020). Thus, the proposed methods mainly focus on the reduction of primary delays and are unable to generalise the findings, such that it can be applied to create initial timetables where no realisation data is available yet. The latter is crucial to cope with delay propagation effects already while constructing an initial timetable.

In this paper, a data-driven approach for determination of buffer time planning rules suitable for usage in an initial timetable is presented. These planning rules are not necessarily generic, but depend on timetable characteristics that have been found by means of a stepwise multiple regression model to influence delay propagation. Two metrics that describe delay propagation are extracted from literature and explained in regression analysis with the use of timetable characteristics related to headway situations of two succeeding trains. The results of the regression analysis are used to determine the amount of scheduled buffer time that would ensure a certain amount of hindrance percentage given a specific headway situation. The proposed approach is data-driven, as it uses historical traffic realisation data, combined with tightly scheduled headway situations in the timetable. This research differs from existing works regarding generating situation specific buffer times by providing an approach that (i) does not require an initial timetable; (ii) predicts delay propagation based on factors known in the timetable design; and (iii) specifies planning rules that can be applied to new situations.

This study contributes to the field of railway timetable design from the following perspectives. First of all, a new data-driven approach to allocate buffer times in an initial timetable is presented. The approach is generic and can be applied to a larger scope and to railway networks in other countries. Additionally, insights are generated into what timetable characteristics contribute to delay propagation in tightly scheduled headway situations.

The remainder of this paper is structured as follows: Section 2 starts with a literature review on timetable design concepts and methods, data-driven delay models and buffer time scheduling and ends with an overview of the scientific gaps this research fills. Section 3 follows to explain the approach of the research in more detail, including the framework used to determine buffer time planning rules. The case study and its results are presented in Section 4. Finally, a discussion of the results is provided in Section 5 and conclusions are presented in Section 6.

2. Literature review

This section provides a literature review of the various theories and methods on buffer times in railway timetable design in scientific literature. Subsection 2.1 presents nominal, robust and multi-level timetable design methods.

Subsection 2.2 discusses previous data-driven methods on buffer time scheduling and delays. Finally, existing research gaps are highlighted in Subsection 2.3.

2.1. Timetable design methods

This section reviews literature on timetable design methods, introducing optimisation methods and analytical approaches for buffer time determination.

2.1.1. Timetable optimisation methods

Given a proposed line plan with desired frequencies and stops, timetabling problems entail determining timetables for a set of lines, satisfying track capacity constraints with the aim of optimising an objective function (Cacchiani and Toth, 2012; Lusby et al., 2011). A distinction can be made between nominal and robust versions. The former has a focus on determining efficient timetables, whilst the latter concerns creating schedules that avoid, in case of disruptions in the railway network, delay propagation as much as possible. Timetabling problems are in literature usually referred to as the Periodic Event Scheduling Problem (PESP) (Serafini and Ukovich, 1989) or Train Timetabling Problem (TTP) (Caprara et al., 2002), optimising various objectives. For an overview of PESP- and TTP-based research, see Cacchiani and Toth (2012).

Multiple approaches to achieving robustness in timetable design have been introduced, such as (i) stochastic programming (Kroon et al., 2008; Jin et al., 2019); (ii) light robustness (Fischetti and Monaci, 2009); (iii) recoverable robustness (Cicerone et al., 2009; Liebchen et al., 2009); (iv) delay management (Liebchen et al., 2010); and (v) bi-objective approaches (Schöbel and Kratz, 2009) An overview of robustness-centred timetable design methods can be found in Cacchiani and Toth (2012); Lusby et al. (2018).

2.1.2. Statistical methods for buffer time scheduling

With regard to distributing time allowances in the timetable, Palmqvist et al. (2017a) identify five strategies from literature: (i) A uniform allocation of running time supplements percentage (Scheepmaker and Goverde, 2015); (ii) Shifting running time supplement towards the beginning or end of the line (Vromans, 2005); (iii) Place dwell time supplement at or near strategic locations (Vromans, 2005); (iv) Based on where disturbances happen most frequently (Yuan and Hansen, 2008; Huang et al., 2019); and (v) Place time allowances at critical points (Andersson et al., 2013; Abid et al., 2017). Critical points here are defined as locations where two trains follow, cross, or overtake, and are most sensitive to delays. Concerning buffer time, always using the same deterministic value as an indicator for the size of the time allowance is the most simple form of distribution. Carey and Kwieciński (1994) suggest to base the scheduled headway on its relation with knock-on delays, delays transferred from a train to its successor. They perform a stochastic simulation to derive an approximate relationship between scheduled headways and knock-on delays.

2.2. Data-driven approaches

Historical realisation data has proven to be beneficial for monitoring operational quality, improving the schedule, diagnosing operational problems, and evaluating countermeasures in transportation science. In this section, literature related to data-driven knowledge on delays and buffer times is reviewed. For a more elaborate overview of big data applications in railway transportation and train dispatching in particular, see Ghofrani et al. (2018); Wen et al. (2019) respectively.

2.2.1. Delay distribution, propagation and recovery

Delays have been proven to be most relevant when arriving or departing from stations (Yuan and Hansen, 2007). Therefore, the probability distributions of arrival and departure delays have been studied in the literature. Wen et al. (2017) investigated the main statistics of primary delays, including (i) delay causes, (ii) delay frequencies, (iii) delay occurrences in time and space, (iv) affected number of trains, and (v) delay recovery patterns. To incorporate secondary delays, more sophisticated prediction models of delays and delay propagation are needed. To this end, Corman and Kecman (2018) present a stochastic prediction of train delays in real-time using Bayesian network. When two trains use the same part of the infrastructure within short time, a delay of the first train can be used to predict the delay of the second.

Alongside prediction models for secondary delays, research has focused on the measurement of this aspect of delay propagation. Weeda and Wiggenraad (2006) introduce a conceptual model to approximate secondary delay from realisation data, based on the delay of two conflicting trains and the scheduled buffer. A more accurate and effective metric is introduced by Goverde and Meng (2011), measuring the time loss associated with a route conflict compared to a reference running time on a block section of unhindered operations. This requires detailed scheduling and realisation data of signal passages.

Lee et al. (2016) introduced a supervised decision tree method that estimates the key factors contributing to knock-on delays. It is found that unscheduled waiting time for meeting or overtaking causes the majority of knock-on delays. More specifically, Zieger et al. (2018) analysed the effects of different buffer time distributions on the formation of knock-on delays. It is shown that the choice of distribution has a significant impact on performance metrics. Furthermore, Huang et al. (2019) provide a methodology to determine how and to what extent running time supplements and dwell time supplements affect delay recovery. They assess the utilisation rate of these two time allowances based on the fraction of the recovered time and the total scheduled time allowances. Similarly, Yang et al. (2020) focus on the effect of running time supplements and dwell time supplements on delay recovery and find that the effectiveness of time allowances to facilitate delay recovery is highly dependent on the size of the delay. Thus, allocation of time allowances should be performed considering the actual impact of the delay.

2.2.2. Timetable design based on realisation data

A conceptual framework is introduced by Weeda and Wiggenraad (2006) to acquire a joint design standard for running times and headway times. With the use of historical realisation data, an empirical relationship between headways and knock-on delays is derived. The buffer time proves to reduce knock-on delays, but the benefits decrease with higher values of buffer time, in line with findings by Yuan and Hansen (2007); Zieger et al. (2018). In line with the work of Carey and Kwieciński (1994), a data-driven method of determining minimum headways for conflicting train movements is proposed by Cerreto and Jonasson (2019). By means of a statistical analysis of the relationship between planned headways and the delay jump, the section specific minimum feasible headway between conflicting train movements in a railway system is estimated with a linear regression.

Hauck and Kliewer (2019) enhance existing solution approaches for the train timetabling problem as discussed in Subsection 2.1 by considering historical delay information. With the use of a Mixed Integer Programming (MIP) model, the difference between the realised running time and the planned running time is minimised. The optimisation is initiated with a given feasible timetable and adjusted in terms of arrival times, departure times and buffer times to adapt the timetable to avoid systematic delays. Huang et al. (2019) address allocation of time allowances in high-speed railway train operations. Based on historical data, a ridge regression model of delay recovery is generated regarding primary delay severity, running time supplements and dwell times. Subsequently, an Integer Linear Programming (ILP) model is presented that maximises the delay recovery by reallocating running time supplements and dwell times in an initial timetable.

2.3. Scientific gaps

To generate situation specific time allowances, in particular buffer times, (i) optimisation approaches, which apply mathematical models; (ii) simulation approaches, which iteratively enhance a current solution; (iii) statistical approaches, which apply statistics to infer relationships between variables; and (iv) data-driven approaches, which indicate recurring bottlenecks and tightly planned headways with the use of historical data, have been introduced. The works have been discussed in the sections above and are summarised in Table 1.

Reference	**Goal**	**Time allowance**	**Approach**	**Initial timetable required**
Burggraeve and Vansteenwegen (2017)	Passenger robustness	BT, RTS, DTS	Optimisation, Simulation	X
Hauck and Kliewer (2019)	Punctuality	BT, RTS, DTS	Optimisation, data-driven	X
Huang et al. (2019)	Delay recovery	RTS, DTS	Optimisation, data-driven	X
Jensen et al. (2017)	Capacity consumption	BT	Simulation, statistics	
Jin et al. (2019)	Average delay time	RTS	Optimisation	X
Jovanović et al. (2017)	Capacity and delay propagation	BT	Optimisation	X
Khoshniyat and Peterson (2017)	Timetable reliability	BT	Optimisation	X
Kroon et al. (2008)	Average weighted delay	BT, RTS	Optimisation, simulation	X
Lee et al. (2017)	Average delay	BT. RTS	Simulation	X
Li et al. (2020)	Cumulative delay time and energy consumption	RTS, DTS	Optimisation, data-driven	
Lu et al. (2017)	Efficiency and robustness	BT	Optimisation	
Meng et al. (2019)	Sum of train delays	BT, RTS	Optimisation	X

Restel et al. (2021)	Delay propagation and energy consumption	RTS	Simulation, data-driven	X
Weeda and Wiggenraad (2006)	Punctuality	BT, RTS, DTS	Statistics, data-driven	
Yang et al. (2020)	Delay expectation at stations	RTS, DTS	Optimisation, data-driven	X
Yuan and Hansen (2008)	Weighted sum of knock-on delays	BT	Optimisation	
This research	Delay propagation	BT	Statistics, data-driven	

Legend: BT = Buffer Time, RTS = Running Time Supplement, DTS = Dwell Time Supplement

Table 1: Situation specific time allowances in previous studies

Previous data-driven studies on generating situation specific buffer times have left various scientific gaps. First, it is unknown to what extent specific timetable characteristics influence delay propagation. The reviewed studies primarily addressed primary delays. Second, the effective allocation of buffer times to alleviate delay propagation based on historical realisation data has not been widely addressed in the tactical planning stage of railway timetables where an initial timetable is constructed. The majority of reviewed research utilises an initial timetable to reallocate buffer time to optimise a chosen objective, but thereby they exclude evaluating the total size of the time allowances and are incapable to change the order of the trains. Third, planning rules for buffer time based on historical realisation data have not been widely explored in literature, even though they are very commonly used in practice.

3. Methodology

Figure 1 summarises the methodology for the determination of buffer time planning rules from realisation data. The framework is explained below and has been divided into four parts: input, processing, statistical analysis and output.

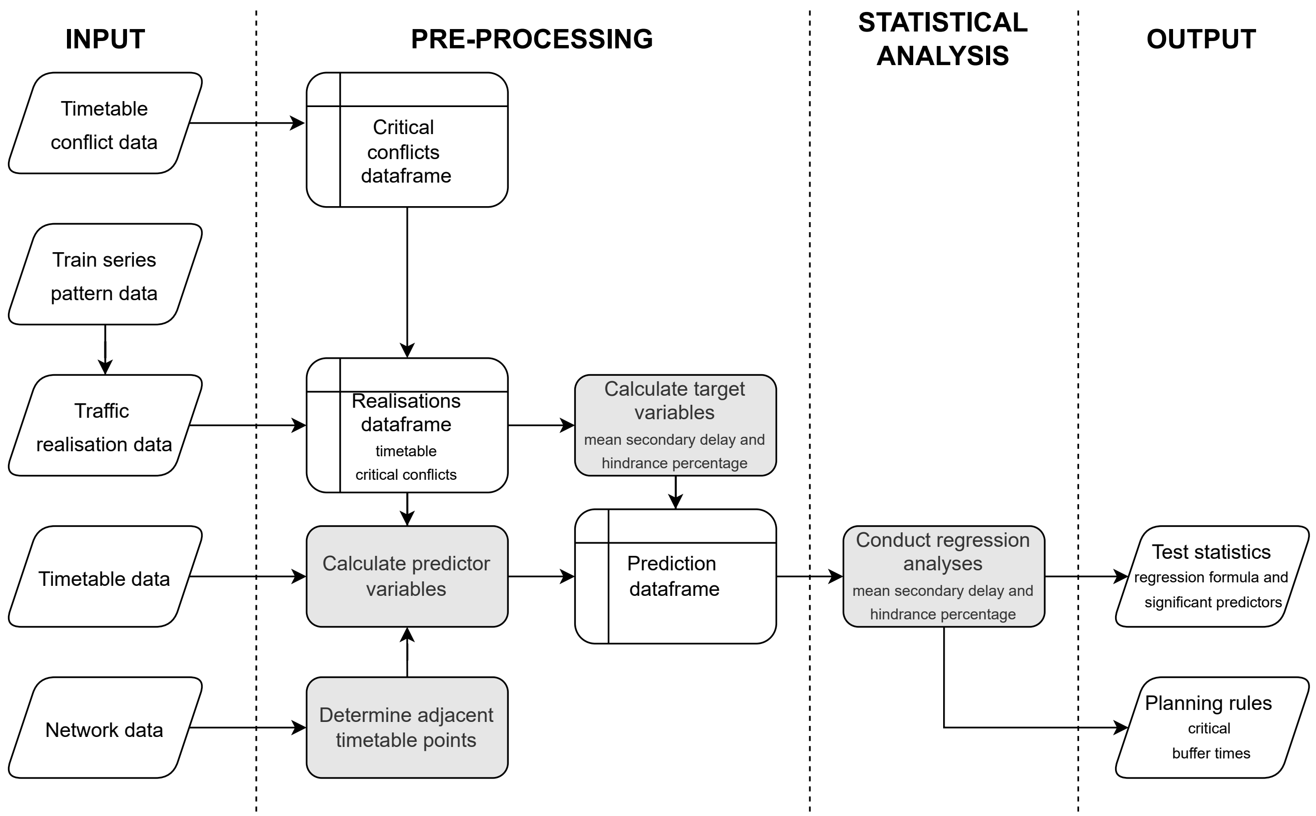


Figure 1: Framework for buffer time rules determination

Various types of input data are collected for this analysis, consisting of *Timetable conflict data*, *Traffic realisation data*, *Timetable data*, and *Network data*. Timetable conflict data specifies the tightly scheduled headways in the timetable, where the extended blocking times of two trains paths overlap. Traffic realisation data should include at least the train series, location, scheduled time and realised time of an event. Network data indicates how each timetable point in the network is connected to neighbouring timetable points. From the *timetable conflict data*, critical conflicts are identified, which indicate a timetable conflict between two specific train series patterns located in a critical point (i.e. the location where the overlap in blocking times is the biggest). The choice to only consider critical conflicts is based on the aim of this research to determine precise critical buffer times for tight headway situations. Then, of these critical conflicts the realisations are found by checking the *traffic realisation data*. With the realisations of the critical conflicts, the target and predictor variables are computed. Subsequently, hindrance distribution and a prediction dataframe with one row for each conflict are created. The dataset is reduced such that (i) delays longer than ten minutes and early arrivals of more than four minutes; (ii) order changes; (iii) shunting movements are excluded from the analysis. Using the prediction dataframe, two regression analyses were performed aiming to predict the two target variables introduced in Subsection 3.1. From these two prediction models, test statistics and buffer time planning rules are extracted. Regression was used as a prediction model due to its transparency and easy interpretation, which makes it straightforward to determine planning rules from. The process blocks in the framework are explained in the remainder of this section. Subsection 3.1 introduces the target an predictor variables and Subsection 3.2 describes the statistical methods applied.

3.1. Selection of target variables and predictor variables

3.1.1. Target variables

The target variables in this study are related to the aspects of delay propagation. Weeda and Wiggenraad (2006) study the mechanism of secondary delays and buffer time by approximating the secondary delay based on the delays of two succeeding train series patterns and the scheduled buffer on a conflict location. To do so, a conceptual model is presented in Figure 2 that shows for approximately 700 fictional realisations the delays of both trains of a fictional headway situation with a buffer of 60 seconds. Each dot represents a fictional realisation of a headway situation between two trains, before the interaction at the conflict point, and falls within a certain category. As long as the delay of train 1 is smaller than the buffer time, train 2 can proceed its normal operation. The diagonal line represents the instant when train 1 just leaves the conflict section when train 2 enters the conflict section. Observations to the right of the line indicate hindrance of train 2 caused by train 1, because train 1 has not left the conflict section yet.

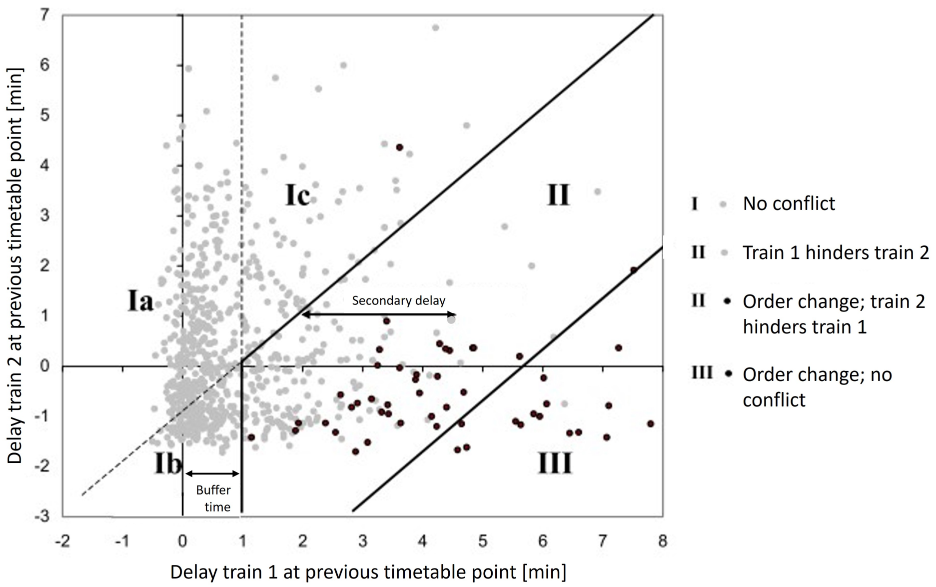


Figure 2: Determining secondary delay from initial delays of train 1 and train 2. Own creature based on Weeda (2005)

The conceptual model enables an estimation of the secondary delay train 2 suffers because of the delay of train 1. The secondary delay is defined as the horizontal distance between an observation in category II and the diagonal left border line of category II and can be calculated with Equation 1, where r_2 is the secondary delay to train 2, d_i is the initial delay to train i, and b_{12} is the scheduled buffer time between both trains. For observations outside category II (i.e. unhindered observations), the secondary delay equals zero.

$$r_2 = d_1 - max(0, d_2) - b_{12} \tag{1}$$

This research utilises the insights from this framework in defining two target variables. Firstly, the observed percentage of secondary delayed trains is utilised as an estimation of the probability of hindrance in a given conflict situation. Secondly, the expected amount of hindrance that occurs in a specific situation is denoted by the mean secondary delay observed in that situation. The hindrance percentage and the mean secondary delay serve as target variables and are both estimated with its own regression model.

3.1.2. Predictor variables

The target variables are impacted by a variety of variables that function as predictor variables in the statistical analysis. An overview of the added predictor variables and its computation are provided in Table 2. Among the predictor variables only factors known in the construction of a timetable are considered, because the buffer time planning rules should be applicable in an early stage of timetable development. Moreover, the reason to include certain predictor variables is substantiated by a source, or in case of own hypotheses explained below.

The numeric predictor variables are accommodated with a hypothesised sign of the correlation with the target variable. For the categorical variables a hypothesised sign does not make sense, since the sign of the relation is expected to differ per category and depends on the reference category used in a dummy coding scheme.

	Variable (source)	**Description**	**Expected sign**
For the headway situation	Direction (Lee et al., 2016)	Whether the two trains run in the same direction	
	Number Preceding Conflicts (Gorman, 2009)	Number of prior conflicts where the first train was involved in on the timetable point of the conflict	+
	Scheduled Buffer (Palmqvist et al., 2017b; Vromans et al., 2006)	The buffer time that was scheduled in the timetable	-
	Scheduled Headway (Gorman, 2009)	The headway that was scheduled in the timetable	-
	Technical Minimum Headway	The minimum headway needed to ensure a conflict-free headway situation	+
	Timetable Point (Lee et al., 2016)	The timetable point the conflict is closest to	
	Type Headway Situation (Lee et al., 2016)	Whether the conflict relates to a succession or a crossing of the two trains	
	Type Timetable Point (Lee et al., 2016)	The type that the timetsble point belongs to	
For both trains	Event (Dietzenbacher, 2021)	The event on the track	
	Carrier	The carrier of the train	
	Driving Characteristic (Lee et al., 2016)	The type of train	
	Previous Timetable Point; Next Timetable Point	The previous / next timetable point the train is scheduled on	
	Rolling Stock Type	The type of rolling stock the train has driven with	
	Rolling Stock Amount Carriages	The number of carriages the train has driven with	+/-
	Scheduled Running Time Supplement (Palmqvist et al., 2017b; Goverde and Meng, 2011)	The percentage of running time supplement that was scheduled in the timetable, determined from the previous stop to the timetable point of the headway situation	-
	Train Series (Lee et al., 2016)	The train series the train belongs to	
	Train Series Pattern (Lee et al., 2016)	The train series pattern the train belongs to	
	Type Previous Timetable Point; Type Next Timetable Point	The type that the timetable point belongs to	

Table 2: Overview of candidate predictor variables for predicting mean secondary delay and hindrance percentage

3.2. Statistical analysis

To predict a target variable based on a wide variety of predictor variables, while maintaining easy interpretability, a regression analysis can be performed. With a regression analysis it could be determined whether the mean secondary delay and hindrance percentage can be inferred from the values of the candidate predictor variables. The intensity of the effects the predictor variables have on the target variables are represented by parameters (i.e. betas) in a regression function. More specifically, the betas indicate the change in the prediction of the target variable for a one unit increase in the predictor variables (Allison, 1999). The parameters are estimated by ordinary-least squares, which minimises the sum of the squares of the deviation between the target variables' observed and predicted values (Berry and Feldman, 1985). This study estimates the linear relationship between the target variable and the predictor variables. The basic formula for such a linear regression is defined in Equation 2, where Y represents the target variable, β_0 refers to the estimated constant, β_1 ... β_n are estimated coefficients for the scores of the target variable, X_1 ... X_n are observed scores for the predictors variables, and ϵ represents the estimated residual, which captures the variance of the target variable not explained by the predictors in the regression model.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \epsilon \tag{2}$$

In case multiple candidate predictor variables have been identified, a regression model including all candidate predictors is often hard to interpret, not parsimonious and suffers from multicollinearity. To avoid this, only the most important candidate predictors are incorporated in the model (Galvão et al., 2008; Ghani and Ahmad, 2010). To this end, this research adopts a variable selection strategy that compares best to the stepwise method.The model estimation procedure, starts with no predictors in the model. Then, the effects of adding a predictor to the model is evaluated for each candidate predictor individually. If adding any predictor to the model appears significant and increases the adjusted R-squared value, then the candidate predictor that yields the highest increase in adjusted R-squared is added to the model. This process continues until adding a candidate predictor does no longer result in a certain minimum increase in adjusted R-squared or when all candidate predictors have been added to or excluded from the model. Simultaneously, already included predictors have to maintain significance and candidate predictors causing multicollinearity are excluded from the model.

The linear regression as presented in Equation 2 can be rewritten in the format of Equation 3, which can be used to determine the amount of scheduled buffer that would ensure a certain amount of hindrance percentage or mean secondary delay. As the acceptable amount of secondary delay is an arbitrary choice, a sensitivity analysis is performed to gain insights into the effects of various acceptable thresholds.

$$X_{scheduled\ buffer} = \frac{Y - \beta_0 - \beta_1 X_1 - \beta_2 X_2 - ... - \beta_{n-1} X_{n-1} + \epsilon}{\beta_{scheduled\ buffer}} \tag{3}$$

4. Case study and results

The methodology is applied to a case study in the Dutch railway network. Subsection 4.1 introduces the case study's geographical and temporal scope. Results were obtained by performing a exploratory data analysis in Subsection 4.2 and a regression analysis in Subsection 4.3. Next, Subsection 4.4 combines all the presented results to create buffer time planning rules.

4.1. Case description

The study area consists of the corridor between Schiphol Airport and Leiden Centraal and the corridor between Haarlem and Leiden Centraal. The four main reasons for choosing this demarcation are (i) the combination of a busy and less busy corridor enables analysing the impact of capacity consumption on mean secondary delay and hindrance percentage; (ii) the (expected) shortage of capacity on the station Schiphol Airport and the corridor Schiphol Airport - Leiden Centraal (Planting, 2016); (iii) a variety of train types make use of the infrastructure and a variety of timetable point types are present in the scope; and (iv) relatively low amount of interaction with freight trains and trains of other carriers of which no data was available. A schematic overview of the geographical scope is depicted in Figure 3, where the names of the timetable points are abbreviated.

The study period is limited to December 15th 2019, until March 15th 2020, because it is the last period before the COVID-19 pandemic and the first timetable to be communicated to drivers in tenths of minutes (as opposed to minutes formerly). This timetable is considered as the benchmark for returning to the regular timetable now that COVID-19 restrictions are lifted.

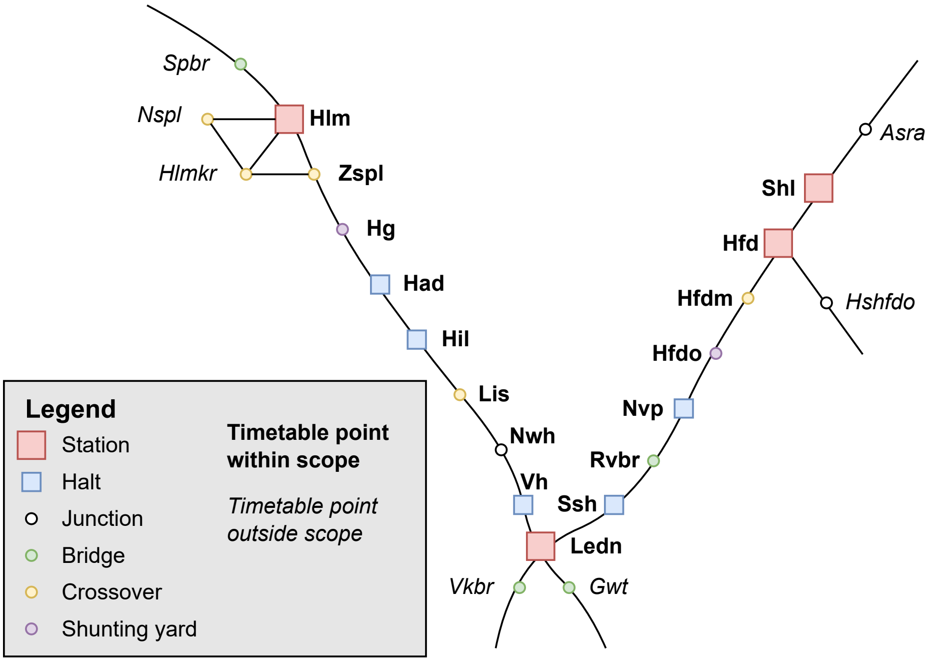
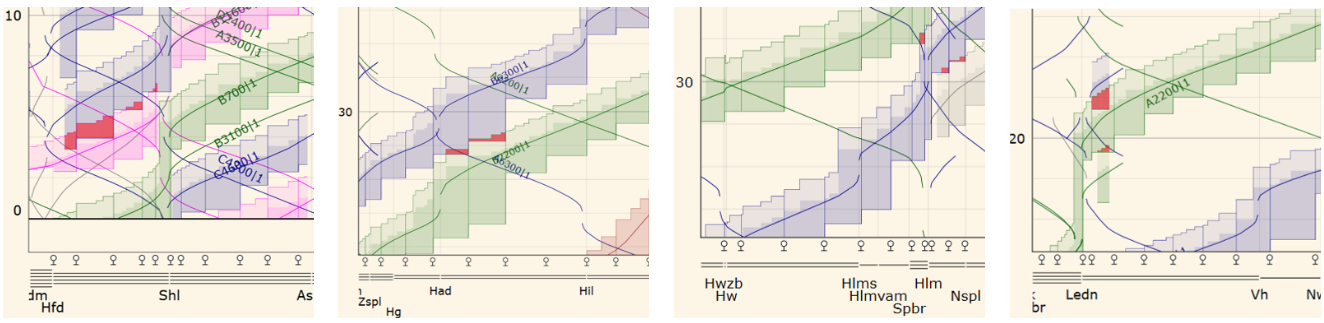


Figure 3: Schematic overview of geographical scope. Own creation based on ProRail (2022)

Some of the results are presented for specific conflicts, to provide the reader with a more intuitive feel on how the results are to be interpreted. These four conflicts have been selected, because they represent a large variety of situations. An overview of the main characteristics of the example conflicts is presented in Table 3 and train traffic diagrams of the conflict are shown in Figure 4.

Train Series Pattern Train 1	Train Series Pattern Train 2	Timetable Point	Event pair	Driving Characteristics	Type Headway Situation	Number of realisations	Hindrance	Order Change
B700	A900	Shl	Arrival - Arrival	IC - IC Direct	Succession	1295	7 (0.54%)	44 (3.40%)
B2200	B6300	Had	Departure - Arrival	IC - SPR	Succession	893	342 (38.30%)	2 (0.22%)
B4800	D2100	Hlm	Departure - Arrival	SPR - IC	Succession	1181	217 (18.37%)	0 (0.00%)
A2200	C3300	Ledn	Departure - Departure	IC - SPR	Crossing	831	359 (43.20%)	32 (3.85%)

Table 3: Overview of example conflicts

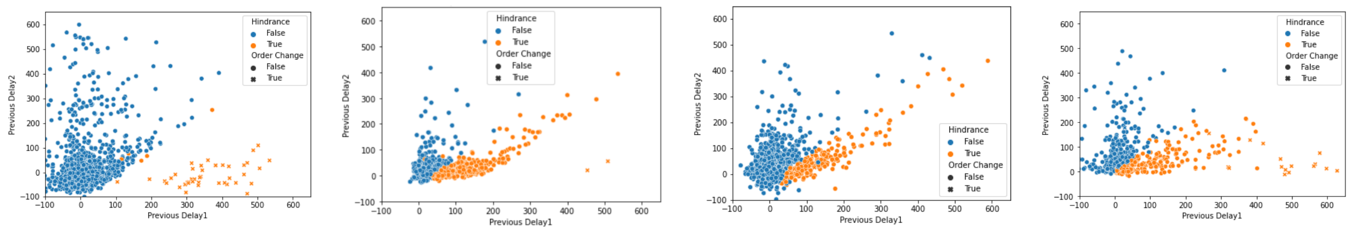


(a) Schiphol (b) Heemstede-Aerdenhout (c) Haarlem (d) Leiden

Figure 4: Train traffic diagram with the critical conflict in red

4.2. Exploratory data analysis

Figure 5 shows the realisations of the four example critical conflicts in the format of the conceptual model discussed in Section 3. The four critical conflicts differ in the number of realisations, hindrance percentage and order change percentage. It can be observed that the extent to which hindrance and order changes occur varies greatly among the critical conflicts that here serve as an example. This implies that hindrance, indeed, can be (partly) contributed to the timetable and network design. One striking observation is that the share of order changes at Schiphol appears considerably larger than at the other example conflicts. This could be caused by the choice that secondary delays at this timetable point are less accepted by traffic control due to the location specific infrastructure (e.g. tunnel). The magnitude of the hindrance is less obvious from this figure, because the scheduled buffer can vary among the observations of a conflict, due to changes in BDu or SD timetable design phases. But in general, observations further to the bottom right have a higher secondary delay, except when both trains have changed order, then the secondary delay as defined in this study is unknown. That is also the reason why observations with order changes are excluded from the remainder of the analysis.



(a) Delay distribution B700 - A900 - Shl

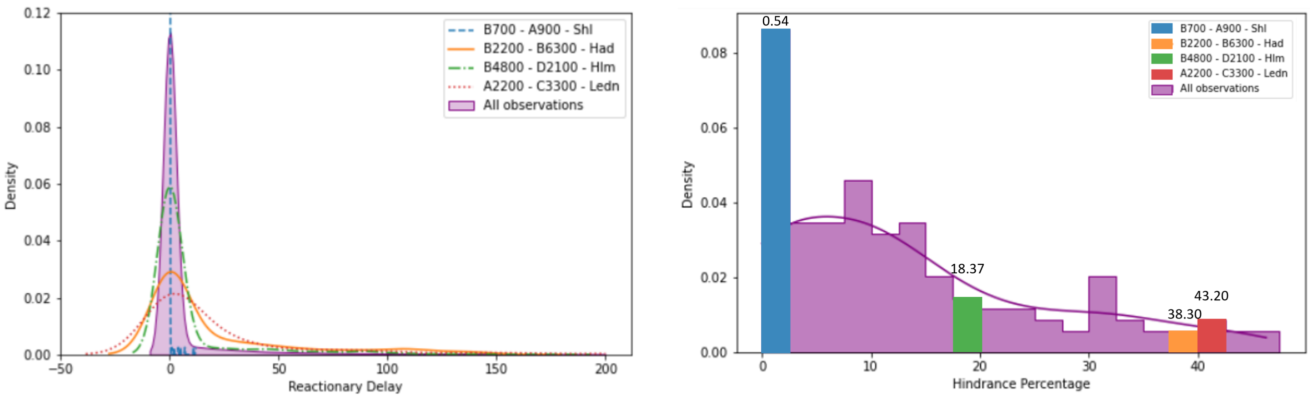
(b) Delay distribution B2200 - B6300 - Had

(c) Delay distribution B4800 - D2100 - Hlm

(d) Delay distribution A2200 - C3300 - Ledn

Figure 5: Delay distributions

The distributions of the target variables over all critical conflicts within the scope are presented in Figure 6, where the values of the four example critical conflicts are highlighted. The graph of the secondary delay follows a right-skewed curve, so the majority of the secondary delay has a small value. This is confirmed by the mean of 7.32 and the median of 0.00 over all observations. The conflicts with a wider distribution indicate both more variance in secondary delay and a higher hindrance percentage. The graph of the hindrance percentage of the conflicts show values predominantly located between 0% and 20%, and there are a few critical conflicts with up to 46.27% of realisations experiencing hindrance. The mean and median over all critical conflicts are respectively 13.32 and 9.58.



(a) Secondary delay distribution

(b) Hindrance percentage distribution

Figure 6: Distributions of target variables

4.3. Regression results

This section presents the results of the regression analyses on mean secondary delay and hindrance percentage. The results have been obtained by following the model estimation procedure explained in Section 3, with a stopping condition of 0.005 as the minimum needed increase in adjusted R-squared.

4.3.1. Mean secondary Delay

The first model aims to predict the mean secondary delay per conflict and is able to explain 90.7% of the variance within this variable (adjusted R-squared = 0.907). Table 4 presents the coefficients of the included variables, which have a significant impact on the secondary delay. What stands out is that the scheduled buffer, nor the scheduled headway is found to have a significant effect on the mean secondary delay. Yet, interesting to see is that the technical minimum headway does have an impact on the secondary delay, suggesting that it might be effective to identify buffer time in percentages of the technical minimum headway rather than absolute values. This furthermore is in line with the way running time supplements are defined.

To conclude, it is stated that the model can sufficiently predict the mean secondary delay and therefore can be used to extract buffer time planning rules from. Although the scheduled buffer has no significant influence and realisations deviate from the mean, the model provides insights into how variables compare to the reference conflict in Haarlem and into the unit that buffer time could be scheduled with.

		Unstandardised coefficients		Standardised coefficients		
Variable	**Category**	**Beta**	**Standard error**	**Beta**	**t**	**Sig.**
Constant		-4.356	0.876		-4.970	0.000
Previous Timetable Point Train 2 (ref = Spbr)	Hfdm	16.854	1.361	2.250	12.379	0.000
	Vkbr	-15.127	1.250	-2.019	-12.098	0.000
Rolling Stock Type Train 1 (ref = SGMM3 SGMM3)	SNG3 SNG4	-5.128	1.400	-0.685	-3.664	0.000
	SNG4	2.587	0.897	0.345	2.885	0.005
Rolling Stock Type Train 2 (ref = VIRM6)	SNG3 SNG3	6.069	1.421	0.810	4.272	0.000
Scheduled Running Time Supplement Train 1		0.167	0.046	0.126	3.644	0.000
Technical Minimum Headway		0.036	0.004	0.318	9.266	0.000
Train Series Pattern Train 1 (ref = B4800)	B1800	-2.652	1.209	-0.354	-2.194	0.031
	D3700	7.985	1.682	1.066	4.749	0.000
	D5800	7.703	2.068	1.028	3.724	0.000
	D6300	6.617	2.047	0.883	3.233	0.002
Train Series Pattern Train 2 (ref = D2100)	B3700	11.724	2.430	1.565	4.825	0.000
	E3300	-6.582	1.805	-0.879	-3647	0.000
	E4800	5.652	1.889	0.769	2.992	0.004
	G4800	5.991	1.696	0.800	3.532	0.001
Train Series Train 1 (ref = 4800)	3400	4.876	1.257	0.651	3.881	0.000
Train Series Train 2 (ref = 2100)	2400	9.914	1.077	1.323	9.204	0.000
Type Next Timetable Point Train 2 (ref = Crossover)	Halt	19.801	0.886	2.643	22.349	0.000

Table 4: Regression results of mean secondary delay

4.3.2. Hindrance Percentage

The second model aims to predict the hindrance percentage per conflict and is able to explain 90.7% of the variance within this variable (adjusted R-squared = 907). Table 5 presents the coefficients of the included variables, which have a significant impact on the secondary delay. For the significant categorical variables it is indicated what the reference level is.

When preceding trains are of the type IC Direct (abbreviation: HSN), the chance on hindrance increases with 13.822% as compared to a sprinter being the preceding train. Moreover, the second train being a starting train at the location of the conflict instead of coming from a bridge results in 23.708% more chance to be hindered. On the brighter side, crossings lead to 9.530% less chance on hindrance as compared to successions, which is explained by the fact the the time both trains are dependent on each other is smaller in case of crossings.

As expected, the scheduled buffer reduces the change on hindrance. Unexpectedly, the percentage of scheduled running time supplement the preceding train has, increases the hindrance percentage. Longer succeeding trains appear to experience less hindrance which might be because they gain a higher priority as more passengers are transported.

To conclude, it is stated that the model can sufficiently predict the hindrance percentage and therefore can be used to extract buffer time planning rules from.

Variable	Category	Unstandardised coefficients Beta	Unstandardised coefficients Standard error	Standardised coefficients Beta	t	Sig.
Constant		12.773	1.280		9.978	0.000
Driving Characteristic Train 1 (ref = SPR)	HSN	13.822	1.458	1.107	9.479	0.000
Next Timetable Point Train 1 (ref = Nspl)	Shl	25.769	1.828	2.063	14.099	0.000
	Vh	9.939	1.986	0.796	5.004	0.000
	Zspl	24.633	1.994	1.972	12.355	0.000
Previous Timetable Point Train 1 (ref = Spbr)	Hg	31.161	2.064	2.495	15.095	0.000
Previous Timetable Point Train 2 (ref = Spbr)	None	23.708	2.988	1.898	7.934	0.000
Rolling Stock Type Train 2 (ref = VIRM6)	SLT6	5.192	1.517	0.416	3.421	0.001
Scheduled Buffer		-0.119	0.015	-0.249	-7.647	0.000
Scheduled Running Time Supplement Train 1		0.337	0.081	0.139	4.144	0.000
Train Series Pattern Train 1 (ref = B4800)	A3300	-6.699	1.695	-0.536	-3.952	0.000
	B2400	12.643	2.841	1.012	4.450	0.000
	C3300	-6.963	2.313	-0.557	-3.010	0.003
	D3700	24.019	2.829	1.923	8.490	0.000
	D6300	11.241	3.197	0.900	3.516	0.001
Train Series Pattern Train 2 (ref = D2100)	B3700	17.067	4.030	1.366	4.235	0.000
	C1000	4.001	1.409	0.320	2.841	0.005
Train Series Train 1 (ref = 4800)	4600	8.352	1.799	0.669	4.644	0.000
Train Series Train 2 (ref = 2100)	2400	13.322	1.601	1.067	8.323	0.000
Type Headway Situation (ref = succession)	Crossing	-9.530	0.911	-0.763	-10.461	0.000

Table 5: Regression results of hindrance percentage

4.4. Determination of case-specific buffer time planning rules

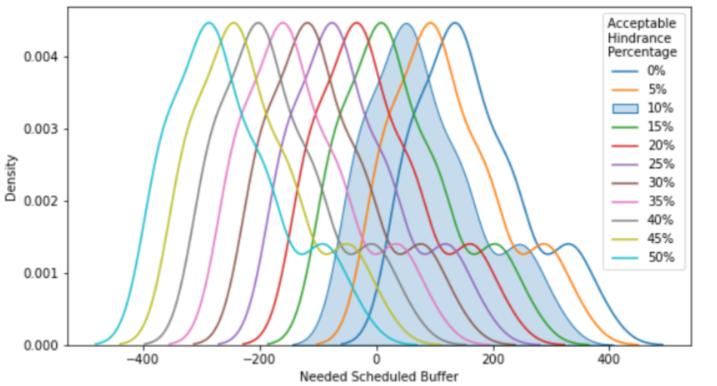
The linear regression as presented in Equation 2 can be rewritten in the format of Equation 3, which can be used to determine the amount of buffer that would ensure a certain amount of hindrance percentage. As the scheduled buffer is not a significant predictor for the mean secondary delay, this method can not be applied to determine the amount of buffer that would ensure a certain amount of mean secondary delay.

Figure 7 shows that the amount of needed scheduled buffer to achieve a certain amount of hindrance percentage differs per conflict. Due to the linear nature of the regression model the distributions in Figure 7a are the same for all hindrance percentages, but are shifted on the x-axis and the lines in Figure 7b are linear. As the effect of one unit buffer does not differ among various conflicts the lines in Figure 7b are parallel.

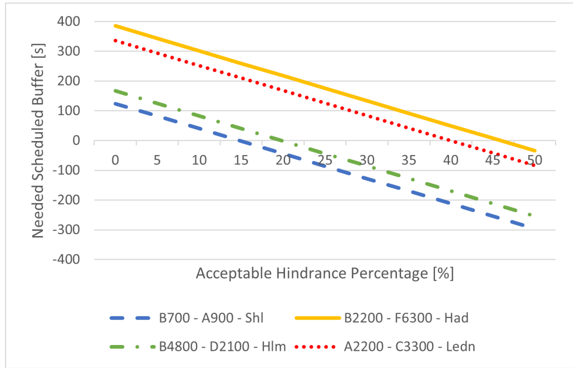
Compared to the currently mean scheduled buffer per conflict (standard deviation = 26) the spread in the predicted values for the needed scheduled buffer per conflict is significantly higher (standard deviation = 93), leading to an increase of 257%.

Essentially, in practice the scheduled headways are composed of a technical minimum headway and a minimum buffer time of 60 seconds. As a rule of thumb, smaller buffer times are acceptable when the gain in headway time is bigger than the loss in running time, which is evaluated for each situation individually. The planning rules presented in this section identify the minimum buffer time case-specifically, while evaluating the percentage of realisations that would be hindered. The planning rules provide more flexibility compared to the current norms, because the size of the minimum needed buffer depends on the preferred hindrance percentage.

Large differences in terms of predicted scheduled buffer are registered between the various conflicts in the case study. To obtain the same hindrance percentage the difference in predicted scheduled buffer between conflicts reaches up to 342 seconds. Thus, compared to the mean observed scheduled buffer per conflict (standard deviation = 26 seconds), the spread in the predicted values for the predicted scheduled buffer per conflict is significantly higher (standard deviation = 93), corresponding with an increase of 257%.



(a) Distribution of needed scheduled buffer for critical conflicts in the geographical scope for different acceptable hindrance percentages



(b) Needed scheduled buffer for the example conflicts for different acceptable hindrance percentages

Figure 7: Plots to show variance in needed scheduled buffer

When the observed mean hindrance percentage of the realised conflicts (=13.32%) is taken as acceptable hindrance percentage for all critical conflicts in the prediction, the hindrance distribution across all conflicts becomes uniform and the sum of all predicted scheduled buffers equals the total observed scheduled buffer, namely 7716 seconds. For every percentage increase in hindrance percentage, the sum of predicted scheduled buffers drops with 1173 seconds. Additionally, the mean of all observed scheduled buffers equals the mean of predicted scheduled buffer across all conflicts. For every percentage increase in hindrance percentage, the mean of predicted scheduled buffers drops with 8 seconds.

5. Discussion

The general impression from the case study in Section 4 is that the delay propagation at different critical conflicts is quite heterogeneous. The regression models were both able to explain 90.7% of this heterogeneity when using aggregated values for conflict situations, being mean secondary delay and hindrance percentage, meaning that structural delay propagation patterns can largely be contributed to timetable design and infrastructure characteristics. In particular, factors related to the route and train series of the trains contributing to the critical conflict appeared to be influential in explaining and predicting mean secondary delay and the probability of hindrance. Still, critical conflicts with the same mean secondary delay and hindrance percentage can have different secondary delay distributions, caused by among others (i) number of realisations of the conflict; (ii) number dispatching options; (iii) proactive attitude of traffic controllers; (iv) delay distributions of the first and second train at the previous timetable point; (v) the variety in scheduled buffer; (vi) the variety in rolling stock; and (vii) time of day.

The planning rules emerging from the regression models can be integrated in timetable design processes as (i) the size of the wake added in the simulation; or (ii) an indication of when slightly hindered headways are acceptable. To reduce the workload for planners further, it would be favourable that the needed scheduled buffer for a specific conflict is automatically shown in the planning application. To this end, all relevant factors explaining these aspects of delay propagation ought to be integrated in planning applications. The standardisation could help streamline data-driven timetable design.

6. Conclusion

In this paper, a data-driven approach for determination of buffer time planning rules suitable for usage in an initial timetable is proposed. The approach involves collecting traffic realisation data of tightly scheduled headway situations in the timetable and constructing two metrics representing delay propagation that occurred in these situations. A case study of the Dutch railway network was conducted to create planning rules regarding buffer time. These planning rules are not necessarily generic, but depend on timetable characteristics that have been found by means of a stepwise multiple regression model to significantly influence delay propagation.

The general conclusion of this paper is that an estimation or prediction of the expected delay propagation between two trains in a given situation could be done quite accurately (adjusted R-squared of both models equals 0.907) is essential to determining suitable buffer time planning rules, because it provides insight into where buffer time can be placed in order to be most effective. This research has operationalised delay propagation as the mean secondary delay and probability of hindrance a succeeding train encounters caused by a delay of a preceding train. Using the prediction model, the planning rules can be determined by setting a desired value for these metrics. This value could possibly be determined by making a trade-off between delay propagation and capacity based on a chosen objective (e.g. passenger punctuality, infrastructure consumption). Interesting to note is that the mean secondary delay is not significantly impacted by the scheduled buffer, but the probability of hindrance is. This indicates that the delay distributions of both trains have a bigger impact on the size of the secondary delay, while the size of the scheduled buffer determines how often hindrance occurs.

The proposed approach in this paper has some limitations. Most importantly, this research used an approximation method for secondary delay introduced by Weeda and Wiggenraad (2006). Although on average this estimation of secondary delay seems to resemble the true secondary delay, the estimation is particularly unsuitable for conflicts including starting trains and successions. A more accurate metric could be the time loss directly related to a route conflict (Goverde and Meng, 2011; Daamen et al., 2009). As time losses are determined locally at the signal of occurrence, more detailed data on signal passages and reference running times of unhindered operations have to be collected to enable this approach. Furthermore, the target variable mean secondary delay is an aggregated value of multiple realisations of one critical conflict situation. So information on the spread of mean secondary delay within the conflict is lost, whilst this could be useful information in assessing the sensitivity of secondary delay with regard to the scheduled buffer time.

One methodological research direction relates to the scope of the analysis. The analysis can be expanded by including freight trains, shunting movements, and trains of other carriers. For freight trains, in particular, other coefficients are expected, because they are often heavier and longer than passenger trains. This implies that they have a slower acceleration and longer braking and thus an unplanned stop has more impact. Moreover, it might be useful to gain insights into what factors may cause order changes and to add factors that can explain temporal variation. Moreover, future research may include concrete simulations of different timetables with varying buffer times, to obtain a better estimate of the effect of tweaking these planning norms on robustness and capacity utilisation. An important question for applying such more specific planning rules in practice, is which timetable structures would be made possible or impossible, and whether the capacity on busy lines will be increased, stay equal, or be decreased. Third, assessing the marginal utility and costs of adding an extra unit of buffer time could probably be explored in a simulation context. Based on this, the maximum performance of the headway situation could be determined based on the economical theory that marginal utility equals marginal costs. Marginal utility of an extra unit of buffer time has been explored by Jovanović et al. (2017). Lastly, among practitioners implicit and tacit knowledge exists about when scheduling tighter buffer times than the planning rules does not result in (too much) secondary delay. This implicit knowledge could potentially be caught in a decision model, where the trade-offs of scheduling buffer times become concrete.

Acknowledgements

The author thanks the members of the graduation committee, Prof. Dr. R. M. P. Goverde, Dr. N. Bešinović, Dr. S. Fazi, and Ir. P. Looij for their support, critical questions and detailed discussions.

References

Abid, M.M., Khan, M.B., Jiang, X., 2017. Ex-ante robustness measures for single track train timetables. Tehnicki Vjesnik 24, 1533–1542. doi:10.17559/TV-20150623101554.

Allison, P.D., 1999. Multiple Regression: A Primer. Pine Forge Press.

Andersson, E.V., Peterson, A., Törnquist Krasemann, J., 2013. Quantifying railway timetable robustness in critical points. Journal of Rail Transport Planning and Management 3, 95–110. doi:10.1016/j.jrtpm.2013.12.002.

Berry, W., Feldman, S., 1985. Multiple Regression in Practice. Sage, Newbury Park. doi:10.4135/9781412985208.

Burggraeve, S., Vansteenwegen, P., 2017. Optimization of supplements and buffer times in passenger robust timetabling. Journal of Rail Transport Planning and Management 7, 171–186. doi:10.1016/j.jrtpm.2017.08.004.

Cacchiani, V., Toth, P., 2012. Nominal and robust train timetabling problems. European Journal of Operational Research 219, 727–737. doi:10.1016/j.ejor.2011.11.003.

Caprara, A., Fischetti, M., Toth, P., 2002. Modeling and solving the train timetabling problem. Operations Research 50. doi:10.1287/opre.50.5.851.362.

Carey, M., Kwieciński, A., 1994. Stochastic approximation to the effects of headways on knock-on delays of trains. Transportation Research Part B 28, 251–267. doi:10.1016/0191-2615(94)90001-9.

Cerreto, F., Jonasson, M.H., 2019. Fact-checking of Timetabling Principles: a Case Study on the Relationship Between Planned Headways and Delays, in: Proceedings of the 8th International Conference on Railway Operations Modelling and Analysis - RailNorrköping 2019, EailNorrköping. pp. 312–323.

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 18, 229–257. doi:10.1007/s10878-009-9247-4.

Corman, F., Kecman, P., 2018. Stochastic prediction of train delays in real-time using Bayesian networks. Transportation Research Part C: Emerging Technologies 95, 599–615. doi:10.1016/j.trc.2018.08.003.

Daamen, W., Goverde, R.M.P., Hansen, I.A., 2009. Non-discriminatory automatic registration of knock-on train delays. Networks and Spatial Economics 9, 47–61. doi:10.1007/s11067-008-9087-2.

Dietzenbacher, D.S., 2021. Scrutinizing headway buffer norms for railway timetabling. Technical Report. Eindhoven University of Technology.

Fischetti, M., Monaci, M., 2009. Light robustness, in: Ahuja, R.K., Möhring, R.H., Zaroliagis, C.D. (Eds.), Robust and Online Large-Scale Optimization, Lecture Notes in Computer Science. Springer-Verlag, Berlin, Heidelberg. volume 5868, pp. 61–84. doi:10.1007/978-3-642-05465-5{_}3.

Galvão, R.K.H., Araújo, M.C.U., Fragoso, W.D., Silva, E.C., José, G.E., Soares, S.F.C., Paiva, H.M., 2008. A variable elimination method to improve the parsimony of MLR models using the successive projections algorithm. Chemometrics and Intelligent Laboratory Systems 92, 83–91. doi:10.1016/j.chemolab.2007.12.004.

Ghani, I.M.M., Ahmad, S., 2010. Stepwise multiple regression method to forecast fish landing. Procedia - Social and Behavioral Sciences 8, 549–554. URL: www.sciencedirect.com, doi:10.1016/J.SBSPRO.2010.12.076.

Ghofrani, F., He, Q., Goverde, R.M.P., Liu, X., 2018. Recent applications of big data analytics in railway transportation systems: A survey. Transportation Research Part C: Emerging Technologies 90, 226–246. doi:10.1016/j.trc.2018.03.010.

Gorman, M.F., 2009. Statistical estimation of railroad congestion delay. Transportation Research Part E: Logistics and Transportation Review 45, 446–456. doi:10.1016/j.tre.2008.08.004.

Goverde, R.M.P., Meng, L., 2011. Advanced monitoring and management information of railway operations. Journal of Rail Transport Planning and Management 1, 69–79. doi:10.1016/j.jrtpm.2012.05.001.

Hauck, F., Kliewer, N., 2019. A Data-Driven Optimization Approach to Improve Railway Punctuality, in: Fortz, B., Labbé, M. (Eds.), Operations Research Proceedings. Springer, Cham, pp. 539–545. doi:10.1007/978-3-030-18500-8{_}67.

Huang, P., Wen, C., Peng, Q., Lessan, J., Fu, L., Jiang, C., 2019. A data-driven time supplements allocation model for train operations on high-speed railways. International Journal of Rail Transportation 7, 140–157. doi:10.1080/23248378.2018.1520613.

Independent Regulators' Group - Rail, 2021. Ninth Annual Market Monitoring Working Document. Technical Report.

Jensen, L.W., Landex, A., Nielsen, O.A., Kroon, L.G., Schmidt, M., 2017. Strategic assessment of capacity consumption in railway networks: Framework and model. Transportation Research Part C: Emerging Technologies 74, 126–149. doi:10.1016/j.trc.2016.10.013.

Jin, B., Feng, X., Wang, Q., Wang, X., Liu, C., 2019. Improving Timetable Robustness through Optimal Distribution of Runtime Supplement, in: 2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019, Institute of Electrical and Electronics Engineers, Auckland, New Zealand. pp. 2803–2808. doi:10.1109/ITSC.2019.8917477.

Jovanović, P., Kecman, P., Bojović, N., Mandić, D., 2017. Optimal allocation of buffer times to increase train schedule robustness. European Journal of Operational Research 256, 44–54. doi:10.1016/j.ejor.2016.05.013.

Khoshniyat, F., Peterson, A., 2017. Improving train service reliability by applying an effective timetable robustness strategy. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations 21, 525–543. doi:10.1080/15472450.2017.1326114.

Kroon, L.G., Maróti, G., Helmrich, M.R., Vromans, M.J., Dekker, R., 2008. Stochastic improvement of cyclic railway timetables. Transportation Research Part B: Methodological 42, 553–570. doi:10.1016/j.trb.2007.11.002.

Lee, W.H., Yen, L.H., Chou, C.M., 2016. A delay root cause discovery and timetable adjustment model for enhancing the punctuality of railway services. Transportation Research Part C: Emerging Technologies 73, 49–64. doi:10.1016/j.trc.2016.10.009.

Lee, Y., Lu, L.S., Wu, M.L., Lin, D.Y., 2017. Balance of efficiency and robustness in passenger railway timetables. Transportation Research Part B: Methodological 97, 142–156. doi:10.1016/j.trb.2016.12.004.

Li, W., Peng, Q., Wen, C., Wang, P., Lessan, J., Xu, X., 2020. Joint optimization of delay-recovery and energy-saving in a metro system: A case study from China. Energy 202. doi:10.1016/j.energy.2020.117699.

Liebchen, C., Lübbecke, M., Möhring, R., Stiller, S., 2009. The Concept of Recoverable Robustness, Linear Programming Recovery, and Railway Applications, in: Ahuja, R.H., Möhring, R.H., Zaroliagis, C.D. (Eds.), Robust and Online Large-Scale Optimization, Lecture Notes in Computer Science. Springer-Verlag, Berlin, Heidelberg. volume 5868, pp. 1–27. doi:10.1007/978-3-642-05465-5{_}1.

Liebchen, C., Schachtebeck, M., Schöbel, A., Stiller, S., Prigge, A., 2010. Computing delay resistant railway timetables. Computers & Operations Research 37, 857–868. doi:10.1016/J.COR.2009.03.022.

Lu, C., Tang, J., Zhou, L., Yue, Y., Huang, Z., 2017. Improving recovery-to-optimality robustness through efficiency-balanced design of timetable structure. Transportation Research Part C: Emerging Technologies 85, 184–210. doi:10.1016/j.trc.2017.09.015.

Lusby, R.M., Larsen, J., Bull, S., 2018. A survey on robustness in railway planning. doi:10.1016/j.ejor.2017.07.044.

Lusby, R.M., Larsen, J., Ehrgott, M., Ryan, D., 2011. Railway track allocation: Models and methods. OR Spectrum 33, 843–883. doi:10.1007/s00291-009-0189-0.

Meng, L., Muneeb Abid, M., Jiang, X., Khattak, A., Babar Khan, M., 2019. Increasing Robustness by Reallocating the Margins in the Timetable. Journal of Advanced Transportation 2019. doi:10.1155/2019/1382394.

Ministerie van Verkeer en Waterstaat, 2010. Priority decision High-Frequency Rail Transport Programme. Technical Report. Ministerie van Verkeer en Waterstaat. URL: https://silo.tips/download/priority-decision-high-frequency-rail-transport-programme.

Pachl, J., 2014. Timetable Design Principles, in: Hansen, I.A., Pachl, J. (Eds.), Railway Timetabling & Operations. second ed.. DVV Media Group GmbH, Hamburg, Germany. chapter 2, pp. 13–46.

Palmqvist, C.W., Olsson, N.O.E., Hiselius, L.W., 2017a. An Empirical Study of timetable Strategies and Their Effects on Punctuality, in: 7th International Conference on Railway Operations Modelling and Analysis. doi:10.1109/HICSS.2015.400.

Palmqvist, C.W., Olsson, N.O.E., Hiselius, L.W., 2017b. Some influencing factors for passenger train punctuality in Sweden. International Journal of Prognostics and Health Management 8.

Planting, T.J., 2016. Ontwerpmethoden van Dienstregelingen. Technical Report. Delft University of Technology.

ProRail, 2020. Startdocument Voorbereidingsfase dienstregeling 2022. Technical Report. URL: https://nsdigitaal.sharepoint.com/teams/Logistiek/KritischeDocumenten/Plannormendienstregeling2022V1.02oktober2020.pdf.

ProRail, 2022. Basiskaart. URL: https://maps.prorail.nl/portal/home/webmap/viewer.html?url=https%3A%2F%2Fmaps.prorail.nl%2Farcgis%2Frest%2Fservices%2FProRail_basiskaart%2FFeatureServer&source=sd.

Restel, F., Wolniewicz, Ĺ., Mikulčić, M., 2021. Method for Designing Robust and Energy Efficient Railway Schedules. Energies 14, 8248. URL: https://doi.org/10.3390/en14248248, doi:10.3390/en14248248.

Scheepmaker, G.M., Goverde, R.M.P., 2015. The interplay between energy-efficient train control and scheduled running time supplements. Journal of Rail Transport Planning and Management 5, 225–239. doi:10.1016/j.jrtpm.2015.10.003.

Scheepmaker, G.M., Goverde, R.M.P., 2021. Multi-objective railway timetabling including energy-efficient train trajectory optimization. European Journal of Transport and Infrastructure Research 21, 1–42. doi:10.18757/ejtir.2021.21.4.5453.

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, Lecture Notes in Computer Science. Springer-Verlag, Berling, Heidelberg. volume 5868, pp. 119–144. doi:10.1007/978-3-642-05465-5{_}5.

Serafini, P., Ukovich, W., 1989. A Mathematical Model for Periodic Scheduling Problems. SIAM Journal on Discrete Mathematics 2, 550–581. doi:10.1137/0402049.

Vromans, M.J., 2005. Reliability of Railway Systems. Ph.D. thesis. Erasmus University Rotterdam. Rotterdam.

Vromans, M.J., Dekker, R., Kroon, L.G., 2006. Reliability and heterogeneity of railway services. European Journal of Operational Research 172, 647–665. doi:10.1016/j.ejor.2004.10.010.

Weeda, V.A., 2005. Plannen met speling, spelen met de planning. Technical Report. Delft University of Technology. Delft.

Weeda, V.A., Wiggenraad, P.B.L., 2006. Joint design standard for running times, dwell times and headway times, in: Computers in Railways X, WIT Press, Southampton, UK. pp. 611–620. doi:10.2495/CR060601.

Wen, C., Huang, P., Li, Z., Lessan, J., Fu, L., Jiang, C., Xu, X., 2019. Train dispatching management with data-driven approaches: A comprehensive review and appraisal. IEEE Access 7, 114547–114571. doi:10.1109/ACCESS.2019.2935106.

Wen, C., Li, Z., Lessan, J., Fu, L., Huang, P., Jiang, C., 2017. Statistical investigation on train primary delay based on real records: evidence from Wuhan-Guangzhou HSR. International Journal of Rail Transportation 5, 170–189. doi:10.1080/23248378.2017.1307144.

Yang, X., Hou, Y., Li, L., 2020. Buffer time allocation according to train delay expectation at stations. International Journal of Rail Transportation 8, 249–263. doi:10.1080/23248378.2019.1655805.

Yuan, J., Hansen, I.A., 2007. Optimizing capacity utilization of stations by estimating knock-on train delays. Transportation Research Part B: Methodological 41, 202–217. doi:10.1016/j.trb.2006.02.004.

Yuan, J., Hansen, I.A., 2008. Closed form expressions of optimal buffer times between scheduled trains at railway bottlenecks, in: IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, pp. 675–680. doi:10.1109/ITSC.2008.4732539.

Zieger, S., Weik, N., Nießen, N., 2018. The influence of buffer time distributions in delay propagation modelling of railway networks. Journal of Rail Transport Planning and Management 8, 220–232. doi:10.1016/j.jrtpm.2018.09.001.

B
Recommendations (in Dutch)

Gestoeld op de bevinden van dit afstudeeronderzoek, worden aanbevelingen gedaan betreft ontwerpmethodieken en data processing binnen NS. De aanbevelingen zijn geformuleerd als incrementele stappen, welke kunnen worden uitgevoerd om het dienstregeling ontwerpproces met betrekking tot buffer tijden gefocust op vervolgvertraging door te ontwikkelen. Het identificeren van tussenstappen tot het meest ideale eindbeeld van dit datagedreven onderzoek komt overeen met de wijze waarop datagedreven organisaties gevormd en geëvalueerd worden in the literatuur (Berndtsson and Svahn, 2020; Davenport, 2018). De genoemde volgorde is een suggestie die momenteel passend en haalbaar lijkt, maar uiteraard kunnen extra stappen worden toegevoegd en overbodige stappen mogen worden overgeslagen.

1. **Gebruik de conclusies uit de case study** - Gebaseerd op het regressie model van gemiddelde vervolgvertraging (i) Vorig dienstregelpunt van trein 2 zijnde Hfdm of Vkbr; en (ii) Type next dienstregelpunt van trein 1 zijnde een halte bleken de factoren met de grootste invloed. De meest invloedrijke factoren op het hinder percentage bleken (i) Volgend dienstregelpunt van trein 1 zijnde Shl of Zspl; (ii) Vorig dienstregelpunt van trein 1 zijnde Hg; en (iii) Trein 2 zijnde een startende trein. Gegeven dat deze factoren een grote invloed hebben op gemiddelde vervolgvertraging en hinder percentage, is het kansrijk om juist wanneer deze factoren zich voordoen in een opvolgsituatie, de geassocieerde voorspelde geplande buffer supplemententen te raadplegen.

2. **Analyseer de verwachte vervolgvertraging van geplande opvolgsituaties** - Gebaseerd op de gevonden regressieformules in the case study, kunnen de verwachte gemiddelde vervolgvertraging en verwachte hinder percentage worden voorspeld voor alle geplande maatgevende conflicten in de concept dienstregeling. Dit kan planners een indicatie geven welke situaties meer buffer tijd behoeven en waar minder buffer tijd toereikend genoeg is. Bijvoorbeeld, een gesorteerde lijst van alle maatgevende conflicten kan worden gecreeërd gebaseerd op de verwachte gemiddelde vervolgvertraging of verwachte hinder percentage. De planner kan tevens een andere controle uitvoeren op (i) een vooraf gekozen aantal maatgevende conflicten met de hoogste waarden; danwel (ii) op de maatgevende conflicten met een waarde hoger dan een vooraf gekozen acceptabele randwaarde.

3. **Een acceptabele randwaarde vaststellen** - Bepaal een acceptabele randwaarde voor gemiddelde vervolgvertraging en hinder percentage per maatgevend conflict, zodat de buffer tijd die hiervoor nodig is bepaald kan worden. Deze waarde kan verschillen per maatgevend conflict, maar het wordt geadviseerd om een alomvattende maximale acceptabele randwaarde vast te stellen. De waarde per conflict kan worden bepaald door een afweging te maken met de beschikbare capaciteit op de locatie van het maatgevende conflict. Een meer straightforward aanpak kan zijn om 30 seconden als maximale acceptabele randwaarde te nemen voor gemiddelde vervolgvertraging, omdat dit de huidige norm is voor acceptabele aankomstvertraging op grote stations veroorzaakt door gehinderde opvolgingen door kleine buffer tijden zoals besproken in Chapter 3.

4. **Integreer alle relevante factoren in planningapplicaties** - Om betere inzichten te krijgen in alle factoren van de geplande opvolgsituatie is het aangeraden om planningapplicaties uit te breiden met details van conflict situaties. Bijvoorbeeld, de geplande buffer tijd, richting van de treinen, en activiteitenpaar van de treinen bijdragend aan het conflict kunnen worden getoond. Wanneer deze werkwijze is ingesleten, is het wenselijk dat de benodigde buffer tijd voor een specifiek conflict automatisch getoond wordt in de planningapplicatie, zodat de werkdruk op planners wordt verlaagd. Deze standaardisatie kan bovendien helpen met het stroomlijnen van datagedreven dienstregeling ontwerp.

5. **Automatisch updaten van regressiemodellen met meer (nieuwe) realisatie data** - Het regelmatig updaten van de regressiemodellen resulteert in meer accurate parameters van de significante predictoren, aangezien

de meest recente realisatie data gebaseerd is op het meest recente dienstregeling ontwerpproces. Bovendien, leidt het analyseren van data gebaseerd op verschillende dienstregelingen tot een holistischer perspectief van de effecten van de verschillende ranges van buffer tijden.

C
Data collection

This appendix provides detailed information about the data collection.

Realisation data

Description	Traffic realisation data
File name	WORK_2019tmjun; WORK_2019TMDEC; WORK_2020TMDEC; WORK_2020TMJUN
File size	3.09 GB; 2.80 GB; 2.37 GB; 2.45 GB
Data format	CSV
Retrieval date	30-11-2021
Retrieval method	Request by e-mail
Contact person	Jorik van Onna
Retrieved columns	SL_VERKEERSDATUM; SL_TREINNR; SL_TREINSERIE; SL_RICHTING; SL_RIJKARAKTERISTIEK; SL_VERVOERDER; SL_DRGLPT; SL_ACT_SRT; SL_SPOOR; SL_BASIC_PLANTIJD; SL_BASIC_UITVTIJD; SL_MATSOORT_TYPE; SL_MATSOORT_BAKKEN; SL_ROUTE_SEINEN; SL_ROUTE_SEINBEELDEN
Used columns	SL_VERKEERSDATUM; SL_TREINNR; SL_TREINSERIE; SL_RIJKARAKTERISTIEK; SL_VERVOERDER; SL_DRGLPT; SL_ACT_SRT; SL_BASIC_PLANTIJD; SL_BASIC_UITVTIJD; SL_MATSOORT_TYPE; SL_MATSOORT_BAKKEN
Time period	01-01-2019 until 31-12-2020

Table C.1: Realisation data

Conflict data

Description	Conflicts in DONS
File name	Conflicten DONS drp meerdere secties v2
File size	1,522 kB
Data format	Excel
Retrieval date	12-01-2022
Retrieval method	Request by e-mail
Contact person	Patrick Looij & Jantine Buren
Retrieved columns	Gebied; Trein1; Trein2; Begin; Einde; Duur; Overlap; Secties
Manually constructed columns	Opvolgsituatie; Zelfde richting; Type opvolging; Opmerking
Used columns	Gebied; Trein1; Trein2; Overlap; Secties; Opvolgsituatie; Zelfde richting; Type opvolging; Opmerking
Time period	15-12-2019 until 15-12-2020

Table C.2: Conflict data

Train series and train numbers

Description	Matching train numbers to their corresponding series
File name	Treinnrs en patronen dec 2019
File size	924 kB
Data format	Excel
Retrieval date	03-01-2022
Retrieval method	Request by e-mail
Contact person	Jantine Buren
Retrieved columns	treinnr; act_plandag; patrooncode
Used columns	treinnr; patrooncode

Table C.3: Train series and numbers data

Timetable data

Description	BU timetable in DONS
File name	dienstregeling BD 2020 348
File size	3,292 kB
Data format	Excel
Retrieval date	24-02-2022
Retrieval method	Request by e-mail
Contact person	Patrick Looij & Jantine Buren
Retrieved columns	Serie; Trein-nr; Actief; Treintype; Vervoerder; Bedrijfsurencode; Regelnummer; Materieel; Drp; Spoor; Lengte; Stopt; R; Stat; SD; VAdvies; RtKaal; RtReal; RtSpeling; TBerCum; StatPlan; DeltaRt; RtPlan; TPlanCum; SpelingPlan; Aankomst; Vertrek; Opmerking
Used columns	Serie; Drp; Spoor; Lengte; RtKaal; RtPlan; Stopt; Aankomst; Vertrek

Table C.4: Conflict data

Network data

Description	Connection of each timetable point to adjacent timetable points
File name	DONNA_78605_VER_1_IAUF_DRGLPT_VERBINDING
File size	211 kB
Data format	Text
Retrieval date	01-02-2022
Retrieval method	Access through NS sharepoint (NO en V&B > NO > Ontwerpen > IPO > Infrastructuur > IA-bestanden (origineel) > _Donna-20191215-BD-20190726.zip
Retrieved columns	
Used columns	

Table C.5: Network data

D
Computation of predictor variables

This appendix provides more details on the operationalisation of some predictor variables.

D.1. Scheduled Buffer

The scheduled buffer time is composed by the difference in scheduled headway time and technical minimum headway time (Pachl, 2014) (Equation D.1), and can deviate slightly from the buffer time planning norms. The technical minimum headway is defined as the minimum interval between scheduled activities of two trains, respecting the requested speed profile of both trains. As explained in Chapter 2, this corresponds to the difference in scheduled times when the blocking times of the two trains touch. In case of extended blocking times the minimum buffer time needs to be subtracted from this value to identify the technical minimum headway. When there is a conflict (i.e. overlap in blocking times), the duration of the conflict indicates the amount of time that is left out of the calculation. To compensate for this symptom, the duration of the conflict should be added to the technical minimum headway calculation.

Thus, from extended blocking time stairways can be extracted that the technical minimum headway time can be calculated by Equation D.2, where h refers to the headway time, b represents buffer time between a pair of trains, and t is the time in seconds. A conflict situation in Leiden (Ledn) for which the technical minimum headway time is calculated in such a way is depicted in Figure D.1 for the train paths in red. Note that there can be a difference between the scheduled headway in the BU phase and the VL phase, because of day specific adaptations. Hence, the difference in scheduled headways in Equation D.1 and Equation D.2.

$$b_{scheduled} = h_{scheduledVL} - h_{min} \tag{D.1}$$

$$h_{min} = h_{scheduledBU} - b_{planning\ norm} + t_{blocking\ time\ overlap} \tag{D.2}$$

D.2. Number Preceding Conflicts

This variable represents the dependence of a timetable conflict on prior conflicts on that location. Figure D.2a shows that the conflict between train G3200|1 and C4300|2 at Schiphol is preceded by 1 conflict. This variable is based on the

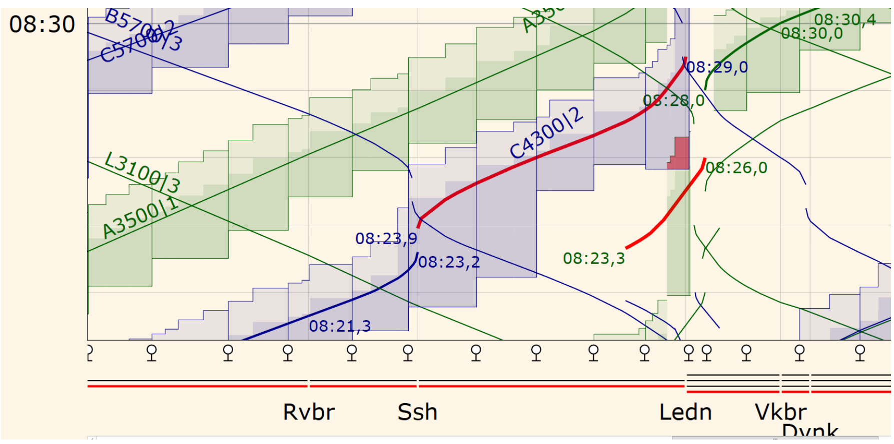
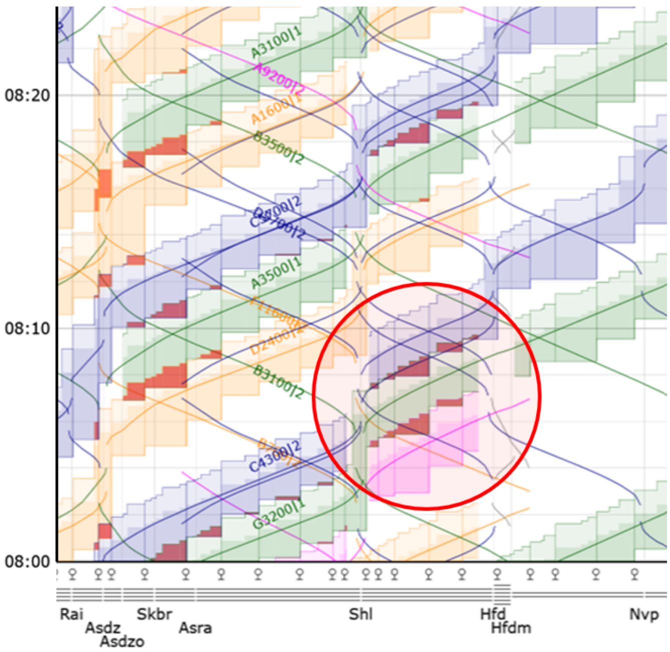
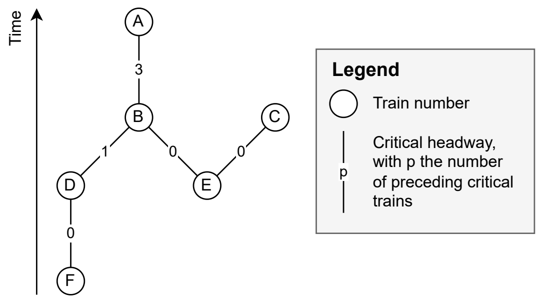


Figure D.1: Visualisation of technical minimum headway and scheduled buffer in a train traffic diagram concept of a conflict tree (Goverde and Meng, 2011). A conflict tree represents complex conflict propagations made up of various conflict chains, which is defined by Goverde and Meng (2011) as *"a linked list of trains that successively hinder one another"*. Adversely, a reversed conflict tree indicates for some train what preceding train(s) have caused hindrance to it. For example, Figure D.2b shows the reversed conflict tree of some train A. The tree implies that train A is hindered by train B, which in its turn is hindered by trains D and E. Train D, moreover, is hindered by train F. This concept can be applied to determine the number of preceding critical trains.

An iterative algorithm was created that checks whether the first train has acted as a second train in one or more prior conflicts on the same timetable point. Accordingly, the number of preceding conflicts equals the number of previous conflicts to which this applies plus the number of preceding conflicts of each previous conflict. If no prior conflicts exist, a value of 0 is given. This concept is visually depicted in Figure D.2b and the implemented code is explained with algorithm 1.



(a) Train traffic diagram showing preceding conflicts, with conflict between G3200|1 (green) and C4300|2 (blue) (circled in red) that has one preceding conflict



(b) Reversed conflict tree of some train A

Figure D.2: Conceptual model of number preceding conflicts

Algorithm 1: Number Preceding Critical Trains

Data: Conflict k; List of all conflicts; List of already visited train numbers
Result: The number of preceding critical trains of conflict k

Number of preceding critical trains[Conflict k] $\leftarrow$ *0;*
Number of previous conflicts $\leftarrow$ *0;*
for *i* $\in$ *All conflicts* **do**
  if *Date[Conflict k] == Date[i]* ***and***
  Timetable point[Conflict k] == Timetable point[i] ***and***
  Train number 2[Conflict k] == Train number 1[i] ***and***
  Scheduled buffer[Conflict k] < 61 ***and***
  Train number 1[i] ***is not in*** *Visited train numbers* **then**
    Number of previous conflicts $\xleftarrow{+}$ 1
    Previous conflicts $\xleftarrow{+}$ i
if *Number of previous conflicts* $\geq$ *1* **then**
  for *n* $\in$ *Previous conflicts* **do**
    Visited train numbers $\xleftarrow{+}$ Train number 1[n];
    Number of preceding critical trains $\xleftarrow{+}$ Number of preceding critical trains of conflict n (algorithm 1)

D.3. Consecutive timetable points

Based on the created adjacency list from the network data, the prior and following timetable points of each train contributing to the conflict are identified. These variables can be deduced by the sequence in which the scheduled times at each adjacent timetable point of the timetable point of the conflict occur. Hence, to find the prior timetable point of a train, the adjacent timetable point with the smallest positive gap in scheduled time compared to the conflict is selected. Similarly, the following timetable point is found by selecting the row with the smallest negative gap in scheduled time compared to the conflict. Then, the prior and following timetable point, its scheduled times and its realised times are added to the dataset. This corresponds with the light blue cells in Table D.1. With the use of these variables and the dictionary created with the network data, the type of the prior and following timetable point are appended to the dataset as well.

When a train starts or ends its journey there is no prior or following timetable point respectively. In this case the prior and following (type) timetable points are set to "None" and all time related columns are set to 0. This process is done for both trains in the headway situation.

Type timetable point	Train Number 1	Timetable Point	Activity 1	Date	Scheduled Time 1	Realised Time 1
Timetable point of the conflict	725	Shl	V	2019-09-01	08:04:00	08:06:08
	725	Hfd	D	2019-09-01	07:58:42	08:00:17
Previous timetable point train 1	725	Shl	A	2019-09-01	08:02:00	08:03:51
Next timetable point train 1	725	Asra	D	2019-09-01	08:08:00	08:09:42

Table D.1: Determine consecutive timetable point

D.4. Scheduled Running Time Supplement

The scheduled running time supplement is defined as the percentage of running time supplement that was scheduled in the timetable, determined from the previous stop of a train to the timetable point of the conflict. For every observation, the previous stop was determined by iterating over the dataframe. If the observation was of a starting train, "None" was taken as value.

In the *timetable data*, the scheduled running time and technical minimum running time are known between two timetable points. Then, the scheduled running time supplement between two timetable points is the difference in scheduled running time and technical minimum running time, which is converted to a percentage by dividing by the technical minimum running time. To achieve the percentage of scheduled running time supplement from the previous stop, the fraction is summed for all prior timetable points in the range between the previous stop and the timetable point of the conflict, excluding the previous stop itself. This process is operationalised by Equation D.3, where i corresponds with a timetable point, and SRT_i and MRT_i are respectively the scheduled and technical minimum running time between timetable point i and its previous timetable point, and m and n refer to the timetable point of the conflict and the previous stop respectively as depicted in Figure D.3.

$$SRTS_i = \sum_{i=n+1}^{m} \frac{SRT_i - MRT_i}{MRT_i} * 100\% \tag{D.3}$$

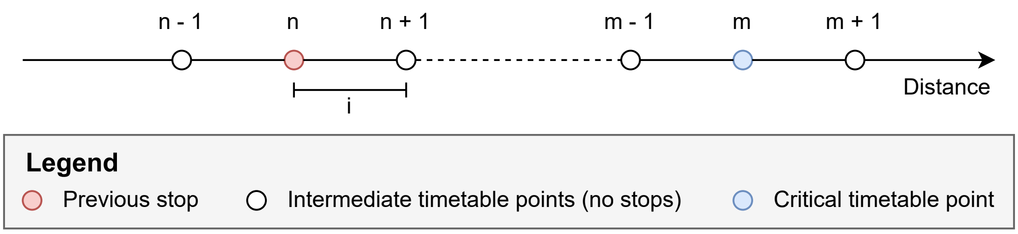


Figure D.3: Summation range of scheduled running time supplement

E
Case Study details

This appendix provides more details on the the case study infrastructure.

E.1. Infrastructure

Figure E.1 shows the infrastructure of the case study area, and Table E.1 presents all timetable points in the case study area with its abbreviation and type. In this table, two categories are distinguished by a line. The top category contains the timetable points within the scope, whilst the bottom category states which timetable points are adjacent to the scope.

	Timetable point	**Abbreviation**	**Type timetable point**
Timetable points inside case study area	Haarlem	Hlm	Station
	Haarlem Goederenstation	Hg	Shunting yard
	Haarlem Zuidelijke Splitsing	Zspl	Crossover
	Heemstede-Aerdenhout	Had	Halt
	Hillegom	Hil	Halt
	Hoofddorp	Hfd	Station
	Hoofddorp Opstelterrein	Hfdo	Shunting yard
	Hoofddorp Middenspoor	Hfdm	Crossover
	Leiden Centraal	Ledn	Station
	Lisse	Lis	Crossover
	Nieuw Vennep	Nvp	Halt
	Noordwijkerhout	Nwh	Junction
	Ringvaartbrug (brug o/d - bij Sassenheim)	Rvbr	Bridge
	Sassenheim	Ssh	Halt
	Schiphol Airport	Shl	Station
	Voorhout	Vh	Halt
Timetable points out-side case study area	Amsterdam Riekerpolder Aansluiting	Asra	Junction
	Galgewater (brug o/h - bij Leiden)	Gwt	Bridge
	Haarlem Noordelijke Splitsing	Nspl	Crossover
	Haarlemkruis	Hlmkr	Crossover
	HSL - Hoofddorp Overloop	Hshfdo	Junction
	Spaarne (brug o/h - bij Haarlem)	Spbr	Bridge
	Vinkbrug (o/d Rijn bij Leiden)	Vkbr	Bridge

Table E.1: Timetable points within the case study area

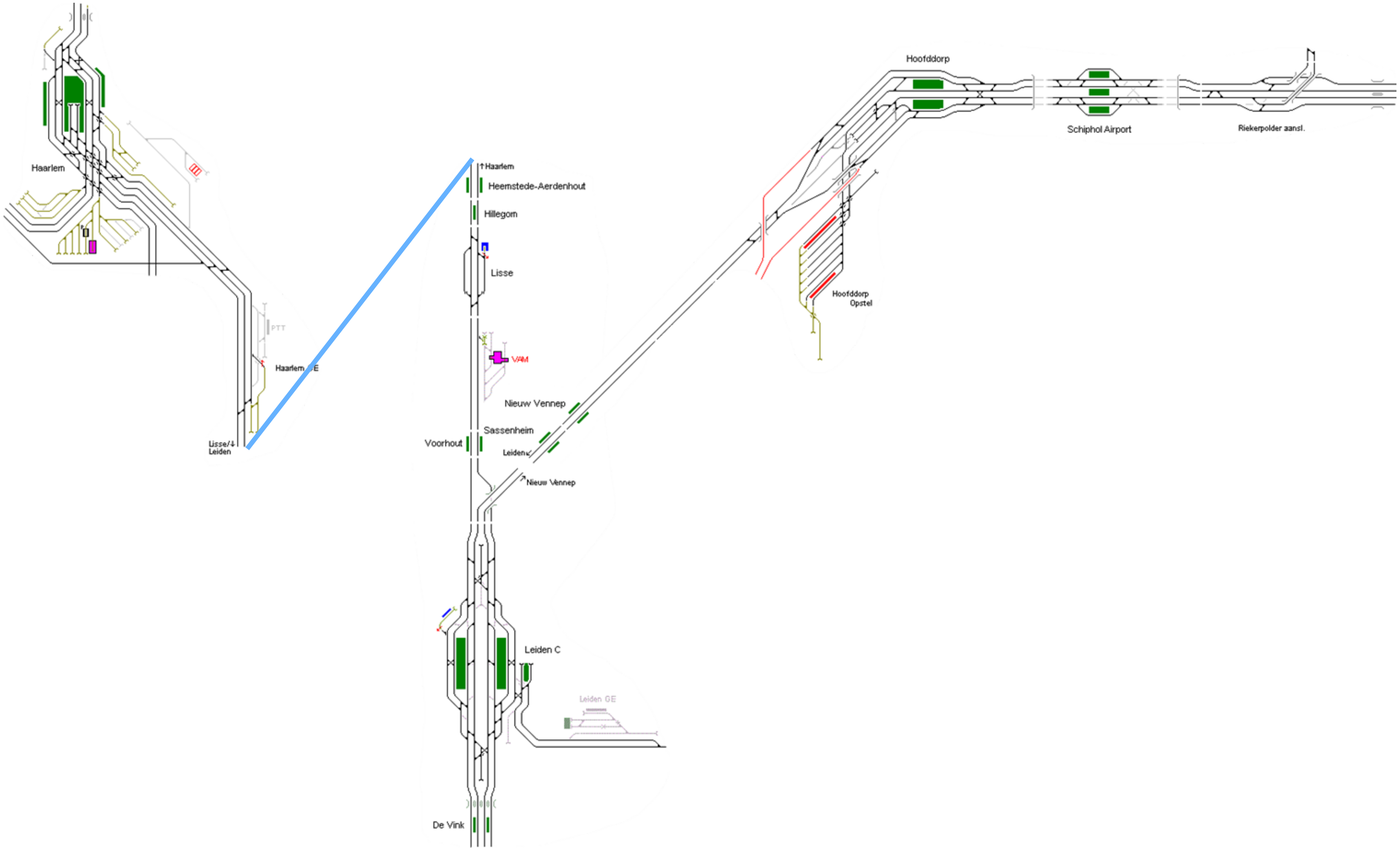
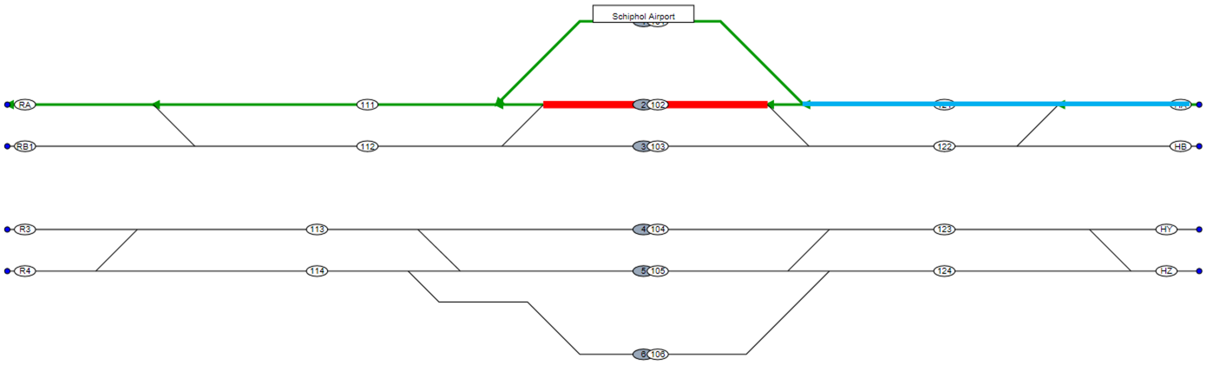
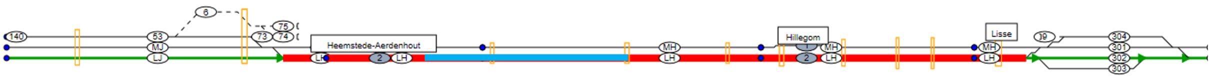


Figure E.1: Track layout of the case study area. Retrieved on SporenplanOnline (2022)

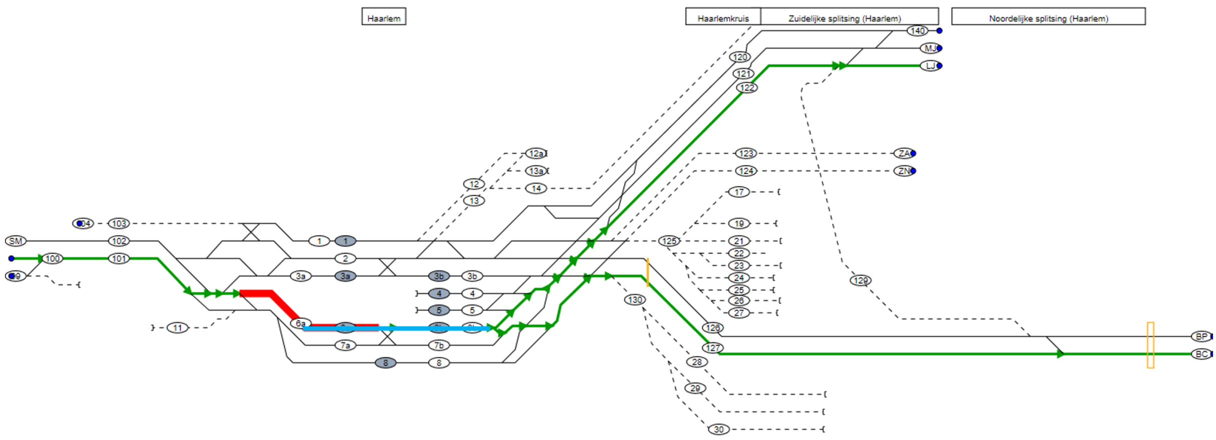
E.2. Track diagrams of example conflicts



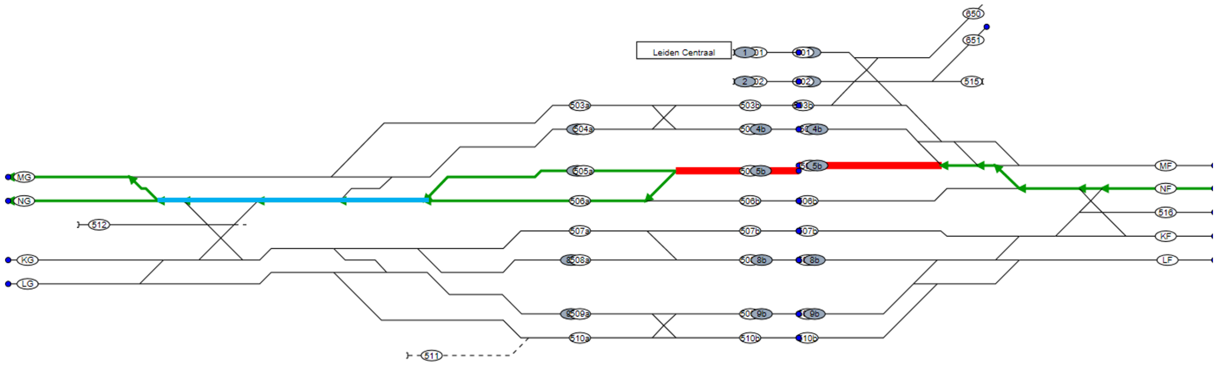
(a) Schiphol



(b) Heemstede-Aerdenhout



(c) Haarlem



(d) Leiden

Figure E.2: Track diagram with the critical conflict section in blue

F
Results

This appendix provides more details on the results of the case study.

F.1. Descriptive statistics of target and predictor variables

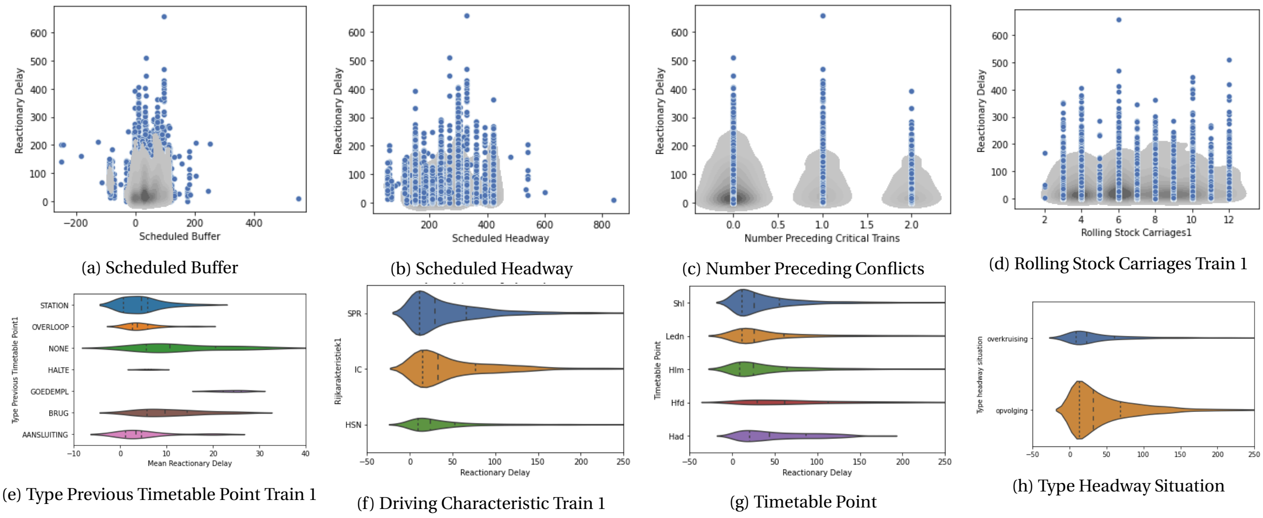
Variable name		Number of observations	Mean	Median	Standard deviation	Minimum	Maximum
Secondary Delay		61753	7.32	0.00	27.45	0.00	657.00
Hindrance Percentage		72047	13.81	10.66	11.79	0.00	46.27
Number Preceding Critical Trains		72047	0.38	0.00	0.69	0.00	3.00
Scheduled Buffer		72047	57.79	56.00	42.16	-250.00	658.00
Scheduled Headway		72047	246.65	240.00	83.35	36.00	840.00
Technical Minimum Headway		72047	188.86	175.00	70.72	-3.00	358.00
Rolling Stock Carriages	Train 1	72019	6.64	6.00	2.42	2.00	12.00
	Train 2	72012	6.39	6.00	2.45	2.00	12.00
Scheduled Running Time Supplement	Train 1	72047	7.03	6.12	4.83	-0.83	20.92
	Train 2	72047	9.37	6.27	8.98	-0.83	35.14

Variable name		Number of observations	Mode	Frequency	Number of unique categories
Same Direction		72047	Yes	65620	2
Type Headway Situation		72047	Succession	55746	2
Type Timetable Point		72047	Station	67805	2
Event	Train 1	72047	Arrival	42722	2
	Train 2	72047	Arrival	65208	2
Carrier	Train 1	72047	NSR	64470	2
	Train 2	72047	NSR	63725	2
Driving Characteristic	Train 1	72047	IC	32988	3
	Train 2	72047	SPR	39783	3
Previous Timetable Point	Train 1	72047	Hfd	16393	11
	Train 2	72047	Hfd	18190	13
Next Timetable Point	Train 1	72047	Asra	17847	12
	Train 2	72047	Asra	18032	12
Rolling Stock Type	Train 1	72019	VIRM6	8629	62
	Train 2	72012	SW7-25KV2+7	6862	62
Timetable point		72047	Shl	32148	5
Train Series	Train 1	72047	2200	10010	17
	Train 2	72047	6300	10200	18
Train Series Pattern	Train 1	72047	B1800	3986	69
	Train 2	72047	B5800	2819	68
Type Previous Timetable Point	Train 1	72047	Station	16393	7
	Train 2	72047	Station	18190	7
Type Next Timetable Point	Train 1	72047	Junction	17847	7
	Train 2	72047	Junction	18032	7

Table F.1: Overall statistics

F.2. Relationships with secondary delay

Figure F.1 shows the relationships between secondary delay and some of the candidate predictor variables. A quick glance at the figures does not immediately show an evident relation. One explanation for this observation is that the secondary delay is influenced by other factors, not included in this research, such as weather information, passenger loads or technical malfunctions. There could simply be too much randomness in the variable secondary delay for the candidate predictors to show a clear relationship. Therefore, an approach similar to one for hindrance percentage is taken where for each conflict, one aggregated value for secondary delay is computed; the mean. It is expected that this leads to a metric that reveals the more structural relationships between secondary delay and its candidate predictors.



(a) Scheduled Buffer (b) Scheduled Headway (c) Number Preceding Conflicts (d) Rolling Stock Carriages Train 1

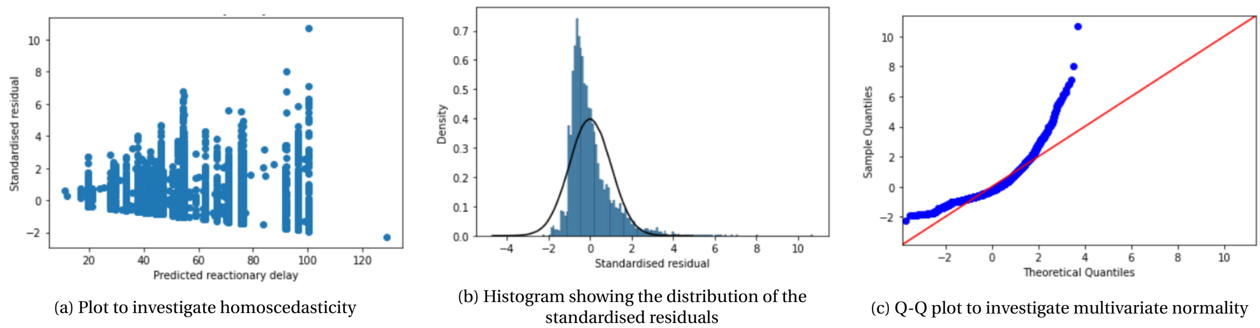
(e) Type Previous Timetable Point Train 1 (f) Driving Characteristic Train 1 (g) Timetable Point (h) Type Headway Situation

Figure F.1: Relationships with secondary delay

F.3. Results of check on assumptions for secondary delay

Figure F.2a shows that heteroscedasticity is present, since the standardised residuals become more spread out as the predicted values get larger. As for the multivariate normality, Figure F.2b and Figure F.2c show that the standardised residuals do not completely follow a normal distribution. However, the mean of the standardised residuals equals 2.35e-13, which is very close to zero, and with a d-statistic of 1.83 the Durbin-Watson test concludes that the observations are independent. Therefore, it is acceptable that the assumption of multivariate normality is not met.

Thus, only the assumption of homoscedasticity is not met. Transforming the target variable - taking the square root, natural logarithm, logarithm with base ten and quadratic were tried - has not solved the problem. For that reason it is decided to create one value for secondary delay per conflict; the mean. From Figure 5.10 and a d-statistic of 2.00 it can be concluded that this data reduction does not violate any of the assumptions.



(a) Plot to investigate homoscedasticity (b) Histogram showing the distribution of the standardised residuals (c) Q-Q plot to investigate multivariate normality

Figure F.2: Plots to check regression assumptions of the model on secondary delay (n = 9166)

F.4. Reference category

The reference levels of all categorical variables as used in the multiple regression analyses is presented in Table F.2.

Categorical predictor variable	Reference level
Same Direction	Yes
Type Headway Situation	Succession
Type Timetable Point	Station

Event Train 1	Departure
Event Train 2	Arrival
Carrier Train 1	NSR
Carrier Train 2	NSR
Driving Characteristic Train 1	SPR
Driving Characteristic Train 2	IC
Previous Timetable Point Train 1	Spbr
Previous Timetable Point Train 2	Spbr
Next Timetable Point Train 1	Nspl
Next Timetable Point Train 2	Zspl
Rolling Stock Type Train 1	SGMM3 SGMM3
Rolling Stock Type Train 2	VIRM6
Timetable Point	Hlm
Train Series Train 1	4800
Train Series Train 2	2100
Train Series Pattern Train 1	B4800
Train Series Pattern Train 2	D2100
Type Previous Timetable Point Train 1	Bridge
Type Previous Timetable Point Train 2	Bridge
Type Next Timetable Point Train 1	Crossover
Type Next Timetable Point Train 2	Crossover

Table F.2: Reference levels of all categorical variables

F.5. Full regression formulas

This section present the full regression formulas based on the output presented in Chapter 5.

$$
\begin{aligned}
\textit{Mean secondary delay} = -4.356 &+ 16.854 * \textit{Previous Timetable Point Train 2}_{Hfdm} \\
&- 15.127 * \textit{Previous Timetable Point Train 2}_{Vkbr} \\
&- 5.128 * \textit{Rolling Stock Type Train 1}_{SNG3} \\
&+ 2.587 * \textit{Rolling Stock Type Train 1}_{SNG4\ SNG4} \\
&+ 6.069 * \textit{Rolling Stock Type Train 2}_{SNG3\ SNG3} \\
&+ 0.167 * \textit{Scheduled Running Time Supplement Train 1} \\
&+ 0.036 * \textit{Technical Minimum Headway} \\
&- 2.652 * \textit{Train Series Pattern Train 1}_{B1800} \\
&+ 7.985 * \textit{Train Series Pattern Train 1}_{D3700} \\
&+ 7.703 * \textit{Train Series Pattern Train 1}_{D5800} \\
&+ 6.617 * \textit{Train Series Pattern Train 1}_{D6300} \\
&+ 11.724 * \textit{Train Series Pattern Train 2}_{B3700} \\
&- 6.582 * \textit{Train Series Pattern Train 2}_{E3300} \\
&+ 5.652 * \textit{Train Series Pattern Train 2}_{E4800} \\
&+ 5.991 * \textit{Train Series Pattern Train 2}_{G4800} \\
&+ 4.876 * \textit{Train Series Train 1}_{3400} \\
&+ 9.914 * \textit{Train Series Train 2}_{2400} \\
&+ 19.801 * \textit{Type Next Timetable Point Train 2}_{Halt} + \epsilon
\end{aligned}
\tag{F.1}
$$

$$
\begin{aligned}
\textit{Hindrance Percentage} = 12.773 &+ 13.822 * \textit{Driving Characteristic Train 1}_{HSN} \\
&+ 25.769 * \textit{Next Timetable Point Train 1}_{Shl} \\
&+ 9.939 * \textit{Next Timetable Point Train 1}_{Vh} \\
&+ 24.633 * \textit{Next Timetable Point Train 1}_{Zspl} \\
&+ 31.161 * \textit{Previous Timetable Point Train 1}_{Hg} \\
&+ 23.708 * \textit{Previous Timetable Point Train 2}_{None} \\
&+ 5.192 * \textit{Rolling Stock Type Train 2}_{SLT6} \\
&- 0.119 * \textit{Scheduled Buffer} \\
&+ 0.337 * \textit{Scheduled Running Time Supplement Train 1} \\
&- 6.699 * \textit{Train Series Pattern Train 1}_{A3300} \\
&+ 12.643 * \textit{Train Series Pattern Train 1}_{B2400} \\
&- 6.963 * \textit{Train Series Pattern Train 1}_{C3300} \\
&+ 24.019 * \textit{Train Series Pattern Train 1}_{D3700} \\
&+ 11.241 * \textit{Train Series Pattern Train 1}_{D6300} \\
&+ 17.067 * \textit{Train Series Pattern Train 2}_{B3700} \\
&+ 4.001 * \textit{Train Series Pattern Train 2}_{C1000} \\
&+ 8.352 * \textit{Train Series Train 1}_{4600} \\
&+ 13.322 * \textit{Train Series Train 2}_{2400} \\
&- 9.530 * \textit{Type Headway Situation}_{Crossing} + \epsilon
\end{aligned}
\tag{F.2}
$$