Master thesis - Applied Mathematics

Evaluating and improving the passenger punctuality of a timetable

Marjan Nelis

Faculty of Electrical Engineering, Mathematics & Computer Science (EEMCS)
Department of Discrete Mathematics and Mathematical Programming (DMMP)

Supervisors:
Prof. dr. M.J. Uetz (University of Twente)
Dr. G. Maróti (Nederlandse Spoorwegen)

Assessment committee:
Prof. dr. M.J. Uetz (University of Twente)
Dr. G. Maróti (Nederlandse Spoorwegen)
Dr. ir. W.R.W. Scheinhardt (University of Twente)

11th September 2017

UNIVERSITY OF TWENTE.

Abstract

Nederlandse Spoorwegen (NS) measures the reliability of their passenger train operations with passenger punctuality. Passenger punctuality is defined as the percentage of passengers whose journey was successful in terms of travel time and is measured after the execution of the timetable. Improving the reliability is a crucial point of focus for NS, and therefore, this research focuses on improving timetables using passenger punctuality as measure. Firstly, we develop an evaluation model to give a measure of the passenger punctuality for any timetable before its execution. The data, obtained through the evaluation model, is also used for further analysis of the timetable on, for example, missed transfers. Secondly, we use the existing Stochastic Optimization Model (SOM) to improve timetables. Originally, the objective of SOM is to minimize the train delay. In this research, we adapt SOM to make customized objective functions possible. By using an objective function based on the passengers in a case study, we obtain timetables which show significant improvements with respect to the passenger punctuality. In the case study, we increased the 5-minute passenger punctuality from 57.31% to 68.90%. Even though the current model is not applicable on the whole railway network, this thesis provides a proof-of-concept for the used approach.

UNIVERSITY OF TWENTE.

Contents

1 Introduction

This first chapter provides an introduction to several subjects which are essential for a full understanding the context of the problem at hand. Firstly, the Dutch Railways and passenger punctuality are briefly introduced in Section 1.1 and Section 1.2 respectively. This is followed by a description of delays in the operation and a relevant model in Section 1.3 and Section 1.4. In Section 1.5, the problem description and research question are presented. Finally, the outline of the thesis is given in Section 1.6.

1.1 Dutch Railways and their performance

In the Netherlands, the train is a common way of traveling. The rail network has a length of more than 3000 kilometers of track and is used daily by over 1.1 million passengers. It is one of the busiest railway networks of Europe. There are several companies responsible for the train operations, including *Nederlandse Spoorwegen* (in English, Dutch Railways, or abbriviated as NS), Arriva, Connexxion and Syntus. The largest operator is NS; which is responsible for operation the passenger trains on the Main Rail Network (HRN). NS aims to deliver high quality services for the passengers including a reliable timetable, sufficient and comfortable materials and adequate response to disturbances (CBS, 2016; NS, 2016; ProRail, 2017).

NS is held accountable for its performance to the Dutch government. The performance is measured by a set of Key Performance Indicators (KPIs) such as general customer opinion, passenger punctuality and seating chance. For all the KPIs, target values for the upcoming years are recorded in an agreement with the Dutch government. If the target values are not achieved, there could be negative consequences for NS, such as fines or even losing its transport concession.

1.2 Passenger punctuality

Passengers punctuality is a KPI, which gives a measure for the reliability of the NS' services. It is defined as the percentage of passengers for whom the journey, in terms of travel time, was successful. There are three variations of passenger punctuality stated in the agreement of NS with the government (NS, 2016).

- Passenger punctuality 5 minutes HRN; i.e. the percentage of passengers arriving at their destination within 5 minutes of their promised arrival times. Here the punctuality is only measured for journeys on the main rail network.
- Passenger punctuality 15 minutes HRN; i.e. the percentage of passengers arriving within 15 minutes of their promised arrival times. Here the punctuality is only measured for journeys on the main rail network.
- Passenger punctuality 5 minutes High Speed Line (HSL); i.e. the percentage of passengers arriving within 5 minutes of their promised arrival times. Here the punctuality is only measured for journeys on the high speed line.

For the trains, a schedule is made which includes the planned arrival and departure times at train stations, this schedule is denoted as the *timetable*. The actual arrival and departure times in the execution of the timetable may differ from the planned times. The *realization data* contains the actual arrival and departure times of the trains, from which the delay of each passenger can be derived. The passenger punctuality is calculated by dividing the number of passengers whose delay is smaller than the threshold of time (either 5 or 15 minutes) by the total number of passengers.

For each of the three passenger punctuality KPIs, Table 1.1 gives an overview of the results of previous years and the target values for 2019. Up until 2017, none of these KPIs have reached their target value. In April 2017, NS made the headlines due to their poor performance on the HSL. NS received a fine of €500.000 from the State Secretary of the Ministry of Infrastructure and Environment, because they had not reached the bottom value for the 5 minute passenger punctuality on the HSL in 2016 (Dijksma, 2017).

The values of both 5 minute KPIs, as noted in Table 1.1, need to increase significantly to reach their target values in 2019. The magnitude of the needed improvement is several tenths of one percent Maróti (2016). As mentioned before, passengers punctuality KPIs give a measure for the reliability of the NS'

Table 1.1: Overview of the passenger punctuality KPIs (NS, 2016).

	5 min. HRN	15 min. HRN	5 min. HSL
Realization 2015	90.0%	97.0%	81.8%
Realization 2016	90.5%	97.1%	80.5%
Progression value 2017	90.5%	97.0%	82.5%
Target value 2019	91.3%	97.3%	84.3%

services. Therefore, improving the reliability of the trains is a crucial point of focus for the following years in order to achieve the target values in 2019.

1.3 Delays in the operation

In order to improve the future passenger punctuality KPIs, analysis of the root cause, delayed trains, is required. The trains either caused the delayed arrivals of the passengers directly or induced missed transfers. Two types of delays are defined:

- *Primary delays* follow directly from disturbances that occur during the execution of the timetable. These disturbances can have many sources, either from within the railway system or its environment. As a consequence, processes, such as movements and stops, take longer to execute than planned and delays may occur.
- *Secondary delays*, also called knock-on delays, are delays caused by earlier delays in the system. In the busy railway network of the Netherlands, the major part of delays are secondary delays.

The design of the timetable influences the impact of both primary and secondary delays on the system by the following reasoning. A movement of a train between locations and the dwelling of a train at a station is noted as a process. Each process within the timetable has a certain *technically minimum process time*, which is the minimal time that is necessary to execute the process. On top of these minimum process times, *time supplements* are added. This additional time is scheduled in the timetable to (partially) absorb delays and reduce knock-on delays. The location and magnitude of the time supplements affects the performance of the timetable (Vromans, 2005).

1.4 Stochastic Optimization Model

The Stochastic Optimization Model (SOM) by Kroon et al. (2008) is designed to redistribute the time supplements in a timetable to improve its performance. Kroon et al. (2008) improve the robustness of the timetable by minimizing the expected train delay. By minimizing the average weighted train delay the authors aim to achieve improvement with respect to train punctuality. Therefore, SOM could be a useful tool to improve the passenger punctuality in our situation.

An improvement in train punctuality could mean improvement in passenger punctuality, but this can not be guaranteed as stated by Nielsen et al. (2009). In Section 2.4 it is shown in several cases that the passenger punctuality can be higher, lower or equal to the train punctuality depending on the delays of the trains.

1.5 Problem definition

Since NS strives for to deliver high quality services and wants to meet the target values of the KPIs in 2019, they should use a timetable which is optimized to reach their goals. Currently, the overall performance of NS' timetable with respect to the passenger punctuality can only be measured after it has been executed. Also, the comparison of the timetable's performances is complicated by external influence, such as the behavior of passengers, weather conditions, accidents and defects in rolling stock. Thus, a timetable can perform better compared to another timetable because there were simply less negative influences during its execution. This makes it practically impossible to test and compare timetables in order to find a better time supplement allocation. These circumstances make it is essential to evaluate the performance of timetables beforehand under similar conditions.

The first goal of this thesis is filling one of these gaps by developing and test a model which can evaluate the performance of timetables under similar conditions. Such a model can provide a measure for the passenger punctuality and will contribute to a simple method of comparing timetables and analyzing where the supplement needs to be allocated for the best performance, without actually executing the timetables. To recreate realistic conditions in such a model, for both the realization data of the trains and the data of the passengers, proper assumptions and predictions have to be made based on data that is available.

With the evaluation model, a second goal is set for this thesis: improve a given timetable with respect to passenger punctuality. SOM will be used to reach this goal. Instead of evaluating the expected average weighted train delay, the performance of a timetable improved by SOM should be evaluated by the new evaluation model. This lead to the following main research question:

How can timetables be improved using passenger punctuality as measure?

In order to answer the main research question properly, several sub questions are formulated. The research is divided into two parts. The first part focuses on developing a measuring tool of passenger punctuality for a given timetable. This part starts with a review about passenger punctuality. The first and second sub questions for this part are therefore formulated as follows:

1. How is passenger punctuality calculated? What input is necessary?
2. How can passenger punctuality be used to develop a new performance measure for an unrealized timetable?

In the second part of the research, focus is on improving timetables using passenger punctuality as performance indicator. SOM will be used as the base of the proposed model. In this part the third and fourth sub questions will be answered:

3. What is SOM and what are its limitations?
4. How can passenger punctuality be used in SOM to improve timetables?

1.6 Thesis outline

The outline of this thesis is as follows. In Chapter 2, the calculation of passenger punctuality, as defined by NS, is elaborated. In this chapter, the difference between train punctuality and passenger punctuality is also illustrated in some theoretical situations. Chapter 3 describes SOM in more detail. Chapter 4 defines the evaluation model which is able to give a measure of the passenger punctuality of a given timetable. Chapter 5 presents the computational results of the evaluation model applied on a case. Chapter 6 describes how SOM is adjusted to improve passenger punctuality. The results of the adjustments are presented and analyzed. Chapter 7 provides a discussion. The thesis is concluded in Chapter 8.

2 Calculation of passenger punctuality

In Chapter 1, passenger punctuality is defined as the percentage of passengers which arrive at their final destination within a certain threshold of time from the promised arrival time. This chapter explains how the passenger punctuality is calculated by NS, and we refer to this system as *NS' system*. NS' system functions as a guideline for the proposed evaluation model in Chapter 4; the evaluation model simulate NS' system for any given timetable to obtain a measure for the passenger punctuality.

To introduce the terminology, we start with an example of the journey of a single passenger in Section 2.1. In Section 2.2, the modules in NS' system are explained. Some of the assumptions made in the calculation are briefly discussed in Section 2.3. In Section 2.4, the difference between train punctuality and passenger punctuality is shown in several theoretical situations. Finally, the chapter is concluded and summarized in Section 2.5.

2.1 Example case

To illustrate how the delay of each passenger is determined, this section provides an example with a single passenger. In the Netherlands, the *OV-chipkaart* is the obligatory payment method for public transport. Passengers check in with this card at their departure station and check out at their destination station. From the data of the Check Ins (CI) and Check Outs (CO), abbreviated to CICO data, the cost of the journey is determined. The same CICO data is used for the calculation of the punctuality of the passengers.

2.1.1 From CICO data to travel promise

The particular passenger in this example travels from station Nuth to station Schiphol Airport in the morning of September 27th. The journey produces the following CICO data, see Table 2.1.

Table 2.1: The CICO data of the passenger.

CI station	Nuth
CI date and time	27 September 2016 07:35
CO station	Schiphol Airport
CO date and time	27 September 2016 10:40

The first step is to find the path the passenger followed: to determine which trains the passenger used in its journey. Taking into account the standard *check in margin* of 1 minute (due to the distance between the check in poles and the platforms), the departure time will be 07:36. The passenger wants to travel as fast as possible from Nuth to Schiphol Airport after 07:36. For this, the travel planner of NS is utilized.

The departure time 07:36, the CI_station Nuth and the CO_station Schiphol Airport are offered to the *travel planner*. The travel planner provides several travel possibilities, each with different departure times or different paths. These travel possibilities are denoted as *travel options*. In Figure 2.1, the option which promises the earliest arrival at the destination is displayed. This travel option is further on denoted as the *travel promise* of the passenger and provides the promised arrival time of the passenger.

2.1.2 Realization of the travel promise

The next step is analyzing the journey. The travel promise can be divided into three *travel parts* where each part is a different train, the parts are separated by the transfers. The travel promise provided the train data as they were planned in the timetable. From the realization data of the trains that morning, we obtain the following data about the travel parts, as represented in Table 2.2.

- Travel part 1 had a arrival delay of 6 minutes, its arrival time is 08:15.
- Travel part 2 had a departure delay of 5 minutes, its departure time is 08:18.
- Travel part 3 had no delay.

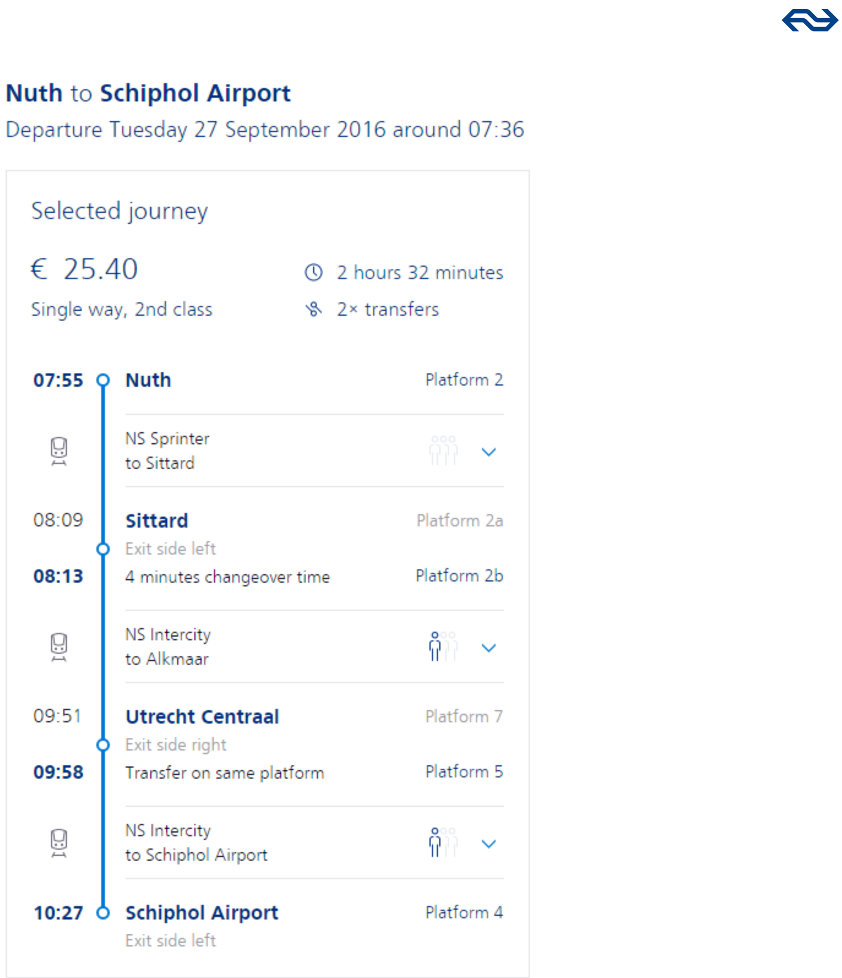


Figure 2.1: The selected travel option: the travel promise (NS, 2017).

Table 2.2: Overview of the three travel parts.

	Departure station	Departure time	Arrival station	Arrival time
1	Nuth	07:54	Sittard	08:09 **+6**
2	Sittard	08:13 **+5**	Utrecht	09:51
3	Utrecht	09:58	Schiphol Airport	10:27

For each travel part, the departure delay is checked. If the departure delay is larger than 15 minutes, the journey is rescheduled. In this example none of the travel parts experience such a delay. What remains is checking the transfers. If a transfer is missed, the journey has to be rescheduled as well. The travel promise of this passenger contains the next two transfers.

- Travel part 1 arrived at 08:15 at Sittard, travel part 2 departed from Sittard at 08:18. With 3 minutes of changeover time this was a successful transfer.
- In Utrecht travel part 2 arrived without delay and part 3 departs without delay. The transfer was executed as promised, and hence, successful.

As both transfers were successful, the arrival time of the passenger at Schiphol Airport equals the arrival time of the final travel part, i.e., at 10:27. The delay of the passenger is zero minutes and the passenger is considered punctual.

2.2 Modules of the NS' system

The system, which calculates the passenger punctuality, determines the delay and therefore the punctuality of all passengers at once. The algorithm behind the calculation of passenger punctuality follows the overview chart depicted in Figure 2.2. The blue blocks represent the input data of the system, the yellow blocks represent the modules in which the calculations take place and the remaining light gray blocks represent output of the modules. In Table 2.3, a summary is given of the modules. The function of each module is explained in the Section 2.2.1 to Section 2.2.5.

Table 2.3: Modules of the NS' system. (van der Berg, 2015)

Module	Function
1	Retrieve and save travel options
2	Determine realization
3	Rescheduling, determine new travel plan
4	Connect CICO travels to travel options
5	Generate Datamart

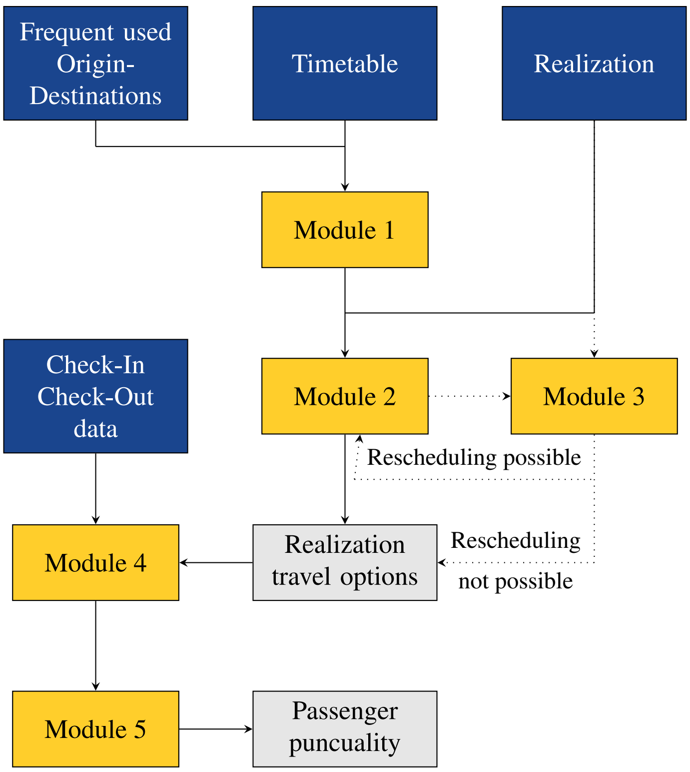


Figure 2.2: A simplified visualization of the system which calculates the passenger punctuality. The blue blocks represent the input data of the system, the yellow blocks represent the modules in which the calculations take place and the remaining light gray blocks represent output of the modules. (Wolters, 2016; van der Berg, 2015)

2.2.1 Module 1: Retrieve and save travel options

The first module is designed to determine all possible travel options. Instead of finding the travel options and a travel promise for each passenger separately, like the example case, it determines all possible travel options at once and connect them later to the passengers.

First, the module determines the list of frequently used *Origin-Destination* combinations (ODs). The system only considers ODs which are frequently used by passengers, ODs which are never or barely used are neglected. The list of frequent ODs is determined based on the CICO data from the previous 100 days. An OD is taken into account if it appear at least 100 times over the last 100 days and on at least 20 different day over the last 100 days. The list is updated four times a year. This way the system includes new stations and neglects stations that are no longer used. It reduces the number of ODs by approximately 70% while preserving over 98% of the journeys (van der Berg, 2015).

Once the frequent ODs are determined, the corresponding travel options are retrieved from the journey planner. The request to the journey planner results in a large number of options.Each unique travel option is identified and saved for the next module.

2.2.2 Module 2: Determine realization

Module 2 determines the realization of all travel options through their corresponding travel parts. This is similar to the process in Section 2.1.2 where the travel is analyzed. The realization is determined with the actual departure and arrive times of the train at stations, this data is denoted as the *realization data*. For each travel part is determined:

- if the train departed and at what time the train departed,
- if the train arrived and at what time the train arrived.

Once the realization of the travel parts is known, the realization of the travel options is determined. Travel options with missed transfers or severe delays require rescheduling before the delay can be determined. These travel options are sent to module 3 and this module returns the rescheduled travel options to module 2.

For every (rescheduled) travel option the realized arrival times are determined by the actual arrival time of the final travel part, like the example case in Section 2.1.2. The outcome of module 2 is a database with every travel option combined with the realization data.

2.2.3 Module 3: Rescheduling

Module 3, as mentioned in Section 2.2.2, takes care of travel options that require rescheduling. These are the travel options with missed transfers or with severe delays. Module 3 returns the rescheduled option to module 2 were its realization is determined.

In the current system, no rescheduling takes place; the checkout time minus a predetermined margin is used to estimate the realized arrival time. This makes the whole rescheduling process easier, but less accurate due to the behavior of the passengers. Passengers may remain longer at a station than the predetermined margin, this results in an overestimation of the delay.

2.2.4 Module 4: Connect CICO data to travel options

In module 4, the passengers from the CICO data are connected to the travel options in order to determine the delay of each passenger. The CICO data contains the origin, the destination, check-in time and check-out time of every trip made. Based on this information, a travel option is selected for each passenger, which is denoted as the travel promise of the passenger. This process corresponds with Section 2.1.1 of the example case, where the fastest journey from Nuth to Schiphol Airport was selected.

The connection of the CICO data to the travel options happens in such a way that the option with the earliest arrival time is selected. This happens according to the following procedure:

i. From the check-in time plus the check-in margin, all travel options of the next hours are collected. The check-in margin is equal to the specific check-out margin per station.
ii. All travel options for which a quicker alternative exists are deleted.
iii. Based on the departure time, the first option after the check-in time is selected as was determined by the journey planner.

After this, all passengers are connected to a travel promise from which the delay is already determined in module 2.

2.2.5 Module 5: Generate Datamart

In the fifth and last module, the data of the previous modules is exported in the correct format for NS. For each OD combination, the number of travels per day and the delay category is saved to determine the passenger punctuality KPIs.

For all ODs the number of passengers is counted and determined if their delay lies within the threshold. There are two KPIs with a threshold of 5 minutes and one with a threshold of 15 minutes. All measurements with a delay less than the threshold is divided by the total number of measurements resulting in the passenger punctuality.

2.3 Assumptions in the calculation

In the method of calculating passenger punctuality, 22 assumptions are made (see Table A.1 in Appendix A). These assumptions are about the behavior of the passengers and the correctness of the data. We are not going to elaborate on all these assumptions here. Note, however, that some of the assumptions made in the calculation are also used in the proposed evaluation model that will be presented in Chapter 4. The fourth item in Table A.1 provides a clear example; this assumption states that passengers want to go as quickly as possible from A to B. Therefore is chosen for a shortest path approach in the new evaluation model to determine travel promises.

2.4 Train punctuality versus passenger punctuality

In order to show that the passenger punctuality can differ form the train punctuality, two theoretical situations are sketched in Section 2.4.1 and Section 2.4.2. The train punctuality is defined as the percentage of trains which arrive at a station with a delay less than the given threshold. By adapting the train delays, the passenger punctuality can turn out to be higher, lower or equal to the train punctuality in each situation. An overview of the results is given in Section 2.4.3

2.4.1 A journey with a transfer

For the first situation, the network of Figure 2.3 is considered. It consists of three stations, two trains and one passenger. The passenger travels from A via B to C and transfers from train t_1 to train t_2 in station B. The transfer time in the schedule is determined as the walking time between the trains at station B plus 2 minutes buffer time. The thresholds for the passenger and train punctuality is set on 5 minutes. Thus the trips are considered unpunctual if their delay is 5 minutes or more. Furthermore, t_2 does not wait on t_1.

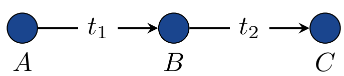


Figure 2.3: The network with a transfer.

In each of the following four cases, the train delay is altered in order to create scenarios with different train and passenger punctualities.

Case I

First, consider the trivial case where both trains have no delay. In this case, the transfer is successful and the passenger arrives at C with no delay. Both train and passenger punctuality are 100%.

Case II

Assume t_1 arrives at B with a delay of 3 minutes and t_2 departs from B as planned. The passenger cannot make the transfer to t_2 and will not arrive at station C. Both trains arrive within the threshold of 5 minutes at their destination, the train punctuality is therefore 100%. The passenger does not arrive at C at all, violating the punctuality threshold of 5 minutes. Therefore, the passenger punctuality is 0%.

Case III

Assume that both trains are delayed. Train 1 arrives at B with a delay of 5 minutes and train 2 departs with a delay of 4 minutes from B and arrives at C with the same delay. The passenger is able to transfer successfully at station B since the transfer time is the walking time plus one minute buffer time. Following the punctuality threshold of 5 minutes, train 1 is considered unpunctual and train 2 is considered punctual. The train punctuality is 50%. The passenger arrives at station C with 4 minutes delay, the passenger punctuality is 100%.

Case IV

Finally, consider for this network the situation where train 1 has no delay and train 2 has a delay of 6 minutes at both departure and arrival station. The transfer is made successfully. The train punctuality is 50% since the delay of train 2 is larger than the threshold. The passenger arrives with a delay of 6 minutes at station C, the passenger punctuality is 0%.

2.4.2 A journey without transfers

The second situation is slightly different from the first. It is illustrated in Figure 2.4. Here, we consider three stations, one train, and one passenger. The train departs from station A and moves via station B to station C, the passenger travels from A to C with this train. The differences in the following two cases are due to the measure points of the train delay. The train delay is measured at each arrival, that is in stations B and C in this situation while the passenger punctuality is only measured at the end of the passengers' journey.

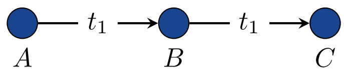


Figure 2.4: The network without a transfer.

Case V

Assume that the train arrives on time at B and with a delay of 6 minutes at C. The train has, with its two measure points a punctuality of 50%. The passenger arrives with a delay of 6 minutes at its destination. Therefore, the passenger punctuality is 0%.

Case VI

In the final case the train arrives with a delay of 6 minutes at B and with a delay of 4 minutes at C. Since the threshold is set on 5 minutes, the train punctuality is 50%. The passenger, who arrives with a delay of 4 minutes at its final destination, has a punctuality of 100%.

2.4.3 Overview of the cases

For the first situation with a transfer, the four cases show the importance and the influence of the transfers on the punctuality. A train punctuality of 100% does not guarantee a passenger punctuality of 100% and vice versa.

Since a large part of the passengers travels without any transfers, a situation without transfers is also discussed. Case V and VI show that, even in a situation without transfers, the passenger punctuality and train punctuality can be different due to the fixed measure points for the train punctuality.

Table 2.4: An overview of the punctualities is the situation with transfer (cases I-IV) and without transfer (cases V-VI).

	Train punctuality	Passenger punctuality
Case I	100%	100%
Case II	100%	0%
Case III	50%	100%
Case IV	50%	0%
Case V	50%	100%
Case VI	50%	0%

2.5 Summary

The main subject of this chapter is passenger punctuality. Using an example case, the punctuality was determined for a single passenger. This is followed by the elaboration of the NS' system, which calculates the passenger punctuality of all passengers. The system consists of the five modules:

- The first module determines all possible *travel options* in the timetable. These are the possible journeys passengers can use to travel between an *Origin-Destination* (OD) pair.
- The second module combines the realization data with the travel options to determine its delay, if possible.
- The third module reschedules the journey in case of large delays or missed transfers, the results of the third module are returned to module 2. The current system does not use dynamic rescheduling, instead the check out times are used to estimate the delay.
- The fourth module connects the passengers to travel option using their *Check-In* and *Check-Out data* (CICO data).
- The fifth and last module categorizes the delays of the passengers and determines the values of the passenger punctuality KPIs.

The chapter is finalized with a section describing two theoretical situations to show differences between the train punctuality and passenger punctuality. Depending on the train delay in the situations, the train punctuality is higher, lower or equal to the passenger punctuality as presented in Table 2.4.

3 Stochastic Optimization Model

The Stochastic Optimization Model (SOM), as shortly addressed in Chapter 1.4, is designed to redistribute the time supplements of a given timetable to improve its robustness against small disturbances. This improvement is established by minimizing the expected weighted train delay. SOM is created by Vromans (2005) and improved by Kroon et al. (2008). The improved model of Kroon et al. (2008) is implemented by the company Niks et al. (2015) and this implementation will be used in Chapter 6 to improve timetables.

In Section 3.1, SOM is described as defined by Kroon et al. (2008). Section 3.2 provides relevant implementation details of the implemented model.

3.1 Model description

SOM is divided into a timetabling part and a simulation part. The timetabling part determines the timetable and shows similarity with the well-known Periodic Event Scheduling Model by Serafini and Ukovich (1989). The simulation part runs independent realizations of the timetable, subject to stochastic primary disturbances, in order to evaluate the expected train delay of the the timetable.

Since SOM only considers small disturbances, we have the following assumptions: The first assumption is that the train order on the tracks is determined by the initial timetable and cannot be modified. Moreover, all planned connections between train are assumed to be maintained in the realizations.

3.1.1 Notation

A timetable consists of *processes* that have to be executed. For example, the movement of a train from one station to another, the dwelling period at a station, safety processes such as headway time between consecutive trains on the same track, and commercial processes like planned passenger transfers. The beginning and completion of a process are called *events*. The events correspond with the departure, arrival or crossing of a certain train at a given location.

NS uses a *cyclic timetable* with a cycle time T of one hour for their operations. Cyclic timetabling means that the timetable of longer periods can be created by copying a timetable of a single cycle. Therefore, the timetabling part of SOM only uses one cycle, the *One-Hour Timetable* (OHT), as input.

In an OHT, the set of processes P and the set of events E are defined. For each process $p \in P$, the events $b(p) \in E$ and $c(p) \in E$ denote the beginning and completion events, respectively. Furthermore, parameter m_p denotes the *technically minimum process time* of process $p \in P$, that is the minimum time necessary to execute process p.

The planned time of an event e is denoted by the parameter V_e. This parameter states the minute in the hour on which event e is planned. The beginning and completion time of a process p is denoted by $V_{b(p)}$ and $V_{c(p)}$, respectively. If a process is completed within an hour, the inequality $V_{b(p)} < V_{c(p)}$ holds. Otherwise, if a process crosses the end of an hour, $V_{b(p)} > V_{c(p)}$ holds. To separate these two cases, the binary parameter Q_p is introduced. This parameter records whether or not the process crosses the end of the hour. This results in the following definition:

$$Q_p = \begin{cases} 0 & \text{if } V_{b(p)} < V_{c(p)} \\ 1 & \text{if } V_{c(p)} < V_{b(p)} \end{cases}$$

In the improved timetable, the event times and time supplements are the decision variables. The new event time of event e is denoted by v_e and the time supplement on a process p is denoted by the decision variable s_p. The improved event times are not restricted to the time interval $[0, T-1]$ in order to prevent unwanted restrictions. The final event times in the optimized timetable are obtained by transferring the initial obtained final event times back to the interval $[0, T-1]$.

SOM evaluates expected average train delay of the timetable under construction through R simulations. In each simulation, the OHT is expanded to a day of H consecutive hours and the timetable is subject to a priori selected independent stochastic primary disturbances. The variable $\tilde{v}_{e,r,h}$ denotes the planned event time of event e in hour h of simulation r. The disturbance parameter $\delta_{p,r,h}$ denotes the primary disturbance on process p that starts in hour h of simulation r, for $p = 1, \ldots, P$, $r = 1, \ldots, R$ and $h = 1, \ldots, H$, so each process is subject to one primary disturbance. The decision variable $D_{e,r,h}$ denotes the delay, including both primary and secondary delays, on event e in hour h of simulation r. Furthermore, the train delay is only measured at the arrival events, the set of arrival events is denoted by E_a.

3.1.2 Timetabling part of the model

Using the notation as introduced in the previous section, the constraints of the time tabling part of SOM are formulated.

$$m_p + s_p = v_{c(p)} - v_{b(p)} + Q_p \cdot T \qquad \text{for } p = 1, \dots, P. \tag{1}$$

The right-hand side of equation (1) describes the time difference between the variables of the planned completion time and begin time of process p, here the possible crossing of the end of the hour is taken into account. Thus the right-hand side is the total planned process time of process p in the new timetable. The total planned process time is equal to the left-hand side: the technically minimum process time m_p, a parameter, plus the added time supplement s_p on process p.

The added time supplements are non-negative, therefore the parameter m_p is a lower bound for the total planned process time as described in the right-hand side of (1). Furthermore, there may be a upper bound u_p specified for process p. This results in the following constraints:

$$m_p \leq v_{c(p)} - v_{b(p)} + Q_p \cdot T \leq u_p \qquad \text{for } p = 1, \dots, P.$$

Let e_1 denote the first planned event in an hour on a location and e_2 denote the last planned event in an hour on the same location. To guarantee that the obtained timetable can be transferred back to the time interval $[0, T-1]$, the time difference between e_1 and e_2 should not exceed the cycle time T. This is important since the event times are not restricted to the time interval, as explained earlier.

$$0 \leq v_{e_2} - v_{e_1} \leq T - 1.$$

The following constraints set a maximum amount of time supplements on selected subsets of processes. Let there be N subsets $A_1, \dots, A_N$ selected, each subset A_n is connected to the maximum amount of time supplement S_n to be allocated in the corresponding subset.

$$\sum_{p \in A_n} s_p \leq S_n, \qquad \text{for } n = 1, \dots, N.$$

Finally, delays $D_{e,r,h}$ are considered non-negative and integrality constraints are imposed on the planned event times v_e and the variables time supplements s_p.

3.1.3 Simulation part of the model

In the simulation part of SOM, days of H consecutive hours are simulated. A process p with $Q_p = 0$ has $V_{b(p)} < V_{c(p)}$, thus it is planned within a single hour and ends in the hours with the same index as the hour it started in. A process p with $Q_p = 1$ has $V_{c(p)} < V_{b(p)}$, it ends in the hour with an index higher than the hour it started in: process p with $Q_p = 1$ starts in hour h of realization r at $\tilde{v}_{b(p),r,h}$, this process ends at $\tilde{v}_{c(p),r,h+1}$

The following constraints simulate the effects of both primary and secondary delays. The event times of the processes are linked to the technically minimum process times and primary disturbances. The primary delays are directly simulated by adding the disturbances $\delta_{p,r,h}$. The secondary delays are indirectly simulated: they follow from interacting processes between trains.

$$m_p + \delta_{p,r,h} \leq \tilde{v}_{c(p),r,h+Q_p} - \tilde{v}_{b(p),r,h} \qquad \text{for } p = 1, \dots P;\ r = 1, \dots, R;\ h = 1, \dots, H. \tag{2}$$

Furthermore, departure events are not allowed to occur too early. Here the set of departure events is denoted by E_d.

$$v_e + h \cdot T \leq \tilde{v}_{e,r,h} \qquad \forall e \in E_d;\ r = 1, \dots, R;\ h = 1, \dots, H,$$

The delay of the arrival events is determined by subtracting the simulated event time from the planned event time, as presented in the following constraints:

$$\tilde{v}_{e,r,h} - (v_e + h \cdot T) \leq D_{e,r,h} \qquad \forall e \in E_a;\ r = 1, \dots, R;\ h = 1, \dots, H.$$

3.1.4 Objective function

The objective of SOM is to minimize the average weighted delays of the trains. The objective is defined by Kroon et al. (2008) as

$$\text{minimize} \sum_{e \in E_a} \sum_{r=1}^{R} \sum_{h=1}^{H} w_e D_{e,r,h} / (|E_a| \cdot R \cdot H). \tag{3}$$

3.2 Implementation by ORTEC

In Chapter 6, SOM is used to improve timetables. We use SOM with both the implemented version of the objective function by Niks et al. (2015) as well as with an customized objective function. Therefore the implementation of the objective function by ORTEC is introduced in this section.

The objective function in the implementation by ORTEC is different from the objective function as presented in (3). In order to explain the differences, several new notations are introduced. Firstly, the subset of arrival events that are located at a predetermined measure points is denoted by E_m. Secondly, the subset of arrival events for which is indicated in the input that extra weight must be added is denoted by E_e. Finally, besides the delay variable $D_{e,r,h}$ an extra variable $DM_{e,r,h}$ is introduced which indicates the delay above a predetermined margin.

The objective function consists of eight differently weighted delays. The eight weighted delays are defined as follows:

$$\text{WD}_1 = w_1 \cdot \sum_{e \in E} \sum_{r=1}^{R} \sum_{h=1}^{H} \frac{DM_{e,r,h}}{|E| \cdot R \cdot H}, \qquad \text{WD}_2 = w_3 \cdot \sum_{e \in E} \sum_{r=1}^{R} \sum_{h=1}^{H} \frac{D_{e,r,h}}{|E| \cdot R \cdot H},$$

$$\text{WD}_3 = w_3 \cdot \sum_{e \in E_a} \sum_{r=1}^{R} \sum_{h=1}^{H} \frac{DM_{e,r,h}}{|E_a| \cdot R \cdot H}, \qquad \text{WD}_4 = w_4 \cdot \sum_{e \in E_a} \sum_{r=1}^{R} \sum_{h=1}^{H} \frac{D_{e,r,h}}{|E_a| \cdot R \cdot H},$$

$$\text{WD}_5 = w_5 \cdot \sum_{e \in E_m} \sum_{r=1}^{R} \sum_{h=1}^{H} \frac{DM_{e,r,h}}{|E_m| \cdot R \cdot H}, \qquad \text{WD}_6 = w_6 \cdot \sum_{e \in E_m} \sum_{r=1}^{R} \sum_{h=1}^{H} \frac{D_{e,r,h}}{|E_m| \cdot R \cdot H},$$

$$\text{WD}_7 = w_7 \cdot \sum_{e \in E_e} \sum_{r=1}^{R} \sum_{h=1}^{H} \frac{DM_{e,r,h}}{|E_e| \cdot R \cdot H}, \qquad \text{WD}_8 = w_8 \cdot \sum_{e \in E_e} \sum_{r=1}^{R} \sum_{h=1}^{H} \frac{D_{e,r,h}}{|E_e| \cdot R \cdot H},$$

The objective function in the implementation by Niks et al. (2015) is defined as

$$\text{minimize } \text{WD}_1 + \text{WD}_2 + \text{WD}_3 + \text{WD}_4 + \text{WD}_5 + \text{WD}_6 + \text{WD}_7 + \text{WD}_8. \tag{4}$$

4 Evaluation model

In this chapter, the model is described that we have built to evaluate timetables. The goal of the evaluation model is to give a measure for the passenger punctuality. This measure is obtained by simulating the delays of passengers under small disturbances. The passenger punctuality is extracted from these delays. The evaluation model aims to remain as close as possible to NS' system to determine passenger punctuality.

After a literature review in Section 4.1, the design of the evaluation model is stated and elaborated upon Section 4.2 to Section 4.6. The chapter is concluded in Section 4.7 by addressing several remarks concerning the evaluation model.

4.1 Literature review

Little research has been carried out that concerns modeling passenger delays. Landex (2008) and Nielsen et al. (2009) give an overview of passenger delay models and present their schedule-based route choice model (see Table 4.1). In the overview, the passenger delay models are categorized into generations. Landex (2008) and Nielsen et al. (2009) classify the simplest models as *0th generation* models, these are models used by railway companies and not reported in international literature. Nielsen et al. (2009) created the descriptions of these models based on interviews with the railway companies. Since this generation of models does not use schedule-based routes, the 0th generation models are disregarded. We start with a description of models Landex (2008) and Nielsen et al. (2009) classified as the *1st generation* passenger passenger delay models, followed by the *2nd* and *3rd generation*.

1st generation

The first models, which use the schedule-based routes, are denoted as the *1st generation* passenger delay models by Landex (2008). The core idea of these passenger delay models is that the passenger delay is modeled by calculating the optimal route in the planned timetable and the optimal route in the realized (potentially delayed) timetable. The passenger delay is calculated by taking the difference in time of the two routes. The optimality of the route in the realized timetable implies the assumption that passengers have knowledge of all present and future delays. This principle of passenger route choice is denoted as the *optimistic principle*.

The advantage of the 1st generation passenger delay models is that they take in account the passengers' route choices. Another advantage is that the models can easily be applied by running the standard route choice model on the realized timetable. Moreover, the entire trip (including transfers) is examined by the model, in contrast to simpler models from the 0th generation.

The disadvantage, as noted before, is that the optimal route choice model, which is used in the realized timetable, assumes that the present and future delays are known to the passengers. This results in an under-estimation of the passenger delay compared with the reality, because in reality, passengers often become aware of the delay during their trip.

2nd generation

The *2nd generation* of passenger delay models uses a number of simulations of the timetable with empirical or simulated delay distributions. The passengers choose their optimal route in each of the simulations. This generation also assumes that passengers have knowledge of future delays, the results of the simulations lead to a route choice taking into account the expected delay distribution. Moreover, a large amount of memory is required as each route choice for each simulation has to be stored.

3rd generation

The *3rd generation* is proposed in Landex (2008) and Nielsen et al. (2009). The authors assume that passengers plan their desired route according to the planned timetable and they argue that passengers may reconsider their route during the trip; the true passenger's route choice lies between the following two extremes.

- *Pessimistic principle*: Passengers may use the same route as planned, they do not act upon better routing opportunities along the route.
- *Optimistic principle*: The optimal path search is applied to determine the paths of the passengers in the realized timetable. This principle assumes that the passengers have knowledge about all present and future delay in the timetable and determine their optimal path accordingly.

The passengers in the 3rd generation model by Nielsen et al. (2009) reconsider their route if they experience a certain amount of delay. Moreover, if the desired route is no longer feasible due to delays, the passengers reconsider their route immediately. The reconsideration takes places at the point threshold is exceeded or the route is no longer feasible. This results in a more realistic model as compared to previous generations, but is more complicated to implement due to the rescheduling along the passengers' journeys.

Table 4.1: Overview of the characteristics of the passenger delay models (Nielsen et al., 2009).

	1st generation	**2nd generation**	**3rd generation**
Consideration of passenger delays	Partly	Yes	Yes
Complexity of the method	Medium	High	High
Needs of information on passengers demand	OD matrix	OD matrix	OD matrix
Passengers may predict delays in the future	Yes	Partly	Can be incorporated
Passengers may arrive before time if a better connection emerges	Yes	Yes	Yes
Accuracy	Low	Medium	High
Bias	Large underestimation of delays	No systematic bias	No systematic bias

4.2 Design of the evaluation model

The design choices of the evaluation model are mostly based on the design choices of NS' passenger punctuality system, because that is the system we want to simulate. In Chapter 2, the necessary input of NS' system is discussed. The only input data we have at our disposal is the timetable. The remaining input (frequent ODs, realization data, and CICO data) is not available, because these follow from the execution os the timetable. Alternative data must be used and that brings along limitations. Furthermore, the input of the simulations (disturbances) is chosen similar to the input of SOM. SOM evaluates the train delay using its own simulations, the evaluation model functions as a extension of SOM by evaluating passengers delay and punctuality in similar simulations. All assumptions, alternative approaches, and limitations are discussed in Section 4.7.

The modules of NS' system return in the following way into the evaluation model:

- Firstly, all possible paths between any combination of train stations are determined using the original timetable and connected to *passenger groups* (see Section 4.3.2). These paths are denoted as the *nominal paths*. This part corresponds with module 1 of NS' system.
- Secondly, a realization is simulated by adding disturbance to the timetable and in this realization all possible paths are determined again and connected to the passenger groups. These paths are denoted as the *realized paths*. This procedure is similar to modules 2 and 3.
- The difference between the nominal and the realized arrival times of each path is noted as the delay of the path. Finally, each path receives a weight based on the expected number of passengers on that path and the punctuality is calculated, like modules 4 and 5.

The next sections are organized as follows: Section 4.3 provides a description of the input data. Section 4.4 describes how this data is used to calculate the nominal paths of the passengers. The calculation of the realized paths is elaborated upon in Section 4.5. This is followed by the output of the model in Section 4.6. In Section 4.7, the limitations of the model are discussed.

4.3 Input data

The input of the evaluation model are the reference timetable, the *Origin-Destination* (OD) matrix, and a passenger distribution over the day. The timetable is used to determine the nominal and realized paths of the passengers. The OD matrix and passenger distribution are used to recreate the frequent ODs and CICO data. Furthermore, *passenger groups* are introduced simulate the passengers.

4.3.1 Reference timetable

The format of the reference timetable in the evaluation model is identical to the format used in SOM; they both use an *One-Hour Timetable* (OHT). This makes it possible to apply both models on the same timetable. We start by repeating the definitions concerning timetables, note that the evaluation model uses a different notation than SOM.

A timetable consists of *events* and *processes*, these are interpreted as follows. An event corresponds with the departure, the arrival, or a crossing of a given train at a given location. The processes of a train are either a movement between two locations or the dwelling at a station. Next to that, there are processes for safety and commercial issues, such as the headway restrictions between trains and transfer opportunities for the passengers.

The reference timetable is presented as a *cyclic timetable*. Cyclic timetables specify a longer period as repeated copies of shorter periods or cycles. In the case of NS, such a short period is an hour, i.e. the cycle time $T = 60$ minutes. All data, events and processes, are only specified for a single hour, the OHT. The length to which the cyclic timetable is extended in the evaluation model is determined by the parameter H, this is the number of cycles in the extension. In this section, a distinction is made between events and processes of the cyclic timetable and the extended *linear timetable*, they are noted as *cyclic events and processes* and *linear events and processes*, respectively.

The cyclic events are the basis of a timetable, they form the set $\mathscr{E}$. The cyclic timetable π assigns the cyclic events to time instants in the cycle. The planned time of cyclic event $\varepsilon \in \mathscr{E}$ in π satisfies $0 \leq \pi(\varepsilon) < T$.

The cyclic events are connected and restricted through the cyclic processes. Let $\mathscr{P}$ denote the set of cyclic processes. The events with which the processes start and end are given. Furthermore, *technically minimum process times* and the binary parameter Q (which indicates whether or not a process crosses the boundary of the hour) are provided. These parameters are used and explained in the construction of the realization in Section 4.5 and in the current section for the extension of the cyclic processes.

From cyclic to linear timetable

Before the travel options can be determined, the cyclic timetable must be extended to length of the desired number of hours, H. The evaluation model extends the cyclic timetable to a day consisting out of H copies, this is denoted as the *linear timetable*. This extension works as follows (Maróti, 2017). The set of linear events are defined as

$$E = \{(\varepsilon, h) \mid \varepsilon \in \mathscr{E}, 1 \leq h \leq H\}.$$

The properties of the linear events, location and event type, are duplicated from their corresponding cyclic events. The planned event times are given by a function $\tilde{\pi} : E \to \mathbb{Z}$ defined as

$$\tilde{\pi}(\varepsilon, h) = \pi(\varepsilon) + (h - 1) \cdot T.$$

Here $h \in [1, \ldots, H]$. Similarly, the set of cyclic processes $\mathscr{P}$ is extended to a set of linear processes P. The set P is defined by

$$\begin{aligned} P = \{((\varepsilon, h), (\varepsilon_*, h)) \quad &| \ (\varepsilon, \varepsilon_*) \in \mathscr{P}, Q(\varepsilon, \varepsilon_*) = 0 \text{ and } 1 \leq h \leq H\} \\ \cup \{((\varepsilon, h), (\varepsilon_*, h+1)) &| \ (\varepsilon, \varepsilon_*) \in \mathscr{P}, Q(\varepsilon, \varepsilon_*) = 1 \text{ and } 1 \leq h < H\}. \end{aligned}$$

As previously mentioned, Q indicates whether or not a process crosses the cycle border. If $Q = 0$, a process is planned within the cycle, here $((\varepsilon, h), (\varepsilon_*, h))$ represents the origin and destination event of the process. When $Q = 1$, the process crosses the border. The destination event of the process takes place in the next hour, $((\varepsilon, h), (\varepsilon_*, h+1))$ are the corresponding events in this case.

4.3.2 Passenger groups, OD matrix and passenger distribution

The evaluation model is designed to simulate traveling passengers. In the calculation of the passenger punctuality, the check in times of the passengers are used to determine the departure time of each passenger separately.

In order to simulate and simplify the check-in times of passengers, all passengers that depart within a given time interval are considered as one *passenger group*. The default length of the time interval is set to 15 minutes. So in the default case, the evaluation model considers four passenger groups per station per hour and this limits the number of nominal and realized paths the model has to save during the the calculations.

The passenger distribution and the OD matrix are used to determine expected size of the passenger groups. This works as follows. The OD matrix provides the expected number of passengers per day who travel between any given OD combination. This number is combined with the passenger distribution over the day. This distribution tells what percentage of the total number of passengers per day is expected to depart in each hour. The combination is denoted as the *origin-destination-time* (ODT) matrix, which gives the expected number of passengers for every hour in a day for every OD pair.

The passenger groups are combined with the ODT matrix to determine the size of each passenger group, this happens by dividing the expected number of passengers per hour over the number of time intervals per hour. For example, for a given OD pair 20 passengers are expected to depart between 12:00 and 13:00. In the case of a departure time interval of 15 minutes, four passenger groups are created with departure times 12:00, 12:15, 12:30 and 12:45 and each group contains five passengers. The evaluation model constructs nominal and realized paths for this set of departure times and gives each path a weight equal to the number of passengers in the corresponding departure group.

4.4 Nominal paths

The nominal paths provide the promised arrival times for a given departure station, departure time and destination. The evaluation model determines the nominal paths for any combination of used departure and destination stations, for selected departure times, as is explained in Section 4.3.2.

The assumption is made that passengers want to get as quickly as possible from origin to destination. This is in line with the assumptions in NS'system, see Table A.1. Thus, the nominal paths are determined by a shortest path algorithm.

Graph representation

The linear timetable can be interpreted as an acyclic directed graph on vertex set E and arc set P as defined in Section 4.3.1, this is denoted as the *nominal graph*. The length of the arcs is determined by the planned process time: the difference in planned event times from the events at the begin and end of a process. The graph is acyclic since all processes go forward in time. Each arc points from an earlier planned time instant to a later one.

The nominal graph representation of the timetable is used to determine all the nominal paths. To make transfers possible, additional arcs are added to the graph between successive arrival and departure events on the same station. These arcs make it possible for passengers to leave a train as it arrives at a station and transfer to any train that departs from that station at a later time. This construction is visualized in Figure 4.1.

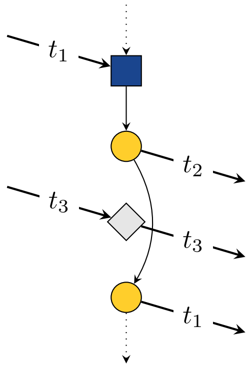


Figure 4.1: Transfer construction at a station. Different vertex types (event types) are highlighted by shape: square (arrival), circle (departure) and diamond (crossing).

In Figure 4.1, all events are located at the same station and sorted in chronological order starting at the top of the figure. The arcs between the arrival events (squares) and departure events (circles) represent potential transfers. In this example train t_1 arrives first at the station. While it waits, a second train t_2 departs and another train t_3 crosses the station without stopping. Passengers from t_1 can transfer to t_2, stay in t_1 or remain at the station. They are not able to transfer to t_3, since this train does not stop at the station.

The model solves the single-source shortest path problem to find the all possible paths. Since the graph is directed and acyclic, this can be done in linear time using topological sorting. From all possible paths, the paths for the chosen departure times are extracted and connected to the departure groups, these are the nominal paths.

4.5 Realized paths

The nominal paths, as described in the previous section, give the paths and corresponding arrival times in case the timetable is executed exactly as planned. In a realization, the processes are subject to stochastic disturbances and hence, processes may take longer than expected. Trains may experience delay due to these disturbances or due to the delay of previous trains, in accordance with the definitions of primary and secondary delays from Section 1.5. To simulate the primary delays in the model, the process time of a process p is defined as the technically minimum process time $l(p)$ plus a disruption $\delta(p)$.

For a given disruption and corresponding primary delays, the secondary delays are included by a dynamic program as displayed in the algorithm below (Maróti, 2017). An event e can be realized as soon as all its predecessors in the graph representation allow it to happen. Additionally, departure events are not allowed to happen before their planned event time. The dynamic program calculates the realized event times, which are denoted by $y(e)$.

forall $e_* \in E$ **do**
  if e_* *is a departure event* **then**
    $y(e_*) := \max\{\tilde{\pi}(e_*);\ \max\{y(e) + l(p) + \delta(p) \mid p = (e, e_*) \in P\}\};$
  else
    $y(e_*) := \max\{y(e) + l(p) + \delta(p) \mid p = (e, e_*) \in P\};$

A new graph is constructed to determine the realized paths, this graph is denoted as the *realized graph*. The realized graph uses the same vertex set E and arc set P as the nominal graph, only the length of the arcs and transfer arcs can turn out to be different: The length of the arcs is determined using the realized event times instead of the planned event times. Again, to make transfers possible, additional directed arcs are added to the graph between successive arrival and departure events on the same station. In the realization the chronological order of events can differ from the order the nominal graph, which results in different transfer possibilities.

In the calculation of passenger punctuality, NS assumes that passengers follow their planned route unless it requires rescheduling (van der Berg, 2015): It requires rescheduling if either a train departs with a delay larger than 15 minutes or when the planned route is no longer feasible because of, for example, a missed transfer. Therefore, the evaluation model uses the pessimistic principle regarding to the passengers' route choice in the realized timetable. The constraint that routes are reconsidered if a train departs with a delay larger than 15 minutes is omitted, since the evaluation model is designed to simulate only small delays and because test runs show that the train delays never exceed the 15 minute threshold. Thus, the route is only rescheduled when the nominal path is no longer feasible.

Rescheduling methods

If it is not possible to follow the nominal path in the realization, an alternative must be chosen. In NS' system, no dynamic rescheduling module is implemented. Instead of rescheduling the path of the passenger, the check-out time of the passenger is used to determine its arrival time at the check-out station, see Section 2.2.3. Unfortunately, the evaluation model cannot simulate the arrival times using the check-out data since such data is not available. The model has to reschedule the path using an alternative way. Due to the absence of a defined rescheduling procedure, various ways of rescheduling are considered below.

- *Optimistic rescheduling*: The easiest way is to determine an alternative for the whole path using the optimistic principle (Nielsen et al., 2009). Here, the shortest path that is found in the realization is chosen as alternative path. This is not the most realistic way, it suggests passengers can predict a missed transfer or delay and choose another path in advance. This method will likely result in an underestimation of the passenger delay (Nielsen et al., 2009).

- *Realistic rescheduling*: It would be more realistic to reschedule the path of a passenger from the point that the nominal path is no longer feasible in the realized graph, as suggested in the 3rd generation of passenger delay models (Nielsen et al., 2009). Therefore, the model follows the nominal path and finds the moment in time and station were a transfer is missed. From this station the shortest path in the realized graph is taken to the final destination to determine the final arrival time.

Both rescheduling methods are implemented into the evaluation model.

4.6 Output of the evaluation model

Once the paths, that are used in the realization, are determined, the delay of each path is calculated. The delay of a path is defined as the difference in arrival times of the nominal path and the used realized path. Given the reference timetable, ODT matrix, and the departure times of the departure groups, the model returns the average delay and punctuality. Besides the desired average passenger delay and passenger punctuality, the model also returns the average delay and train punctuality.

Train delay and punctuality

The average train delay is defined as the average delay of arrival events. The planned times $\tilde{\pi}(e)$ and realized event times $y(e)$ of these events are compared and the difference is noted as D_e, the delay of that event. So $D_e = \max\{0, y(e) - \tilde{\pi}(e)\}$. Let E_{arr} denote the subset of arrival events. The average train delay $\bar{D}_{\text{train}}$ is defined as

$$\bar{D}_{\text{train}} = \frac{\sum_{e \in E_{\text{arr}}} D_e}{|E_{\text{arr}}|}.$$

The train punctuality is determined for a 5 and 15 minute threshold. In order to calculate the punctualities, the binary function B is introduced.

$$B(D_e, x) = \begin{cases} 0 & \text{if } D_e \geq x, \\ 1 & \text{if } D_e < x. \end{cases}$$

The number of trains which arrive within the thresholds can be counted using B. The train punctualities, denoted by $P_{\text{train},5}$ and $P_{\text{train},15}$ respectively, are defined by

$$P_{\text{train},5} = \frac{\sum_{e \in E_{\text{arr}}} B(D_e, 5)}{|E_{\text{arr}}|} \quad \text{and} \quad P_{\text{train},15} = \frac{\sum_{e \in E_{\text{arr}}} B(D_e, 15)}{|E_{\text{arr}}|}.$$

Passenger delay and punctuality

The evaluation model calculates the delays per passenger group, as discussed in Section 4.3.2. Let the set of departure groups be denoted by G. For every group $g \in G$ the arrival delay D_g, where $D_g \geq 0$, and weight or group size w_g are stored. The average passenger delay is formulated as

$$\bar{D}_{\text{passenger}} = \frac{\sum_{g \in G} w_g \cdot D_g}{\sum_{g \in G} w_g}.$$

Here, the delay of a passenger group is weighted with corresponding number of passengers in that group. The sum of all weighted delays is divided by the total number of passengers. The punctualities, denoted by $P_{\text{passenger},5}$ and $P_{\text{passenger},15}$, are defined by

$$P_{\text{passenger},5} = \frac{\sum_{g \in G} w_g \cdot B(D_g, 5)}{\sum_{g \in G} w_g} \quad \text{and} \quad P_{\text{passenger},15} = \frac{\sum_{g \in G} w_g \cdot B(D_g, 15)}{\sum_{g \in G} w_g}.$$

In these equations the function B is used to determine whether or not each group delay lays within the specified threshold. The sum of the weighted outcomes of B corresponds with the total number of passengers who arrived within the threshold. Dividing this sum by the total number of passengers results in the passenger punctuality.

4.7 Limitations and assumptions

The evaluation model is a simplified simulation of the execution of a timetable. As mentioned in Section 4.2, most of the input data is not available. Moreover, the journey planner in NS' system is different from the shortest path design in the evaluation model, due to the transfer construction. In this section, all limitations and assumptions, that are made in the evaluation model, are discussed.

4.7.1 Input data

CICO data and frequent ODs

The number of passengers between each combination of stations is based on historic data; it is an estimation. Firstly the daily expected number of passengers is given by the OD matrix. The OD matrix, which is used in the evaluation model, is not the original classified matrix of NS but a modified version of it. This modification gives a rough idea of the expected number of passenger per day. The distribution of the passengers over the hours of a day is also based on data from the past. The combination of the two estimations provides the CICO data for the evaluation model: the check in time, check in station, and the check out station. Moreover, the frequent ODs are determined: stations without any passengers are neglected.

Note that in both the OD matrix and the passenger distribution make no distinctions are made between different days of the week. The number of passenger and distribution during workdays is different from weekends and holidays.

Disturbances

It is complicated to determine the size or even the distribution of disturbances. All available realization data from the trains shows the total delay, this includes both primary and secondary delays. An analysis of the total delays to find a distribution is already complicated, but it is even more complicated to separate primary and secondary delays within the total delay, and to extract the disturbances.

In accordance with SOM, the evaluation model only adds small scale disturbances. Therefore, the effects of large scale disturbances are not visible in the simulations. Large scale disturbances are not relevant for robust timetabling, this would include traffic control decisions in the model.

The distribution in the evaluation model is equal to the distribution in SOM; which is an exponential distribution. Detailed analyses of realization data in the Netherlands showed that late arrivals, departures, and dwell time prolongations, all fit well into exponential distributions (Vromans, 2005). However, these results are based on a limited period of realization data and on one location only. Due to the lack of further knowledge about disturbances, exponentially distributed disturbances are used in both SOM and the evaluation model.

The means of the exponentially distributed disturbances are determined and a maximum is set using the default settings for SOM. In the default setting of the case study, which is analyzed in Chapter 5, the disturbances over movement processes are exponentially distributed with a mean of 5% of the technically minimal process time and a maximum of 5 minutes. The disturbances over the dwelling processes are exponentially distributed with a mean of 30% of the technically minimal process time with a maximum of 2 minutes. The question remains whether or not these choice of parameters will result in realistic delays, this question is out of scope for this thesis.

4.7.2 Journey planner

Passengers' path choice

Another assumption in the evaluation model traces back to an assumption made in the calculation of the passenger punctuality, namely: the passenger wants to go as quickly as possible from A to B. Therefore, the model always chooses the shortest path, even if the passenger has a preference for another path (with less transfers for example).

Furthermore, NS has not defined a dynamic rescheduling method. Therefore, in this thesis, two rescheduling methods as stated in Section 4.5.1 are proposed and implemented into the evaluation model.

Zero-minute transfers

Another remarkable feature of the evaluation model is transfer construction. The used construction allows passengers to make a transfer as long as the transfer time is nonnegative. Zero-minute transfers are possible in the evaluation model. This simplifies the model but also makes some transfers possible that are impossible to make in real life.

An expansion of the zero-minute transfer structure is considered during the construction of the evaluation model. The following construction, described by Bast et al. (2016) and Müller-Hannemann et al. (2007) as a realistic time-expanded model, makes transfers with minimal transfer times possible. Instead of direct arcs between the arrival and departure vertices, as is illustrated in Figure 4.2, additional *transfer vertices* are added to the graph for each departure event, see Figure 4.3. Note here that the minimum transfer time is different per station, per transfer and even per passenger.

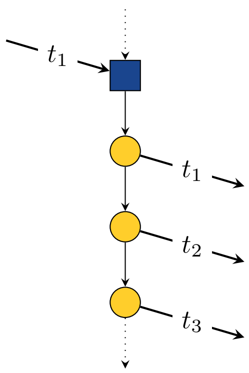


Figure 4.2: Current transfer construction at a station. Different vertex types (event types) are highlighted by shape: square (arrival) and circle (departure).

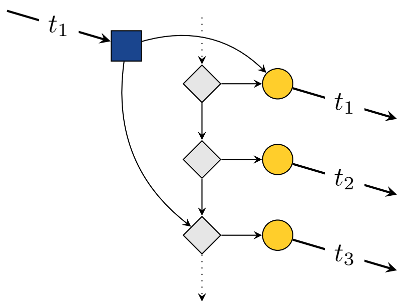


Figure 4.3: Expanded transfer construction at a station. Different vertex types (event types) are highlighted by shape: square (arrival), circle (departure) and diamond (transfer vertex).

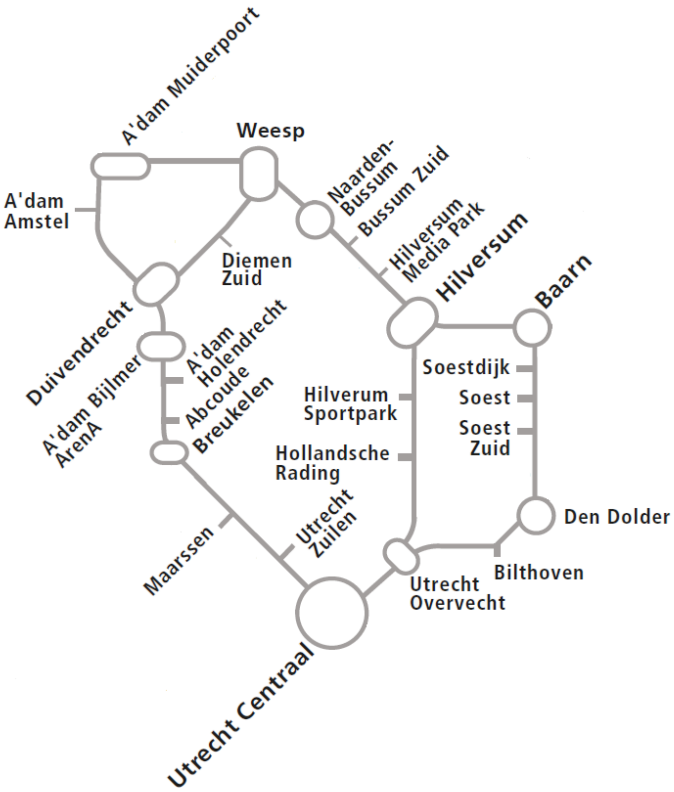
Both figures illustrate the same situation at a station. Train t_1 arrives at the station, followed by the departure of trains t_1, t_2, and t_3, in this order. Since the departures happen after the arrival, passengers are able to transfer to each of the three trains in the current model, see figure 4.2. The expanded transfer construction in Figure 4.3 makes it possible to add minimal transfer times to stations. Each arrival vertex is connected to the first transfer vertex that obeys the minimum change time constraints, that is the transfer node corresponding to the departure of t_3 in this example.

This expansion of the evaluation model will result into more realistic paths of the passengers, but creates a larger graph and additional research is necessary to determine the minimal transfer times. In the evaluation model is therefore chosen for the simple transfer construction.

5 Computational results and data analysis

In this chapter, a timetable is evaluated using the evaluation model (see Chapter 4). The used timetable contains only a small part of the Dutch railway network due to the limitations of the evaluation model as we describe in Section 7.

The computational results are obtained by implementing the evaluation model in MATLAB R2016b running on Windows 10. The hardware was an Intel Core i7 processor with a clock speed of 2.4 GHz and 8 GB internal memory.



We consider an area around Utrecht and a timetable constructed in 2013. The area contains only two intercity stations; some large passenger flows that are not fully contained in the area are excluded. As a consequence, the area will experience relatively more transfers in relation to the whole Dutch railway system.

For this case study, the evaluation model creates days with a length of 12 hours. The mean size of the disturbances on the processes is determined by the default parameters of the timetable. The results in this chapter are obtained by running 1000 simulations of 12-hours days, each with independent stochastic disturbances. These simulations took over 40 hours of calculation time, running four simulations parallel at all times.

In Section 5.1, the results of the Utrecht case are presented. In Section 5.2, the data from the model is further analyzed. Moreover, the results of simulations with different disturbance vectors are presented in Section 5.3. Finally, the chapter is concluded in Section 5.4 with a summary.

5.1 Results Utrecht case

This section discusses the results obtained for the Utrecht case described above. First, the results regarding the train delay and punctuality are presented in Section 5.1.1. Followed by the results regarding the passenger delay and punctuality in Section 5.1.2. Furthermore, the difference between the two rescheduling methods is discussed in Section 5.1.3.

5.1.1 Train delay and punctuality

The delay of the trains is measured at each arrival event, each measured with the same weight. In Table 5.1 the average delay and average punctualities are given together with the corresponding 95% confidence intervals (CI). The confidence intervals are determined using the Batch Means Method (Alexopoulos et al., 1997). Additionally, a histogram of the distribution of the train delays is presented in Appendix B, see Figure B.1.

Table 5.1: Results of trains at arrival with default parameters.

	Average	**95% CI**
Train delay (minutes)	2.1092	[2.1050, 2.1133]
Train punctuality 5 minutes	89.14%	[89.07%, 89.20%]
Train punctuality 15 minutes	100%	[100%, 100%]

5.1.2 Passenger delay and punctuality

In the results of the passenger delay and punctuality, two variables are distinguished: the rescheduling method and the journey type. Two rescheduling methods are included into the evaluation model. The first method is denoted as *optimistic rescheduling*, the complete journey is rescheduled using the realized timetable. The second way to reschedule is denoted as *realistic rescheduling*, this method reschedules the journey of the passengers from the point they missed their transfer. Journeys are divided into several types and are assigned labels. We refer to the journey types using these labels throughout the remainder of this chapter. The labels' descriptions are given below.

- ***Overall*** contains, as the name suggests, all journeys.
- The journeys with a direct connection between their origin and destination station carry the label ***without transfers***. Since they cannot miss a transfer, the results for the two methods of rescheduling will be identical.
- The remaining group of journeys, those who do have a transfer, are labeled: ***with transfer***.

Within the group with transfers, a distinction is made between journeys which include a missed transfer, and those which do not.

- The group ***with transfers, without missed transfers*** were able to follow their promised path, the results of the two methods of rescheduling will again be identical.
- The group ***with transfers, with missed transfers*** did miss a connection and had to reschedule their journey.

In Figure 5.1, the relative size of each journey type is displayed. The ratio between the number of journeys with and without transfers is independent of the simulations. However, the relative size of the number of journeys with missed transfers follow from the 1000 simulations. In the Utrecht case, the number of the journeys without transfers is slightly smaller than the group with transfers.

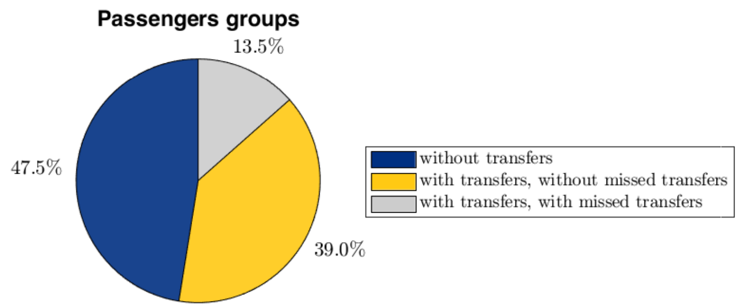


Figure 5.1: The relative size of the passenger groups.

For each combination of journey type and rescheduling method, the results are stated in Table 5.2. Similar to the results of the trains, the average values and their corresponding 95% CI are given for the passenger delay, 5-minute passenger punctuality, and 15-minute passenger punctuality.

Table 5.2: Results of the passenger delay and punctuality for 1000 simulations. Some remarkable results are highlighted in bold (see Section 5.2.1).

		Optimistic rescheduling		Realistic rescheduling	
		Average	**95% CI**	**Average**	**95% CI**
Overall	Delay (minutes)	5.3009	[5.2884, 5.3134]	6.0405	[6.0263, 6.0548]
	Punctuality 5 min.	57.65%	[57.50%, 57.81%]	55.74%	[55.58%, 55.90%]
	Punctuality 15 min.	96.50%	[96.47%, 96.52%]	94.43%	[94.40%, 94.45%]
Without transfers	Delay (minutes)	**5.0137**	[5.0026, 5.0248]	**5.0137**	[5.0026, 5.0248]
	Punctuality 5 min.	**49.95%**	[49.72%, 50.17%]	**49.95%**	[49.72%, 50.17%]
	Punctuality 15 min.	100%	[100%, 100%]	100%	[100%, 100%]
With transfers	Delay (minutes)	**5.5613**	[5.5460, 5.5766]	**6.9730**	[6.9544, 6.9916]
	Punctuality 5 min.	**64.61%**	[64.48%, 64.75%]	**60.97%**	[60.84%, 61.11%]
	Punctuality 15 min.	93.34%	[93.29%, 93.39%]	89.40%	[89.35%, 89.45%]
With transfers, without missed transfers	Delay (minutes)	3.3083	[3.3003, 3.3162]	3.3083	[3.3003, 3.3162]
	Punctuality 5 min.	78.38%	[78.24%, 78.52%]	78.38%	[78.24%, 78.52%]
	Punctuality 15 min.	100%	[100%, 100%]	100%	[100%, 100%]
With transfers, with missed transfers	Delay (minutes)	12.1903	[12.1583, 12.2223]	17.6772	[17.6580, 17.6964]
	Punctuality 5 min.	24.44%	[24.34%, 24.55%]	10.56%	[10.51%, 10.61%]
	Punctuality 15 min.	74.19%	[74.05%, 74.32%]	59.37%	[59.28%, 59.47%]

5.1.3 Comparison of the rescheduling methods

As explained above, there are two rescheduling methods implemented into the evaluation model: the optimistic rescheduling and realistic rescheduling method. In this section the results of the different methods are analyzed. As described in Section 4.5, we expect an underestimation of the passenger delay as result of optimistic rescheduling.

First of all, note the significant difference between the overall passenger punctuality of the two rescheduling methods in Table 5.2. This was expected and here we quantify the effect. The optimistic rescheduling method performs better. Especially, the journeys with transfer, with missed transfer show the effect of the different rescheduling methods: The journeys using the optimistic and realistic rescheduling experience an average delay of 12.1 minutes and 17.7 minutes, respectively.

In Figure 5.2 and Figure 5.3, the distributions of the delays are visualized. These figures show the cumulative representations of the delays for optimistic and realistic rescheduling methods, respectively. The results of the journeys without transfers (dark blue) and without missed transfers (yellow) are equal since they do not require rescheduling. The remaining journeys, with missed transfers (light gray), show differences. This difference is displayed in Figure 5.4. As expected, Figure 5.4 shows that many passengers experience more delay with realistic rescheduling. The increase in delays starting from 10 and 30 minutes is due to the frequencies of the train lines in the timetable. There are two used frequencies: either a train departures every 10 minutes or every 30 minutes from a station. A passenger has the option to wait for the next train of the same train line if the transfer was missed, this option provides the shortest rescheduled path more often in realistic rescheduling than in optimistic rescheduling.

Although the number of passengers with a different rescheduled path is small (3.2%) in comparison to the total number of passengers, the impact is large on the overall averages. In Table 5.2, the overall average delay shows a difference of 0.7396 minutes (an increase of 14%), the 5-minute punctuality a difference of 1.91% and the 15-minute punctuality a difference of 2.07%.

The scale on which NS strives for improvement of the punctuality is several tenths of a percent: from 90.5% in 2016 to 91.3% in 2019 for the 5-minute punctuality, and from 97.1% to 97.3% for the 15-minute punctuality on the HRN. The significant underestimation of the optimistic method makes it an unreliable method. The method assumes that passengers are able to predict the future, since passengers alter their complete journey, knowing they will miss a transfer in their nominal shortest journey. Therefore, the realistic rescheduling method must be considered in the evaluation model as its most accurate rescheduling method.

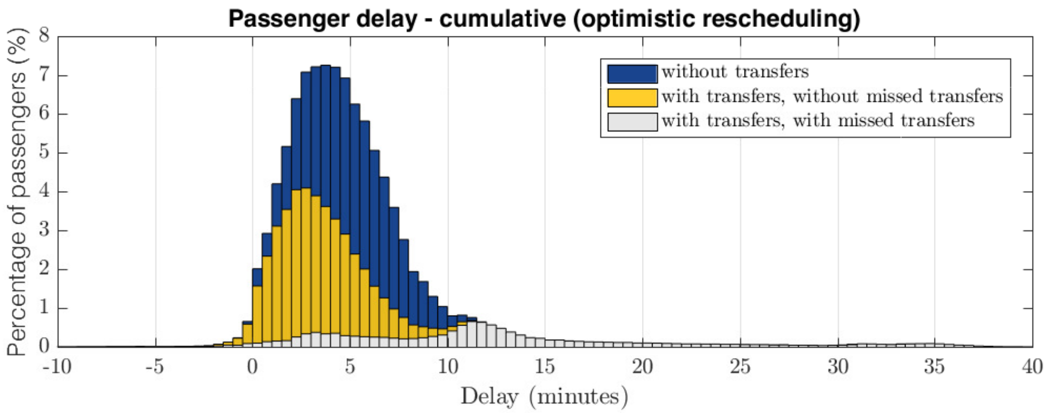


Figure 5.2: The normalized cumulative distribution of the passengers delay with optimistic rescheduling.

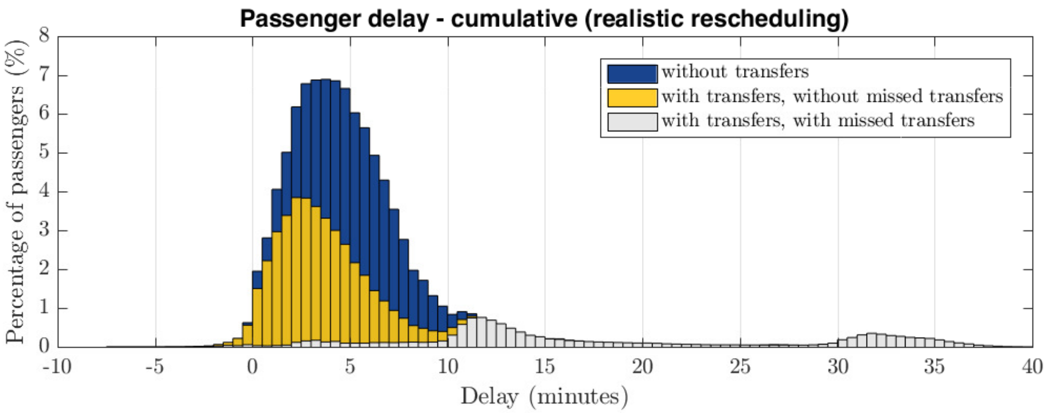


Figure 5.3: The normalized cumulative distribution of the passenger delay with realistic rescheduling.

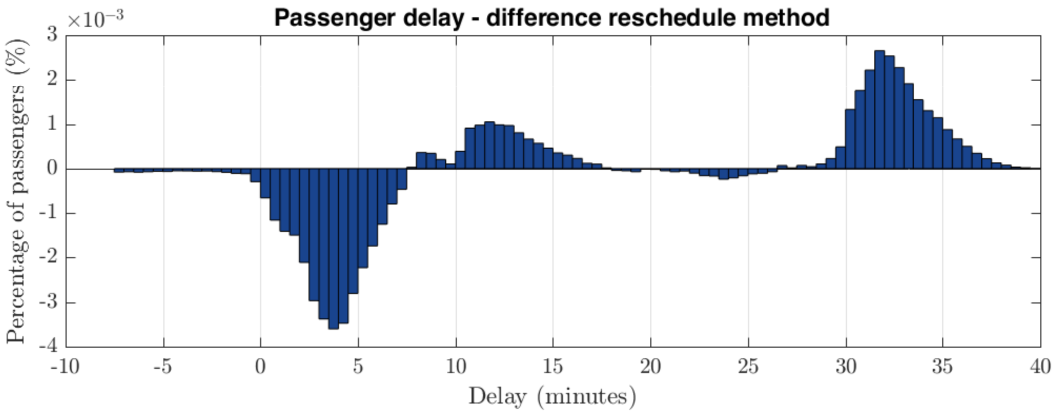


Figure 5.4: The difference in the passenger delay distribution between the two rescheduling methods, this is defined by the realistic rescheduling method minus the optimistic rescheduling method.

5.2 Data analysis

In this section, the results are further analyzed. We analyze the unexpected results regarding the passenger delay and punctuality in Table 5.2, and we look into the locations of the (missed) transfers.

5.2.1 Unexpected results in Table 5.2

In Table 5.2, there are some unexpected results; these are highlighted in bold. For both rescheduling methods in the journeys without transfers, the average delay is 5.02 minutes and the corresponding 5-minute punctuality is 49.87%. We compare this with the results for the journeys with transfers. The average delay is, as expected, larger due to missed transfers: 5.56 and 6.98 minutes for the optimistic and realistic rescheduling, respectively. Since the average delay is larger, we expected the 5-minute punctuality to be lower. However, the 5-minute punctuality is higher: 64.68% and 61.01% respectively. The results show a decline in delay and simultaneously, an improvement in punctuality.

The cause of the odd results are presented in Figure 5.5; this figure shows the distribution of the passenger delay for the groups with and without transfers in front of each other. The distribution of the delay without transfers (blue) has a single peak around its average, while the distribution of the group with transfers (yellow) shows multiple peaks. The first peak is around 3 minutes, and a second and third peak starting at 10 and 30 minutes, respectively.

The results show a higher 5-minute punctuality with transfers, because the first peak with transfers is located more to the left than the peak without transfers: relatively more journeys with transfer experience a delay less than 5 minutes. The large delays in the second and third peak have a large weight in the average delay, but not in the punctuality. This explains the values we see in the results.

It is still remarkable that the peak of transfer is shifted to the left compared to the peak without transfers. We looked further into the passengers without transfer, to find the cause of the shift. The passengers without transfers turn out to be concentrated in a few trains: over 60 percent of these passenger are located in 1.2% of the trains. These 1.2% of the trains have an average delay of 6 minutes, resulting in a high passenger delay compared to the journeys with successful transfers. We can not explain why the average delay is high for these trains, it might be an artifact of the given data.

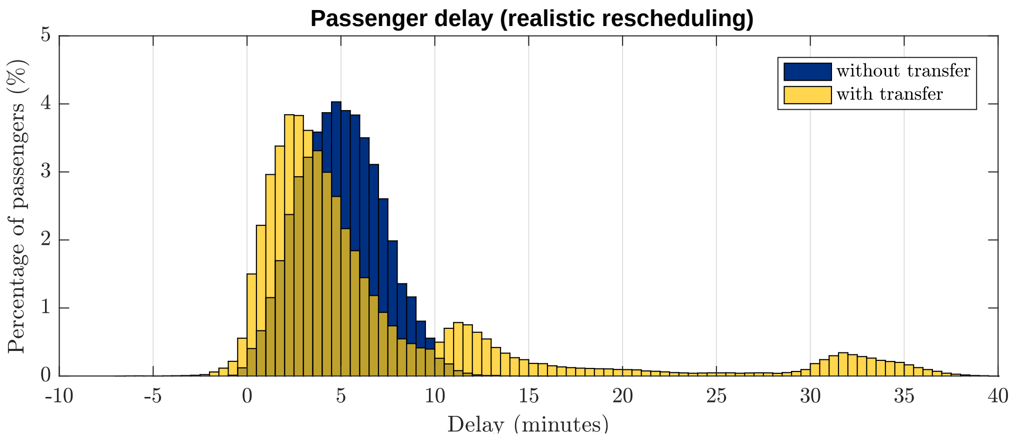


Figure 5.5: The normalized distribution of the delay of passengers for the passengers with (yellow) and without transfers (blue).

5.2.2 Locations of the (missed) transfers

The missed transfers are further analyzed using the nominal and realized paths of the passengers. From these paths, the locations of all transfers and missed transfers are determined. In Appendix B, these results are displayed in Figure B.2 and Figure B.3. In Table 5.3, the percentage of missed transfers per station is given, expressed as the number of missed transfers with respect to the total number of transfers in a station. The stations, that are not mentioned in Table 5.3, did not have any missed transfers in the simulations.

Table 5.3: An overview of the percentage of missed transfers per station.

Location	Percentage of missed transfers
Utrecht	60.81%
Utrecht Overvecht	33.32%
Weesp	17.80%
Baarn	12.85%
Bussum Zuid	10.84%
Duivendrecht	4.86%
Amsterdam Bijlmer	1.24%

Remarkable in Figure B.2 and Figure B.3, are station Weesp and station Utrecht because of the following results. At station Weesp the most transfers occur, 28% of the total, followed by Utrecht with 20% of the total number of transfers. The number of missed transfers for these two stations is not proportional: 62% of the total number of missed transfers are missed in Utrecht and 25% in Weesp. This reflects back in the results in Table 5.3.

In Utrecht, over 60% of the passengers missed their transfer. This large percentage is caused by the transfer times at the station. As discussed in Section 4.7, it is possible to have zero-minute transfers in the evaluation model. In Figure 5.6, the distribution of the transfer times at station Utrecht and station Weesp are presented. A majority of the passengers who transfer in Utrecht have a transfer time less or equal to 3 minutes; zero-minute transfers occur at station Utrecht as well. At station Weesp, all passenger have a transfer time greater or equal to 3 minutes. As a result, the short transfer times cause many missed transfers in Utrecht.

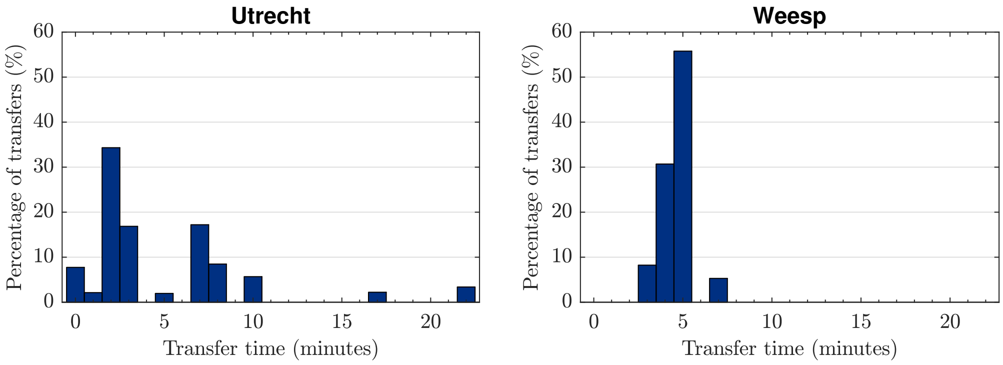


Figure 5.6: The planned transfer times at stations Utrecht and Weesp.

5.3 Varying the disturbance parameter

For all results of the previous sections, the disturbances were determined with the default set of parameters. In this section, these parameters are altered and the corresponding results analyzed. We discuss the results of the train and passenger delay. Other results, concerning the train and passenger punctuality, are added to Appendix B.

The parameters represent the mean of the exponential distribution, from which the disturbances on the processes are determined. In the default case, the mean of the disturbances over movement processes is set on 5% of the technical minimal process time and 30% over dwelling processes. In Table 5.4, an overview of the tested disturbance parameters is given. The upper row gives the ratio of the parameters with respect

Table 5.4: Overview of the tested disturbance parameters.

Factor	0	0.2	0.4	0.6	0.8	1	1.2	1.4	1.6	1.8	2	3	4
Movement	0%	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%	15%	20%
Dwelling	0%	6%	12%	18%	24%	30%	36%	42%	48%	54%	60%	90%	120%

to the default parameters: the *disturbance factor*. For each pair of parameters, 100 simulations are executed by the evaluation model.

The results of the average train and passenger delay are presented in Figure 5.7. As expected, the delay increases as the disturbance factor increases. Moreover, the difference between the two different rescheduling methods increases. In Section 5.1.3, we have shown that the optimistic rescheduling method results in a significant underestimation of the passenger delay, with a disturbance factor of 1.

The relation between results of the optimistic and realistic rescheduling methods is nearly linear, this is presented in Figure B.6 in Appendix B. This linear relation between the passenger delays is remarkable: over 85% of the journeys are identical for each of the rescheduling methods and therefore linearly related with disturbance factor 1. For smaller disturbance factors, the percentage of identical journeys is even larger. For larger disturbance factors, the percentage of identical journeys is smaller, because the number of rescheduled journeys increases. Here, the rescheduling methods have more influence on the average delay. Even though the journeys are rescheduled by completely different methods, the results show a nearly linear relation between the average delays.

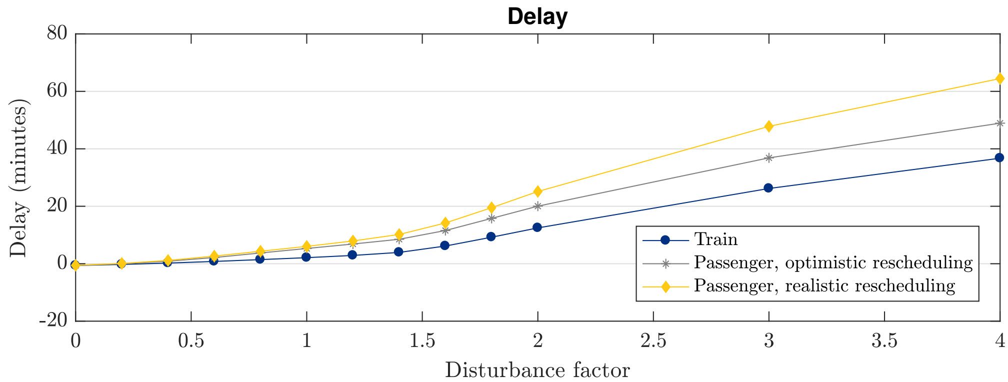


Figure 5.7: Average train and passenger delay over the disturbance factor.

In Figure 5.8, the train delay is plotted against the average passenger delay: again, we see a nearly linear relation. Here, the results of the realistic rescheduling method are used. The linear relation can be explained for the majority of the journeys: for both the journeys without transfers and with successful transfers, the passenger delays are directly related to the train delay. The journeys with missed transfers are not directly related; these journeys use another train than planned to arrive at their final destination, and additional delays occur.

The relation between the train delay and the journeys with missed transfers is presented in Figure 5.9. The results shows a stagnation in around 18 minutes of passenger delay; this is due to the frequencies of the train lines. Most passengers use the same train line to continue their journey after a missed transfer, this results in an additional delay of 10 or 30 minutes (see Figure 5.3). The ratio between these additional delays remains almost equal for small train delays (between 0 and 3 minutes on average), only the number of journeys with missed transfers differs.

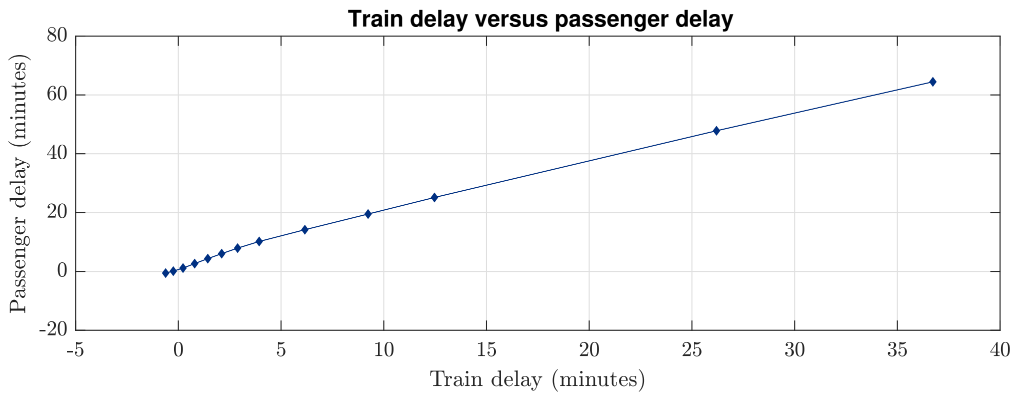


Figure 5.8: The average train delay plotted against the average passenger delay. The passenger delay is determined with the realistic rescheduling method.

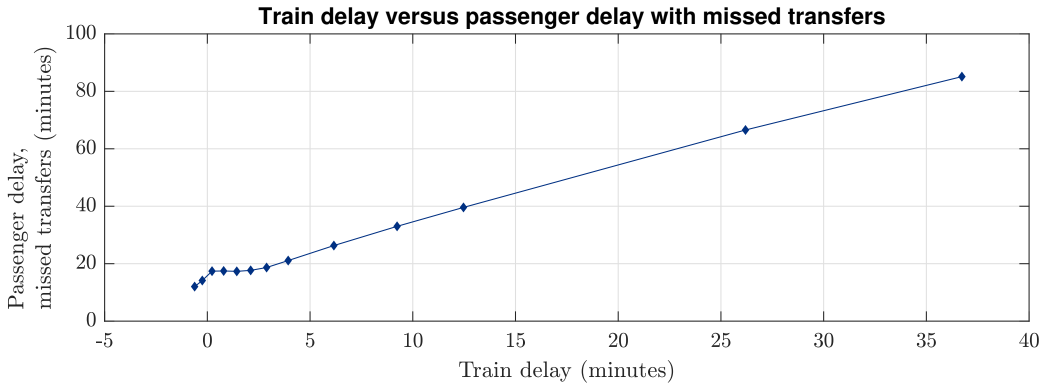


Figure 5.9: The average train delay plotted against the average passenger delay of the journeys with missed transfers. The passenger delay is determined with the realistic rescheduling method.

5.4 Summary

The conclusions of the results in this chapter are summarized below.

- The optimistic rescheduling method gives, as expected, a significant underestimation of the passenger delay. The underestimation increases as the disturbance factor increases.
- The journeys with transfers perform better on punctuality but worse on average delay than the journeys without transfers. This is due to the distribution of the delays, see Figure 5.5.
- The journeys without transfers have on average more delay than the journeys with successful transfers. This is due to the delay of the 1.2% of the trains: these trains experience larger delays than average and contain a majority of the passengers without transfers.
- Most transfers are missed in Utrecht due to short transfers times.
- Varying the disturbance factor reveals a nearly linear relation between the train delay and realistic passenger delay.

6 Improving timetables with SOM

In this chapter, we apply SOM to the reference timetable of the Utrecht case. SOM is designed to allocate the time supplements in a timetable in order to improve the expected train delay (see Chapter 3). We aim to improve the expected passenger punctuality of the reference timetable with SOM.

In Section 5.3, the results of the different disturbance factors revealed a nearly linear relation between the train delay and the passenger delay. According to this relation, we expect that the passenger delay will improve with the train delay. Therefore, we initially apply SOM with its default objective function in Section 6.1. Furthermore, we will adapt the objective function of SOM by introducing customizable weights to it. This adaption is made in order to target more directly our actual objective, which is passenger punctuality and not train punctuality. Firstly, we test two weight sets which are derived from the passenger behavior in the simulations in Section 6.2. In Section 6.3, we test several variations of these weights. The chapter is concluded with a summary in Section 6.4.

6.1 Benchmark: results with default settings

To obtain the results in this chapter, we run 100 6-hours simulations of each timetable with the evaluation model, i.e. $R = 100$ and $H = 6$ in the evaluation model. For each timetable, the same seed is used to determine the random disturbances on the processes. In order to make an honest comparison, new results of the original timetable are obtained under identical circumstances. In Table 6.1, the results of the original timetable are presented, and the result of SOM using its default objective function. For the results of SOM, we use the implementation as provided by ORTEC, it is implemented in C$\#$ and solved using Visual Studio 2015 and CPLEX 12.5. SOM is used with $R = 30$ and $H = 12$ in each of its applications and a fixed seed is used for the disturbances. The two sets of results in Table 6.1 function as a benchmark for the upcoming results.

Table 6.1: Results of the original timetable and the improved timetable by SOM. SOM is used with its default settings and weights. For both timetables, the same seed is used to determine the disturbances in the evaluation model.

	Original		SOM, default	
	Average	**95% CI**	**Average**	**95% CI**
Train delay	2.0603	[2.0425, 2.0781]	2.0329	[2.0133, 2.0525]
Train punctuality 5 min.	89.56%	[89.17%, 89.95%]	88.81%	[88.53%, 89.09%]
Train punctuality 15 min.	100%	[100%, 100%]	100%	[100%, 100%]
Passenger delay (*Optimistic*)	5.1466	[5.1114, 5.1818]	5.2984	[5.2358, 5.3610]
Passenger punctuality 5 min.	59.14%	[58.48%, 59.79%]	58.54%	[57.88%, 59.21%]
Passenger punctuality 15 min.	96.65%	[96.55%, 96.75%]	96.11%	[96.00%, 96.22%]
Passenger delay (*Realistic*)	5.8478	[5.8089, 5.8867]	6.0388	[5.9627, 6.1150]
Passenger punctuality 5 min.	57.31%	[56.62%, 57.99%]	56.59%	[55.92%, 57.26%]
Passenger punctuality 15 min.	94.59%	[94.48%, 94.70%]	94.14%	[93.96%, 94.32%]

The default settings of SOM optimize the timetable on train delay: each arrival event has the same weight in the objective function, i.e. $w_4 = 1$ and the remaining weights $w_i = 0$ in function (4) in Section 3.2. Indeed, the results show a non-significant decrease in train delay. Furthermore, we see that the passenger delay significantly increases. This is not unexpected, since the timetable is optimized on train delay, not passenger delay.

6.2 Adjusted weights based on the passengers

We apply SOM to the reference timetable with customized weights. Altering the weights results in a change of the objective function that is used in SOM. We chose to do this because of its simplicity in the implementation, and the idea that the passenger delay is related to the train delay.

The objective function of SOM is adapted to make the customized weights possible: instead of fixed weights for sets of events with certain properties (see Section 3.2), the weight of each arrival event $e \in E_a$

is given by the parameter $w_e \in \mathbb{N}$. Thus, the customized objective function is defined as

$$\text{minimize} \sum_{e \in E_a} \sum_{r=1}^{R} \sum_{h=1}^{H} w_e \frac{D_{e,r,h}}{(|E_a| \cdot R \cdot H)}. \tag{5}$$

Here $D_{e,r,h}$ denotes the delay variable of event e in hour h of simulation r. Any point in the $|E_a|$-dimensional space $\mathbb{N}^{|E_a|}$ is a possible weight choice. Due to the excessive calculation times of the models (both SOM and the evaluation model), it is not possible to apply excessive local search methods to find the optimal weights, that result in the best passenger delay or punctuality.

We propose two options to determine the weights for the arrival events. In both options, we give more weight to the arrival events that are important for the passengers: their arrival at their transfer station and arrival at their final destination. At their transfer station, because a successful transfer is crucial to arrive within the 5-minute time window of the passenger punctuality, and at their final destination, because the delay of the passengers is measured at that point. The two weight options are defined as follows:

- **Weight option 1:** The weight of each arrival event is equal to the number of passengers leaving the train at that moment. This includes passengers who reach their destination and passengers who leave the train to transfer to another train.
- **Weight option 2:** The weight of each arrival event is equal to the number of passengers reaching their destination, plus the expected delay per transfer per passenger. The expected delay of a transfer is defined as the chance that the transfer is unsuccessful times the waiting time until the next train of the same train line. Let X denote the set of passengers who enter a transfer on arrival event e. Then, the weight of arrival event e is defined as:

  $$w_e = \sum_{x \in X} \mathbb{P}(\text{passenger } x \text{ misses transfer}) \cdot \mathbb{E}(\text{delay}|\text{transfer is missed})$$

  Here, the probability that a transfer is missed, $\mathbb{P}(\text{passenger } x \text{ misses transfer})$, is determined by the 1000 simulations used in Chapter 5: the probability is defined as the percentage of missed transfers in those simulations. The expected delay in case of a missed transfer, $\mathbb{E}(\text{delay}|\text{transfer is missed})$, is defined as the time between the planned departures of the missed train and the successive train of the same train line at that station.

Table 6.2: Results of the timetables improved by SOM. The objective functions were determined according to weight option 1 and weight option 2, respectively.

	SOM, option 1		SOM, option 2	
	Average	**95% CI**	**Average**	**95% CI**
Train delay	1.9829	[1.9714, 1.9944]	2.0267	[2.0136, 2.0398]
Train punctuality 5 min.	91.34%	[91.06%, 91.62%]	91.28%	[91.02%, 91.54%]
Train punctuality 15 min.	100%	[100%, 100%]	100%	[100%, 100%]
Passenger delay (*Optimistic*)	4.5238	[4.4370, 4.6105]	4.6781	[4.5777, 4.7784]
Passenger punctuality 5 min.	69.61%	[68.67%, 70.56%]	66.84%	[65.53%, 69.14%]
Passenger punctuality 15 min.	96.56%	[96.39%, 96.73%]	96.81%	[96.66%, 96.97%]
Passenger delay (*Realistic*)	5.2259	[5.1323, 5.3194]	5.3605	[5.2482, 5.4729]
Passenger punctuality 5 min.	67.55%	[66.58%, 68.53%]	64.84%	[63.50%, 66.18%]
Passenger punctuality 15 min.	94.05%	[93.91%, 94.18%]	94.28%	[94.13%, 94.44%]

In Table 6.2, the results of the two options are presented. The train delay of both options show a significant improvement compared to the original timetable. Weight option 1 even shows a significant improvement compared to the timetable, which was optimized using the default objective function.

Furthermore, both weight options show impressive, significant improvements with respect to the passenger delay and passenger punctuality for both optimistic and realistic rescheduling method. In Figure 6.1, an overview of the realistic results regarding the passenger delay and 5-minute punctuality is presented.

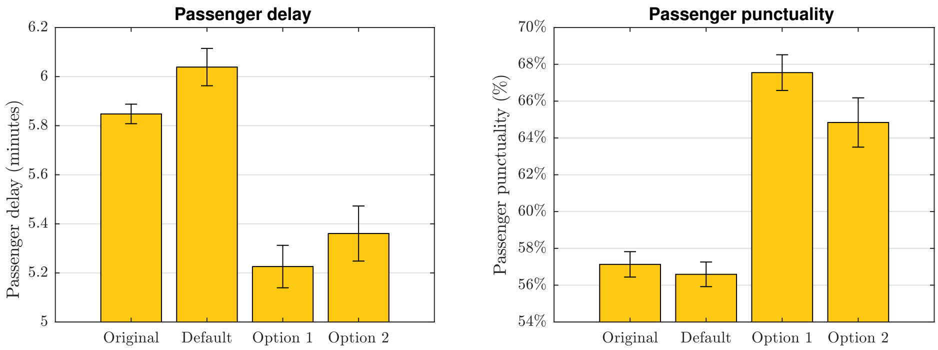


Figure 6.1: The average passenger delay and 5-minute passenger punctuality, including their 95% confidence interval, for all timetables of Section 6.1 and Section 6.2. This includes: the original timetable, the timetable optimized on train delay by SOM, and the two timetables resulting from the weight options 1 and 2.

6.3 Variations of weight option 1

In Section 6.2, weight option 1 shows the best results and is easier to determine for a timetable than weight option 2. Therefore, we test the effects of several variations of weight option 1:

- The square root of the weights: the differences between the weights will decrease.
- The square of the weights: the difference between the weights will increase.
- Iterative adjusted weights: determine the weights according to weight option 1 on the optimized timetable instead of the original timetable.

In Section 6.3.1, the results of the squared and square rooted weights are presented, and in Section 6.3.2, the results of the iterative adjusted weights. An overview of the results is provided in Section 6.4.

6.3.1 Squared and square rooted weights

The weights of option 1 are contained within the interval $[0, 1376]$. Taking the square root of these weights, limits the range of the weights to an interval of $[0, 37]$. Taking the square, extends the range to $[0, 1.9 \cdot 10^6]$. In Table 6.3, the results of both variations are presented.

Table 6.3: Results of the timetables improved by SOM. The objective functions were determined according to a square rooted and squared version of weight option 1.

	SOM, $\sqrt{(\text{option } 1)}$		SOM, $(\text{option } 1)^2$	
	Average	**95% CI**	**Average**	**95% CI**
Train delay	1.9574	[1.9465, 1.9683]	2.0027	[1.9892, 2.0163]
Train punctuality 5 min.	91.49%	[91.24%, 91.75%]	91.41%	[91.15%, 91.68%]
Train punctuality 15 min.	100%	[100%, 100%]	100%	[100%, 100%]
Passenger delay (*Optimistic*)	4.6038	[4.5173, 4.6903]	4.4987	[4.4151, 4.5822]
Passenger punctuality 5 min.	68.02%	[67.41%, 68.62%]	70.07%	[69.56%, 70.57%]
Passenger punctuality 15 min.	96.60%	[96.43%, 96.77%]	96.55%	[96.37%, 96.73%]
Passenger delay (*Realistic*)	5.3148	[5.2204, 5.4092]	5.1775	[5.0883, 5.2668]
Passenger punctuality 5 min.	65.91%	[65.28%, 66.54%]	68.05%	[67.54%, 68.56%]
Passenger punctuality 15 min.	94.03%	[93.88%, 94.17%]	94.04%	[93.88%, 94.21%]

Compared to the results of weight option 1, the square rooted version shows improvement with respect to train delay and train punctuality, but performs worse with respect to the passenger delay and passenger punctuality. The squared version shows opposite results: worse train delay and train punctuality and improved passenger delay and passenger punctuality. However, the improvements of the squared version are not significant compared to option 1.

6.3.2 Iterative weights

Another variation of weight option 1 is obtained by determining the weights (according to weight option 1) on an optimized timetable. Therefore, we analyze the optimized timetable and determine the weights according to the new paths of the passengers. This new set of weights is denoted as (option $1)_1$. The results of (option $1)_1$ are presented in Table 6.4.

Table 6.4: Results of the timetables improved by SOM. The weights of the objective functions were determined by applying weight option 1 again on the optimized timetable by weight option 1.

	SOM, (option $1)_1$	
	Average	**95% CI**
Train delay	1.9842	[1.9725, 1.9958]
Train punctuality 5 min.	91.44%	[91.16%, 91.72%]
Train punctuality 15 min.	100%	[100%, 100%]
Passenger delay (*Optimistic*)	4.4959	[4.4073, 4.5846]
Passenger punctuality 5 min.	70.38%	[69.40%, 71.37%]
Passenger punctuality 15 min.	96.55%	[96.38%, 96.72%]
Passenger delay (*Realistic*)	5.1979	[5.1025, 5.2934]
Passenger punctuality 5 min.	68.32%	[67.30%, 69.34%]
Passenger punctuality 15 min.	94.03%	[93.89%, 94.17%]

The results in Table 6.4 show further improvement in passenger delay and passenger punctuality compared to weight option 1. However, the improvements are not significant. If we determine the weights for the optimized timetable with weight option 1_1, we obtain the exact same weight. Therefore, more iterations will result in the same timetable.

In Section 6.3.1, the squared version of option 1 showed improvement in passenger delay and passenger punctuality: we applied the recursively adjusted weights also on this timetable. We consider two variations of the iterative adjustments: apply weight option 1 on the timetable obtained from (option $1)^2$, denoted as (option $1)_1^2$, and apply the squared weight option 1 on the timetable obtained from (option $1)^2$, denoted as ((option $1)_1^2)^2$. In Table 6.5, the results of those two new weight options are presented.

Table 6.5: Results of the timetables improved by SOM. The weights of the objective functions were determined by applying weight option 1 and the squared version of it, again on the optimized timetable by (option $1)^2$.

	SOM, (option $1)_1^2$		SOM, ((option $1)_1^2)^2$	
	Average	**95% CI**	**Average**	**95% CI**
Train delay	1.9851	[1.9734, 1.9968]	1.9964	[1.9828, 2.0100]
Train punctuality 5 min.	91.45%	[91.18%, 91.73%]	91.64%	[91.37%, 91.91%]
Train punctuality 15 min.	100%	[100%, 100%]	100%	[100%, 100%]
Passenger delay (*Optimistic*)	4.5017	[4.4128, 4.5905]	4.4610	[4.3774, 4.5446]
Passenger punctuality 5 min.	70.47%	[69.54%, 71.39%]	70.91%	[70.37%, 71.46%]
Passenger punctuality 15 min.	96.37%	[96.37%, 96.72%]	96.55%	[96.37%, 96.73%]
Passenger delay (*Realistic*)	5.2038	[5.1080, 5.2996]	5.1396	[5.0503, 5.2289]
Passenger punctuality 5 min.	68.40%	[67.44%, 69.36%]	68.90%	[68.34%, 69.45%]
Passenger punctuality 15 min.	94.03%	[93.89%, 94.17%]	94.04%	[93.88%, 94.21%]

The results of (option $1)_1^2$ are slightly worse than (option $1)^2$ and ((option $1)_1^2)^2$ slightly better than (option $1)^2$, but again, the differences are not significant. The results of ((option $1)_1^2)^2$ are the best we obtained in this research.

6.4 Summary

In this chapter, we have tested several weight options in the objective function of SOM (see equation (5)). In Table 6.6, an overview is given of all average values with respect to the train and passenger delay and

5-minute punctuality of all timetables. Additionally, a histogram with the average values and the 95% CI of the passenger delay and realistic 5-minute punctuality is presented in Figure 6.2. A summary of the results is listed below:

- The train delay of $\sqrt{(\text{option } 1)}$ performs significantly better than all other timetables. However, the differences are small: The average train delay of the best timetable (by $\sqrt{(\text{option } 1)}$) equals 1.9574 minutes and of the worst timetable (by the original timetable) equals 2.0603 minutes.
- The 5-minute train punctuality of any of the customized weight options perform significantly better than both the original timetable and the timetable by the default setting of SOM. Within the customized weight options, there are no significant differences. It is remarkable that the default timetable by SOM performs worse than the customized, because this timetable is optimized on train delay.
- The passenger delays of the optimistic rescheduling method are not discussed, since they provide an underestimation of the passenger delay. The results of the realistic rescheduling method provides us with several impressive results (see Figure 6.2): all customized weight options result in timetables which perform significantly better than the original and default SOM timetables. Within the customized weight options, no significant differences are measured.
- The passenger punctuality of the realistic rescheduling method shows results similar to the passenger delay: all customized weight options perform significantly better. The difference within the customized weight options do show significant differences: option 2 and $\sqrt{(\text{option } 1)}$ perform significantly worse than the remaining customized weight options.

Overall, the customized weight options are successful in the Utrecht case. The best results are obtained in $((\text{option } 1)_1^2)^2$. This weight option improves the passenger delay from 5.8478 minutes on average (by the original timetable) to 5.1396 minutes, and the corresponding 5-minute passenger punctuality from 57.31% to 68.90%. That is a relative decrease in passenger delay of 12%, and relative improvement of the 5-minute passenger punctuality of 20%.

Table 6.6: An overview of the results of all tested timetables. Here, the average values of the train delay and 5-minute punctuality, and the passenger delay and 5-minute punctuality (with realistic rescheduling) are displayed. For each measure, the best result is highlighted in bold.

	Train delay	**Train p. 5 min.**	**Passenger delay**	**Passenger p. 5 min.**
Original	2.0603	89.56%	5.8478	57.31%
Default	2.0329	88.81%	6.0388	56.59%
Option 1	1.9829	91.34%	5.2259	67.55%
Option 2	2.0267	91.28%	5.3605	64.84%
$\sqrt{(\textbf{Option 1})}$	**1.9574**	91.49%	5.3148	65.91%
$(\textbf{Option 1})^2$	2.0027	91.41%	5.1775	68.05%
$(\textbf{Option 1})_1$	1.9842	91.44%	5.1979	68.32%
$(\textbf{Option 1})_1^2$	1.9851	91.45%	5.2038	68.40%
$((\textbf{Option 1})_1^2)^2$	1.9964	**91.64%**	**5.1396**	**68.90%**

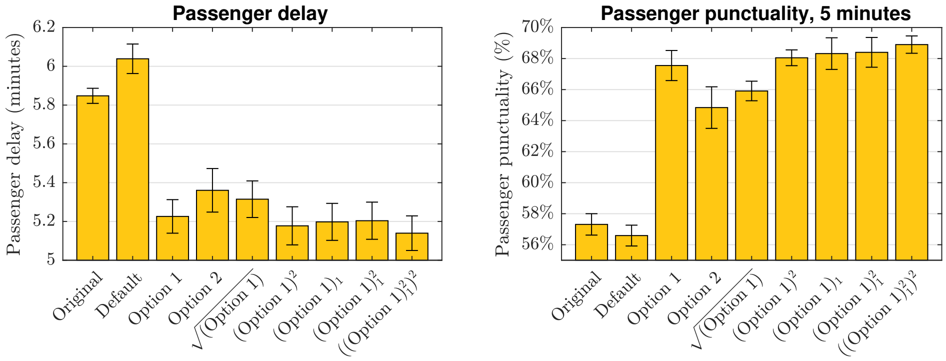


Figure 6.2: The average passenger delay and 5-minute passenger punctuality, including their 95% confidence interval, for all timetables of this chapter.

7 Discussion

In this chapter, we discuss the limitations and assumptions of the model and results. After a general notion, the limitations of the evaluation model and its results are discussed in Section 7.1. In Section 7.2, we discuss SOM and its results.

Both models, the evaluation model and SOM, are designed to simulate small disturbances, because the goal of the models is to make the timetable more robust against smaller delays. Therefore, the results of both models cannot be compared with actual results of passenger punctuality.

7.1 Limitations and assumptions of the evaluation model

In Section 4.7, the limitations and assumptions of the evaluation model are discussed. A major part of the input data is not available and estimations have been made based on historic data. The rescheduling methods make assumptions on the passenger behavior, the optimistic rescheduling method results in a significant underestimation of the passenger delay.

Furthermore, the transfer construction allows any transfer with a non negative transfer time, this reflects back in the results. In Section 5.2.2, we analyzed the transfer per station. For example, in station Utrecht, over 60% of the transfers were missed due to the short transfer times, causing many rescheduled journeys with large delays.

Finally, the computation time limits the usage of the model. As mentioned in Chapter 5, the 1000 12-hour simulations took over 40 hours to compute. Therefore, we limited the simulation size in Chapter 6; here we used 100 simulations of 6-hour days, which could be computed in 35 minutes. All these simulations were conducted on the small Utrecht case, simulations of the whole railway network of the Netherlands would lead to even larger computation times.

7.2 Limitations of SOM

SOM is mostly limited by its computation time and memory usage. In Chapter 6, we used $R = 30$ and $H = 12$ to obtain the new timetables: this was the limit the hardware could handle. A larger value for either of the parameters resulted in an out of memory error. With these settings for R and H, the optimized timetable was calculated in 11 minutes.

Together with the evaluation model, it takes 46 minutes to test a single weight option without analyzing its results. In this thesis, only a couple of different weights are tested for this reason. The computation time is too long to be useful to find the optimal weight option with local search methods.

Additionally, the small value of R and H could lead to overfitting, as the timetable is optimized to perform the best on a small set of simulations and therefore, a small set of disturbances. A test with another seed with the default SOM objective resulted in exact the same optimized timetable. However, this gives no guarantee that the results of SOM are optimal for any random seed. The higher the number of simulations that is taken into account, the better the approximation of the true minimizing distribution of the time supplements of that instance (Kroon et al., 2008).

SOM, with the customized objective function, shows impressive results in Chapter 6. However, due to the small part of the railway network that is used and the limited number of simulations, we cannot conclude anything about the effect of SOM on the whole railway network. In order to apply SOM and the evaluation model on the whole network, additional research is necessary to improve both models.

As mentioned in Section 1.2, the magnitude of the improvement in passenger punctuality we need to achieve on the HRN is several tenths of one percent: from 90.5% in 2016 to 91.3% in 2019 for the 5-minute punctuality, and from 97.1% to 97.3% for the 15-minute punctuality. In the Utrecht case, the original passenger punctualities of the timetable are 57.31% and 94.59% for the 5-minute and 15-minute measure, respectively. We showed that an optimized timetable can achieve passenger punctualities of 68.90% and 94.04%. These are improvements on another scale than required on the HRN, and therefore, cannot be compared with each other.

8 Conclusion and future research

In this chapter, the conclusions and recommendations of research are provided. In Section 8.1, the thesis is concluded with the answers of the sub questions and research question. In Section 8.2, some recommendations for further research are provided.

8.1 Answering the sub questions and research question

In Chapter 1, we formulated the research question of this thesis and four sub questions. We provide the answer of each of the sub questions and finally answer the research question.

1. **How is passenger punctuality calculated? What input is necessary?**

Passenger punctuality is a measure of the reliability of a timetable. It is defined as the percentage of passengers whose journey was successful in terms of travel time. A journey is considered successful if the passenger arrives within a given threshold of time from the promised arrival time. NS calculates the passenger punctuality with a threshold of 5 and 15 minutes.

The passenger punctuality is calculated using the following input data: the timetable, the set of frequently used origin-destination (OD) pairs, the realization data of the trains and the Check In and Check Out (CICO) data of all the passengers. For all passengers, who travel between frequently used ODs, the travel promise is determined using the timetable and the CICO data. This travel promise contains the promised train(s) the passenger used in its journey and the promised arrival time. The realization data of the train is used to determine the actual arrival time of the passengers. Together with the promised arrival times, the passenger punctuality is calculated.

2. **How can passenger punctuality be used to develop a new performance measure for an unrealized timetable?**

In this thesis, we successfully developed the evaluation model to calculate the passenger punctuality of an unrealized timetable in the same way as the actual passenger punctuality is calculated. Since the timetable is unrealized, all input data, except the timetable, are unavailable and estimated using historic data. Furthermore, two dynamic rescheduling methods are developed to reschedule the journey of passengers whose planned journey was no longer feasible.

Besides providing a measure for the passenger punctuality, the data of the simulations is useful for further analysis of the timetable. Specifically, in this thesis, we looked into the passenger delay distributions of different journey types, the train delay and punctuality, the difference between the two rescheduling methods, and the locations of (missed) transfers. Additionally, simulations with different disturbances revealed a nearly linear connection between the train delay and passenger delay in the Utrecht case.

3. **What is SOM and what are its limitations?**

The Stochastic Optimization Model (SOM) is a model by Kroon et al. (2008), which is developed to improve the robustness of a timetable against small disturbances. It allocates time supplements in a timetable such that the expected train delay is minimized.

The implementation of SOM by ORTEC is mostly limited by the computation time and memory usage. The implementation can only cope with small cases with a small number of simulations, resulting in an approximation of the true minimizing distribution of the time supplements. The higher the number of simulations taken into account, the better the approximation. Furthermore, the long computation times make the model not useful for large scale testing.

4. **How can passenger punctuality be used in SOM to improve timetables?**

The default objective function of SOM allows the user to determine the weight of the delays of fixed groups of events. In this thesis, the objective function is altered, such that we can determine the weight of each event separately. Motivated by the nearly linear connection between the train and passenger delay, we chose weights based on the expected number of passengers leaving the trains at arrival events. For the Utrecht case, this resulted in impressive improvements of the passenger delay (from 5.8478 minutes to 5.1396 minutes) and passenger punctuality (from 57.31% to 68.90%).

Finally the research question of this thesis:

How can timetables be improved using passenger punctuality as measure?

In this thesis, we showed that the passenger punctuality of a timetable can be improved by combining the evaluation model with SOM. The evaluation model gives a measure for the passenger punctuality of a timetable and provides the data which is used to determine the weights for the objective function of SOM. SOM improves the timetables using the sets of alternative weights that we have proposed, and the improved timetable is returned to the evaluation model to analyze the improvement with respect to passenger punctuality.

8.2 Recommendations for future research

The ultimate goal for future research is to develop a model that can optimize the passenger punctuality of the timetable for a complete railway network. With this goal in mind, we recommend further research in the following directions:

- Firstly, both models can be improved. It is interesting to improve the quality of the input of the simulations. Especially, the stochastic disturbances on a railway system affect the quality of the simulations. The evaluation model can be improved in several ways: the model can be implemented in a more efficient manner, to make the evaluation model useful on a larger scale. Furthermore, the realistic transfer construction, as discussed in Section 4.7, can be added. The realistic transfer construction requires extra research on transfer times at stations.

  Additionally, a better implementation of SOM is necessary to obtain better results, and to apply SOM on larger scale. Merging the evaluation model with SOM is also interesting: it would be more efficient to evaluate the passenger delay and punctuality within SOM. In the current situation, the models use similar simulations. SOM uses the simulations to measure the train delay, and the evaluation model to measure the passenger delay (using the train delay). Therefore, the total number of simulations can be reduced by merging the models, while maintaining the same quality of the results.
- Furthermore, the analysis of the data from the evaluation model can be extended. By improving or merging the two models, it may be possible to apply local search methods to find the optimal weights for minimizing the expected passenger delay. In this thesis, we have provided proof-of-concept for such an approach.
- Finally, more research on the effects of the improvements is interesting: the passenger punctuality increases but what does that mean for the quality of the journeys? It is possible that the overall travel time of the passenger increases and therefore the quality decreases. We believe, a careful consideration of this trade off is necessary for practical implementation.

9 Bibliography

Alexopoulos, C., Fishman, G. S., and Seila, A. F. (1997). Computational experience with the batch means method. In *Proceedings of the 29th conference on Winter simulation*, pages 194–201. IEEE Computer Society.

Bast, H., Delling, D., Goldberg, A., Müller-Hannemann, M., Pajor, T., Sanders, P., Wagner, D., and Werneck, R. F. (2016). Route planning in transportation networks. In *Algorithm Engineering: Selected Results and Surveys*, pages 19–80. Springer International Publishing.

CBS (2016). Lengte van Spoorwegen. http://cbs.overheidsdata.nl/71024ned. Last accessed on 17-03-2017.

Dijksma, S. (2017). Prestaties NS en ProRail 2016 [kamerbrief]. https://www.rijksoverheid.nl/documenten/kamerstukken/2017/04/13/prestaties-ns-en-prorail-2016. Last accessed on 14-04-2017.

Kroon, L., Maróti, G., Helmrich, M. R., Vromans, M., and Dekker, R. (2008). Stochastic improvement of cyclic railway timetables. *Transportation Research Part B: Methodological*, 42(6):553 – 570.

Landex, A. (2008). *Methods to estimate railway capacity and passenger delays*. PhD thesis, Technical University of Denmark (DTU).

Maróti, G. (2016). Personal communication, December 14, 2016.

Maróti, G. (2017). A branch-and-bound approach for robust railway timetabling. *Public Transport*, 9(1):73–94.

Müller-Hannemann, M., Schulz, F., Wagner, D., and Zaroliagis, C. (2007). Timetable information: Models and algorithms. In *Algorithmic Methods for Railway Optimization*, pages 67–90. Springer.

Nielsen, O. A., Landex, O., and Frederiksen, R. D. (2009). *Passenger delay models for rail networks*, pages 1–23. Springer US, Boston, MA.

Niks, F., Schuurman, S., and Leemhuis, S. (2015). Functioneel Ontwerp SOM. Unpublished internal document, ORTEC B.V.

NS (2015). Toetsdocument Aannames RPUN. Unpublished internal document, N.V. Nederlandse Spoorwegen.

NS (2016). Vervoerplan 2017. https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2016/12/13/bijlage-1-vervoerplan-ns-2017/bijlage-1-vervoerplan-ns-2017.pdf. Last accessed on 15-06-2017.

NS (2017). Reisplanner. http://www.ns.nl/reisplanner. Last accessed on 02-03-2017.

ProRail (2017). Onze klanten. http://http://www.prorail.nl/vervoerders/onze-klanten. Last accessed on 15-06-2017).

Serafini, P. and Ukovich, W. (1989). A mathematical model for periodic scheduling problems. *SIAM Journal on Discrete Mathematics*, 2(4):550–581.

van der Berg, B. (2015). Functioneel ontwerp RPUN v1.1. Unpublished internal document, N.V. Nederlandse Spoorwegen.

Vromans, M. (2005). *Reliability of Railway Systems*. PhD thesis, Erasmus Universiteit Rotterdam.

Wolters, G. (2016). Passenger punctuality: An analysis of the method of calculation and describing models. Master's thesis, Delft University of Technology.

A Assumptions passenger punctuality

The assumptions differ from passenger behavior to the accuracy of the used data. The absence of module 3 is covered by assumptions 17 to 20. Instead of rescheduling the check-out time is used to determine the realized arrival time. This paragraph merely mentions this assumptions. No further analysis of the impact of the assumptions on the passenger punctuality is given since this is not relevant for the remaining of the thesis.

	Assumptions
1	The selected CICO pairs represent the average passenger.
2	OVCP-passengers do not travel with Thalys, Eurostar and/or ICE.
3	The closing of OVCP terminals at stations do not have an influence on the representativeness of the CICO pairs.
4	The passenger wants to go as quickly as possible from A to B, even if this does not concern a train of the NS.
5	The check-out station is the desired destination of the passenger.
6	The check-in station is the desired origin of the passenger.
7	The check-in time is the desired starting time of the passenger.
8	In case of planned roadwork and trips (partly) conducted by bus are not taken in account.
9	The passenger follows the advise from the travel planner and hereby uses the default settings.
10	The passenger bases its expectations on the advise given by the journey planner two days prior to the journey.
11	Every trip from the OVCP data can be connected to a travel promise.
12	The used check-in margin per station is correct.
13	The clock in the card reader is exact.
14	The realized arrival and departures times are measured exactly on each station.
15	The determination of 45 seconds as stopping time at a shortstop is accurate.
16	The transfer margins from the realization model correspond with the promised transfer options.
17	If the first train has a delay of more than 15 minutes at departure, the check-out time minus the walking time is used as realized arrival time.
18	If a train does depart from its departure station, the check-out time minus the walking time is used as realized arrival time.
19	If a train does not arrive on its destination station, the check-out time minus the walking time is used as realized arrival time.
20	If a transfer is not made, the check-out time minus the walking time is used as realized arrival time.
21	The calculated check-out margin per station is correct.
22	Change in the location of the card readers has no effect on the KPI.

Table A.1: The 22 assumptions in the calculation of the passenger punctuality. NS (2015)

B Additional results evaluation model

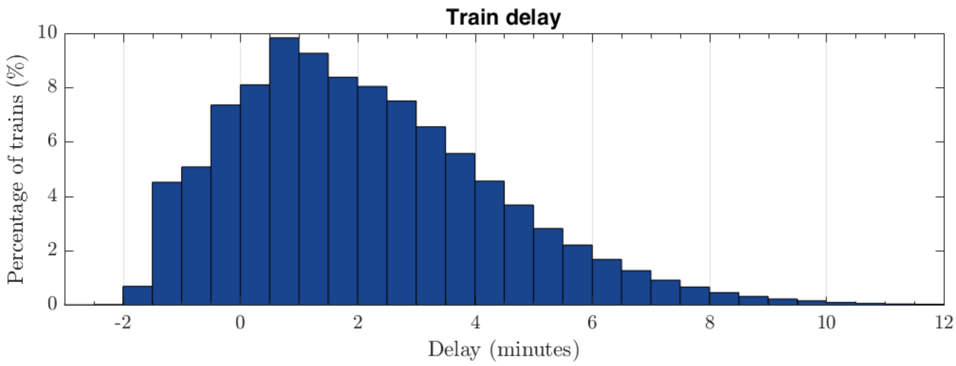


Figure B.1: The normalized distribution of the delay of trains at arrival.

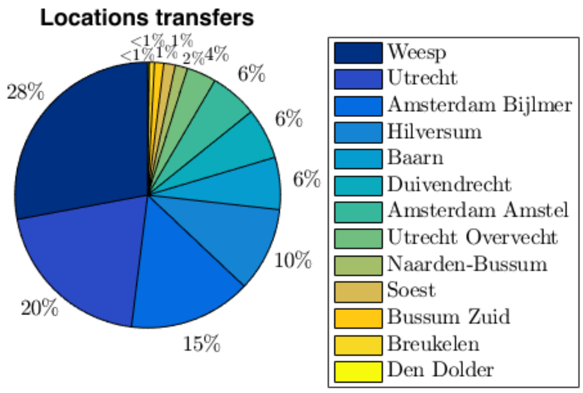


Figure B.2: The distribution of the transfers over the station.

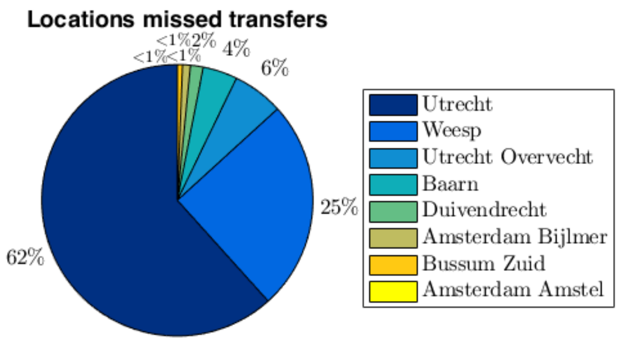


Figure B.3: The distribution of the missed transfers over the stations.

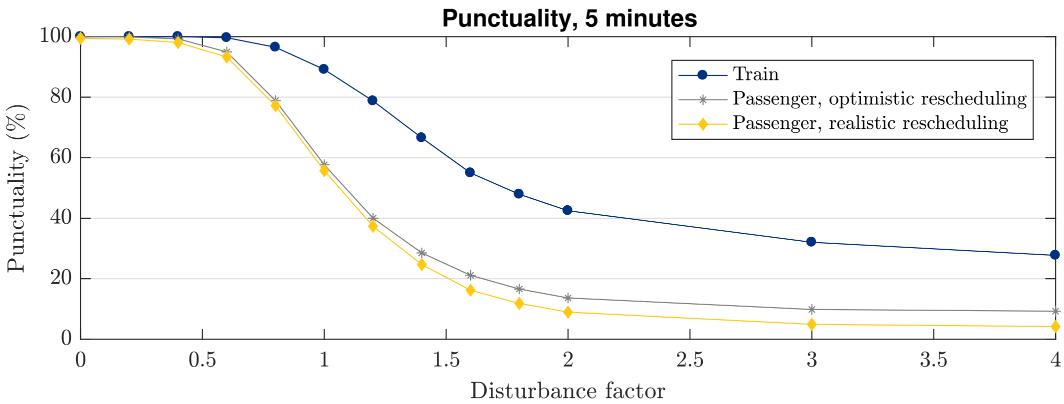


Figure B.4: The results of the 5-minute punctualities over different disturbance factors.

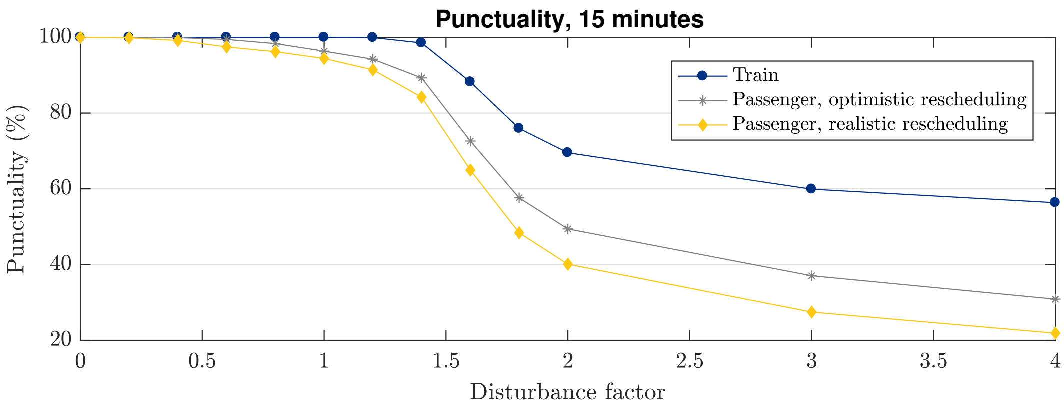


Figure B.5: The results of the 15-minute punctualities over different disturbance factors.

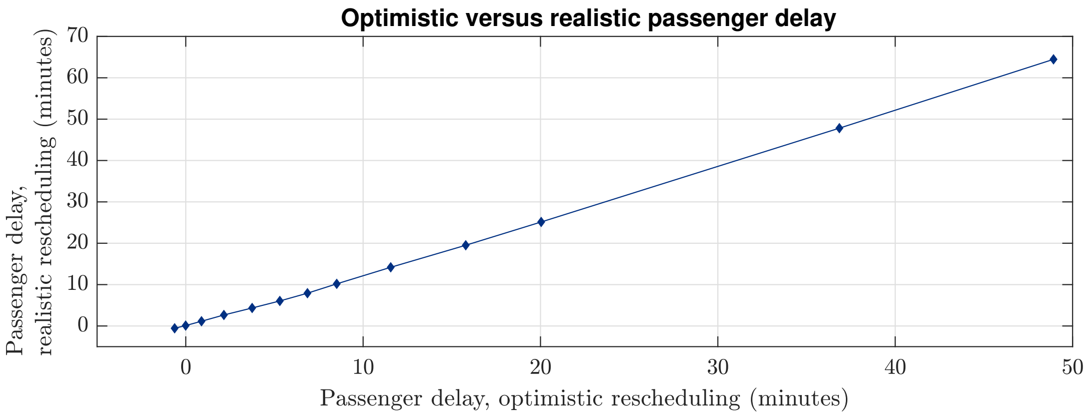


Figure B.6: The results regarding the passenger delay according to the optimistic rescheduling versus the passenger delay according to realistic rescheduling.