

From vehicle-based punctuality to passenger-based reliable rapid transit

Measuring network-wide effects in public transit

MSc Transport, Infrastructure and Logistics Thesis

Thieme Brandwagt

From vehicle-based punctuality to passenger-based reliable rapid transit

Measuring network-wide effects in public
transit

Thieme Brandwagt

to obtain the degree of Master of Science in Transport, Infrastructure and Logistics
at the Delft University of Technology,
to be defended publicly on Monday August 31, 2026 at 15:30 PM.

Student number:	5232910	
Project duration:	February 2, 2026 – August 31, 2026	
Thesis committee:	Dr. ir. Niels van Oort,	TU Delft, chair
	Prof. ir. Wijnand Veeneman,	TU Delft, supervisor
	ir. Erwin Stoker,	OV-bureau Groningen Drenthe

Cover: busposities.nl © MHVentura

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

Preface

First and foremost, I would like to thank my supervisors from the TU Delft. Niels van Oort and Wijnand Veeneman provided good guidance throughout the development of this thesis. While Niels was always honest about the state of my work, he always ended with a positive note, which I truly appreciated. Wijnand helped me step back from the mathematical modelling process and assisted me in designing the stakeholder session, which better placed my research into both practical and academical perspectives.

Second, I would like to express my gratitude to my friends and family. They were able to distract me when necessary, and leave me to it when a deadline was closing in. Special thanks to my girlfriend. Although the combination of moving in and writing a thesis was sometimes challenging, she was always there for me.

Third, during my internship at the OV-bureau Groningen Drenthe, I was inspired by many stimulating discussions about transit, which really motivated me to progress. From the first day, the team made me feel welcome at the office in Assen. My room mates Karsten, Arent-Jan, Roy, Jorne and Flip were always ready to help and provided great company for writing a thesis. We all arrived at 9 AM, and took the same train to Groningen at 17:07. Special thanks to Kees and Leon for supporting my thesis with all the necessary datasets.

Last but not least, I would like to thank my supervisor from the OV-bureau Groningen Drenthe, Erwin Stoker. During our weekly update meetings, we would first discuss the national state of transit for half the meeting before diving into my thesis. He provided extensive feedback and was often able to temper my overenthusiastic opportunism with the reminder: "In der Beschränkung zeigt sich der Meister". In the end, he was right...

*Thieme Brandwagt
Groningen, August 2026*

Executive summary

0.1. Research gap

Literature describes vehicle punctuality or line regularity, while a fair share of the passengers use multiple vehicles or lines in their journey. This neglects the transfer time that is needed for waiting from one vehicle to the next. Furthermore, the trade-off between holding time and the different components of a full passenger journey are often unclear. These shortcomings in assessing transit timetables is the research gap this thesis aims to close.

0.2. Methodology

To assess the influence of including transfers in evaluating timetable quality from a passenger perspective, this thesis developed a passenger-centric methodology that explicitly models transfer time, as well as waiting time and in-vehicle time.

This thesis is unique in terms of the availability and combination of data sources. It uses three different data sources. Automatic Vehicle Location (AVL) was used for the simulation of vehicle trips. The other two data sources originate from smartcard data. Origin-destination (OD) data defined the number of direct and through passengers, while a dedicated transfer dataset provided the amount of transfers between the case study area and the cordon across multiple operators.

For estimating the experienced journey time of transferring passengers, this thesis introduces a network-time graph. The network-time graph consists of temporal nodes in which the simulated vehicle trip can be stored, as well as the transfer lines that connect the study area to a wider cordon. Passengers are assigned to their shortest paths using dynamic programming, which makes it possible to efficiently calculate the journey time of all passengers.

To include both travel time and reliability into one metric, a new network-wide metric is introduced. RRR aims to capture the *Rider-based Rapidness and Reliability*, by selecting the 80th percentile of passenger-specific journey time over multiple simulation runs. The indicator is introduced to combine speed and reliability into a single passenger-oriented metric, using the 80th percentile of passenger-specific journey times over multiple simulation runs to account for both average travel time and the risk of delays.

0.3. Case study

To demonstrate the use of the methodology, a case study is introduced. The city of Groningen, The Netherlands, has diametric bus lines which share a common corridor with multiple transfer points that connect to a wider transit network. For the case study, it is unclear how the allocated timetable slack propagates over the network and how it is experienced by direct, through and transferring passengers.

To investigate the trade-off between holding time and the experienced travel time by all types of passengers, alternative holding strategies were developed that were compared in terms of their RRR performance. The alternative holding strategies were altered in three ways. First, the percentile that determines the schedule-based departure time was changed. Second, the amount of control points at which holding is applied was altered. Finally, a maximum holding time was set for headway-based holding strategies.

0.4. Results

Including transfers in the analysis revealed that the traditional assumption of a uniform passenger arrival distribution at transfer stops is incomplete. Instead, passenger arrivals at transfer stations are clustered around scheduled departure and arrival times. As a result, line regularity becomes less critical for the total experienced travel time in a transfer-dependent network. The inclusion of transfers negatively im-

pacted the performance of holding strategies, as the synchronized arrival and departure times at transfer stations reduced the effectiveness of holding interventions. Results showed that schedule-based holding could improve network performance, with the schedule tightness influencing the effectiveness of holding control. However, when transfers were included, holding strategies with less scheduled slack generally performed better, as the actual passenger arrival patterns at transfer stops deviated from the assumed uniform distribution.

This comparison could be made, because the simulation was conducted with three different setups. At first, only direct and through passengers of the OD dataset were included. Then an assumption was made that overestimated the amount of transfers. Finally, the transfer dataset revealed the case-specific best performing holding strategies. For the specific case study in this thesis, including transfers in a demand-oriented indicator favoured holding strategies with less scheduled slack. The analysis showed that neglecting transfer times risks selecting a holding strategy that appears optimal on paper but fails to account for unobserved passenger experiences in practice.

Stakeholder perspectives on a timetable quality indicator that includes transfers were gathered through a focus group session. Stakeholders acknowledged the importance of transfer time and agreed that a demand-oriented indicator is needed to better capture passenger experiences. However, they noted that such an indicator is complex to implement and may be more suitable as a guideline for timetable development rather than a strict concession requirement. There was consensus that transfers should be more carefully monitored, with greater attention to the probability of successful transfers and the synchronization of heavily used connections.

0.5. Discussion

The major critique with regards to the used methodology is that passengers are not constrained by the vehicle capacity. Hence, denied boarding is not taken into account. Additionally, buses do not dwell longer once there are more passengers on the stops, which prevents from properly modelling bunching effects. A recommendation to future research is the incorporate crowding constraints within the proposed methodology.

The other way around, a recommendation to future research is to include transfers to a wider network at all cost, since it is a critical component of a passenger journey. This is strengthened by the fact that passengers experience a transfer more negatively, as compared to in-vehicle time or waiting time. If data is missing on the amount of transferring passengers, assumptions should be made that prevent uniform arrival patterns at transfer stations. Further research is needed to refine the integration of transfers into agent-based models, allowing for the calculation of elasticity effects and a more comprehensive understanding of passenger behaviour.

0.6. Conclusion

The main takeaway from this thesis is that transfer times should not be neglected in timetable-related research, as they significantly influence results. Including all transfers that connect a study area with a cordon can yield different outcomes for the performance of holding control strategies. The root cause of these differences lies in the assumed passenger arrival patterns at transfer stops, which are typically clustered around scheduled times rather than uniformly distributed. Since every network differs in terms of transfer dependency, the generalizability of these findings is limited. Future research should focus on integrating transfers into agent-based models to account for capacity constraints, bunching, and elasticity effects, enabling a more dynamic analysis of transit systems over time.

The network-time graph that is described in the methodology can aid decision makers in assessing a transit network in terms of reliability and speed as experienced by the passenger. This research demonstrates that a transit agency can operationalize the collection and analysis of journey-level data to create a demand-oriented metric that is able to quantify the level of reliability and speed in of transit timetables.

Contents

Preface	i
Executive summary	ii
0.1 Research gap	ii
0.2 Methodology	ii
0.3 Case study	ii
0.4 Results	ii
0.5 Discussion	iii
0.6 Conclusion	iii
Nomenclature	ix
1 Introduction	1
1.1 Problem statement	1
1.2 Research questions	2
1.3 Scope	2
1.4 Scientific contribution	5
1.5 Practical relevance	5
1.6 Report outline	5
2 Literature Review	7
2.1 From supply-based punctuality to demand-based reliability	7
2.1.1 Vehicle punctuality	7
2.1.2 Line regularity	8
2.1.3 Network reliability	8
2.1.4 Combining speed and reliability	9
2.2 Simulation	10
2.2.1 BusMezzo	10
2.2.2 Monte Carlo	11
2.3 Graph Theory	11
2.4 Holding control	11
2.4.1 Schedule-based holding versus headway-based holding	12
2.4.2 Transferability	13
2.4.3 Single-line holding versus multi-line holding	13
2.5 Synthesis	13
3 Methodology	16
3.1 Holding strategy selection	18
3.1.1 Holding types	18
3.1.2 Assumptions	18
3.2 Simulation of transit vehicle trips	19
3.2.1 Drawing from the probability function	19
3.2.2 Assumptions	22
3.3 Assigning passengers in a network-time graph	23
3.3.1 Path finding	23
3.3.2 Dynamic Programming	24
3.3.3 Assumptions	26
3.4 Network indicator	27
3.4.1 Assumptions	28
3.4.2 Testing for significance	28
3.4.3 Flow chart	29

3.5	Stakeholder perspectives	30
4	Case Study	32
4.1	Network characteristics	32
4.2	Challenges	32
4.3	Data availability	34
4.3.1	Case-specific sets and parameters	36
4.4	Vehicle patterns	37
4.5	Passenger patterns	40
5	Experimental setup	44
5.1	Selection of holding control strategies	44
5.2	Selection of simulation setups	44
5.3	Implementation of the case study in the network-time graph	46
6	Results	51
6.1	Timetable cost	51
6.1.1	Tension between drivers, passengers, operator and authority	51
6.2	Rider-based Rapidity and Reliability	52
6.2.1	Excluding transfers	52
6.2.2	Including transfer assumption	54
6.2.3	Including transfer data	55
6.2.4	Comparison of results	55
6.2.5	Extended analysis	56
6.2.6	Replications needed for significance	56
6.3	Stakeholder relations	57
6.3.1	The punctuality paradox	59
6.3.2	Ridership spirals	59
6.3.3	Bridging the gap between short and long-headway services	60
7	Discussion	61
7.1	Dependency between demand and supply	61
7.2	Deterministic passenger assignment	62
7.3	Time elapsing parameters	63
7.4	Limited strategies	63
7.5	Elasticity	64
7.6	Equity	65
7.6.1	Alternative equity principles	65
7.6.2	Exclusion types	66
7.7	Practical implications of the RRR indicator	67
7.7.1	Incorporating a demand-oriented indicator as concession requirement	68
7.8	Research generalizability	68
8	Conclusion	70
8.1	Answers to the research questions	70
8.2	Scientific recommendations	71
8.3	Practical recommendations	72
	References	73
A	Pseudo code	78
B	Network maps	80
C	AI statement	83
D	Indicator significance	84
E	Bus-bus transfer numbers	87

List of Figures

1.1	Elements of a passenger journey (based on Van Oort (2011))	2
1.2	Diametric bus lines with multiple transfer stations that connect the study area to the outer cordon	3
1.3	Holding control positioned in the transit network planning process (based on Vuchic (2005), Van Oort (2011) and Ibarra-Rojas et al. (2015))	4
2.1	Passenger requirements for using transit (Peek and Van Hagen 2002)	7
2.2	Theoretical framework on reliability	9
2.3	The relation between speed, reliability and schedule tightness	10
2.4	Directed graph for calculating shortest paths over transfer connections (Nguyen and Pallottino 1988)	12
2.5	Conceptual model	14
3.1	Methodological Framework	17
3.2	Controlling the bunching phenomenon; U: uniform headways stage, B: bunching stage, S: schedule-based holding intervention result, H: headway-based holding intervention result	18
3.3	Calculation of vehicle runs based on historic AVL data	19
3.4	Schedule-based holding criterion based on the chosen segment percentile n	21
3.5	The headway-based holding criterion in a line-time graph (left: no holding intervention, right: headway-based holding intervention)	21
3.6	Transfer problem example for passengers traveling between A and D	24
3.7	Comparison of journey feasibility in the network-time graph	24
3.8	Dynamic programming for shortest path calculation	25
3.9	Integrating speed and reliability for a single OD-pair (based on Uniman et al. (2010) and Van Oort (2011))	27
3.10	Flowchart of the complete algorithm (inspired by Cats, Burghout, et al. (2010))	29
4.1	Q-link network of Groningen (source: Van den Brink, 2022)	33
4.2	Vehicle exterior matching the line color. In this case: a blue bus serving Q-link 4 (source: busposities.nl).	33
4.3	Train station directional categorization (source: Translink)	35
4.4	Bus stop directional categorization (source: Translink)	36
4.5	Q-link 3 eastbound trip time distribution per hour	38
4.6	Q-link 3 westbound trip time distribution per hour	38
4.7	Q-link 4 eastbound trip time distribution per hour	39
4.8	Q-link 4 westbound trip time distribution per hour	39
4.9	Holding time distribution at <i>Hoofdstation</i>	40
4.10	Eastbound load profile of an average Q-link 3 or 4 bus between 8 and 9 AM	41
4.11	Assumed (vice versa) transfer relations at <i>Hoofdstation</i> (with corresponding train operator)	41
4.12	Shankey diagram of the average daily amount of transferring passengers at <i>Hoofdstation</i> from train to bus and vice versa	42
5.1	Schedule-based holding strategies	45
5.2	Network graph with all line directions	47
5.3	Network-time graph with two vehicle trips per line direction running between control points	48
5.4	Network-time graph with two vehicle trips per line direction running between all stops	49
5.5	Network-time graph including bus-train transfers between 8:00 and 8:30 am	50
6.1	Strategy performance of 50 runs excluding transfers	53

6.2	Strategy performance of 50 runs including transfer assumption	54
6.3	Strategy performance of 50 runs including transfer data	55
6.4	Extended strategy performance of 20 runs including transfer data	57
6.5	The influence of speed and reliability on the inter-stakeholder relationships	58
A.1	Pseudo code of the algorithm used	79
B.1	Metro-style Q-link network and its transfer opportunities to train and long-distance bus services (source: Van den Brink, 2025)	81
B.2	Geographical network representation	81
B.3	Metro-style network map from the city of Groningen (source: Van den Brink, 2025) . . .	82

List of Tables

2.1	Literature on holding	12
4.1	Case-specific data availability	34
4.2	Top 10 most used bus on bus transfers (daily average amount of passengers)	43
5.1	Strategies to be simulated	45
6.1	The daily cost of timetable strategies (based on $m=50$ initial runs)	51
6.2	Comparative analysis: Dominating holding strategy per hour of day type and per demand data source	56
6.3	Extended analysis: dominating holding strategy per hour of day type (OD + TLS demand dataset)	56
7.1	The adaptability of the methodology to minimize bias caused by assumptions	61
D.1	Significance of the RRR indicator with the transfer assumption ($m=50, b=10$)	85
D.2	Significance of the RRR indicator with transfer data ($m=50, b=10$)	86
E.1	Top 50 most used bus on bus transfers (daily average amount of passengers)	88

Nomenclature

Abbreviations

Abbreviation	Definition
AVL	Automated Vehicle Location
CP	Control Point
DP	Dynamic Programming
GTFS	General Transit Feed Specification
HB	Headway-based holding
OD	Origin-destination
OTP	On-Time Performance
RBT	Reliability Buffer Time
RRR	Rider-based Rapidity and Reliability indicator
SB	Schedule-based holding
TTH	Time table hours, <i>Dienstregelinguur (DRU)</i>

1

Introduction

1.1. Problem statement

While transit plays an important role in urban mobility, accessibility, liveability and sustainability, it is faced with operational problems. There is a high variation time it takes to cover segments between stops, which varies over the course of a day. As the abundance of other road users leads to congestion, transit vehicles often end up in traffic jams, when dedicated infrastructure is missing. Furthermore, the amount of passengers varies over the day too, which results in overcrowded buses in peak hour, while there is a low demand in off-peak hours.

This is problematic for passenger satisfaction, since key determinants of passenger satisfaction are reliability and speed (Peek and Van Hagen 2002). Specifically, passenger travel time and its associated punctuality is a catalyst for long-term ridership. Traditionally, the assessment of transit punctuality comprised of measuring the performance of individual transit vehicles, which tends to overlook the passenger experience (Rietveld 2005). Ait-Ali (2024) suggests that adopting passenger-centric metrics is essential for improving transit performance monitoring. These metrics can provide valuable insights that traditional vehicle-centric approaches may overlook, particularly in understanding the impact of passenger demand variations in space and time. By combining passenger-centric and vehicle-centric metrics, transit stakeholders can make more informed, data-driven decisions that enhance service performance, operational efficiency, and ultimately, passenger satisfaction.

Currently, vehicle-centric metrics are often preferred when designing transit timetables, despite the excessive amount of theory towards more passenger-centric measures such as Van Oort, Wilson, et al. 2010; Wolters 2016, Van der Werff et al. 2019 and Yap et al. 2022. For example, an improvement to better cover more elements of the passenger journey is to study the line regularity as a consequence of headway variability (Van der Werff et al. 2019). Instead of the vehicle punctuality, it now includes the waiting time and its variation for a single line. To further assess the time passengers spend in the network, an approach is missing that explicitly includes more elements of which a whole transit trip comprises.

An aspect that is often neglected within passenger-centric measures is the transfer time. Although studies exist that take into account transfers within a studied network itself (Hadas and Ceder 2010; Cats, Koutsopoulos, et al. 2011; Nesheli and Ceder 2017), connecting a study area to a wider network is often not considered. This thesis explores an alternative speed and punctuality metric with more focus on the passenger perspective that more accurately reflects the journey of all passengers, including transferring passengers.

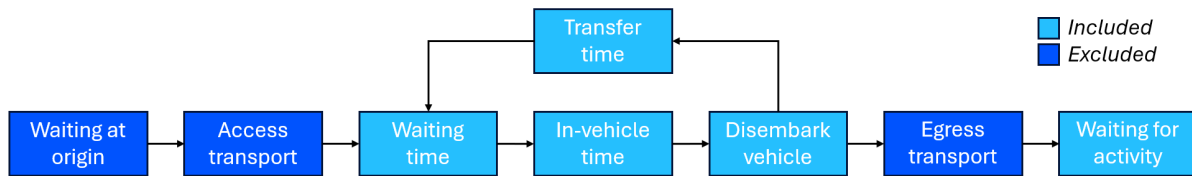


Figure 1.1: Elements of a passenger journey (based on Van Oort (2011))

1.2. Research questions

The goal of this thesis is to find a suitable method that includes transfer time in the appreciation of holding strategies. The main research question is therefore formulated as follows:

What is the influence of including transfers in assessing the quality of the timetable as experienced by passengers?

To answer the main research question, the following sub research questions are formulated:

- A) *How can the speed and reliability of a transit timetable be measured if transfers are included?*
- B) *What is the effect of including transfers on the performance of holding strategies?*
- C) *What are the perspectives of stakeholders on a timetable quality indicator that assesses passenger journeys including transfers?*

1.3. Scope

The section defines the bounds of this thesis. It underlines what lies within scope, and how this relates to variables that are out of the thesis scope.

This thesis focuses on incorporating transfers to make it able to assess the reliability and speed of more complete passenger journeys from origin stop to destination stop. Looking at an entire journey, this thesis takes into account the elements given in Figure 1.1. Waiting at the origin until the journey begins is not considered. Access transport, for instance biking or walking to a transit stop, is not considered. The first variable that is included in the analysis, is the waiting time until the first vehicle arrives. The in-vehicle time is included. After the passenger disembarks the vehicle, the passenger decides either to conclude their journey and wait for the activity to begin, or walks to the transfer point. At the transfer point, it waits for the next vehicle. This feedback loop is sustained until the final stop is reached. Egress transport, the inverse of access transport is not considered. Waiting at the destination before the activity starts is the last variable that is included in the analysis. The latter is chosen over waiting at the origin, due to the available type of demand data during this thesis.

Within the shift of vehicle-oriented to passenger-oriented, this thesis gasps on multiple facets. Rietveld (2005) found that vehicle-oriented measures of quality lead to a systematic overestimate of quality as experienced by travellers. To cover the unobserved quality elements, the study suggests shifting towards passenger-oriented measures. The study recommended shifting the focus for the following elements, from supply-oriented towards demand-oriented:

1. from inter-arrival times of buses towards the **mean waiting time of passengers**
2. from seat occupancy rate towards **experienced crowding**
3. from vehicle punctuality towards **passenger punctuality**
4. from missed vehicle transfer connection probability towards **passenger transfer probability**
5. from distance between stops towards **average walking distance for passengers**

This thesis is specifically scoped on the transfer probability of passengers (4). Often, transfer time is excluded from analysis, which poses a fundamental problem. The probability of successful transfer connections is closely related to all other above mentioned numbers. When a transfer connection is neglected, part of the passenger journey variables are missing. For example, this prevents calculating the proper punctuality of passengers (3), since it is unknown if the final destination stop has been reached. The other way around, an analysis could miss a prior vehicle trip, neglecting the transfer time

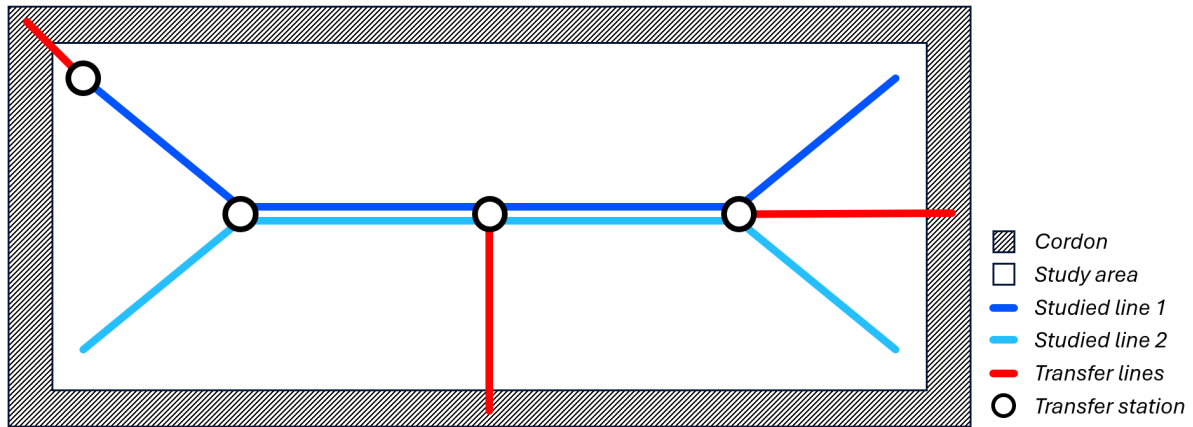


Figure 1.2: Diametric bus lines with multiple transfer stations that connect the study area to the outer cordon

that is needed before reaching the study area itself. Thus, explicitly including transfers into research on transit is necessary for calculating the quality performance of a single line, or a wider transit network.

In this thesis, the role of the other numbers are as follows. The mean waiting time of passengers (1) is explicitly included in this thesis. As this thesis aims to include the full passenger journeys, the passenger punctuality (3) is within scope, too. On the contrary, the experienced crowding (2) and average walking distance (5) for passengers lie out of the scope of this thesis.

This thesis is scoped on the tactical and operational planning process within the transit network planning process. Considering the whole process, Figure 1.3 shows the concepts within scope, and gives a wider framework that connect to the problems at hand, but lie out of scope for this thesis.

Starting at the long-term decisions, the strategical planning step lies out of scope. The strategic planning focuses on network development, infrastructure and the line configuration. The network of the strategical planning is forwarded towards the tactical planning phase. The network that is used is assumed fixed and will not be changed during the course of this thesis. The lines and stops are taken as a given. The methodology adapted by this thesis that is used to calculate the reliability and speed of passengers is carried out on a network that consists of branched diametric bus lines (Vuchic 2005). This thesis specifically focuses on a high-frequency network with a transfer-dependent central hub. At transfer points, the studied network connects to other lines in a given wider network, which is not altered. Two branch lines that share the same bus-dedicated corridor are under study, where overtaking is possible. The timetables of the two lines are the only ones to be manipulated. A conceptualization of the network in this thesis can be found in Figure 1.2. The study area is connected to the cordon area by transfer lines, which are connected to the studied lines at transfer stations. The transfer time at the transfer stations is used to make a more realistic estimate on the experienced passenger journey time.

The tactical planning step lies within scope. In the tactical planning phase, schedules are created for the service. Crew rosters and vehicle cycles are created in this phase, too. The operational, short-term planning, manages the day-to-day operation, for example managing fleet and crew assignment. Another operational action is the execution of control strategies. The two decision variables within this thesis are designed in the tactical step. For estimating the role of transfers in timetable assessment, this thesis generates alternative schedules. In this thesis, three variables are altered. The first is the tightness of the schedule. The second is the amount of control points. The third is the maximum holding time, specific to headway-based holding.

The operational step lies within scope. For this thesis, the alternative schedules are created by applying different holding control strategies. Both schedule-based holding and headway-based holding strategies are introduced in this thesis. Holding belongs to a family of station control interventions. The other station control intervention is stop-skipping (Eberlein et al. 1998), useful when a delayed vehicle is followed too closely by another vehicle having the same destination. Stop-skipping comprises of dead-heading, when passengers are asked to leave the vehicle so that a vehicle can make up for a delay. Expressing prevents new boardings, since buses do not stop to pick up new passengers. Short-turning

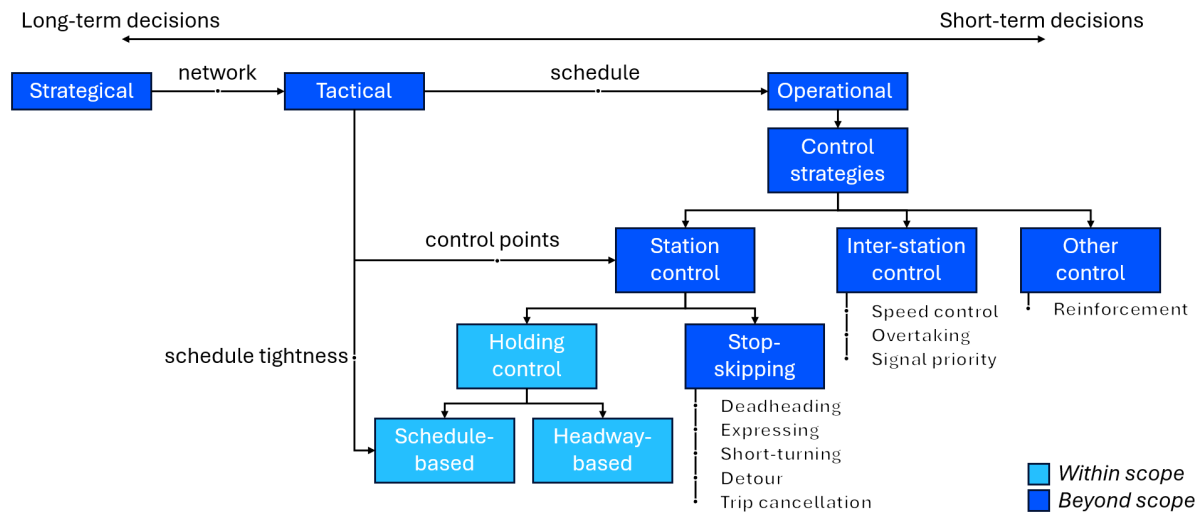


Figure 1.3: Holding control positioned in the transit network planning process (based on Vuchic (2005), Van Oort (2011) and Ibarra-Rojas et al. (2015))

terminates a trip sooner than planned, allowing for a quicker return trip. Detours are used when routes are blocked or a short-cut is available. In extreme unfavourable circumstances, trip cancellation can be used to align a vehicle to the original schedule. During this thesis, the only station control strategy that is used, is holding control. Station control itself is part of a wider collection of control strategies, that could also be applied inter-station. Between stops, the speed can be controlled, overtaking can be initiated and transit vehicles can be prioritized at signalized intersections. Although the network in this study has both instances equipped, the effect of these control strategies lies out of the scope of this thesis. The odd one out, is reinforcement control, where an additional supply of vehicles can be deployed to cope with sudden demand increase or disruptions. This latter control strategie lies beyond scope, too.

1.4. Scientific contribution

The scientific contribution of this thesis is threefold:

1. **Methodological:** A new transit network graph representation is introduced, the *network-time graph*, which connects pairs of time-dependent nodes based on simulated service runs. This graph can be used in determining the travel time of journeys that consist of transfers.
2. **Applied:** Based on the topology of the network-time graph, a demand-oriented performance indicator is defined that aims to capture the speed and punctuality of passengers, including transfers. This performance indicator utilizes the set of simulation runs. While a mean or median travel time is able to represent the speed, reliability is not taken into account. Instead, this thesis uses a higher percentile of journey time that is able to include both the speed and reliability into one metric.
3. **Empirical:** The methodology is applied to a real-world network and compares the passenger performance indicator for different holding control strategies. The case study network is characterized as diametric, consisting of a transfer-dependent central hub. Providing this case-study, this thesis contributes to body of academical work on assessing holding strategies.

1.5. Practical relevance

In the field of transportation, there is a need to assess the quality of strategic, tactical and operational measures. Often, decisions are based on the performance of transit vehicle speed and punctuality. This says something about the reliability of service of individual vehicle runs. However, in reality passengers do not only experience in-vehicle time, but also waiting time, as well as transfer time from other connecting services. Thus, vehicle punctuality can give a distorted picture of the experienced passenger punctuality (Hendren et al. 2015). If a vehicle departs on time, it does not tell anything about the success of a transfer.

There is a lack of assessment tools that can assist in finding the optimum schedule that is able to provide a sufficient level of robustness, speed, transfer synchronization and exploitation cost. Specifically, an indicator is missing that is able to quantify the needs of a passenger. This thesis aims to cover this missing variable by introducing a network-wide indicator that is able to quantify the speed and reliability of passengers in a network. A corresponding methodology that uses graph theory is provided for estimating this indicator.

Although graph theory and network science try to fulfil the role of these assessment roles, in practice they are barely used for decision-making. There has been a widely recognized concern that in network science, transit networks are primarily used as real-world examples for theoretical developments (Šfiligoj et al. 2025), rather than problem-solving solutions. The lack of implementation studies has led to a weak benchmark for utilizing graphs in the field of transportation.

Data becomes increasingly accessible. Automatic Vehicle Location (AVL) data and minute-by-minute passenger demand are now widely available data sources. This offers the opportunity to combine all that is known into an indicator that is better able to cover the needs of the passenger. Furthermore, passengers are better able to actively plan their journeys both static and real-time, due to the standardized data format GTFS: General Transit Feed Specification (Google 2024). These emerging data streams provide a richer and more precise understanding of passenger demand and transit operations (Ghofrani et al. 2018).

1.6. Report outline

This thesis is structured as follows: Chapter 2 presents an overview of related literature. In the qualitative review, traditional supply-oriented punctuality studies are discussed. Then, articles on line regularity are reviewed. Subsequently, literature is reviewed that gasps on network reliability. Literature on simulation, graph theory and holding control is provided, too. Chapter 3 focuses on the methods that are used to answer the research questions. This includes the methodological steps to construct the network-time graph, the definition of the used demand-oriented performance indicator and how the selection of holding strategies is decided upon.

This thesis uses a case study to be able to apply the methodology in a contextual setting. The case study is introduced in Chapter 4. The network characteristics are described, as well as the main challenges and the available data sets. Furthermore a brief supply-oriented analysis is conducted, as well as some background on the passenger loads on the lines. Chapter 6 presents the results, of which the outcomes are discussed in Chapter 7. In this chapter, the assumptions made in the methodology are discussed, too. The thesis conclusion is to be found in Chapter 8.

By the means of assumptions, the actual system can be theoretically modelled. Analysis and computations are done that lead to results. Consequently, they lead to the findings and recommendations. It is important to take into account the interpretation by the actors in order for a future system to be implemented (Vuchic 2005). To bridge the gap between academia and practice, this thesis is constructed in close corporation with transit stakeholders: the transit authority, transit operator, passengers and two levels of political actors. From now on, the stakeholder contributions to this thesis are highlighted by a gray box, as the one this text is part of. This distinction is made so that the stakeholders were able to perform a final check on their contributions more easily, before the thesis was published.

2

Literature Review

Literature is reviewed on four subjects to gain theoretical insights on the research questions and to enhance the base for empirical research. It starts with an overview of academic works on punctuality and reliability. Then, the use of graph theory in the field of transit is elaborated on. Subsequently, simulation methods are discussed and finally, a brief overview is given of literature on holding control.

2.1. From supply-based punctuality to demand-based reliability

As transit systems are expected to have reliable service, punctuality is often used as an indicator for qualifying a transit system. Transit punctuality appears in different forms in the passenger satisfaction pyramid in Figure 2.1. First with the pyramid base: reliability, which can be the greatest dissatisfier if not managed well. Reliability is the base of good service. A perfectly reliable system is inherent to the speed. When a schedule is designed with a lot of slack or layover time, the reliability may be maximized (Furth and Muller 2008; Y. Lee et al. 2017). This does not necessarily mean that passengers are satisfied most. If a loose schedule means that vehicles idle a lot, it takes a toll on the speed and increases operational cost (Mazloumi et al. 2012). Speed can be a dissatisfier to a lesser extent, compared to reliability (Peek and Van Hagen 2002). The optimal amount of slack and layover time is decided by the desired balance between reliability and speed (Friedrich et al. 2018).

Measuring and monitoring reliability and speed are complex, as it depends on multiple factors, including network characteristics, demand patterns, service design, traffic conditions, and operational management (Ait-Ali 2024). Therefore, this section reviews previous research that can give theoretical background into supply-oriented punctuality measures and gradually moves towards theory on demand-oriented reliability.

2.1.1. Vehicle punctuality

Traditionally, punctuality is often measured in *On-Time Performance*, in short OTP (Rudnicki 1997). OTP represents the proportion of trips that meets a set punctuality threshold. This threshold is often lower for short-headway services, compared to long-headway services. The vehicle punctuality is

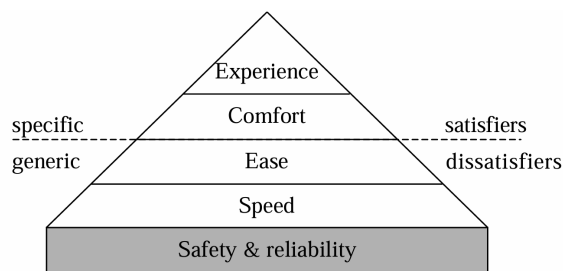


Figure 2.1: Passenger requirements for using transit (Peek and Van Hagen 2002)

mostly measured at the beginning or end of a trip, but it can also be measured along the way (Denti and Burroni 2023).

Vehicle punctuality measures can give tangible advice to transit operators on where service is lacking. Especially when a vehicle is tracked during a lot of passing points. Joborn and Ranjbar (2022) demonstrate how punctuality data can explain the causes of delay. For example, bottlenecks can be tracked down at which transit vehicles experience delay more often.

Problematic with OTP is the binary nature: a vehicle is either on time or too late. This neglects the magnitude of lateness. If the threshold is 3 minutes, a 5 minute delayed vehicle is valued similar to a 30 minutes delayed vehicle. Furthermore, OTP has difficulties with cancelled vehicle trips (Barabino et al. 2015), although suggestions were made on how to include cancelled trips into OTP (Parbo et al. 2016) to better describe the needs of the passenger. In practice, cancellations often have their own metric, independent to punctuality.

In long-headway services, vehicle-based punctuality can give proper insight on how well the timetable is adhered to. This suggest that the punctuality is highly dependent on the timetable quality (Zhao et al. 2013). Thus, vehicle-based punctuality is more suitable when passengers rely on published schedules, compared to a network where passengers arrive at random (Cramer et al. 2009).

The main pitfall of supply-oriented vehicle punctuality metrics arises with short-headway services. If all vehicles of the same line are delayed. Say, there is a planned headway of 10 minutes and all trips are delayed by 10 minutes, the OTP measure will define the system as under-performing. While in reality, the passengers experience no difference, as all vehicles are now driving in the path of their downstream predecessor. This indicates that headway regularity in short-headway services, is better able to explain the needs of a passenger, than vehicle punctuality (Van Oort 2011; Diab et al. 2015).

2.1.2. Line regularity

Then, other research includes the variation in headway that can explain the actual waiting time. For this, an assumption of the arrival pattern of passengers needs to be done. For small headway services, passengers are assumed to arrive at random. For long headway services, the arrival time is more clustered towards the actual departure. Critical note on this latter assumption is that it misses on one of the elements of a complete journey (Figure 1.1), which is the waiting time at the origin, before the trip starts or the waiting time on destination before the activity starts.

When looking at a complete line, more can be said about the waiting time of passengers at stops. Research considering waiting time is divided by the headway length. Studies about long-headway services (e.g. Furth and Muller (2007)) typically consider passenger arrivals at the stops to be determined on the schedule, while studies about short-headway services typically assume an uniform arrival pattern. There is a academic consensus that short-headway services occur with headways equal to or smaller than 10 minutes (Van Oort 2011; Ibarra-Rojas et al. 2015). Since the scope of this thesis is on short-headway services, only the latter group of articles is reviewed.

Barabino et al. (2015) introduced a binary variable that couples the line punctuality, based on AVL, with passenger demand. In the study, the stop-specific punctuality was defined as the fraction of passengers who will be served within an acceptably short interval after they arrive.

Rudnicki (1997) has introduced the covariance of the vehicle headway as a performance indicator of a line. From this indicator, the additional waiting time (Van Oort 2011) can be calculated as a consequence of reliability. More recent studies use this indicator, too (Van der Werff et al. 2019; Yap et al. 2022). Although the additional waiting time is now estimated, it does not tell something about the transfer time.

2.1.3. Network reliability

When the transfer time is included into the punctuality metric, more can be said about the performance of a wider network. Multiple studies try to specifically include the demand patterns of passengers, combined with a full trip including transfers.

Since 2015, the Dutch Railways use their Automated Fare Collection system to calculate passenger punctuality. The metric is calculated by dividing the number of journeys that arrived as promised, by the

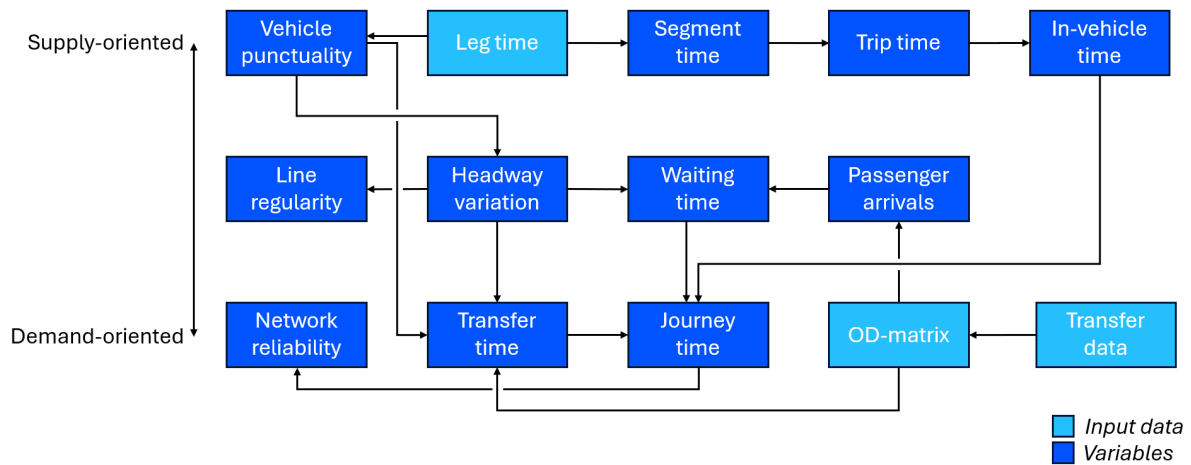


Figure 2.2: Theoretical framework on reliability

total amount of passenger journeys. The actual journey time is calculated by the difference between the tap-in and tap-out time. The promised time is the shortest scheduled path between the origin and destination, plus a predefined threshold. This metric directly translates into an OTP for passengers, instead of for vehicles and can adequately define the reliability of passengers (Wolters 2016). While it can be seen as demand-oriented punctuality, it is still binary by nature. There is no difference between a passenger experiencing a 10 or a 60 minute delay.

In transit systems that require passengers to both tap-in and tap-out in vehicles, the transfer time can be derived directly without the need of additional assumptions. Bagherian et al. (2016) demonstrate how this excessive data environment can lead to precise calculations on the duration and reliability of transfer movements.

Figure 2.2 gives an overview of the reliability types discussed. It is found that traditionally, vehicle punctuality is used. The On-Time-Performance is based on the leg time or a derived travel time over a segment or entire trip. This aids in finding the in-vehicle time, one of the elements of the total journey time. Then, waiting time is typically based on the headway variation, resulting in the line regularity. The waiting time itself is combined with the assumed passenger arrival pattern that is derived from the origin-destination matrix. The waiting time further adds up to the journey time. The transfer time can be determined on either the vehicle punctuality in long-headway services or headway variation in short-headway services, telling something about the reliability of passengers passing through a wider network.

Although passenger demand can be coupled to any punctuality or regularity metric, the ultimate passenger reliability metric can be reached when the demand is associated with the complete journey. If the resulting journey time is combined with temporal demand data, network reliability is the ultimate demand-oriented passenger reliability metric. Typically, temporal demand data comes in the form of a time-dependent origin-destination matrix. Network reliability is based on the OD-matrix, albeit with additional transfer data. The OD-matrix be manipulated with transfer data, in order to connect with the cordon (Figure 1.2). This can increase the explanatory value of the reliability indicator, since more complete journeys can be quantified.

2.1.4. Combining speed and reliability

Figure 2.3 gasps on the trade-off between reliability and speed. Both are related via the schedule tightness. If there is a loose schedule with plenty of slack and layover time, lines can operate perfectly reliable. The allocation of slack happens along the route, often at major time points, while layover time is the additional time built into the schedule at the end of the route that results in the total scheduled trip or cycle time (Desautniers and Hickman 2007). A common way to allocate slack and layover is the use of percentiles. Often, a higher percentile is chosen for segment (slack) or trip (layover) time determination, so that the majority of vehicles is able to arrive on time (Van Oort, Boterman, et al. 2009). However, due

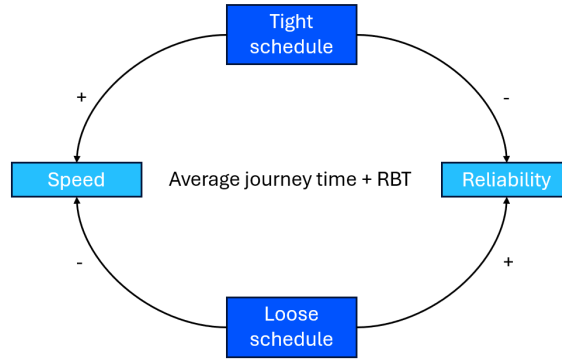


Figure 2.3: The relation between speed, reliability and schedule tightness

to the artificially extended schedule, the speed is lower. The other way around, when there is a tight schedule with little to none layover or slack time, the theoretical service can be sped up. In practice, not all trips will then be able to arrive or depart punctual, which leads to unreliable service and unstable journey times. Hence, holding is a trade-off between reliability and travel speed. Increased holding can improve schedule reliability while simultaneously increasing passenger in-vehicle time and operational costs (Furth and Muller 2009; Y. Lee et al. 2017).

Multiple studies combine speed and reliability into one metric. To include the reliability in the travel time, a Reliability Buffer Time (Uniman et al. 2010; Van Oort 2011) can be used to add to the total trip length. This is visualized in the middle of Figure 2.3. Average journey time represents the speed, while the RBT represents the buffer used by passengers in order to arrive on time at the destination. With this RBT variable, speed and reliability can be included into one metric. The advantage of combining the two is that the effect of speed increase or reliability reductions can be made monetarily, as illustrated by Van Oort (2016). The study provided evidence that the benefits of a tram project were underestimated, because reliability was not taken into account in the cost-benefit analysis. The study assigned a willingness to pay to both the speed and the reliability, which more adequately covered the complete benefits of the project.

2.2. Simulation

Simulation can be used for recreating an average day of service. This is needed to estimate the network travel time for an alternative timetable that not yet exists. Although the scheduled travel times are deterministic and symmetric due to the assumed equal headways, the driving and boarding times of buses are stochastic. Driving behaviour is different for every driver and the amount and type of passengers is always heterogeneous. For this reason, simulation is used for estimating the total passenger travel time over the network. Simulation is preferred over optimization, when stochastic run and dwell times are at stake (Van Oort 2011). Optimization is more suited when deterministic run and dwell times are used. For instance, common practice is to optimize the transfer synchronization (K. Lee et al. 2022). A combination between the two is possible too, for instance in minimizing the passenger delay in case of disruption (Yap et al. 2022).

2.2.1. BusMezzo

For simulating vehicle runs, BusMezzo is often used within literature. BusMezzo is a mesoscopic multi-modal traffic simulator, which dynamically simulates public transport operations and passenger assignment. BusMezzo makes it possible to create hypothetical scenarios and line setups. It is able to calculate spill-over effects when transfers are missed in a multi-level network (Yap et al. 2022), consisting of different public transit types.

The way BusMezzo assigns passengers to the network is by using preset alternative paths with an unique set of line and transfer locations. Of each path, the utility is calculated according to the weights given of different time elapsing types. In-vehicle time, waiting time, walking time and transfer time are included in the utility function as separate components. For every decision stage, the utility function is calculated according to the predicted future state of the network. Then, the relative utility is deter-

mined by using a multi-nominal logit model, which is able to estimate the market share of any path and the corresponding passenger travel time (Cats, Koutsopoulos, et al. 2011). Because passengers are modeled as decision-makers, BusMezzo can be seen as an agent-based model.

The assignment of passengers within BusMezzo can also be dependent on specific passenger experience components, such as crowding levels, reliability and real-time information (Cats and West 2020). This can even include elasticity dynamics, in which passengers are assigned dynamically. Iteratively, an equilibrium can be found in which all passengers are close to having reached their perfect journey. The explicit congestion modeling can increase the associated benefits of improving a transit system is estimated 60 per cent higher than in conventional cost-benefit analysis, indicating that including reliability and the associated elasticity effects should not be overlooked (Cats, West, and Eliasson 2016).

2.2.2. Monte Carlo

With standardized data formats as AVL (Ghofrani et al. 2018) and GTFS (Google 2024) being widely available, the driving times between stops can be based on historical data. For a Monte Carlo simulation to be possible, the vehicle driving time distribution has to be known.

Monte Carlo is a computational method that is used to describe complex systems, processes or uncertainties. The name originates from the famous casino of the same name, because the method was often used for estimating chance, similar to gambling games. The Monte Carlo simulation is organized as follows;

1. **Problem definition:** there is a given process triggered by uncertain variables
2. **Building of mathematical model:** a model is build that explains the relationship between variables
3. **Randomized input:** generation of a large number of uncertain values, based on probability distribution function
4. **Executing simulation:** for every set of input, execute the simulation and store the result
5. **Analysing results:** The distribution of the resulting parameter is key, because the mean, median and standard deviation can be derived. Most importantly is the confidence interval, which tell how repeatable the results are. This can be calculated by a significance metric.

An example study that draws from the historical driving time distribution is Gavrilidou and Cats (2018). Still, this study uses BusMezzo to further simulate passenger flows. Similar to BusMezzo, the number of required runs for any type of simulation can be determined by checking whether the results lie in the confidence interval (Burghout 2004).

2.3. Graph Theory

For the assignment of passengers, graph theory can be used. Graph theory makes it possible to calculate the shortest path between any pair of nodes. Nguyen and Pallottino (1988) suggested representing transit networks as directed arcs, making it possible to include transfer interactions. Figure 2.4 shows how a transfer connection can be modeled. From both lines, the stop node is connected via boarding arcs. This makes it possible to specify which transfers are feasible, and which are not. Furthermore, it enables for calculating shortest paths in larger transit networks.

Šfiligoj et al. (2025) combine the L and P-spaces, respectively representing the infrastructure and service graphs, to measure accessibility. The newly introduced *access graph* only draws the edges that are reachable within a maximum threshold. When it comes to serving passengers, fairness comes to play. For an overview of all other fairness standpoints with regards to transportation policies, see Dijkstra (2025). For the assessment of equity Šfiligoj et al. (2025) suggest a Gini coefficient that strives for a minimum level of accessibility in a transit network. Consequently, time-based exclusion in the transport system (Luz and Portugal 2021) can be decreased.

2.4. Holding control

Extensive research is conducted on the relationship between public transit holding control and the travel time on the held line. Although transfer time is sometimes included, it is never investigated how the

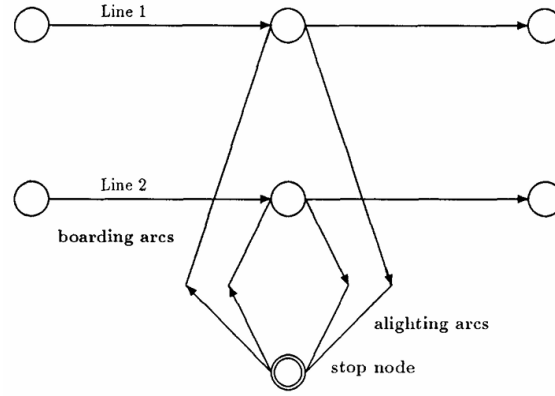


Figure 2.4: Directed graph for calculating shortest paths over transfer connections (Nguyen and Pallottino 1988)

Table 2.1: Literature on holding

Publication	Parameter	Transfer	Holding Type	Line	Tightness	Passenger Indicator	Overtaking	Capacity	Method
Hadas and Ceder (2010)	stochastic	✓	sched, head	multi	–	wait	–	–	DP, sim.
Van Oort, Wilson, et al. (2010)	stochastic	–	sched, head	single	var.	wait, ride, dwell, hold	–	–	sim.
Cats, Rufi, et al. (2014)	stochastic	–	sched, head	single	var.	wait, ride, dwell, hold	–	✓	sim (Busmezzo), opt.
Hernández et al. (2015)	stochastic	–	head	multi	–	wait, in-veh	–	✓	sim.
Schmöcker et al. (2016)	deterministic	–	–	multi	–	wait	✓	–	sim.
Nesheli and Ceder (2017)	stochastic	✓	head	multi	–	wait, in-veh, trans	–	✓	sim., MILP
Wu et al. (2017)	stochastic	–	sched, head	single	–	wait, in-veh	✓	✓	sim.
Wu et al. (2018)	stochastic	–	sched, head	single	–	wait, in-veh	–	✓	sim.
Imran et al. (2019)	stochastic	–	sched, head	single	85th	wait, ride, dwell, hold	–	✓	sim (Busmezzo)
Van der Werff et al. (2019)	stochastic	–	sched, head	single	80th	wait, in-veh, stand	–	✓	sim (Busmezzo)
Laskaris et al. (2021)	stochastic	–	sched, head	multi	–	wait, in-veh	–	–	sim (Busmezzo)
current thesis	stochastic	✓	sched, head	multi	var.	wait, in-veh, trans	✓	–	sim, DP

artificially extended travel time as a consequence of holding propagate over the network. Little is known about the holding policies and their resulting transfer interactions to a wider transit network. Table 2.1 gives an overview of the articles that measure the effects of holding.

The similarities across the considered literature is that most articles that include holding, maintain a stochastic approach to the trip time distribution (besides Schmöcker et al. (2016), who assume static exogenous delays). This is not surprising, since one of the goals of holding is to mitigate the effects of trip time variation. Different passenger metrics are used to capture the travel requirements for using transit. Hadas and Ceder (2010) and Schmöcker et al. (2016) only consider the waiting time on the studied lines. For all of the other studies in Table 2.1, a similar passenger indicator is used. Namely, a combination between waiting time and in-vehicle travel time. For articles that included transfer connections, the former is split into waiting time for the first vehicle and transfer time for the second vehicle. The latter is often split in dwelling time, driving time and holding time. The different components closely relate to the supply-demand scale in Chapter 2.1.

Most holding-related articles include both schedule-based and headway-based holding. For schedule-based holding, often an existing schedule is taken as a given. From Figure 2.3, it becomes clear that it highly influences the reliability and speed of a transit network. Most studies do not specify how the schedule is determined, besides Imran et al. (2019), and Van der Werff et al. (2019), who use respectively the 85th and 80th percentile of trip time distributions. There are studies that actively analyze the impact of the schedule tightness, which are Van Oort (2011) and Nesheli and Ceder (2017).

2.4.1. Schedule-based holding versus headway-based holding

Traditionally, schedule-based holding is applied in transit operation, because of its easy implementation (Van der Werff et al. 2019). Recently, real-time holding is applied, too. Although both holding types can use the same set of holding points, they differ in terms of holding duration. For schedule-based holding, the holding time is predetermined. While for headway-based holding, the decision on the amount of waiting time is made ad-hoc, depending on temporal distance to the successor and predecessor vehicle

(Ibarra-Rojas et al. 2015). In case of major disruptions, the buses can be held longer than usual, in order to prevent bunching and overcrowding. The costs of headway-based holding do not always outweigh the benefits. Based on a case study in The Netherlands, Van der Werff et al. (2019) made a decision tree in which real-time holding is compared to schedule-based holding, given the line characteristics. In general, headway-based control on high frequent routes score better than schedule-based holding, unless one of two following scenarios occur. Scenario 1: A lot of early arrivals, in combination with little disruptions. Scenario 2: Late arrivals in combination with enough turn around time and a high demand at the terminals. Note that both scenarios are dependent on the schedule tightness, indicating a large dependency on the set schedule.

2.4.2. Transferability

Research shows that people are disadvantaged by having to transfer in multi-segment trips (Ceder et al. 2013). Research on transfer coordination is mostly optimization-based, to calculate the best possible deterministic timetable. Mostly, a mathematical programming framework to calculate the optimal vehicle holding time to synchronize transfers for transferring passengers (Delgado et al. 2009). It is hard to research transfer coordination, while respecting the stochasticity of service. Furthermore, the selection of where control point should be located, is a complex process. Control points are typically chosen by the operator or transport authority, in order to coordinate transfers at busy stations (Nesheli and Ceder 2017). Using an optimization framework for that is possible too, in which the best location for control points can be determined (Cats, Rufi, et al. 2014).

Most articles excluded transfer interactions, besides Hadas and Ceder (2010). The study used dynamic programming to maximize the probability of transferring. This to reduce the total trip time, also if it contains multiple trip legs. In a later study, Nesheli and Ceder (2017) looked at the synchronization of transfers too, but only focused on headway-based holding.

2.4.3. Single-line holding versus multi-line holding

Most research focuses on single-line holding (Ibarra-Rojas et al. 2015). A handful of studies focus on multi-line holding, in which multiple low frequency branch lines share the same corridor for part of the route. Together, they contribute to a higher frequency on the corridor. The joint frequency can be coordinated, too, making the branch lines dependent on each other. For example, line 1 arrives on time, but has to wait for delayed line 2, which has just left, in order to maintain even headways. Laskaris et al. (2021) underline that the demand context matters: if more than 20% of the passengers travel on both the branch and the corridor, it is better to implement single-line control. A recommendation to future research they posed was to include a transfer criterion.

When considering a multi-line setup, the ability to overtake is of importance. In the study of Wu et al. (2017), overtaking was allowed, so that empty buses overtake full buses on purpose. Although this approach is only applicable with really high frequencies, notably the service regularity was improved. For multi-line operation using the same corridor, overtaking shows positive effects to the average waiting time. In contrast, no possibility of overtaking with heterogeneous lines has a negative impact on average waiting time (Schmöcker et al. 2016). The article suggest that infrastructure and network design that allows for overtaking is important to reduce the negative effects of delay-prone lines on the overall network performance.

2.5. Synthesis

Within literature and practice, there is an ongoing shift from supply-oriented punctuality metrics that are based on vehicle trips, towards demand-oriented metrics that are better able to cover the needs of passengers (Ait-Ali 2024). Over the course of a passenger journey, recently more elements are included that contribute to the quantification of the complete passenger experience in terms of speed and reliability.

Graph theory provides opportunities in connecting vehicle trips and to include transfers. In practice, graph theory is not used often, but is promising for connecting a case study area with a wider cordon. For respecting the stochasticity of service, simulation is preferred over optimization. A common tool used is BusMezzo, having the ability to include a lot of the complexity typical to the dynamic environment of transit operations. It can properly model bunching, crowding and elasticity phenomena.

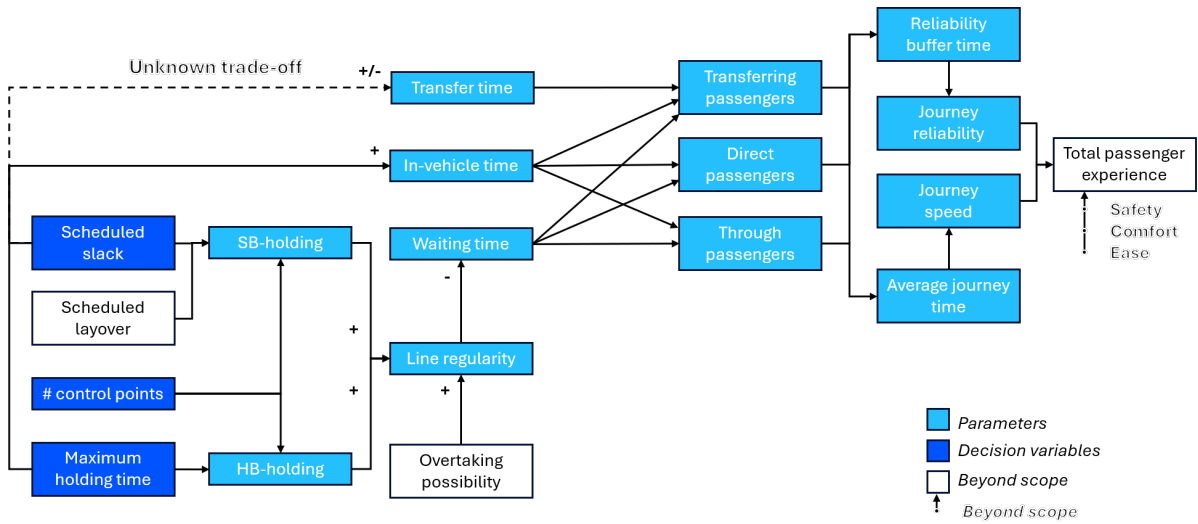


Figure 2.5: Conceptual model

For holding control, the main takeaway from the literature is that there is no generally applicable best holding strategy. The best holding strategy is primarily determined by network characteristics, overtaking rules, and the distribution of demand across the network. To further discover the impact of the network characteristics, case studies are commonly used.

To conceptualize all dependencies that are of relevance to this thesis in a model, Figure 2.5 is constructed. This thesis explores different holding strategies, in which the decision variables of scheduled slack, the amount of control points and the maximum holding time are altered. The scheduled layover is not considered in this thesis. Schedule-based and headway-based holding aim to have a positive impact on the line regularity. Line regularity itself is positively influenced if overtaking is possible. The influence of the overtaking possibility on line regularity is not further discovered during this thesis.

The three elements of the considered journey time, transfer time, in-vehicle time and waiting time, are influenced by the decision variables of holding. Literature shows that waiting time is, in general, positively influenced by improved line regularity as a consequence of holding. In-vehicle time in general increases by the additional slack and additional holding time. Table 2.1 has shown that the trade-off between in-vehicle time and waiting time is often included in other studies. The unknown trade-off which is often neglected and this thesis is interested in, is that of the scheduled slack and the maximum holding time to the transfer time.

The three journey time elements of waiting, in-vehicle and transfer time have an impact on the journey time of the different types of passengers. Transferring passengers experience all the time elements, while direct and through passengers only experience waiting time and in-vehicle time. Direct passengers (that end or start their journey at a transfer point) experience less additional in-vehicle time than through passengers do, as they do not experience holding time from inside the vehicle. This suggests that the total benefits of holding depend on the group size of the passenger types

For all passengers, all journey time elements add up to the total journey time. If multiple journeys are simulated, the total journey time has an average and a reliability buffer time, that is added to the experienced trip time and reflects on the additional time passengers reserve for the usual unreliability. The other factors of the total passenger experience, namely safety, comfort and ease, are not considered in this thesis.

Concluding, in related literature there is an ongoing shift from supply-oriented punctuality measures towards demand-oriented measures that can more adequately capture the experience of passengers. Still, a lot of studies exclude transfers and assume uniform passenger arrival, even at transfer stations. Research that makes use of simulations is able to quantify passenger experience metrics for different timetable setups, also for transferring passengers. The choice behavior is key in assigning passengers to a network with transfer opportunities. Yet it is unclear how a simulated study area can be connected

to a wider transit network. Graph theory is used in theoretical transit problems and can aid in connecting transfer stations within the study area, to outside connections. Holding control is thoroughly researched. The majority of literature focuses on headway-based holding, while in practice, schedule-based holding is applied more often. This suggests a missing bridge between academia and practice and underlines the importance of involving all transit stakeholders into research. This way, the probability of successful implementation of for example, headway based-holding and graph theory, can be increased.

3

Methodology

This chapter elaborates on the methods used to answer the research questions. After the methodological framework is discussed in general terms, this chapter explains all elements of the framework in detail. Additionally, this chapter features the assumptions made that are needed to model the real-life situation. The assumptions are described and are elaborated on why these are made. In Chapter 7, the impact of the assumptions to the validity of the results are further discussed. In this methodological chapter, direct passengers refer to passengers that can use a single line to get to their destination, without transferring and experiencing little to no additional in-vehicle time due to holding. Through passengers refer to passengers that do not transfer, but experience relatively high amount of additional holding time due to holding. Transferring passengers refer to passengers having a journey that contains a transfer.

Figure 3 presents the methodological framework. With the help of the supply-oriented Automatic Vehicle Location (AVL) data, an analysis can be made on the context-specific problems and challenges. For example, which route sections or service hours experience the most variation in driving time or which transfers are likely to be missed (Mozzoni et al. 2022). Since every transit problem is different, no general methodology is applicable here. Historical data is dependent on the timetable or strategy that used to be in place when the data was created. Hence, in Figure 3.1, arrows are drawn from the current strategy to all data sources. The OD-matrix is supported by transfer data to include transfer time from lines that are connected with the cordon. Chapter 4.3 further describes how these data sets are combined. With the help of AVL data, alternative holding strategies are developed that result in different timetable setups. Then, both the current strategy and alternative strategies are simulated with historical driving times. After, for every simulated run, the passengers are assigned to the network according to the shortest path. Dynamic Programming is used to efficiently calculate these. The resulting travel times of the passenger assignment are combined and for each of the strategies leads into the performance of the network. The network performance is then validated by the means of a focus group session with the involved transit stakeholders.

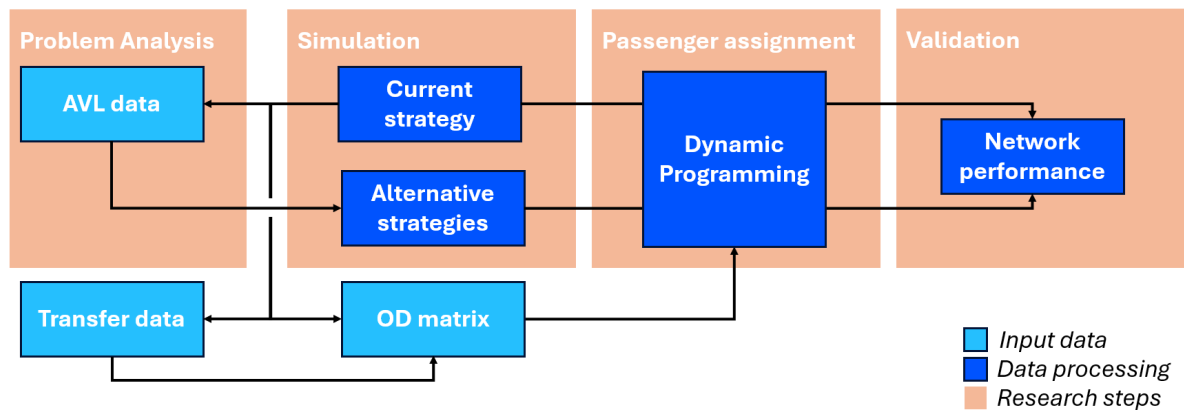


Figure 3.1: Methodological Framework

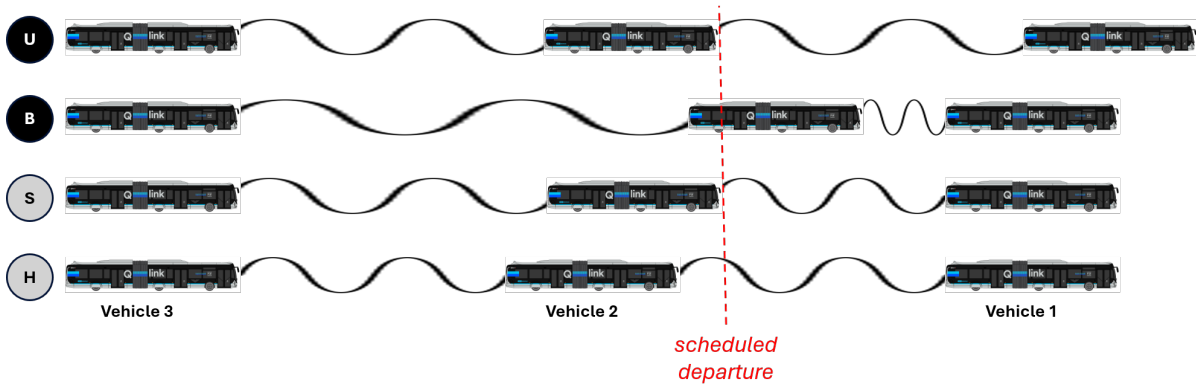


Figure 3.2: Controlling the bunching phenomenon; U: uniform headways stage, B: bunching stage, S: schedule-based holding intervention result, H: headway-based holding intervention result

3.1. Holding strategy selection

This section elaborates on the generation process of alternative holding strategies. Two kinds of holding are considered, which are schedule-based holding and headway-based holding. For schedule-based holding, the slack is allocated differently for each strategy. Since the benefits and magnitude of holding interventions depend on the tightness of the schedule (Van Oort 2011; Cats, Larijani, et al. 2011), multiple percentiles are tested. Furthermore, the schedule-based holding strategies have a different amount of control points, since they also influence the effectiveness of holding (Cats, Rufi, et al. 2014). For headway-based holding, only the mean headway strategy is considered, since Cats, Larijani, et al. (2011) provides evidence that this strategy works best. This headway-based holding strategy strives for the same headway for both the up and downstream vehicles. For the different strategies, the amount of control points is altered, as well as the maximum holding time. This section explains the different holding strategies. More information on how the variables are altered and the formulation used for holding is given in the next section.

3.1.1. Holding types

Figure 3.2 illustrates the bunching phenomenon and the possible holding interventions. Three consequent vehicles are drawn, with waves that represent the vehicle arrival frequency at intermediate stops. In the upper row (U), perfectly uniform headways between vehicles 1, 2 and 3 can be observed. Row B experiences bunching, due to an exogenous delay to leading vehicle 1. As a result, the headway between vehicle 1 and 2 is smaller than between vehicle 2 and 3. This effect is worsened by vehicle 2 experiencing less passenger friction. Fewer passengers are waiting at stops, since passengers have just been collected by vehicle 1. No holding control allows for this bunching phenomenon from row B to happen during the entirety of the line.

In row S and H, a control point was able to control the departure time of vehicle 2. Row S is the resulting situation after which schedule-based holding was applied. In that case, the early-running vehicle 2, can be brought back into its original scheduled headway. As a result, the headway between vehicle 2 and 3 is brought back into its original length. Still, vehicle 1 is delayed, causing heterogeneity in headways. Headway-based holding in row H can more adequately restore headway uniformity. Note that passengers may experience more holding time within the vehicle in row H than in S. Boarding passengers experience more regularity at a stop, which reduces waiting time. To reduce the negative impacts for direct passengers, a maximum allowable holding duration can be set for headway-based holding.

3.1.2. Assumptions

- **Drivers fully comply with the holding strategy:** a maximum holding time can be set to increase the likelihood of drivers respecting the holding rules. Furthermore, it can support passenger acceptance. For this thesis, it is assumed that the maximum holding time never exceeds 4 minutes for headway-based holding strategies. The maximum holding time for schedule-based holding is

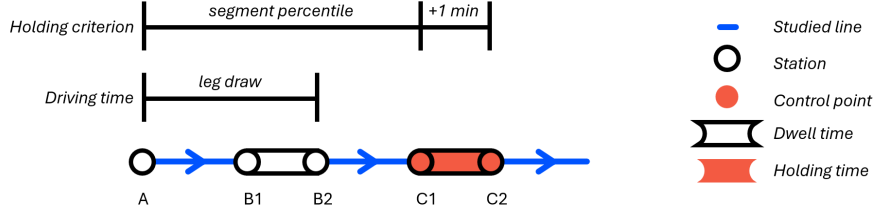


Figure 3.3: Calculation of vehicle runs based on historic AVL data

determined by the chosen percentile.

- **Historical data for leg times is available:** the historical driving duration distribution between each pair of stops is used to draw from. This results in accurate simulated driving and dwell times, while there is a risk that the correlation between crowding and the duration of dwell time is neglected. For schedule-based holding, this distribution is used to determine the scheduled departure time at control points.
- **Real-time data of all vehicles is available for headway-based holding:** for any transit vehicle, the actual headway to the preceding vehicle, as well as the estimated headway to the following vehicle is known. Both these headways are used to calculate the mean headway for determining the optimal holding duration during headway-based holding, once it is actually operationalized in real-life.
- **Speeding up is not possible:** which can be a controlled inter-station mechanism that can reduce headway variation. It is unfavourable that buses drive over the speed limit or skip stops. This leads to when the opposite of row B (Figure 3.2) is happening, namely when the following bus is 4 minutes away and the preceding 6, no action is taken for the middle bus for headway-based holding.

3.2. Simulation of transit vehicle trips

This section describes how a day of service is simulated. It is explained how the driving time, dwelling time and holding time of transit vehicles are determined.

3.2.1. Drawing from the probability function

For determining the driving time between each stop, the simulation uses the leg time distribution as a probability density function. In this implementation, vehicle leg times between stops were simulated by sampling from given leg distributions (similar to Gavrilidou and Cats (2018)). The dataset of leg times is retrieved from historical AVL data. To also include dwell time, the time difference between the departure of the first stop, until the departure of the second stop is taken into account. In Figure 3.3, this corresponds with the temporal distance between point A and B2. A randomization function is used to draw from the probability density function, similar to step 3 of the Monte Carlo simulation (see Chapter 2.2.2). If the historical sample contains a large number of service days, the probability density function is normally distributed around the mean, as formulated by Equation 3.1.

$$R_{ijh}^{\text{leg}} \sim \mathcal{N}(\mu_{ijh}^{\text{leg}}, (\sigma_{ijh}^{\text{leg}})^2) \quad (3.1)$$

where:

R_{ijh}^{leg} : drawn leg time from station i to j at service hour h

μ_{ijh}^{leg} : mean leg time between i and j at service hour h

$(\sigma_{ijh}^{\text{leg}})^2$: variance of leg time between i and j at service hour h

$\mathcal{N}(\mu_{ijh}^{\text{leg}}, (\sigma_{ijh}^{\text{leg}})^2)$: normal distribution of $(\mu_{ijh}^{\text{leg}}, (\sigma_{ijh}^{\text{leg}})^2)$

For determining the schedule-based holding criterion, another draw is done, but now from the segment

time probability distribution function as formulated by Equation 3.2. In Figure 3.3, the temporal distance between A and C1 is used to draw from. Drawing the time from A to C2 would include the current holding strategy. Thus, an additional 1 minute is set as a fixed dwell time at control points, instead of drawing the dwell time from the distribution, too. Figure 3.4 visualizes how the schedule-based holding criterion is calculated. After percentile n is chosen, the percentile value over all vehicle run observations between A and C1, plus the 60 seconds between C1 and C2, is taken as the earliest departure time at the control point. The control point percentile can be specifically set for each control point, which determines the tightness of the schedule. The holding condition for schedule-based holding is formularized by Equation 3.4: if the scheduled departure time (Equation 3.3) lies in the future, the vehicle holds until the scheduled departure time. If the vehicle lies behind schedule, it continues right away with a holding time of 0.

$$R_{coh}^{seg} \sim \mathcal{N}(\mu_{coh}^{seg}, (\sigma_{coh}^{seg})^2) \quad (3.2)$$

$$T_{SB}^{dep} = T_{SB}^{dep-1} + P_p(R_{coh}^{seg}) + 1 \quad (3.3)$$

where:

R_{coh}^{seg} : drawn segment time of control point c to o at service hour h

μ_{coh}^{seg} : mean segment time between c and o at service hour h

$(\sigma_{coh}^{seg})^2$: variance of segment time between c and o at service hour h

$\mathcal{N}(\mu_{coh}^{seg}, (\sigma_{coh}^{seg})^2)$: normal distribution of $(\mu_{coh}^{leg}, (\sigma_{coh}^{leg})^2)$

p : schedule-based holding strategy percentile

$P_p(R_{coh}^{seg})$: p -th percentile value of the segment time R_{coh}^{seg}

T_{SB}^{dep} : scheduled departure time

T_{SB}^{dep-1} : departure time of the same vehicle from the previous checkpoint

$$T_{SB}^{hold} = \max(T_{SB}^{dep} - k_r, 0) \quad (3.4)$$

$$T_{HB}^{hold} = \min\left(\frac{T_{HB}^{dep-1} + T_{HB}^{dep+1}}{2}, T_{HB}^{max}\right) \quad (3.5)$$

where:

T_{SB}^{hold} : schedule-based holding time

T_{SB}^{dep} : scheduled departure time

k_r : simulation clock of simulation run r

T_{HB}^{hold} : headway-based holding time

T_{HB}^{dep-1} : departure time of preceding vehicle

T_{HB}^{dep+1} : departure time of following vehicle

T_{HB}^{max} : maximum headway-based holding time

For headway-based holding, the holding criterion was applied after the simulation of every line. This methodological decision was made, because for a bus to hold, both the departure times of the following and leading vehicle have to be known. After each simulated line, the line is subject to a loop that checks for interval regularity by executing Equation 3.5. This loop starts at the beginning of service and iterates until the end of service. The loop checks for both the preceding and following vehicle. When the following vehicle is 6 minutes behind and the preceding bus is 4 minutes in front, the loop will hold the middle vehicle, so it departs from a control point exactly when both the following and preceding vehicle are 5 minutes away. A typical action the loop performs is visualized in Figure 3.5. If holding is applied, a holding edges are created until the holding criteria is met. Consequently, the remainder of

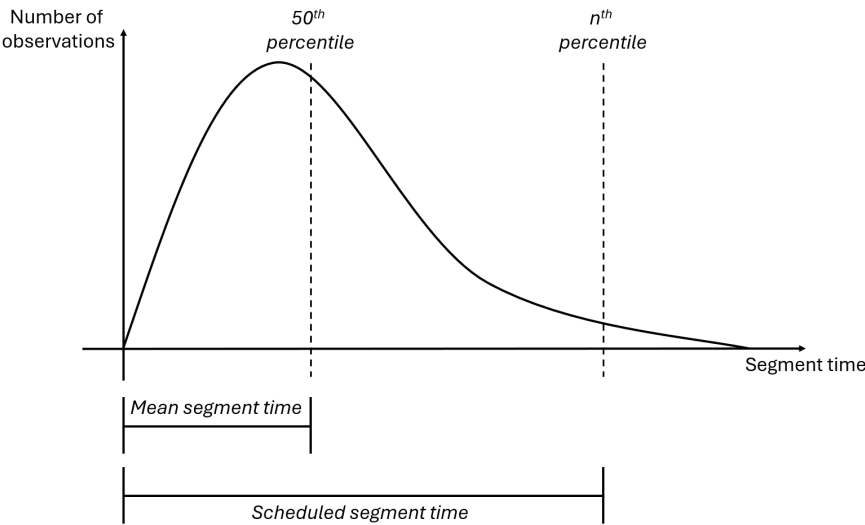


Figure 3.4: Schedule-based holding criterion based on the chosen segment percentile n

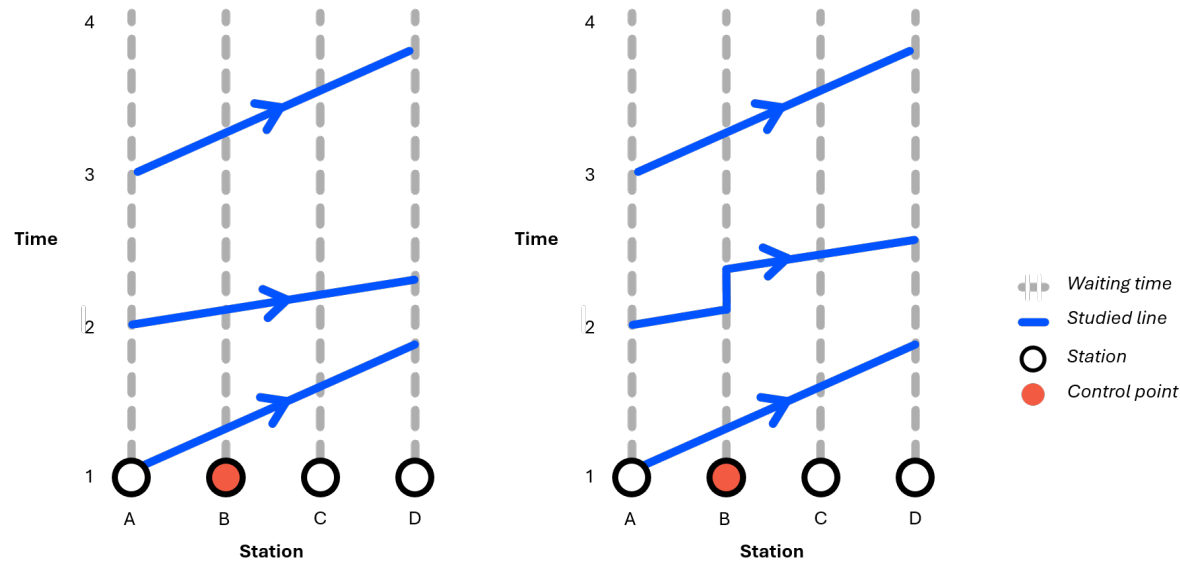


Figure 3.5: The headway-based holding criterion in a line-time graph (left: no holding intervention, right: headway-based holding intervention)

the line is lifted so that more equal intervals are established for the coming stops B, C and D. Maximum holding time is defined into the loop, so that through passengers and drivers will not be held for too long. This maximum holding is defined by an maximum amount of minutes that is either set to 2 or 4, depending on the holding strategy.

3.2.2. Assumptions

- **The estimated arrival time is equal to the simulated arrival time:** this is needed for the headway-based strategies in the simulation. This could result in an overestimate in the performance of headway-based holding strategies. To overcome this, an error for can be assumed for the estimated arrival time. This is not done for this thesis, to minimize modelling complexity.
- **Independent vehicle runtime:** draws from the distribution function are independent. They are not dependent on the previous draw of the same vehicle. The simulated runtime is determined solely by historical runtime, specific to the operating hour between a set of stops. The impact of this assumption is that driver behaviour is not taken into account, neither are specific congested trips.
- **Dwell times are not influenced by the amount of passengers:** irregularity of transit vehicles can lead to more passengers at stops when the headway is longer than usual. For this simulation, vehicles do not experience any friction by the amount of passengers at the stops. Thus, crowding does not lead to additional delays. The dwell time is dependent on historical AVL data only, because passengers are assigned after the simulation. This means that bunching is not internally modelled. It does happen externally, because for intermediate stops, the dwell time is part of the draw (Figure 3.3). The impact of this assumption depends on the amount of passengers at stops where vehicle bunching is likely to happen.
- **Homogeneous dwell times for control points:** the dwell times are assumed equal for control points. For major transfer points, the dwell time is assumed 60 seconds. This assumption is made, because no pure data is available on the needed dwell time. Historical idling time at stops is influenced by the current holding regime. The impact of this assumption is that in peak hours, there could be an underestimation of dwell time, while in off-peak hours, it may leads to an overestimate.
- **Deterministic arrivals and departures for transfer lines:** only the transit lines from within the study area are simulated. The arrival and departure time of lines that interact with the study area are fixed for every simulation run. These vehicle runs can be simulated too for increased realism. The added value of simulating these trips too, depends on the arrival and departure punctuality of these trips. Furthermore, data has to be available for the vehicle punctuality of the connecting lines, or a deviation has to be assumed.
- **Punctual and fixed terminal departures:** as cycles are not modelled, it is not included how delay from a previous vehicle trip spills over to the next vehicle trip. The reason for it is the unavailability of fleet and crew assignment data.

To be able to include cycle constraints, data is needed on the fleet composition. For example, the battery state of the electric vehicles. The fleet allocation is the responsibility of the transit operator, and is therefore not considered in this thesis. During the focus group the province acknowledged that a shift in departure time of any starting terminal would trigger multiple transfer relations across the whole network. The synchronization of transfers can be subject to tailor-made optimization studies. The operator added that if tight schedules are set, opportunities should be researched in providing more direct connections, removing the need of transferring. This would shift the focus from mostly operational to the synchronization of transfers in the tactical solution space (Figure 1.3).

To keep within the scope set in Chapter 1.3, the starting terminal departure times are taken as a given. A deterministic optimization framework that synchronizes transfers can be subject to other research.

3.3. Assigning passengers in a network-time graph

This section elaborates on how the network-time graph is composed. It is explained how the simulated vehicle trips are represented in the graph and how passengers are able to move through the network-time graph. Then, it is expounded how Dynamic Programming can efficiently calculate the experienced travel time of all passengers within a network.

3.3.1. Path finding

Figure 3.6 describes the problem that arises for passengers that are willing to travel from A to D. A direct line is missing between A and D, which makes it necessary to transfer at station B. From Figure 3.6, it is unclear whether the transfer can successfully be made. This is caused by the missing time dimension. Among the longitude and latitude, a third axis is introduced that represents the time. The nodes still represent the stations, but can now be split by time. As a result, station B in Figure 3.6 can appear multiple times, now in the form of B1 and B2. The edges that connect two constitutive nodes are equipped with the corresponding travel time: the runtime for the transit line edges and a single minute for a waiting time edge. This minute is multiplied with the waiting/transfer time parameter.

The introduction of the temporal axis makes it possible to tell if a transfer is feasible. Figure 3.7a represents a missed transfer. The passenger starts at A1 and tries to catch the connection from line 2 to line 1 at station B. The passenger arrives at B2. Since the graph is directed, the passenger is not able to travel from B2 to B1. Which makes sense, since a passenger is not able to travel back in time. The passenger is stuck at B2 and has to wait for another vehicle of line 2 to reach the final destination stop D. Figure 3.7 represents a situation that is more beneficial for the transferring passenger that travels from A to D. Line 2 is held stationary at B1 so it arrives at B2. After A1 has arrived, it continues towards D. Note that due to the holding time, a through passenger on studied line 1 arrives at D3 instead of D2.

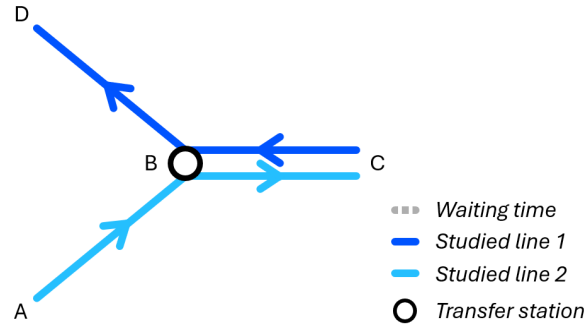


Figure 3.6: Transfer problem example for passengers traveling between A and D

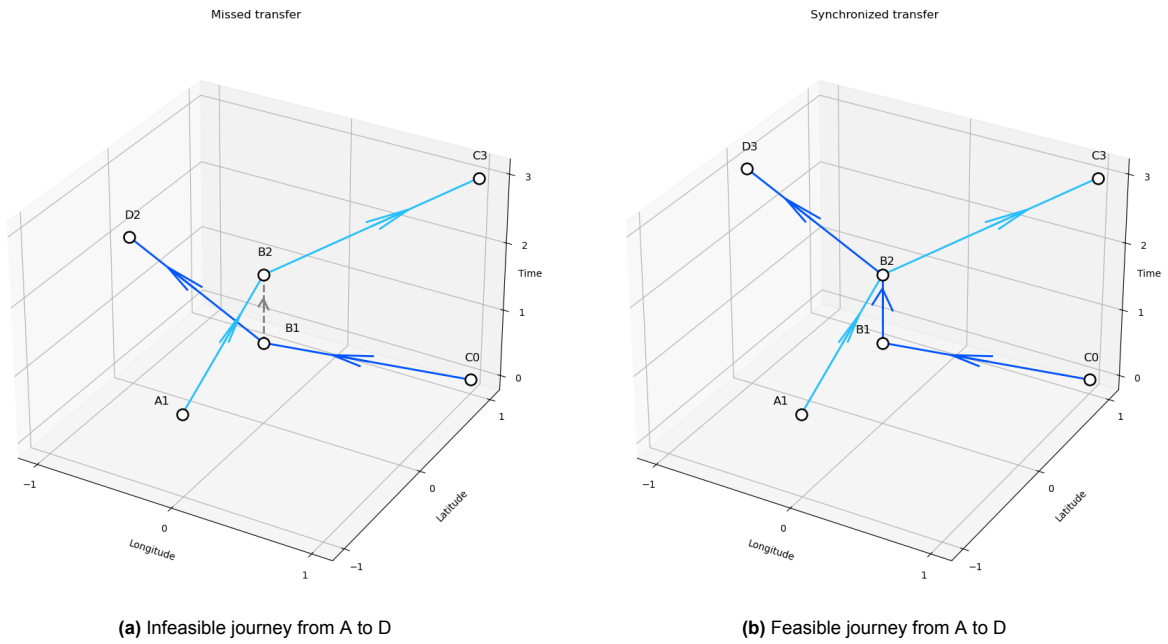


Figure 3.7: Comparison of journey feasibility in the network-time graph

3.3.2. Dynamic Programming

Dynamic programming assigns the passengers to the simulated transit lines, according to the shortest path. The python package NetworkX, in which the network-time graph is constructed, provides a shortest path algorithm. However, when each temporal origin-destination demand cell is to be ran for this algorithm, it would take an enormous amount of computational time. For every origin-destination pair at each minute, it would have to calculate all possible paths. Instead, the methodology used in this thesis utilizes a memory dictionary that is able to efficiently store shortest paths given a destination node. This approach is also known as dynamic programming, in which a problem is subdivided into sub problems, given a terminal state. In this case, the terminal state is the destination node at a specific time. Although it is possible to find all possible paths origins within milliseconds, the algorithm must still be executed for every destination-time node. To calculate all of the shortest paths in a whole day, the algorithm recursively seeks all possible shortest paths, until all origin nodes are found. The mathematical formulation of the Bellman equation used for the dynamic programming algorithm is given by Equation 3.6.

$$f(s_n, a_n) = c(s_n, a_n) + \min_{a_{n+1}} [f(s_{n+1}, a_{n+1})] \quad (3.6)$$

where:

$f(s_n, a_n)$: shortest path of the immediate action and all future actions,

s_n : state at stage n

a_n : candidate passenger action taken at stage n ,

$c(s_n, a_n)$: immediate travel time as a consequence of action a_n at state s_n .

If the terminal state matches the temporal destination node of the demand cell as well as the origin node, travel time tt_{ijn} is assigned to a specific route between origin i and destination j arriving at m , according to Equation 3.7. If the origin node is not yet found, the route is not defined and set to infinity. If the lines are modelled correctly, the $\min$ component in Equation 3.6 is always able to find a route faster than infinity.

The minimization component in 3.6 is also responsible for selecting the fastest route, if multiple are found. Given an example the example network in Figure 3.8a, there exists two alternative paths. From the example network visualization, it is not apparent what the departure time and trip duration are. Figure 3.8b shows the stages of time on the x-axis, and the node name on the y-axis.

In the algorithm, a terminal state is selected. This corresponds with the destination-time in a OD-matrix. In this case, a passenger is travelling from A to D, arriving at time stage 4. The algorithm searches backwards from terminal state D4 to the initial state A. In the first iteration, it does so by looking for candidate passenger actions that deliver the passenger at D4. In this case, the passenger candidate actions is waiting at transfer node D. The resulting action is D3 to D4. The other actions are using the two lines: B3 to D4 with line 1 and C2 to D4 with line 2. Exploring these candidate actions is done for every stage. Constantly, the associated cost is saved in a memory, making the algorithm computationally efficient. At A1 something interesting happens. As A was already visited by line 1 at A2 at a lower cost, the path from A to D4 via C is not stored in the memory. The resulting shortest path for this passenger is 2 time stages.

Figure 3.8b is plot in a two-dimensional plane. The algorithm is build to work on three-dimensional node data. This makes it possible to calculate shortest paths in the network-time graph. Consequently, transfers are included, as illustrated by Figure 3.7.

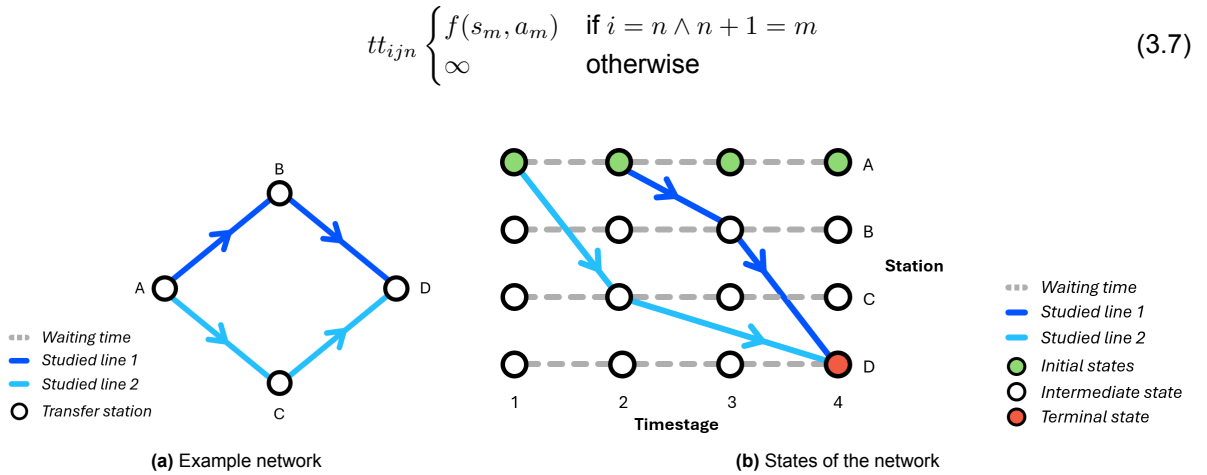


Figure 3.8: Dynamic programming for shortest path calculation

The advantage that a network-time graph has by using a 3-dimensional space is that multiple transfer lines can be used for searching alternative routes. If a transfer has taken place, the edge weight is altered according to 3.8. If the origin stop i differs from destination stop j for a given action $c(s_n)$, it

means that the passenger action was utilizing a line. Thus, a weight of 1.0 is used for determining the travel time cost $c(s_n, a_n)$. This number is multiplied by the edge duration $(m - n)$, representing the in-vehicle time. When the origin stop i is equal to the destination stop j , it means the passenger has not moved in the network and has waited at the same stop. In this case, a weight of 2.0 is assigned to the waiting edge. Finally, when a line change took place, an additional 15 minutes is added to the experienced travel time.

$$c(s_n, a_n) = \begin{cases} 1.0 \cdot (m - n) & \text{if } i \neq j, \\ 2.0 \cdot (m - n) & \text{if } i = j \\ 2.0 \cdot (m - n) + 15 & \text{if } i = j \wedge \text{line change} \end{cases} \quad (3.8)$$

where:

i : departure station i

j : arrival station j

m : departure minute m

n : arrival minute n

3.3.3. Assumptions

- **Historical data for passengers with direct and through journeys is available:** the origin-destination demand is known. This data type demands a system where both tap-in and tap-out is registered to determine the origin and destination of a journey.
- **Historical data for transferring passengers is available:** for passenger journeys that contain a transfer, the interacting line is known. The transfer time can be estimated by the scheduled arrival time of the incoming line, or the scheduled departure time of an outgoing line. If the OD-matrix already contains transferring passengers, additional data on transfers seems redundant. Still, it can be useful in connecting a study area to a cordon (Figure 1.2).
- **Passengers are assigned deterministically to their shortest path:** the path is chosen that yields the minimal experienced travel time. This assumption reduces computational time, because solely one route is stored in the memory.
- **Direct passengers follow an uniform departure distribution:** it is assumed that the lines have a frequency of 6 times per hour each (headways up to 10 minutes), so they can be considered high frequency lines. For passengers performing no transfers, an uniform departure distribution is assumed. In the pseudo code in Appendix A, in section 1.1 it is assumed that direct passengers follow an uniform arrival distribution at stops. The available origin-destination dataset consist of data every 15 minutes. For the demand per minute, the data is divided by 15.
- **Transferring passengers follow deterministic walking times:** when multiple vehicles are to be used in a journey, passengers need to cover a distance to reach the connecting stop at a transfer station. To reduce model complexity, a homogeneous walking time is assumed for all passengers, which is specified for each transfer station.
- **The transfer penalty is 15 minutes:** for each transfer, 15 minutes of additional travel time are added to the total travel time. This is equivalent to 15 in-vehicle minutes or 30 waiting/transfer minutes. This value is based on Garcia-Martinez et al. (2018), who used a combination of revealed and stated preference to determine the experienced transfer penalty.
- **Waiting/transfer time parameter is 2:** the weight that is set to the edges that self-connect the station nodes in the time dimension, is set to double the in-vehicle time. Hence, a minute spend waiting at a station is experienced twice as long as spending it in-vehicle. This parameter is needed alongside the transfer penalty, so that a longer transfer is a larger burden for the passenger.
- **Capacity constraints are not binding:** denied boarding due to crowding is not considered. Passengers are always able to board the vehicle and do not experience travel time more negative when crowding occurs. The decision is made, because it is not compatible with the network-time

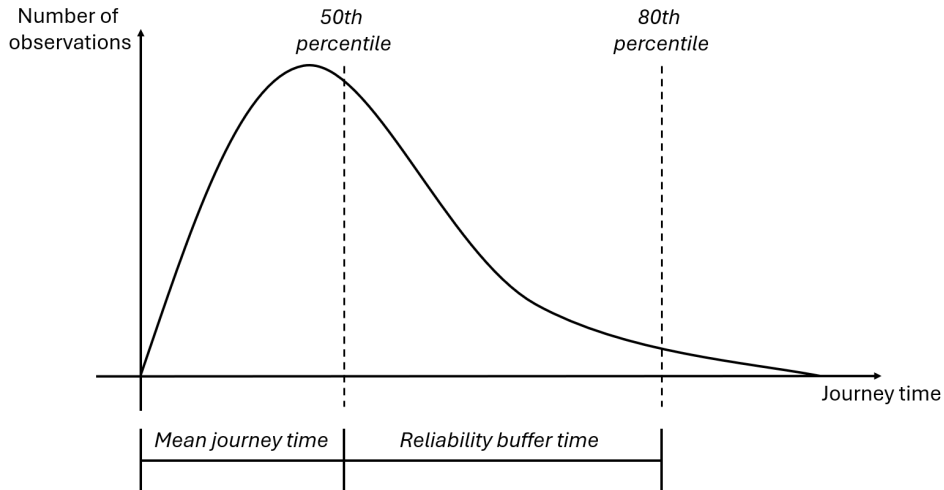


Figure 3.9: Integrating speed and reliability for a single OD-pair (based on Uniman et al. (2010) and Van Oort (2011))

graph. Since the passengers are assigned one by one, the future state of the network has to be known by passengers for the current passenger assignment model. A failure-to-board status on an edge will disturb this known future state.

- **The schedule can be altered every hour block:** the schedule can be adapted ever hour, which is done to analyse the effect of different holding strategies during the day.

The focus group revealed practical limitations in determining the number of schedules over the course of a day. In constructing the definite real-life schedule, five or less time blocks are chosen for a day at maximum. The operator pointed out that while making too much temporal blocks with a dedicated scheduled trip time may increase the punctuality of buses and passengers, it is important to take into account that a schedule should be somewhat logical to passengers. The transit authority added that the transition period between two scheduled vehicle trip time blocks should prevent immediate bunching. For example, when coming out of rush hour enables for a tighter schedule that yields a better passenger experience on paper, the preceding vehicle should be released from the line starting terminal at a longer headway than the previous departure. Otherwise, the preceding vehicle will meet its predecessor at a timing point, leading to irregular service.

3.4. Network indicator

After each shortest path is calculated by Equation 3.6 for each demand row in the OD-matrix, it is combined into a network-wide indicator. This value is used as a quality indicator to compare different holding strategies in terms of reliability and speed that reflects on the experience of passengers. The following section elaborates on this indicator in more detail.

For integrating the level of reliability and speed into one network-wide metric, a new indicator is introduced. Rider-based Rapidity and Reliability, in short RRR, reflects on the experienced journey time of all passengers in a network. Figure 3.9 illustrates how the RRR is determined for every journey. For each simulation run r , a journey time is observed for every per-minute origin-destination pair. After an R amount of runs, there exists variability within the sample. The mean or median can be chosen to elaborate on the average time spend in the network as a measure of centrality. However, passengers value their reliability too (Peek and Van Hagen 2002), which depends on the type of trip, the time of day and varies from person to person (Van Oort 2016).

Equation 3.9 shows the calculation steps for the new indicator. Across all simulation runs, the 80-th percentile value of journey time is picked that reflects on both the speed and reliability. Then, it is multiplied with the amount of travellers that historically travelled from origin i to destination j , arriving at

minute n . To calculate the total experienced daily travel time, all of the origin-destination-time demand data is summed. This can also be done for specific intervals, such as for every hour. This can lead to a more specific analysis of the indicator for different times of day. Although the unit of the RRR indicator can be understood as the total experienced time spend in the network, the value of RRR is no stand-alone indicator. It is all about the relative difference between timetable setups. Other holding strategies can be plotted as absolute percentages around a base strategy. An example of a base line could be the current strategy, that can act as a reference. Then, the resulting RRR will be better readable and for every hour, the performance of the different strategies can be easily derived.

$$RRR = \sum_{i \in I} \sum_{j \in J} \sum_{n \in M} d_{ijn} \cdot P_{80}(tt_{ijn}) \quad (3.9)$$

where:

RRR : total experienced travel time spend in the network by all passengers [minutes],

d_{ijn} : number of passengers travelling from origin i to destination j arriving at minute n ,

I : set of origins i ,

J : set of destinations j ,

M : set of service minutes n ,

tt_{ijn} : travel time from i to j arriving at minute n ,

$P_{80}(tt_{ijn})$: 80-th percentile of tt_{ijn} over all simulation runs r

3.4.1. Assumptions

- **Passengers are willing to risk a 20 per cent chance of arriving too late at the destination (P_{80}):** This means that the experienced travel time is equal to the 80 per cent fastest trips for a time-specific origin-destination pair. This translates to passengers choosing for a journey that is able to arrive at a destination on time 4 out of 5 times. In light of transfers, when a passenger misses a connection more than 2 in 5 times, it is assumed that passengers start their journey earlier in order to arrive on time at the destination stop. This decision was deliberately made. Increasing this percentile is possible, but requires more computational time for additional simulation runs for a significant RRR indicator.
- **An utilitarianism standpoint is chosen as equity principle:** in this fairness standpoint, the overall welfare is maximized by ensuring transport policies benefit the most people. It prioritizes efficiency over individual needs, resulting in cost-effective policies and maximized benefits. For this thesis, this translates to all passenger travel minutes being considered equally important. Chapter 7.6.2 reflects on the impacts and possible ways to overcome exclusion due to utilitarianism.

3.4.2. Testing for significance

In order to return significant results, multiple runs are needed for a simulation process. Although every hour consists of the amount of buses equal to the frequency, it does not automatically mean that a simulated day of service is representative. Testing for significance makes it possible to test the representativeness of the sample. If the simulation is ran once, there is a chance that extreme leg times are drawn from the distribution. Although this can occur in a real day of service, for instance when there are problems for a certain vehicle run, it does not tell much about the average performance of a certain timetable.

For determining the significance of the RRR indicator, bootstrapping is to be used for the set of produced simulation runs. For bootstrapping, a large number of samples is taken from the total set of simulation runs ($m=50$). This drawing happens randomly and results in a subset of simulation runs of size $m=43$. For every subset, the percentile value is calculated. Then, the mean for the percentile value can be derived, as well as the standard deviation. This mean and standard deviation can determine if the percentile indicator is consistent over all simulation runs.

To compare sample means and their associated standard deviation with each other, a t-test (Equation

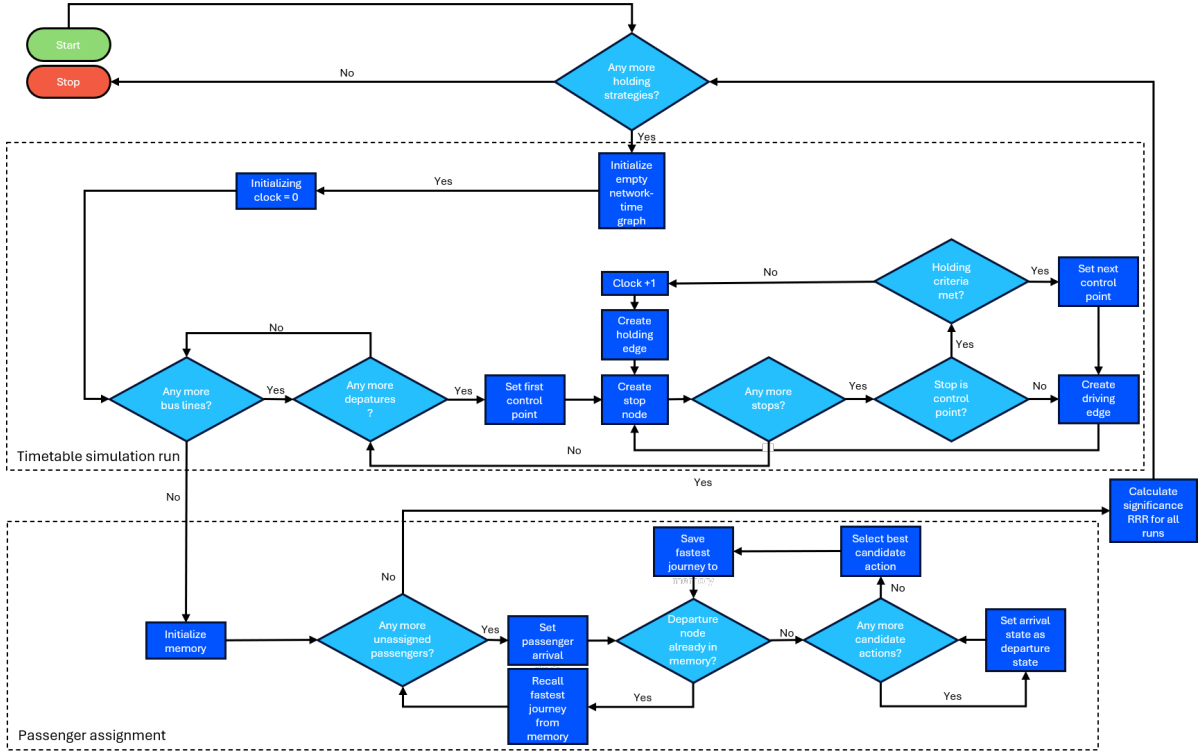


Figure 3.10: Flowchart of the complete algorithm (inspired by Cats, Burghout, et al. (2010))

3.10) is used to calculate if the p-value according to Equation 3.11. If the p-value stays below the threshold of 0.05, there is a significant difference between the indicator of two holding strategies. Each hour, the t-statistic can be used for comparing a sample mean to another sample mean. The significance of each strategy to the base (SB-multi-85) is calculated as:

$$t = \frac{\bar{X}_i - \bar{X}_{\text{base}}}{s_i / \sqrt{m}} \quad (3.10)$$

where:

$\bar{X}(m)_i$: mean of strategy i ,

$\bar{X}_{\text{base}}$: mean of the base strategy,

s_i : standard deviation of strategy i ,

m : sample size.

After, the p-value for a two-tailed test can be calculated as follows:

$$p = 2 \cdot \left(1 - t_{m-1, \frac{(1-\alpha)}{2}}(|t|) \right) \quad (3.11)$$

3.4.3. Flow chart

To summarize the above methodology, a flow chart is constructed that connects all elements discussed above, which can be found in Figure 3.10. In this subsection, the arrows of the flow chart are carefully followed, beginning at start, and ending at stop. First, it is explained how a single simulation run works, after which the passenger assignment procedure is explained.

The first question that is raised by the flow chart, is if there are any remaining holding strategies to be tested. Since none are tested from the start, this question is to be answered with "Yes". We move one

block lower into the first simulation run for the holding strategy that is to be tested. Note that instead of a simulation run, also a full historical service day can be used, to see the performance of a network in the past. In the current scope of this thesis, we only look at hypothetical future holding strategies. Every simulation run, an empty network-time graph is initiated that only contains a standard set of edges that connect the nodes in time. This ensures the waiting and transfer time possibilities, which corresponds with the edge from node B1 to node B2 in Figure 3.7a. In the empty network-time graph, all the bus lines are saved one by one, starting from the beginning of service. Hence, the clock is set to zero. Then, the algorithm checks for remaining scheduled bus line departures. While there are still departures left, a new vehicle run is initiated by setting the first control point with its associated minimum departure time. Every stop, it is checked whether the stop is a control point. If not, after the vehicle has dwelled, it continues right away by creating a driving edge that connects to the next stop node. If it is a control point, the holding criteria is checked and holding occurs by creating a holding edge until the holding criteria is met. If the criteria is met, the new control point is set and a driving edge is created.

Once all transit lines are plotted in the network-time graph in a simulation run, the passengers are assigned to the drawn network. First, a memory is initiated that can be used for storing shortest paths. For every origin-destination-time demand row, the algorithm first checks if the route between origin and destination is already known for the given arrival time. If it is, the fastest route can be simply recalled from the memory and no calculation is needed. If the algorithm is faced with a new departure node, it can use the shortest paths in the memory to efficiently calculate entire routes that consist of an unknown candidate action and a already known shortest path from a previous calculation. If all fastest journeys are found, which means that there are no remaining unassigned passengers, the significance of the RRR is determined for every hour across all conducted runs. Once all strategies are simulated, the algorithm stops and the results can be analysed.

A full pseudo code of the algorithm is included in Appendix A. The coding approach is shown and gives more detail on how the separate methodological elements are connected. The code in section 1.1 divides the 15-minute demand blocks by 15, for the assumption of equal arrival distribution for direct and through passengers. The code in section 1.2 simulates one day of service for a chosen holding strategy. It is assumed that every bus is able to depart on time from a terminal at a perfect interval. The run times are determined on a draw of the individual stop to stop times, including boarding passengers. The latter is not included for the control points, since the idling time is not equal to the dwell time. For this, it is assumed that at control points, one minute is needed for boarding passengers. Finally, section 1.3 describes how the experienced travel time is calculated by means of the RRR indicator.

3.5. Stakeholder perspectives

During the early stages of the thesis, individual interviews were held with the authority and operator. This contributed in a sharper problem statement, clearer solution spaces and matched expectations for the project. Near the end of the modelling process, all the stakeholders were invited in a focus group session once more for a verification of the methodology.

Furthermore, the focus group was used for validating the preliminary results of the simulation. The main goal of a focus group is to assess if the research fulfils the needs of the stakeholders. For this research, the following actors were invited: the transit authority, operator, passenger interest groups, municipality and province. The focus group is installed instead of separate expert interviews, so that the independent actors could form a meaningful dialogue that is better able to highlight the multi-actor complexity on transit problems.

Organizationally, the focus group consists of multiple rounds. First, the methodology and preliminary research results were presented. Then, in the first introductory round, the roles of all stakeholders were elaborated on one by one, in light of passenger-oriented indicators. In the second round, the methodology and its modelling assumptions were discussed. This can be seen as a form of verification. If only the method was verified, it does not tell anything about the usefulness of results. All actors will value results differently. That is why in the third round, the results were validated. The main goal of this round was to see how useful the results are and if the approach meets the needs of the stakeholders. A final round contained of statements aiming for a common action plan that could benefit the experience of passengers for the future of the studied lines, while

respecting the point of view of all stakeholders.

After the focus group interview has been conducted, the model was improved according to feedback that arose during the focus group. Additionally, the complexity of interests will be brought together into a brief stakeholder analysis. With consent of all stakeholders, the focus group session was recorded. The text derived from the focus group was sent to the members in order to check for inaccuracies, before the thesis is published.

4

Case Study

The methodology presented above can be used for practical transit scheduling problems, such as holding control. To illustrate the application of the model, a real world case study is used. First, the network characteristics of the case study are given. Then, the major challenges in terms of scheduling are given. These are the problems that are aimed to be solved by the proposed methodology. The availability of data is described, after a brief analysis is conducted on vehicle and passenger patterns.

4.1. Network characteristics

For a contextual setting, this thesis uses a case in the city of Groningen, namely Q-link bus lines 3 and 4. With a population of 240,000, the municipality of Groningen forms the major urban cluster in the north of The Netherlands. The responsible public transportation agency *OV-bureau Groningen Drenthe* strives for a metro-style branded network for its high frequency buses. The so-called Q-link network has a metro-style network map (Figure 4.1) and vehicle exteriors matching the line color (Figure 4.2). The vehicles used for Q-link lines have a higher occupancy compared to a standard 12 meter sized bus, with a total vehicle capacity of 110 passengers. Respectively 45 and 65 for seating and standing. This high capacity is achieved by the longer vehicles, which are articulated and have a minimum length of 18 meters. During rush hour, Q-link lines have a frequency between 4 up to 12 times per hour each.

Apparent from Figure 4.1, Q-link lines have common corridors, contributing to an even higher joint frequency if headways are properly coordinated. The network mainly consists of diametric lines, starting at an outskirt or nearby village and connecting to the central station *Groningen Hoofdstation*, after they continue in the same wind direction they came from to reach the other side of Groningen. Along the way, Q-link lines connect to various transfer points, called *hubs*, where transfers to other (Q-link) bus lines, regional train lines and shared mobility can be made. Additionally, these *hubs* often include Park+Ride facilities, offering free parking for cars (and bicycles) and have a special P+R return ticket with a reduced fare. The specific Q-link lines in this thesis, line 3 and 4, serve the diametric southwest-northeast branches (Figure 4.1).

The line infrastructure is characterized by a high degree of separation. The infrastructure used by Q-link 3 and 4 consists of respectively 60% and 44% dedicated bus ways on delay-sensitive sections (*OV-bureau Groningen Drenthe*, 2023). This minimizes the interaction with other road users and decreases the variability in trip times (Van der Bijl and Van Oort 2024). Figure 4.2 illustrates the essence of the dedicated bus lane. The Q-link bus uses the dedicated lane near *Hoofdstation* to skip the queue of car traffic.

4.2. Challenges

The bottom line problem of the case study is that the speed and reliability as perceived by the passenger are unclear. The current schedule is designed loose, making the network reliable in terms of vehicle punctuality. Transfers seem to be robust when all vehicles arrive and depart punctual. However, since the studies lines have a high and shared frequency, it is questionable whether the synchronization of

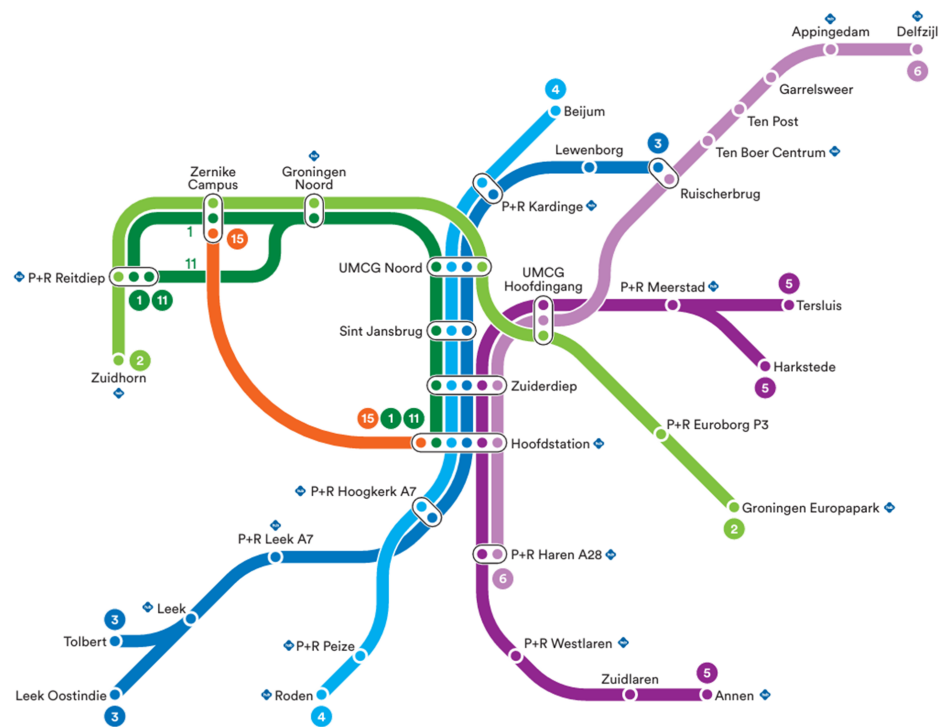


Figure 4.1: Q-link network of Groningen (source: Van den Brink, 2022)



Figure 4.2: Vehicle exterior matching the line color. In this case: a blue bus serving Q-link 4 (source: busposities.nl).

source	stop		line (+ direction)		time			#pax	
	current	future	current	future	window [mins]	departure	arrival	through	trans
AVL	✓	✓	✓		1/60	✓	✓		
OD	✓	✓	✓		15		✓	✓	
TLS	✓		✓	✓	60	✓			✓

Table 4.1: Case-specific data availability

transfers into the high frequency lines actually improves the passenger experience. Furthermore, the speed for through passengers is low, due to the excessive holding time. A quantitative measure is needed that represents the full journeys of all passengers, including transfers.

The main challenges regarding service design for Q-link are reliability and overcrowding issues. Scheduled trip time is designed loose and timing points are in place, so delays can be coped with. However, a large share of buses arrives early, which increases the in-vehicle waiting time for through passengers. This asks for quantitative research for tactical strategies regarding the percentiles used to determine the schedule-based holding duration. Furthermore, the benefits of headway-based holding are not known. Q-link lines 3 and 4 already contribute to a joint headway of 5 minutes at the busiest times of day. Between *Hoogkerk* and *Zuiderdiep* (city center), more than 30 buses per hour per direction serve the corridor of line 3 and 4. This results in headways under two minutes for the shared corridor. Even if the reliability of Q-link 3 and 4 are decreased, the waiting time remains low. Note that Q-link 3 and 4 serve some unique pairs of stops, since the other lines on the corridor make use of stop-skipping and terminate where Q-link 3 and 4 continue.

The Q-link lines 3 and 4 have a per-direction frequency of 6 buses per hour each during the busiest times of day, aggregated 12 for the shared line section between *P+R Hoogkerk* and *P+R Kardingse*. At the moment, the trips are evenly distributed over the hour. To ensure this "perfect" distribution, the buses are released from timing points at equal intervals. This schedule-based holding happens at *P+R Hoogkerk*, *Hoofdstation* and *P+R Kardingse* on the shared corridor. Additionally, several timing points are in place on the branches, before the shared corridor is reached at *Leek Centrum*, *P+R Leek A7* and *Roden Centrum*.

The transit authority implemented a lower control point percentile for a while. This has been reverted, since the operation was influenced too much. For example, it led to unreliable driver changes at control points. The trade-off was unclear too, because the resulting absolute travel time gains (or losses) were unknown.

4.3. Data availability

This section describes the case-specific parameters that couple the general methodology to the case study. First, the availability and characteristics of the data sources are discussed. The attributes of each dataset are first described separately, after which it is explained how the differences in attribute levels is coped with. Table 4.1 gives insight into the data availability of the case-study. Three different data sources are used, AVL, OD and TLS. They vary in terms of stop, line, time and passenger demand attributes, which are described in more detail below.

AVL, Automated Vehicle Location data, provides the temporal information on transit vehicles. AVL data is provided by the operator and is delivered to the authority that conducts quality checks on the historical service. The AVL data consists of the current stop only. However, the future stop that is visited next, can be easily derived due to the known line direction. Only the current line is known, because a vehicle can be assigned to one line simultaneously. The time windows at which data is available every second (1/60 minutes) This high level of detail is useful when determining the runtime between consecutive stops. At each stop, both the departure and arrival time are given. For this thesis, this proved useful in determining the holding time distribution, and the general dwell time at stops. AVL can be seen as a pure supply-oriented data source, because it only contains data on transit vehicles without describing the amount of passengers (#pax in Table 4.1).

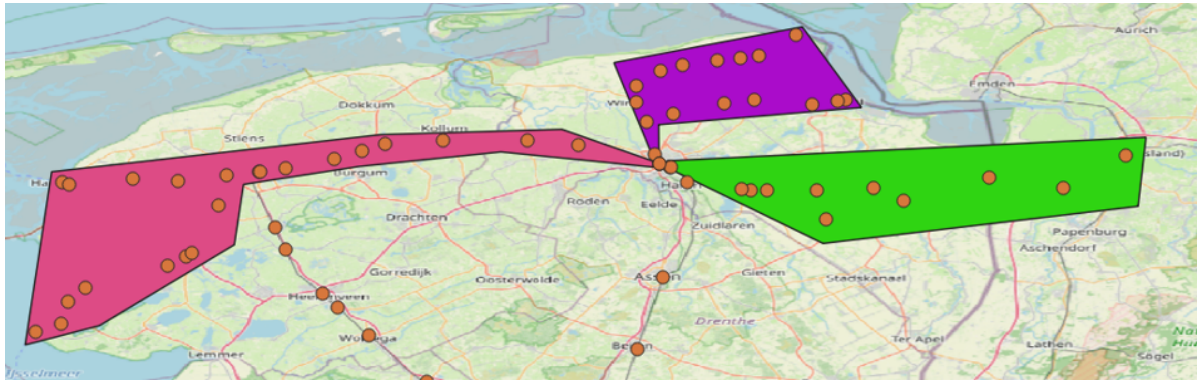


Figure 4.3: Train station directional categorization (source: Translink)

On the contrary, the OD (Origin-Destination) dataset provides demand-oriented information, as it contains the temporal origin and destination locations passengers have used in their trip, based on their tap-in and tap-out transactions. This dataset too originates from the transit operator and is handed to the transit authority, that conducts service quality assessment. The OD data consists of both the current stop attribute, the origin, and the future stop attribute, which represents the destination. The OD dataset provides the in-vehicle time as highlighted in Figure 1.1 that is spend in a specific line. For this case-study, the OD did not consist of the future line. Hence, for the transfer time to be included. Thus, passengers are binned to the time windows according to their arrival time at the final stop on the case study lines. Only through passengers within the case study lines are included in the OD dataset.

For the transfer time to be included in the journey, the TLS dataset was used. TLS stands for *Trans Link Systems*, which is the only authorized party in the Netherlands that provides numbers of transfers across different operators. For this thesis, a data request was placed at TLS, which provided an anonymized dataset on the number of transfers between case study and cordon. The included attributes are the current stop, in this case the station at which a transfer took place. Both the incoming line (current) and outgoing line (future) are specified. The last included variable is the amount of transferring passengers, which is specified per month and per hour. The line direction of the cordon stops are determined according to Figure 4.3 and 4.4, which respectively show the categorization of the line directions of connecting train and bus services. The planes represent the directions, while the node colours represent the line number. The train line categories were more homogeneous than the bus categories, because of the fare system. Trains in the Netherlands use a tap-in and tap-out system at the stations, while the bus services have an onboard tapping system. Consequently, it is hard to say which specific train was used, while the used bus vehicle directly shows up in the data. As a result, for trains, only the wind direction is known. For buses, both the wind direction and line are known. The time window at which transferring passengers were binned was set to 60 minutes. This was done so according to the departure of the connecting (second) vehicle.

Across datasets, the attributes differ in their existence or units. This mismatch demands data transformation for combining the datasets into a comprehensive model, especially for the stop, time and passenger attributes.

Starting with the stop attributes, from Table 4.1 it becomes visible that the TLS dataset misses the future stop. Only the current stop, which is the stop at which a transfer happens, is included. A transfer dataset with full origin-destination pairs does exist, but is not available due to privacy constraints. The national rail operator Dutch Railways (*Nederlandse Spoorwegen*) requires at least 50 per data bin, before the data can be exported. Hence, the future stop is not included in the dataset. Instead, the line plus the direction is included.

As explained above, all datasets differ in their time window. For this reason, the datasets should be converted into the same temporal unit. Because holding has small-scale temporal impacts (i.e. a schedule deviations of single minutes), the chosen time unit is minutes. The AVL data becomes less precise, since it is specified in seconds. Rounding up is used to arrive at the next minute. For the OD dataset, the bins of 15 minutes were divided by 15 to also get the minutes. For similar privacy reasons,

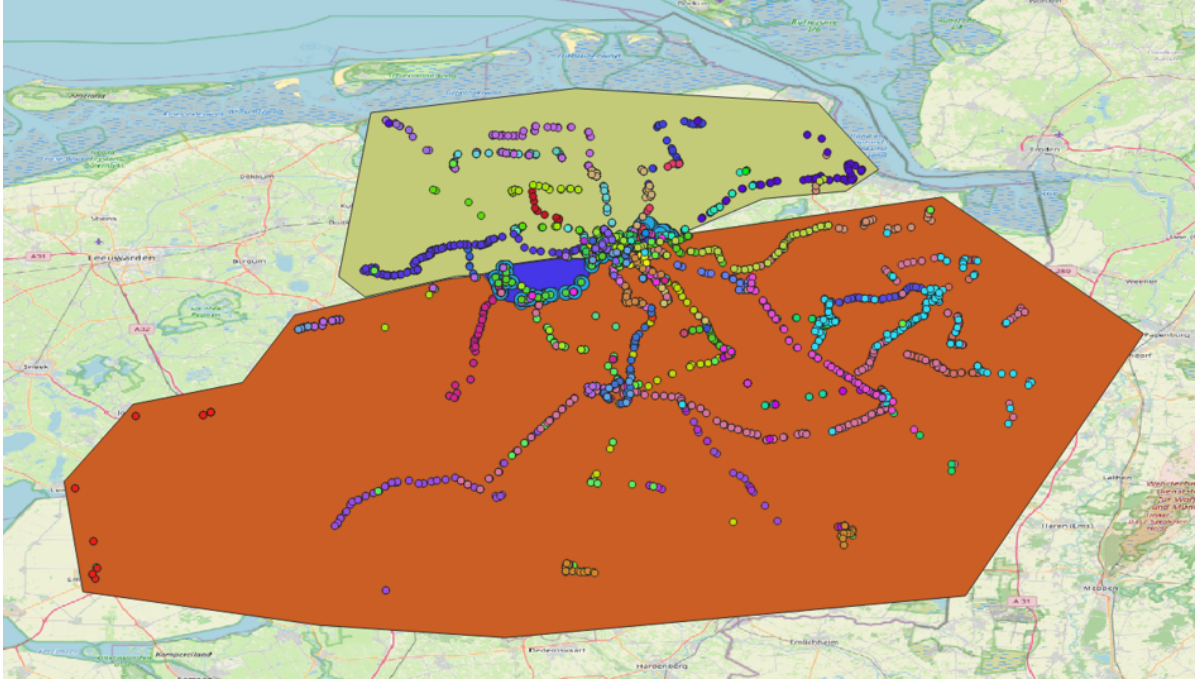


Figure 4.4: Bus stop directional categorization (source: Translink)

the time window for the TLS dataset was less detailed and set to 60 minutes, instead of a smaller window, which could have made it possible to assign passengers to specific vehicles. For example, if a train departs twice per hour, a bin of 30 minutes allows for better estimates on which vehicle was used, compared to the current 60 minute bin. To cope with this limitation, this thesis assumes that the passengers follow an uniform arrival distribution at their final stop. This made it possible to use a common time window of one minute. The OD dataset was divided by 15 to match this unit, too.

Finally, to connect the OD and TLS datasets, an additional assumption was needed to determine the origin for outgoing passengers, and the destination for incoming passengers. It is assumed that transferring passengers are distributed over the stops similarly as through passengers.

4.3.1. Case-specific sets and parameters

The following section elaborates on the sets, parameters and variables that lets the case study use the methodology from Chapter 3.

The set for all stations N defines the case study area, together with the fictional cordon stops. the case study lines in set S consist of 5 terminals, 6 control points and 54 intermediate stops, contributing to a total of 65 case study stops. Additionally, the cordon is represented by imaginary nodes. Each line category from Figure 4.3 and 4.4 created another stop that transferring lines of set T fictionally connect to. Altogether, the total set of real and fictional consists of around 100 stations. The network-time graph inserts these stations and equips them with a time component from set M .

The set of studied lines S consist of the bus lines Q-link 3 and Q-link 4. Apparent from 4.1, there are a total of 6 different possible directions. Q-link 3 contains four directions, since it is bidirectional and has a branch in *Leek* with two different terminals in *Leek Oostindie* and *Tolbert*. Q-link 4 is not branched and thus has two directions only. The set of transferring lines T are categorized according to Figure 4.3 and 4.4. It results in 55 transferring lines.

The set of service hours H starts at 6 AM and ends at 2 PM, which corresponds with the first and last buses. Between 5 and 6 AM, there are (partial) trips too which are drawn in the network-time graph. However, the passenger demand in this early hour is excluded, since the passenger assignment algorithm is not able to find some paths to particular journey origins. Consequently, the total set of hours H consists of 20 elements. As the network-time graph has a time component, every minute of service

M creates a unique node for each stop of set N . A full day of service results in a total of 360.000 nodes per network-time graph.

For the set of service days D , the period between September 1st 2025, until the May 29th of 2026 was used for all the data types. The former was chosen, because that day marked the end of the summer break, as well as a new train schedule that could be implemented after the *Hoofdstation* reconstruction works. The latter is chosen, because it was the final day in the TLS dataset. In the chosen period, 15 vacations days were excluded, as well as 5 national holidays and 4 days of extreme winter whether. All these exceptions saw alternative schedules and lower demand, which poses difficulties in data representativeness. Only weekdays were considered. Hence, the results are based on an average weekday.

4.4. Vehicle patterns

A brief supply-oriented analysis is conducted that gives an insight on the vehicle-based speed and reliability within the case study. This is done to build a deeper understanding of the case and can provide additional evidence in describing the results in the next chapter. The data that is needed for the vehicle patterns is the Automatic Vehicle Location (AVL) dataset. The historical AVL data is used for a vehicle trip time analysis, specifically on the hourly variation. An important side note is that the holding time is excluded from the total trip time. This way, the current holding strategy does not influence the analysis. An additional minute is added up for each of the control points to cover for the dwell times.

The graphs used for the analysis explore the trip times for every hour. It can be read as consecutive box plots, with quartiles set at a 35th percentile distance from the median. The 15th, 50th and 85th percentiles of trip times are given. The 50th percentile can be seen for a measure of centrality and therefore reflects the speed. The distance between the 15th and the 85th percentile is separately calculated and plot as a dotted line. This line indicates the variation in trip time for the associated hour. This could be an explanatory factor for line (ir)regularity.

The similarities between the lines is that typically, a trip takes somewhere between 45 and 60 minutes over the course of a full day. For both Q-link lines in both directions, the peak hour can be clearly found back in the delta percentile value. As this describes the reliability, it can be observed that for any of the trip distributions, the lines are most unreliable between 7 and 9 am. Additionally, a more gentle and wider peak is visible at the afternoon rush, ranging from 14 pm to 18 pm.

The differences between the graphs is in the magnitude of reliability levels. Most obvious is the morning in peak of Q-link 3 in eastbound direction (Figure 4.5). The delta percentile reaches a level of 12 minutes. A possible explanation is the geographic position of the city of Groningen. It is one of the only urban clusters in the wider region that attracts a lot of home-work travelers for its daily urban system. As a result, there typically is traffic congestion in the direction of Groningen on the A7 highway. The negative effects of traffic congestion is partly mitigated, due to the exemption buses have to use the emergency lane with a speed between 30 and 50 km/h. The normal cruising speed of the dedicated Q-link fleet is 80 km/h. Q-link 4 experiences similar delays in the morning rush hour (Figure 4.7).



Figure 4.5: Q-link 3 eastbound trip time distribution per hour

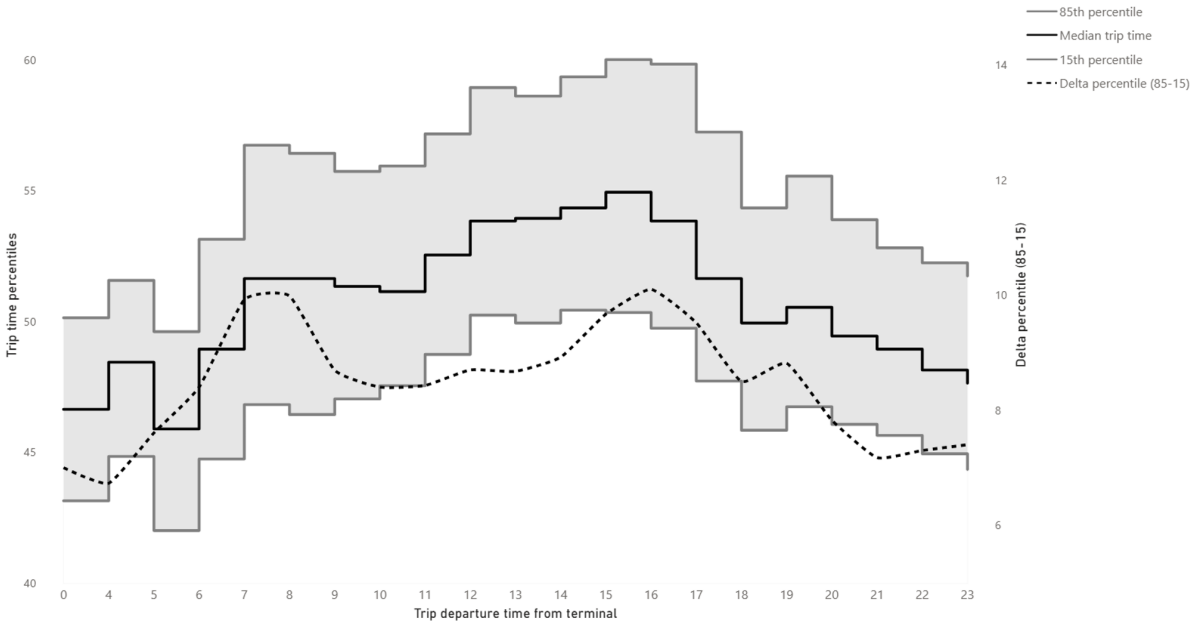


Figure 4.6: Q-link 3 westbound trip time distribution per hour

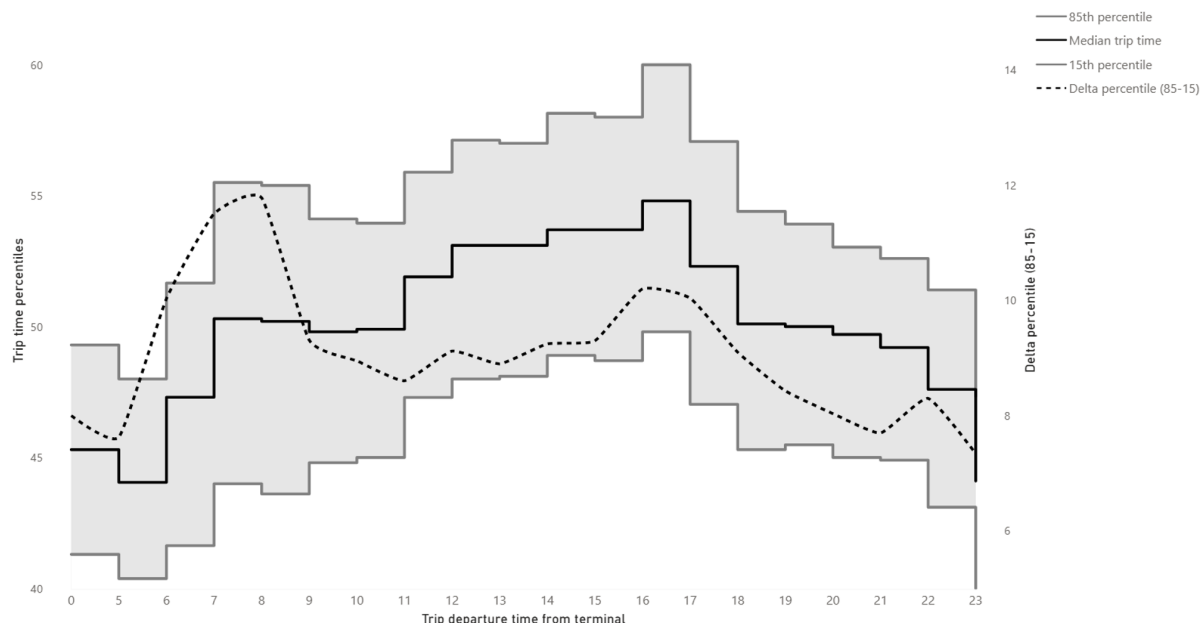


Figure 4.7: Q-link 4 eastbound trip time distribution per hour

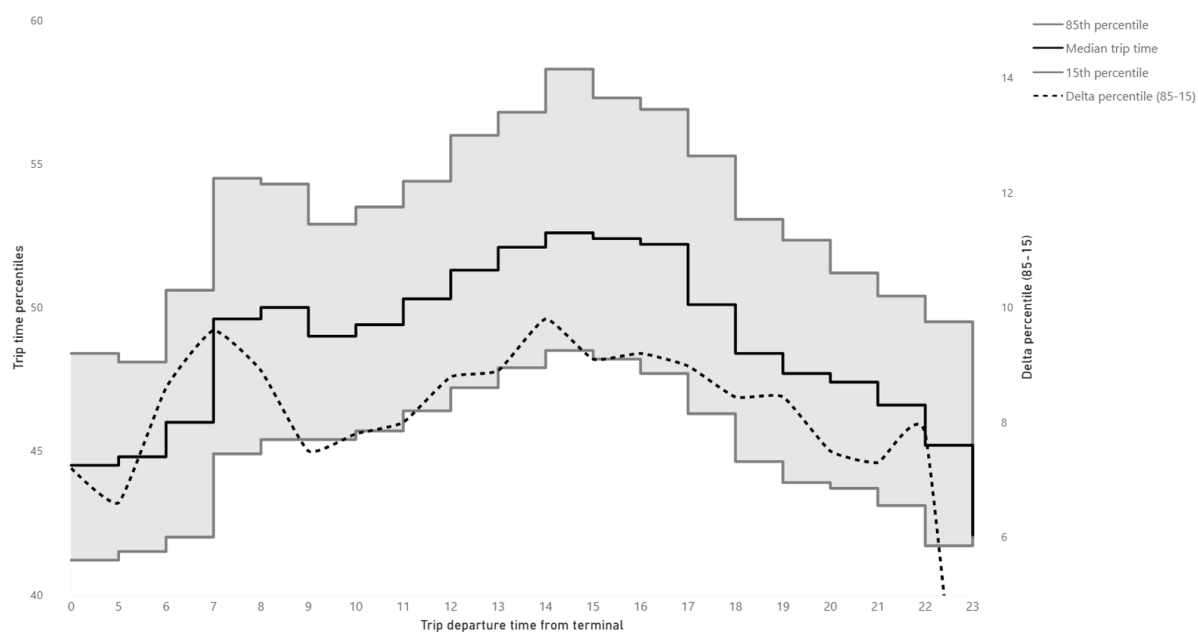


Figure 4.8: Q-link 4 westbound trip time distribution per hour

The holding time itself is separately analysed for the major transfer station *Hoofdstation*. Included are all trips in 2025. Figure 4.9 shows the distribution of holding time there, as it is the station where the most holding time occurs. The holding time that occurs most is five minutes. The distribution is right-skewed, with more observations to the left of the mean. Holding longer than 8 minutes barely happens, while holding 1 or 2 minutes occurs more often. In these cases, after the required dwell time, the bus is able to continue right away. This distribution gives an indication on the amount of slack that is (under)utilized at *Hoofdstation*. It suggests that the schedule is set too loose.

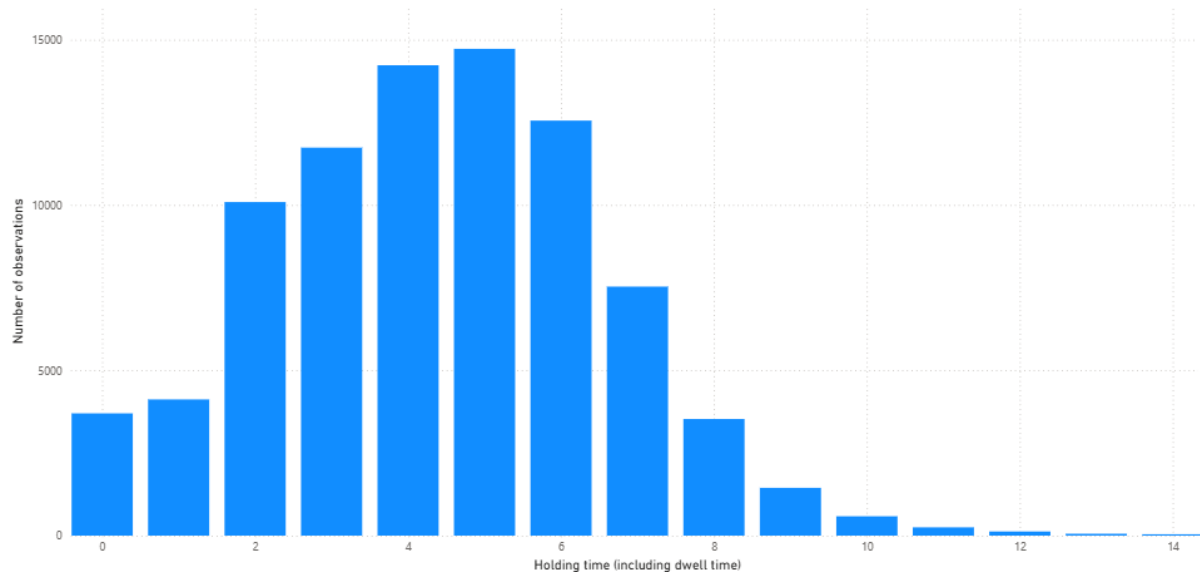


Figure 4.9: Holding time distribution at *Hoofdstation*

4.5. Passenger patterns

The diametric nature of the Q-link 3 and 4 lines shows a typical loading profile. Figure 4.10 shows that most passengers board at the branches, of which half alights at *Hoofdstation*. As a result, about half of the passengers can be defined as through passenger, while the remainder is a direct or transferring passenger. The busiest alighting stop is *UMCG Noord*, after the tail of the route shows the opposite of the start, in which the passenger load gradually decreases.

Before the TLS dataset was available, a guess was made based on personal experience on the share of passengers transferring at the main transfer station *Hoofdstation* in Figure 4.11. The shares are used vice-versa, for both incoming and outgoing trips and is similar for each hour of day. The largest share of passengers is assumed to have a travel relation with the southbound trains, operated by the Dutch Railways (NS). The southern branch namely connects to the *Randstad*, the largest metropolitan area in The Netherlands. The other train lines are assumed to have a share of 10 per cent, each. Although most destinations are closer than to the *Randstad*, the lines mostly serve regional villages and minor cities. For 30 per cent of outgoing and incoming passengers, no transfer is assumed. In this thesis, this passenger is referred to as direct passenger. This means either their journey starts or ends at *Hoofdstation*, without using other transit vehicle trips.

In reality, *Hoofdstation* provides transfers to local and regional buses, too. Furthermore, it is not the only transfer station on the route of Q-link 3 or 4. All transfer possibilities are to be found in Appendix B. The assumption acts as a starting point from which it is able to add transfer time to the journey time for a large group of transferring passengers. The assumption had to be made in order to test the methodology. Later, the assumption was lifted and replaced by the TLS dataset. This did not alter the methodology of this thesis.

Specific to this case study, alterations have to be done to make the transfer assumption work in the network-time graph. The assumption is processed by manipulating the original OD-matrix, which is based only on the demand within the study area. Whenever a passenger was observed having ended its trip at *Hoofdstation*, the demand is split. In this case, 4 additional demand cells are added to the matrix, contributing to a total of 5 cells. Then, the amount of passengers is multiplied by the assumed share. If the original demand cell consisted of 10 passengers, it is split to 1, 1, 1, 3 and 4 passengers, respectively for eastbound, northbound, westbound, no transfer (which remains at *Hoofdstation*) and southbound connections. These five new cells replace the former singular cell. Now, the OD-matrix forces passengers to transfer in the network-time graph.

4.12 presents a Shankey diagram in which the OD and TLS datasets are combined. The OD is used

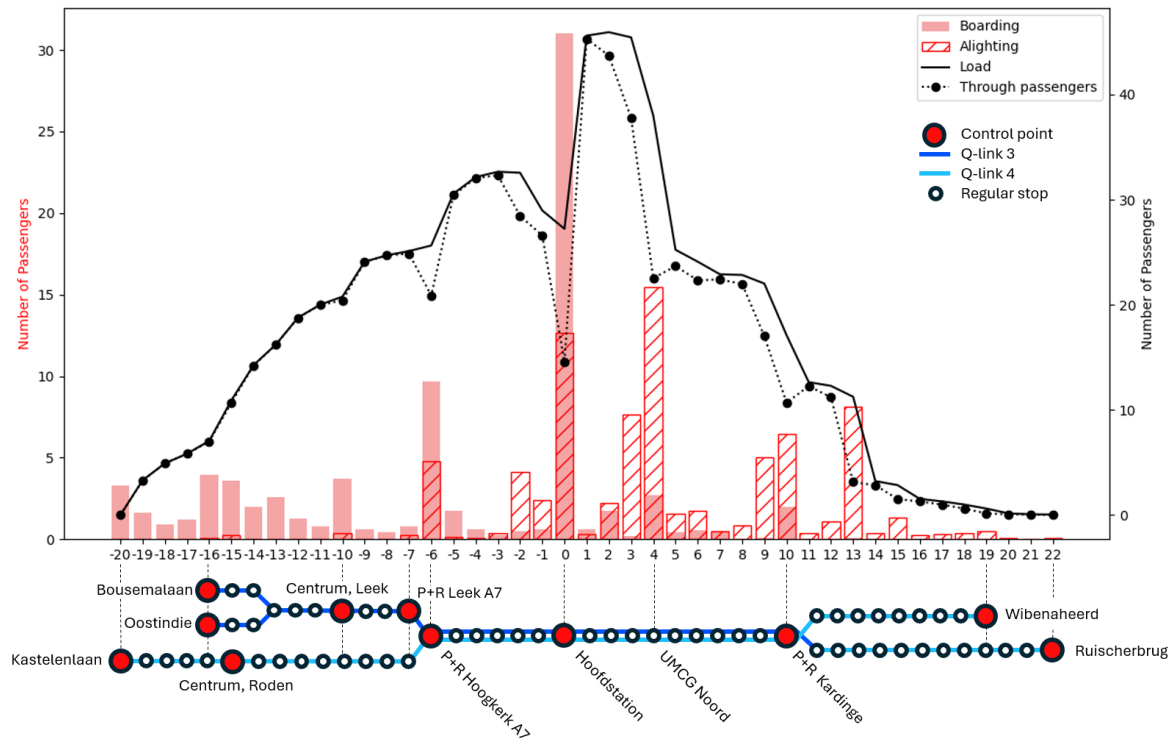


Figure 4.10: Eastbound load profile of an average Q-link 3 or 4 bus between 8 and 9 AM

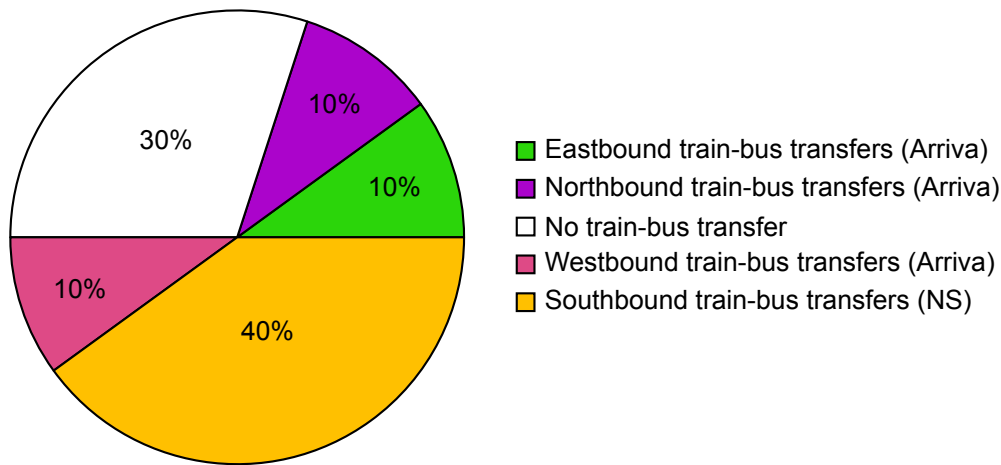


Figure 4.11: Assumed (vice versa) transfer relations at Hoofdstation (with corresponding train operator)

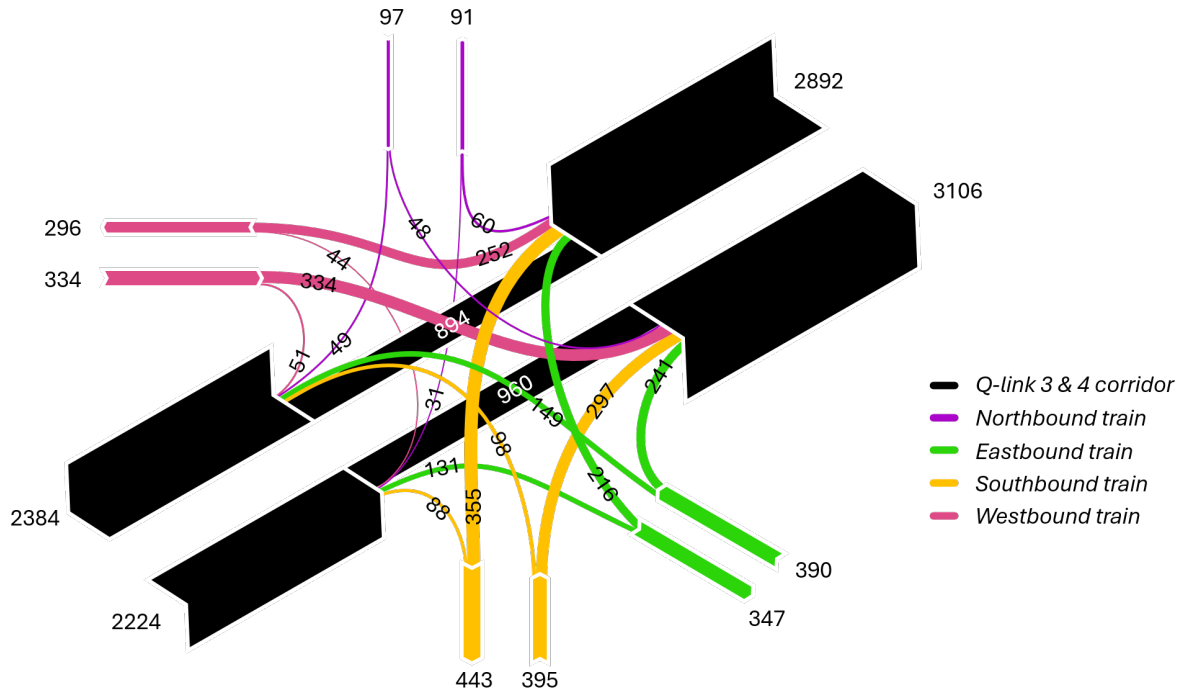


Figure 4.12: Shankey diagram of the average daily amount of transferring passengers at *Hoofdstation* from train to bus and vice versa

to determine the passenger load on the corridor, while the TLS is used to determine the amount of transferring passengers. It gives insights on the busiest transfer station *Hoofdstation*. The width of the lines corresponds with the absolute value labels. It is visible that around 30 to 40 percent of passengers are through passengers. After the addition of train-bus and bus-train passengers, all corridor directions have 50 percent or more observed passenger behaviour. There are two possibilities for the unrepresented passenger behaviour. The first possibility is that their journey either start or ends at *Hoofdstation*. The second possibility is a bus-bus transfer. The TLS data includes these numbers too, but for time limitations it will not be part of this thesis. Note that the Shankey diagram is a daily average over all service days d . During the different service hours h , this Shankey diagram will have different absolute and relative values. Because the TLS dataset consists of all the service hours h , the mathematical model in this thesis uses this hourly variation.

The transfer assumption can be regarded as an overestimate of transfers. If we compare the assumption of Figure 4.11 to the real data as presented by the Shankey diagram in Figure 4.12, it can be seen that each separate train lines contributes less than 12 per cent of the total passenger flow. Note that this is a daily average, so the percentage can be a fraction different over the course of a day. The assumption is included in the analysis, because it can give hypothetical context in an extreme amount of transfers. In total, three scenarios can be compared: no transfers, data-supported transfers and overestimated transfers.

Concluding, the case of the Q-link lines in Groningen can be a perfect testing field for the proposed methodology. The quantitative effects of holding are unknown, because the current supply-oriented metrics are unable to meaningfully describe the high-frequent Q-link system in terms of passenger experience on speed and reliability.

Bus on bus transfers are not considered due to (computational) time limitations. The methodology however, is suitable to include these transferring passengers too. To give an idea of the scale of transfers, the 10 most frequent transfers are given in Table 4.2. An extended version of this overview that includes up to 50 transfers can be found in Appendix E. It is expected that bus-bus transfers have a smaller impact on the results, than train-bus transfers. Reason being the smaller scale: the busiest bus-bus transfer has an equal amount of passengers as the least used train-bus transfer opportunity.

	Transfer station	Current line	Future line	#pax
1	Hoofdstation	Q-liner 309	Q-link 3/4 (eastbound)	102
2	Hoofdstation	Q-liner 300	Q-link 3/4 (eastbound)	98
3	Hoofdstation	Q-link 3/4 (westbound)	Q-liner 309	93
4	Hoofdstation	Q-link 3/4 (westbound)	Q-liner 300	84
5	P+R Hoogkerk A7	Q-link 3/4 (eastbound)	Bus 10	71
6	Ruischerbrug	Q-link 6 (eastbound)	Q-link 3/4 (westbound)	69
7	Verzetsstrijderslaan	Q-link 15 (southbound)	Q-link 3/4 (westbound)	67
8	UMCG Noord	Q-link 1 (southbound)	Q-link 3/4 (eastbound)	60
9	P+R Hoogkerk A7	Q-link 3/4 (eastbound)	Bus 18	58
10	Hoofdstation	Bus 50	Q-link 3/4 (eastbound)	52

Table 4.2: Top 10 most used bus on bus transfers (daily average amount of passengers)

Another reason is the high frequency of most connecting bus lines. Consequently, bus-bus transferring passengers tend to arrive less clustered, compared to the train-bus transferring passengers. Still, it is recommended to include bus-bus transfers, for a better estimate of the transfer time at all transfer stations.

5

Experimental setup

This chapter explains which variables are changed for the selection of the holding control strategies. Then, it is explained which simulation setups are used. Finally, it is visualized how the case study is implemented into the network-time graph.

5.1. Selection of holding control strategies

The observation of the current timetable is disadvantageous for through passengers, because the buses are held at control points. The hypothesis of the case study problem is that a tighter timetable will be beneficial. Since the relative loss of travel time due to holding is unknown, a quantitative assessment is needed to express the minutes gained as opposed to the minutes lost associated with holding. Thus, strategies were developed where theoretically, less holding is expected to be needed. The duration of holding is dependent on the percentile used for the timetable determination. With schedule-based holding, a bus cannot depart until this percentile value has elapsed. For headway-based holding, a maximum holding duration can be set.

The set of the tested strategies can be found in Table 5.1. SB indicates schedule-based holding. In Figure 5.1, it is illustrated that the amount and location of control points can be varied. "single" means there is just one control point at *Hoofdstation* and "multi" indicates that all control points are used. The number in the strategy name determines the strictness of the schedule. For SB-multi 50 for example, the 50th percentile of segment running time is used to calculate the earliest time of departure at control points.

SB-multi-85 is set as the base strategy. Reason being the representativeness for the current situation. The segment times between control points are most similar to the strategy that is in place as of July 2026. In real-life, the strategy in place is more complex, with different segment time percentiles in place that determine the strictness of the schedule.

For the finalized dataset including the addition of transfers, 8 additional strategies are analysed in terms of their RRR performance. As the former strategies contained of schedule-based holding only that was applied at all control points, the extended strategies alter in the amount of control point. Furthermore, the extended strategies contain headway-based holding.

5.2. Selection of simulation setups

The simulation is ran with different setups that differ in terms of their demand data availability.

1. **Without transfers:** first, the "SB-multi-..." strategies were compared based on the travel time of the passengers in the origin-destination dataset only.
2. **Overestimated transfer assumption:** second, the transfer assumption as visualized by Figure 4.11, was used to determine the travel time including theoretical transfers. The "SB-multi-..." strategies were then compared in terms of their RRR performance.

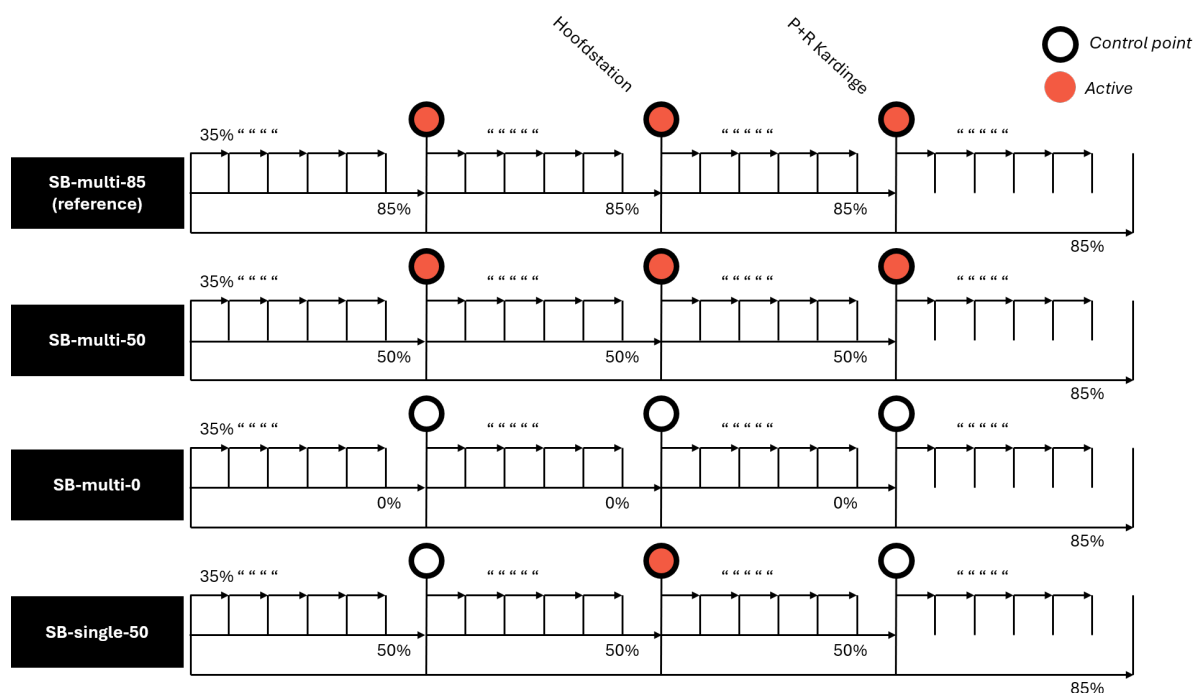


Figure 5.1: Schedule-based holding strategies

holding strategy		decision variables		
name	condition	control points	segment percentile	max. holding time (mins)
SB-multi-85 (base)	schedule-based	6	85	∞
SB-multi-65	schedule-based	6	65	∞
SB-multi-50	schedule-based	6	50	∞
SB-multi-35	schedule-based	6	35	∞
SB-multi-0	none	0	0	∞
<i>Extended analysis</i>				
SB-single-85	schedule-based	6	85	∞
SB-single-65	schedule-based	6	65	∞
SB-single-50	schedule-based	6	50	∞
SB-single-35	schedule-based	6	35	∞
HB-multi-4	headway-based	6	0	4
HB-multi-2	headway-based	6	0	2
HB-single-4	headway-based	1	0	4
HB-single-2	headway-based	1	0	2

Table 5.1: Strategies to be simulated

3. **Data-supported actual transfers:** finally, the transfer assumption is replaced by the TLS dataset, which includes data on the amount of passengers that utilizes specific transfers. As explained in Chapter 4.5, only train-bus transfers are considered.

This comparative analysis was possible only for the "SB-multi-..." strategies, due to computer memory limitations, the "SB-single-..." and "HB-..." strategies were simulated with a reduced origin-destination matrix, that only included the demand cells that showed on average 0.58 passengers per 15-minute OD-window. In the comparative analysis, all passengers were included. This reduction resulted in 33.000 to 4.000 unique origin-destination-time pairs. This does not mean the same reduction took place in the amount of passengers that were included in the analysis, because most passengers travel on popular origin-destination pairs. To illustrate, the demand load function showed no visible difference between the full OD dataset and the reduced OD dataset at 8 AM. The impact is more visible in off-peak hours. This makes the results of the extended analysis more representative in peak hours, compared to off-peak hours.

5.3. Implementation of the case study in the network-time graph

Every simulation run initiates a network-time graph. In the graph, the simulated vehicle trips that have driven under a holding regime are stored. This section showcases the network-time graph in action.

To visualize the graph creation process, the network-time graph can be plotted. Figure 5.2 shows all control points on a plane. A third dimension is added so the network transforms into a time-space diagram (Figure 5.3). The X and Y axis represent the longitude and latitude, respectively. The newly introduced Z-axis represents the time. The edges that connect the nodes are directed so that passengers cannot travel back in time. This means that passenger journeys can only progress if the next node has a higher Z-value than the previous. From any given node in the network-time graph, the resulting candidate actions for passengers then become taking a bus (blue lines), waiting at a stop (red line). All possible candidate actions are recursively explored, after which the fastest route is decided upon according to Chapter 3.3.2.

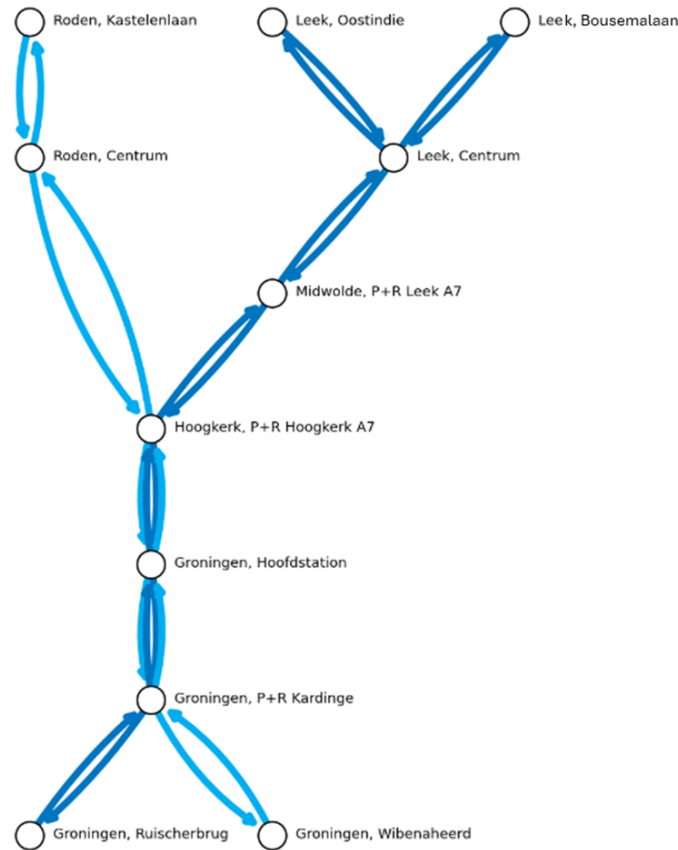


Figure 5.2: Network graph with all line directions

Now with the time-axis as a third dimension, the whole network can be plot. Figure 5.4 shows the studied lines in a network-time graph. In contrast with Figure 5.3, the node labels are not displayed anymore for viability reasons. Furthermore, the control points are complemented by all intermediate stops. This leads to the possibility of inserting an origin-destination demand dataset, which is specified on the stop level, too. Note that only two deterministic vehicle trips are plot, because the edges are drawn perfectly in parallel. This figure illustrates a synchronized schedule. The lines are synchronized at some points, as can be observed at long 6.50, lat 53.22. From that coordinate, which represents *P+R Hoogkerk*, Q-link 3 and Q-link 4 are spaced evenly in northeastern direction.

Finally, the study area has to be connected with transfer lines that connect to the cordon (Figure 1.2). Figure 5.5 is a visualization of the actual simulation run used in the passenger assignment. The figure only visualizes a subset of the run between 8:00 and 8:30 am to make it somewhat readable. Consequently, the top and bottom lines are cut before having reached the terminals. Additionally, the simulated lines run less parallel, which is visible with Q-link 3 at 6.45, 53.14. The highest and second highest edges, which have the same direction, indicate line irregularity.

The transfer lines, now only consisting of train lines connecting at *Hoofdstation* are present as well. The connection time to the study area transfer point is shifted by two minutes compared to the actual arrival or departure of a train, in order to respect the deterministic walking time from bus to train and vice versa. Consequently, a passenger has to arrive at least two minutes in advance to have a feasible transfer to an outgoing train. The other end of the train lines are connected to the cordon, which is fictional. This makes it irrelevant where the edges lead to in the figure. An arbitrary node location is picked, mostly on one of the first stations the train line serves after having left the study area. In contrast, the edge weight does matter a lot. Due to the scope of this thesis, the travel time of the train lines are not to be included in the RRR indicator. Hence, the weight of the train line edges are set to zero. Another critical attribute of the transfer lines are the directions. If they are not set properly, passengers are unable to

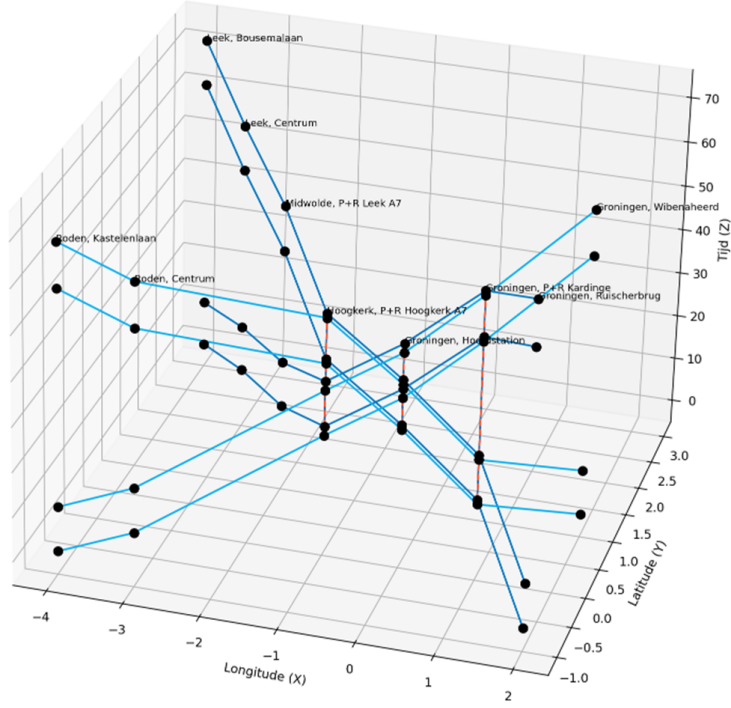


Figure 5.3: Network-time graph with two vehicle trips per line direction running between control points

use the lines in their path to reach the origin or destination of their journey. As a final remark, during the simulation runs, the network-time graphs as the one in Figure 5.5 are not drawn each iteration. This saves computational time. It is solely used as a verification tool, whether a single simulation run is able to properly include the complete set of desired lines.

For transfers specifically at *Hoofdstation*, it is assumed that 2 minutes are needed to cover the walking distance between all types of transfers. To reduce modelling complexity, this duration is assumed homogeneous across all passengers. For incoming and outgoing transfer lines, respectively Equation 5.1 and 5.2 are used to make journeys with transfer times below 2 minutes infeasible.

$$n_{ct} = a_{ct} + 2 \quad (5.1)$$

$$n_{ct} = v_{ct} - 2 \quad (5.2)$$

where:

n_{ct} : network-time node at control point c representing transfer line t

a_{ct} : arrival time of transfer line t at control point c

v_{ct} : departure time of transfer line t at control point c .

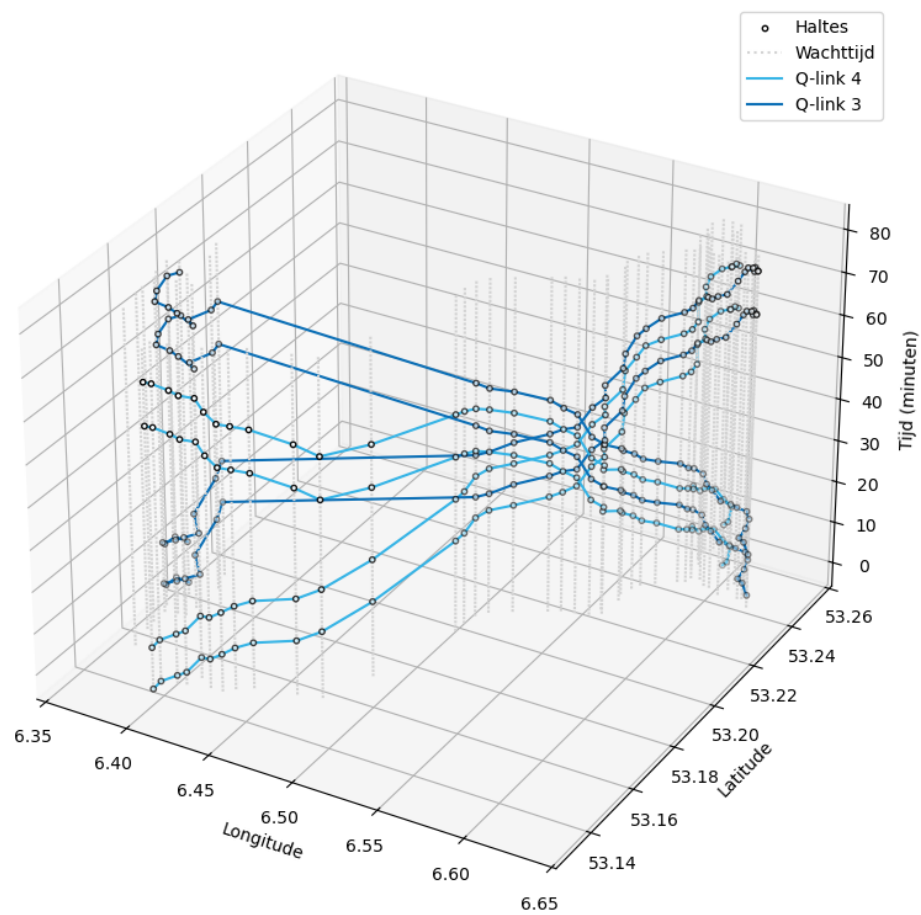


Figure 5.4: Network-time graph with two vehicle trips per line direction running between all stops

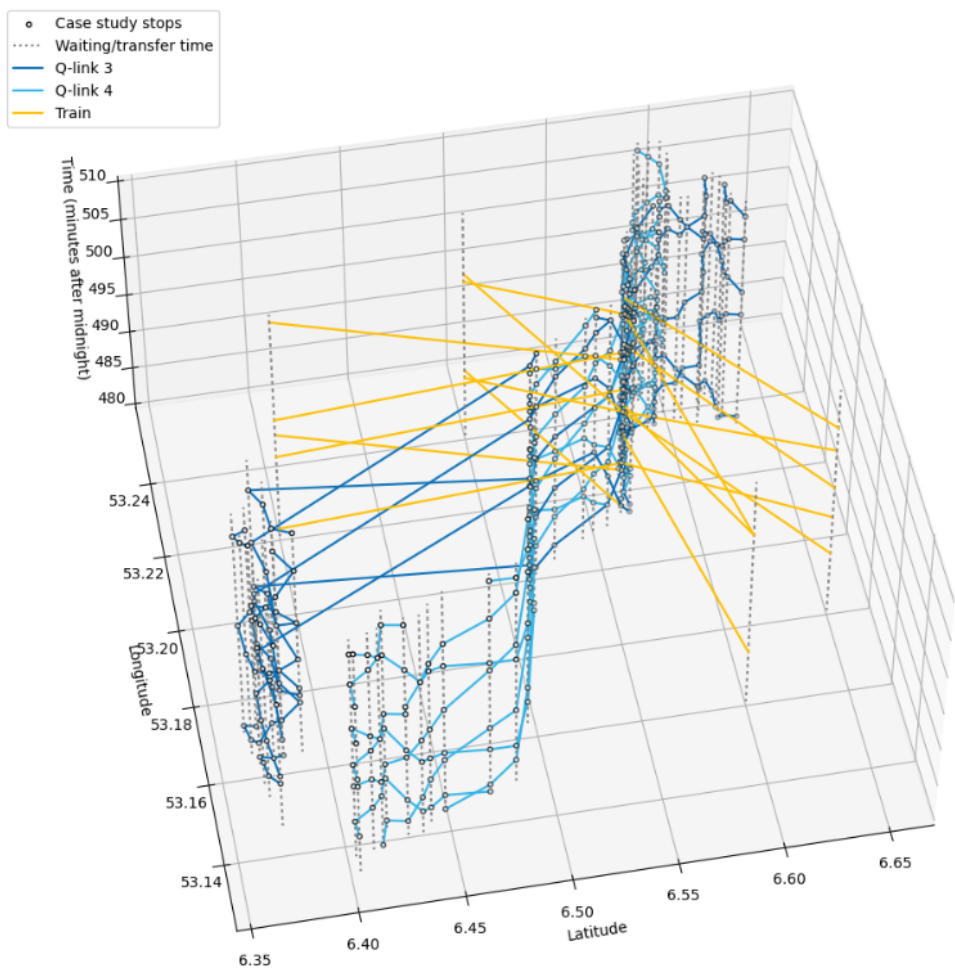


Figure 5.5: Network-time graph including bus-train transfers between 8:00 and 8:30 am

6

Results

This chapter gives an overview of results. First, the cost of the holding strategies is given, which is followed by the results on how the holding strategies perform under the new RRR indicator. Finally, there is an elaboration on the stakeholder roles, as a result of the focus group session.

6.1. Timetable cost

The advantage of saving the vehicle trips in a network-time graph, is that the total runtime can be calculated over all vehicles of all lines. This tells something about the associated cost of the chosen strategy. The simulated timetables can be compared in terms of the estimated costs in terms of timetable hours (TTH).

Table 6.1 shows that there is no significant difference in the sum of vehicle runtime. A t-test is performed to compare the base strategy with the other strategies in search of significant differences. This is calculated according to Equation 3.10 and 3.11. The p-value never falls below 0.05, meaning that the total TTHs are statistically indifferent to each other. For this reason, no additional trips are added to strategies with a tighter schedule. In reality, the determination of the cycle time is based on the 85th percentile of trip times for chosen hour blocks. Thus, for future analysis it is recommended that the TTHs are calculated based on the 85th percentile of total trip time over a chosen block of hours.

6.1.1. Tension between drivers, passengers, operator and authority

Three members of the focus group were in possession of a bus driver license and have driven the studied Q-link lines. It is acknowledged that there is a consensus among drivers to drive between perfectly punctual or a bit on the early side, in order to secure transfer connections and match with their crew roster (e.g. driver shift brakes). On the contrary, passengers expect buses to be on time, or slightly on the late side, so they do not miss the planned journey. Furthermore, transferring passengers expect a punctual arrival at their transfer stops, too. If a schedule is designed with enough holding time, the wishes of drivers and transferring passengers are fulfilled. Furthermore,

strategy	TTH		significance	
	mean	standard deviation	t-statistic	p-value
SB-multi-85 (base)	245.65	0.8567	-	-
SB-multi-65	245.59	0.9504	0.333	0.739
SB-multi-50	245.51	0.7519	0.878	0.382
SB-multi-35	245.57	0.8357	0.456	0.649
no holding	245.79	0.9063	-0.792	0.429

Table 6.1: The daily cost of timetable strategies (based on $m=50$ initial runs)

the operator is better able to minimize cost, since there is less risk in overtime driver shifts and to compensate for cost associated with delays, such as service quality punishments set in the concession agreement.

However, direct through passengers are less content with a loosely design schedule, because they experience additional in-vehicle time. Another actor that is against too much slack time, is the transit authority. Since a fixed hourly timetable unit is set in the concession agreements, unused slack time is reimbursed by the authority, without the passenger benefiting from it. If the same budget can be used to accommodate for for example a higher frequency or additional lines, more ticket income can be generated and more demand can be satisfied. The operator demands less strict service quality punishments in the concession agreement, when a tighter schedule is chosen. For high-frequent lines, another indicator than solely vehicle punctuality should be considered then.

As explained above, the breaks are an important aspect of setting the definite scheduled trip times. It is thereby recommended to conduct a full analysis on the scheduled trip time given chosen blocks of hours, since they form the main determinant on how much time is embedded into the schedule. In turn, this leads to the cost for the authority and the risk for the operator. It can yield larger returns on investment if the amount of TTHs is set high due to a loose schedule. On the contrary, there is a higher chance the drivers are satisfied with their breaks, which makes them perform better for the remainder of the shift. Furthermore, the chance of stop-skipping (Figure 1.3) decreases, benefiting both the operator and authority.

6.2. Rider-based Rapidity and Reliability

In this section, the holding strategies are compared in terms of RRR. Three simulation setups are used. First, the simulation is ran without transferring passengers. Second, the transferring passengers are included in the form of a passenger assumption, according to Figure 3.6. Finally, transfer data is replaced for the transfer assumption, similar to Figure 4.12. The results are then compared in Table 6.2. For the demand including transfers, 8 additional

Figure 6.1 and 6.2 work as follows: the horizontal axis shows the hours of the day in which passengers arrive. For example, an 8 represents all the arriving passengers between The vertical axis is the relative difference to a set base line. If absolute levels were chosen, it is hard to draw conclusions. Plus, the absolute RRR level does not have a stand-alone value. SB-multi-85 is set as the baseline, since this strategy has most similarities with the strategy that is in place in Groningen. The other strategies are plot around the base line, so that they can be compared to the current strategy. A strategy that sits lower than the base line means it performs better at the chosen hour. A strategy that sits higher than the base line is a worse performing strategy in terms of RRR.

6.2.1. Excluding transfers

First, the model is ran without traking into account transfers. Figure 6.1 shows the strategy performance for each hour of service. The key observation is that during the day, the current strategy scores relatively well. Installing another strategy will improve the RRR by no more than 3 per cent. Note that the recommendations following this first figure are only based on the demand within the case study lines, excluding transfers.

The results suggest that in the morning rush hour between 7:00 and 10:00, the current strategy (SB-multi-85) performs average compared to other strategies. Although some strategies outperform SB-multi-85, it is only by a small margin of less than 1 per cent. This is possibly explained by the large trip time variation in the morning rush hour (Chapter 4.4), which makes holding effective in ensuring line regularity. Therefore, it is not recommended to tighten the schedule in the morning rush hour

After the morning rush hour however, from 10:00 to 13:00, several strategies outperform SB-multi-85. If we compare the demand-oriented situation with rush hour, the scheduled headways have remained at she same level. At the same time, the trip time variety decreases. Hence, holding becomes less necessary to ensure line regularity. This is a plausible explanation for the well performing tight schedules. It is recommended to schedule according to SB-multi-65 or SB-multi-50 until at least 13:00.

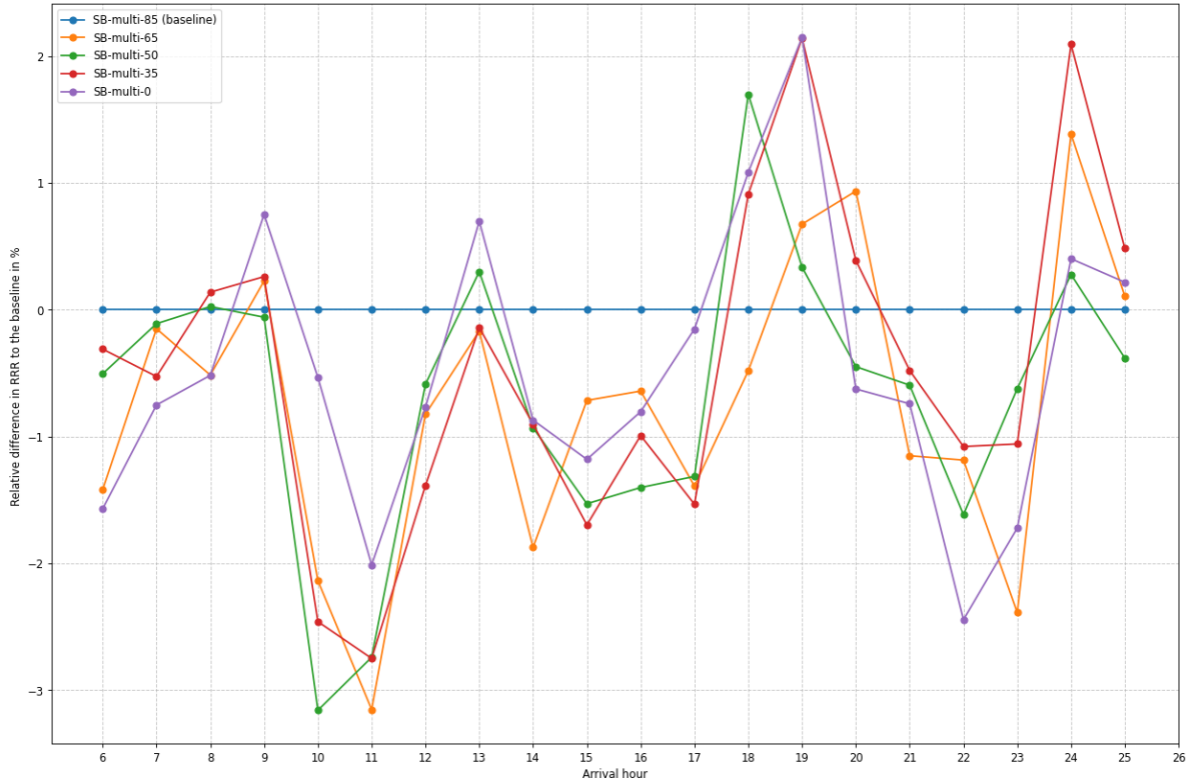


Figure 6.1: Strategy performance of 50 runs excluding transfers

At 13:00, SB-multi-65 is still the best performing strategy. Not by much, as SB-multi-85 scores similar. If we look once more at supply-oriented analysis, the sudden performance difference change could be related to the start of the second, more gentle local maxima. At 13:00, all lines in all direction experience an increase in trip time variation. Because the schedule determination is often determined over a block of multiple hours, SB-multi-65 is still recommended from 13:00 until 15:00, at which the strategy is outperformed by even tighter schedules. If SB-multi-65 is chosen upon directly after leaving the rush hour, the strategy can be sustained until 19:00, without being outperformed too much by other strategies.

In the actual afternoon rush hour from 15:00 up until 18:00, it can be observed that SB-multi-35 outperforms SB-multi-65. If we compare the morning with the afternoon rush hour in the supply-based background, It can be observed that the trip time variation is more stable over the period and on average, lower. This could explain why in the morning rush hour, SB-multi-85 scores better than in the afternoon rush hour. The block where SB-multi-35 dominates over SB-multi-85, consists of 3 consecutive hours. Thus, the strategy could be considered for the afternoon rush hour, in favor of SB-multi-65.

After 18:00, the frequency of the lines drop. Still, there is demand spread over the hour that can benefit from good line regularity. After 20:00, the results are expected to be biased, since the frequency is at its lowest. This can be explained by the following: for low-frequency hours, the historical demand is dependent on historical supply. The OD-matrix in this case study is specified in hour quarters. When a trip was previously scheduled in a quarter bin, whereas when no trip had been scheduled in the following trip, this directly translates into OD-matrix values. It is recommended to further analyse the effects of different binning methods of passengers during low-frequency service hours, to prevent status quo bias.

Concluding from the analysis without transfers, the performance seems dependent on a combination between supply and demand oriented variables. The trip time variation has the most explanatory power, since holding becomes more necessary when the trip time variation is high. Lower trip time variation leads to better performing tight schedules. The trip time variation could be caused by passenger volumes. The passenger volumes may have a explanatory power itself, if the demand is clustered on

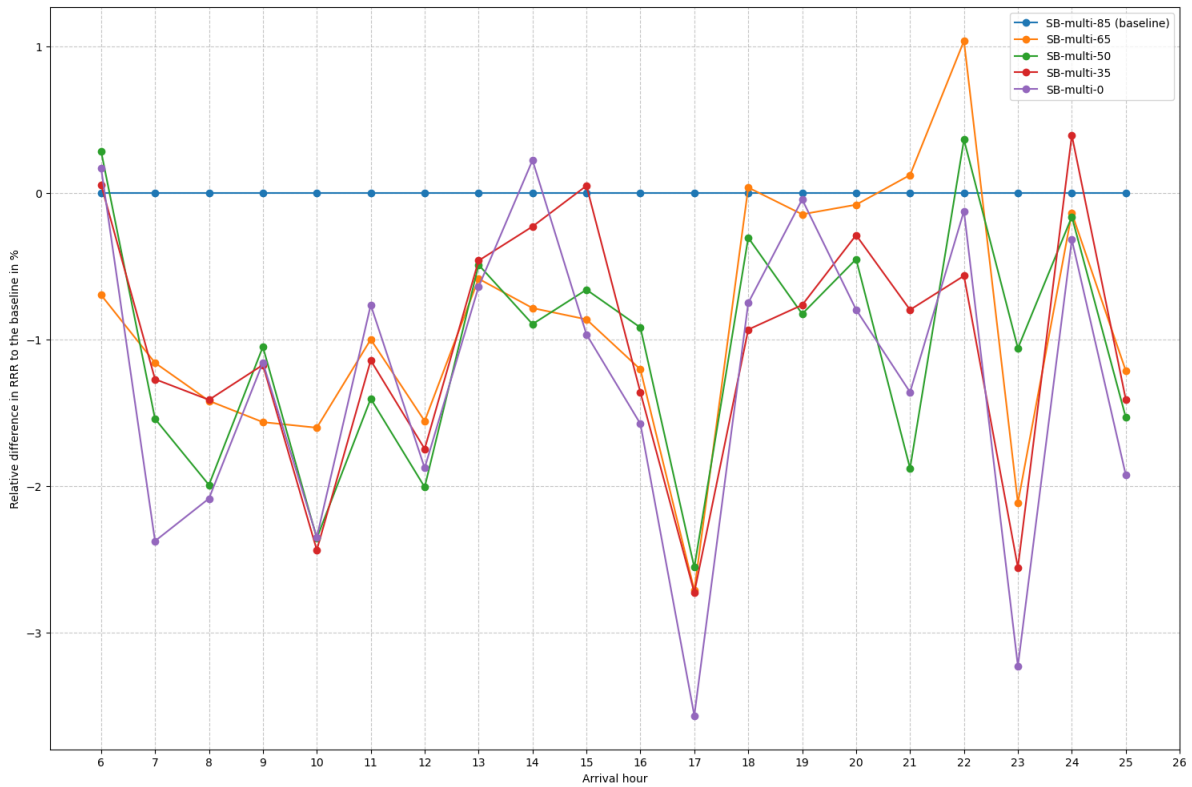


Figure 6.2: Strategy performance of 50 runs including transfer assumption

stops where there is line irregularity.

6.2.2. Including transfer assumption

After the OD-data was manipulated in order to facilitate transfers to the cordon area, the simulation was ran for another 50 initial runs. This time, passengers were forced to transfer according to the transfer assumption in line with Figure 4.11.

Overall, what immediately becomes clear is the poor performance of the base strategy. At none of the passenger arrival hours, SB-multi-85 is outperformed by other strategies. Over the course of the day, a mix of strategies performs best, which are SB-multi-0, SB-multi-35 and SB-multi-50. The hour at which the highest percentile outperforms the rest of the strategies is SB-multi-65, only from 9:00 until 10:00. What is similar to the simulation without transfers is the maximum relative difference between two strategies. Choosing one strategy over another will result at a maximum of 3.5 per cent of total travel time improvement for all the passengers in the system. Although this is a minor improvement, only slight alterations in the timetable are needed in order to execute another holding strategy. If the right tightness is chosen, a quick win can be made in favour of passenger experience.

If we look once more at the rush hours, it is interesting to see that in both the morning and the afternoon rush hour, SB-multi-0 outperforms all other strategies. The only variable that was changed between the two simulations was the ability of transferring. The difference thus has to be explained by the decrease in transfer time in the rush hours, when the most tight schedule is chosen. This means that arrival distribution of passengers must be centred around the departure of trains, instead of the assumed uniform arrival distribution in the previous simulation.

Between rush hours, the transfer time has resulted in a dominant strategy that is less tight than in the rush hours. Between 9:00 and 10:00, SB-multi-65 scores best. Between 10:00 and 11:00, SB-multi-35 performs best. Between 10:00 and 16:00, SB-multi-50 performs as one of the best strategies. To maintain larger hour blocks, this strategy is recommended between these hours. After the rush hour, it is recommended to stick with SB-multi-0 until the end of service, with a gap between 19:00 and 23:00,

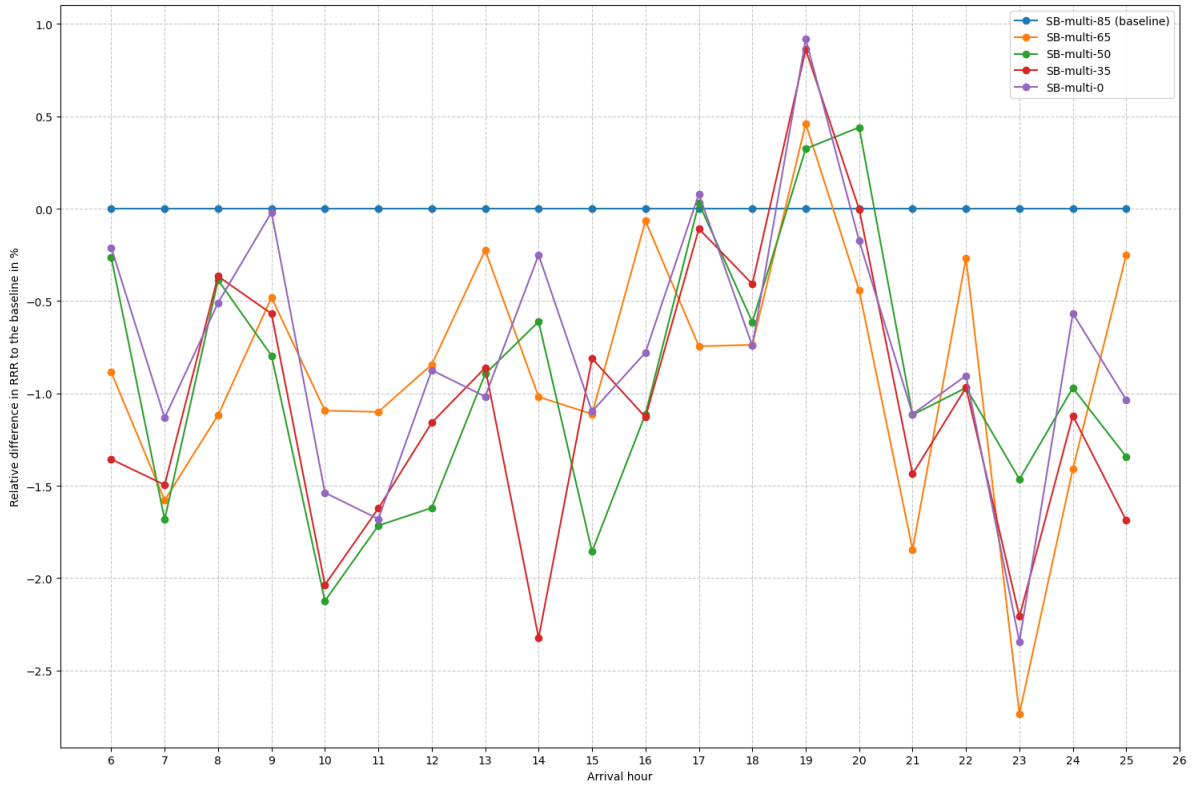


Figure 6.3: Strategy performance of 50 runs including transfer data

in which SB-multi-50 performs better.

Concluding, the inclusion of the transfer assumption made a tighter schedule more favourable, then in a situation where an uniform passenger arrival pattern is assumed. The only explanatory variable is the transfer time, which is now explicitly included in the total journey time. These results underline the importance of including transfers in timetable related research questions.

6.2.3. Including transfer data

When the transfer data is included, it results in an intermediate strategy performance, in between the overestimated transfer assumption and the sole OD-dataset.

In rush hours, SB-multi-65 performs relatively well. In between rush hours, SB-multi-35 and SB-multi-50 perform better. Interestingly, the no holding strategy SB-multi-0 does never outperform another strategy. In contrast, SB-multi-85 is the best performing strategy for arrival hour 19. The overall difference between strategies is no more than 2.7 per cent.

6.2.4. Comparison of results

When all the above results are compared, it can be concluded that the exclusion of transfers favours schedule-based holding strategies with a higher percentile for segment time determination. In comparison, the overestimated transfer assumption favours strategies with lower percentiles. When the amount of transfers is based on actual transfer data, which contains of less transfer movements, the dominating strategy is a mix between the former two simulation setups. Before and in the morning peak, a higher percentile than in the simulation setup excluding transfers dominates. After the rush hour, the dominating percentile is at least equal or lower than the dominating percentile for when transfers are excluded. Concluding, the addition of transfer time has made lower percentile strategies perform better, compared to when transfers were excluded.

		dominating strategy		
		excluding transfers	including transfers	
group name	window	OD	OD + assumption	OD + TLS
early morning	06:00 - 06:59	SB-multi-0	SB-multi-35	SB-multi-35
morning peak	07:00 - 08:59	SB-multi-0	SB-multi-0	SB-multi-65
morning off-peak	09:00 - 11:59	SB-multi-50	SB-multi-50	SB-multi-50
lunch	12:00 - 13:59	SB-multi-65	SB-multi-50	SB-multi-50
noon peak	14:00 - 15:59	SB-multi-65	SB-multi-50	SB-multi-35
evening peak	16:00 - 17:59	SB-multi-50	SB-multi-0	SB-multi-35
evening	18:00 - 23:59	SB-multi-65	SB-multi-0	SB-multi-65
night	00:00 - 01:59	SB-multi-50	SB-multi-0	SB-multi-35

Table 6.2: Comparative analysis: Dominating holding strategy per hour of day type and per demand data source

time		dominating strategy
group name	window	
early morning	06:00 - 06:59	SB-single-65
morning peak	07:00 - 08:59	SB-single-65
morning off-peak	09:00 - 11:59	SB-multi-50
lunch	12:00 - 13:59	SB-multi-35
noon peak	14:00 - 15:59	SB-single-65
evening peak	16:00 - 17:59	SB-single-50
evening	18:00 - 23:59	SB-single-85
night	00:00 - 01:59	SB-multi-0

Table 6.3: Extended analysis: dominating holding strategy per hour of day type (OD + TLS demand dataset)

6.2.5. Extended analysis

When all the available strategies are tested against OD + TLS demand data, different results are visible. Figure 6.4 shows the strategy performance when all the available strategies are used. Note that due to time and computer memory limitations, $m = 20$ initial runs are plot.

The most noticeable observation is the poor performance of the headway-based holding strategies. This is most likely caused by an assumption in the methodology. Since the departure time from terminals is assumed fixed, vehicles hold while there is no delay to preceding vehicle but a mismatch in planned headways. The interaction between vehicles is most likely the problem, since at the beginning and to the end of service, the HB-single-... strategies seem to perform similar to other strategies. Another plausible explanation could be an error in the coding process, which extremely influences the holding time. Hence, the HB-... strategies do likely not represent their true performance and are therefore not considered in the conclusions of this thesis.

When looking at the valid results, the SB-single-... strategies are able to outperform SB-multi-... strategies. Especially after arrival hour 17 until arrival hour 23, the SB-single-85 outperforms all other strategies. Furthermore, SB-single-65 outperforms all other strategies between arrival hour 6 and 8, 9 and 11, 14 and 15. This suggests that the SB-single-... strategies are suitable for off-peak hours. Table 6.3 shows the updated dominated strategies, now including both SB-multi-... and SB-single-... strategies.

6.2.6. Replications needed for significance

Bootstrapping is used to test the variation in the sample of $m=50$ initial simulation runs. Within each bootstrap, a subset is drawn of 42 simulation runs. For every subset, the mean RRR was determined, as well as the standard deviation. A t-test was performed according to equation 3.10, with the SB-multi-85 as reference sample. A total of 10 bootstraps ($b=10$) was used, which is normally used for data exploration only. This low number was chosen due to time limitations. It is recommended to increase

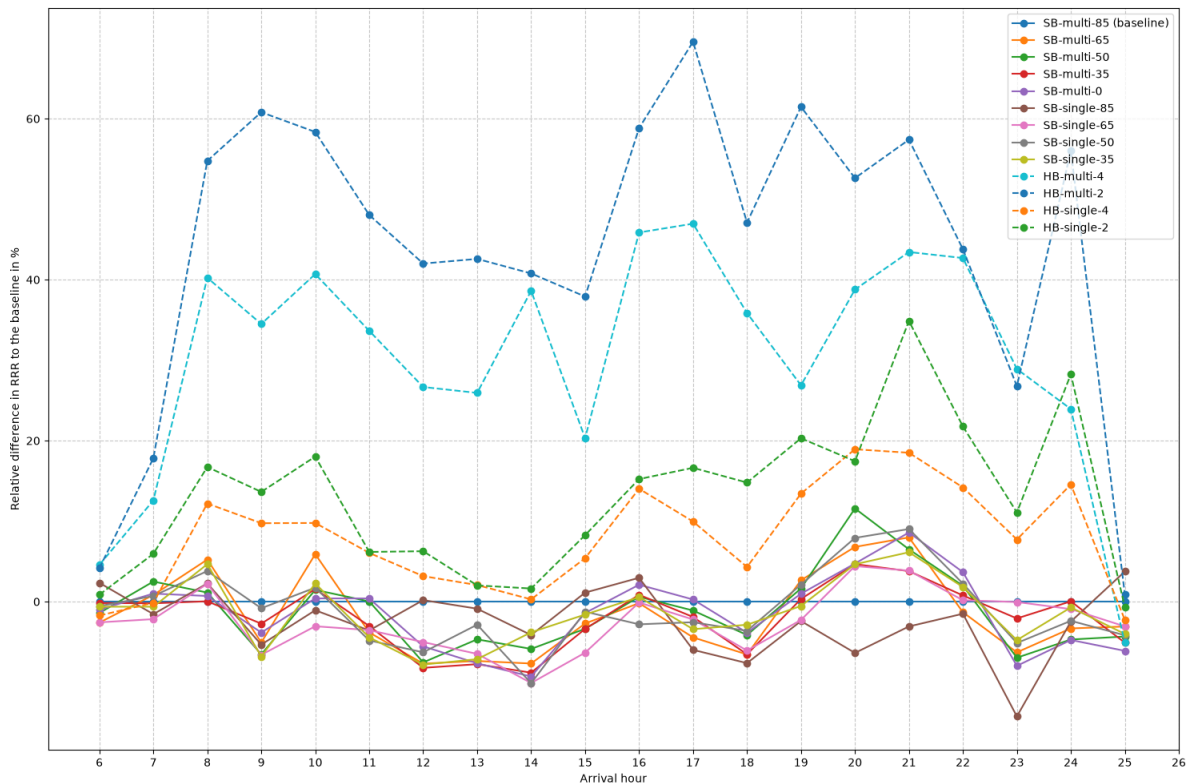


Figure 6.4: Extended strategy performance of 20 runs including transfer data

this number to at least 20 and ultimately 100. The mean and standard deviation are not representable yet for the whole sample. Still, it can give insights on how the significance behaves within the sample. In the Appendix, Table D.2 presents the standard deviation and the mean of the RRR indicator within the set of 50 simulation runs. The amount of needed additional simulation runs can be decided on by

For every hour, the best performing strategy in Table D.2 is marked in bold. In general, this corresponds with the best performing strategy in Figure 6.2, with similarities between table and figure. Differences can be observed too, because the table is based on bootstrapped data, while the figure is based on the full set of 50 runs. Most p-values stay below the 0.05 threshold, which means that most of the strategies perform significantly different compared to the base strategy. The values that are not either lie too close to the base strategy, or a larger number bootstraps or simulation runs is needed. It is recommended to gradually increase these numbers. Note that the p-values only tell something about the difference between SB-multi-85 and the other individual strategies. For example, the p-value does not describe the difference between SB-multi-0 and SB-multi-50.

An error is visible at arrival hour 8 for SB-multi-85. The table shows a 50 per cent strategy performance difference, compared to SB-multi-65. However, the figure . It is recommended to investigate the error. Most probable is a faulty bootstrap algorithm, as it is heavy on computer memory. The rest of the table is consistent with the figure.

6.3. Stakeholder relations

This section elaborates on the stakeholder relationships that were observed during the focus group. The invited focus group members are representatives of different organizational actors of the bus network. All members were present, besides the representative of the *Municipality of Groningen*.

- **Transit authority:** The *OV-bureau Groningen Drenthe* is responsible for managing the public transport contract. The agency mainly checks if the contractual agreements are legally complied with during operations. Unique to the Netherlands, the OV-bureau receives ticket income. This explains their pro-active role in passenger marketing and network development,

with this thesis contributing to the latter. In the concession agreement, a fixed cost per time table hour unit is set. Colleagues from the development department, as well as a colleague from concession management were invited.

- **Municipality:** The city of Groningen lies in the *Municipality of Groningen*. Although Q-link 3 and 4 crosses other municipalities too, Groningen is the only one invited, since the majority of both lines is within the municipality. Furthermore, the most driving time variability happens within the city of Groningen, suggesting that more reliable infrastructure is needed there most. An employer is invited that decide on the accessibility strategy and infrastructural decisions. Unfortunately, the employer was not able to attend.
- **Province:** The lines mainly serve the *Province of Groningen* and utilize multiple provincial roads. Besides these infrastructural role, the invited employee also has a say in the accessibility policy direction of the greater region.
- **Passenger representative group:** The *Consumer Platform Groningen and Drenthe* stands up for the interests of the PT user in the provinces of the concession. Two employees are invited to protect these interest during the focus group rounds, of which one was present.
- **Transit operator:** Since 2010, *Qbuzz* runs the bus service operations, agreed upon in the concession Groningen Drenthe. *Qbuzz* is expected to run the legally binding trips in this contract. Multiple employees are invited that can debate on the timetable feasibility changes, caused by specific holding strategies.

During the focus group section, the relations between stakeholders became more apparent. Acknowledging the existence of interaction between all elements, an effort is made to visualize the complexity of relationship between the actors in Figure 6.5. The relational diagram helps explaining the observation in the sections below. An important disclaimer is that all actors involved strive to deliver good quality public transit and the diagram is simplified and shows only the relations within the scope of the thesis, i.e. holding has a prominent place in the diagram, while in reality it is just tool in the wider toolbox of network design as described by Figure 1.3.

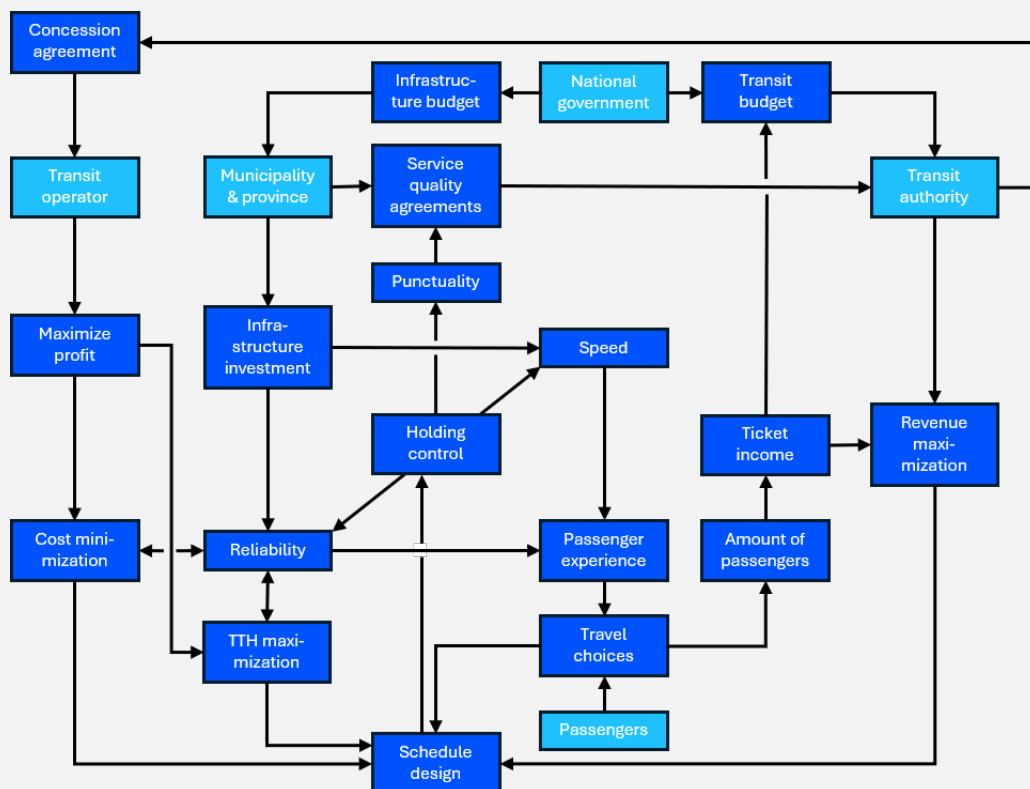


Figure 6.5: The influence of speed and reliability on the inter-stakeholder relationships

Common practice is that a transit operator is accountable for both the operational cost and the

ticket revenues. However, unique to the concession agreement in this case study, the transit authority has the responsibility of maximizing revenues. The concession agreement sets constraints to the transit operator, who in its turn tries to maximize profit by optimizing a cost minimization function on crew and fleet assignment. Furthermore, the operator can maximize profit by agreeing on more timetable hours (TTH). The passenger (and their representative group) strive for the best passenger experience, ultimately leading to passengers satisfyingly choosing for transit as a travel option. The municipality in province have a stake in the transit authority, to which they outsource all regulatory and operational responsibilities. This political construction is controlled via the service quality agreements. Although the national government was not part of the focus group, it is included, because the monetarily resources originate there. For example, the transit budget is awarded to the transit authority to partly subsidize the cost of the concession. Furthermore, the infrastructure budget is awarded to the municipality and province, so that municipal and provincial (transit) infrastructure can be build commissioned by the local political actors.

6.3.1. The punctuality paradox

Current practice in the concession Groningen Drenthe, is that the punctuality levels are leading in evaluating the reliability of the transit system. These levels are one of the elements of the overall service quality agreements made with the municipalities and provinces. Since the tightness of the schedule is dependent on historical trip times, vehicle-based punctuality metrics actually defend the status quo. When a lot of external congestion is happening on the road, the punctuality of vehicles decreases. Hence, more slack and layover time is build into the existing schedule by the operator, so that the punctuality is restored. This underlines the importance in investing in both rapidity and reliability. Then, the schedule can be set tight, while the punctuality is maintained at a sufficient level. These investments should be done by the actors operator, agency, municipality and province. Examples of investments that positively influence reliability, are dedicated bus lanes, modal filters and intersection priority. Ultimately, this paradox should be understood by citizens and their representatives, for supporting sustainable political decisions in the field of transportation.

In Figure 6.5, the punctuality paradox can be seen as follows: The municipality and province have agreed on service quality agreements, such as the level of punctuality. While the transit authority strives for a maximization of ticket revenues, it takes into account the punctuality agreements in their scheduling design decisions. Together with the stake of the operator a cost-minimization, the amount of slack and layover time is included as holding time. The reliability increases, while the speed decreases. The punctuality increases too, contributing to meeting the service quality agreements. The critical problem here is that speed is not part of the service quality agreements as an indicator. It is not specified how fast certain origin destinations are connected with each other.

6.3.2. Ridership spirals

Another complexity that came to light in the focus group, is how the ridership can end up in a positive or negative spiral. Positive spirals occur when a line is successful. The line gets crowded and more supply is needed to cope with the demand. Consequently, more trips are added to the schedule. The new trips again attract more passengers. If this feedback loop is sustained, transit can be constantly improved, leading to ever-growing ridership. For other lines, the situation looks more like a negative spiral. When the demand is low for low-frequency services, a line is relatively expensive to run. To an extend that lines have to be cut in length or frequency. The resulting service becomes less attractive and the amount of passengers will be ever-declining.

The spirals can be found back in Figure 6.5. If schedule design causes increased reliability and speed due to a "perfect" holding control strategy, the passenger experience is improved. More people will choose to use transit more often and for greater distances. The amount of passengers increases, so does the ticket income. This directly supports the goal of the transit authority, which is the maximization of revenues. Also, the transit budget increases, which makes better schedules possible, with for example, higher frequencies. This can further improve the passenger experience and a positive feedback loop can be sustained. The contrary can happen when a bad schedule design leads to less passengers, lower incomes, less budget and again, worse schedules.

Current practice on (over)crowding and induced demand at the public transit authority OV-bureau is the post-event adaption to overloaded buses. Each month, a sheet is exported from operator to authority in which all driver observations of full buses are automatically summarized. Tailor-made solutions are carried out, such as the deployment of higher capacity buses or a (temporal) increase in frequency. Thus, the supply mostly adapts to demand. There are exceptions. Policy-wise namely, the authority tries to create Transit Oriented Development (TOD) in new residential areas, by introducing a high quality bus line before the first people move in. For forecasts in the longer term, a combination of historical data, expert judgment and rules of thumb are used. For example, a frequency increase from 1 to 2 times per hour could yield a 60% increase in the amount of passengers.

6.3.3. Bridging the gap between short and long-headway services

A concern was raised by the passenger representative group and the province. An indicator that describes the speed and reliability of passengers could miss minority groups that travel on infrequent and sparsely loaded lines. This underlines the importance of not only aggregating a value for all individuals, but also looking at the minority groups. This can for example be done by a constraint or indicator that aims to improve the situation for the worst-off first, before optimizing the total passenger benefits.

The new RRR indicator may be beneficial for short-headway services, but it may not be implementable for long-headway services. Thus, supply-oriented metrics may still be used for lines that are scheduled a few times per hour, while a new indicator is promising for the Q-link network. The gap between short and long-headway services should always be taken into account, to prevent that a minority group is suffering too much for the greater good.

7

Discussion

In Chapter 3, a detailed overview was given of all the assumptions made in this thesis. In the current chapter, the main assumptions are discussed that contribute to the validity of the results and may be improved in future research. Table 7.1 gives an overview of these main assumptions made during this thesis. This chapter discusses the impact to the research outcomes and potential solutions to overcome bias caused by the assumptions. The order in Table 7.1 is based on the predicted magnitude of the bias. If adaptations can be made to overcome the assumptions, the potential solution directions are explained. Otherwise, a recommendation is made to future research. After the assumptions are discussed, the implications of the proposed indicator are discussed, after which the generalizability of methodology and results is elaborated on.

7.1. Dependency between demand and supply

Independent trip legs are assumed. In other words, the driving time between stops does not depend on the driving time of the entire trip. This could lead to an underestimation of external effects, such as weather conditions, crowding levels and driver behaviour. This could be partially solved by drawing from the total trip time instead of the leg distribution. If done so, it is recommended that trips are drawn from the specific time of day. In the morning peak, there is more variation in the inbound direction of Groningen, while in the afternoon, there is more variation in the direction of Groningen. This is true for most urban-rural regions. A possible way to control for all external effects, is machine learning. Machine learning is able to deal with an unlimited amount of variables, making it possible to test the dependency of both trip time and the time of day, as well as any other external effects.

In the thesis methodology there are no constraints with regards to vehicle capacity. An unlimited amount passengers may utilize a certain vehicle. In reality, the buses on Q-link 3 and 4 have a capacity of 100

section	assumption	impact	adaptable in methodology
7.1	deterministic path choice	overestimation RRR	×
7.1	vehicle capacity excluded	underestimation T^{wait} , T^{trans}	×
7.1	independent trip legs	underestimation external effects	—
7.1	punctual terminal departures	overestimation RRR	—
7.5	elasticity excluded	status quo bias	—
7.4	schedule-based only	limited strategies	✓
7.4	homogeneous control point percentiles	limited strategies	✓
7.2	uniform departure pattern passengers	overestimation T^{wait}	✓
7.6.2	deterministic transfer walking time	underestimation T^{trans}	—
7.3	$T^{\text{trans}} = T^{\text{wait}} + 15$	underestimation T^{trans}	—
7.3	$T^{\text{trans}} = 2 \cdot T^{\text{in-veh}}$	overestimation T^{wait}	✓

Table 7.1: The adaptability of the methodology to minimize bias caused by assumptions

passengers. At peak hour occasionally, buses exceed capacity leading to denied boarding, which is no different to other successful high-frequent systems. The current experimental setup underestimates the waiting time, since passengers can always board the first vehicle. An iterative approach could assign passengers for every time stage. Theoretically it would be possible to update the edges of the network-time graph every minute. This will significantly increase computational time, since the dynamic programming path finding cannot recognize the network, since the availability of buses is constantly changing due to crowding. A failure-to-board edge destroys dynamic programming path finding, since it makes the future state of the network unknown. Hence, future research on how to incorporate capacity constraints to a network including transfers, is recommended.

This also means passengers are not tracked during their trip. It is not saved which stops are served, and the passengers are assigned one by one. This poses problems for the bunching phenomenon, as showcased in Figure 3.2. As vehicles do not experience passenger friction, there is no internal cause of bunching. In other words, the vehicle dwell time does not depend on the amount of passengers on the stops. Bunching is happening within the simulation, but is only caused externally by the independent draws of the distribution function that includes the stochastic dwell time into the runtime. This may lead to an underestimation of holding strategy performance.

Fixed terminal departure times are another dependency-related problem, since consecutive trips that exist in one vehicle cycle may have spillover effects to one another. More operational information is needed from the operator with regards to charging time, driver shift breaks and timetabling to further analyse the sensitivity to the RRR indicator when the terminal departure times are variable. For this research, it was out of scope, but the same methodology can be applied for this sensitivity analysis. Note that the amount of passengers is dependent of the way transfers are synchronized, which makes the shifting of existing terminal departure times more complex.

Including the cycle time into the simulation would make the model unnecessarily complex, since it would require data on vehicle charging, fleet composition and driver shifts. The feasibility of the schedule is a responsibility of the operator. Still, this is the main reason to respect the 85th percentile for the trip time, to prevent spill-forward delay from previous trip at terminal turn around. The other assumption made regarding terminals is that the departure times are fixed.

All dependency-related constraints can be modelled by more sophisticated simulation software, such as BusMezzo (Cats, West, and Eliasson 2016), which can mimic vehicle behaviour. Especially with completely new line setups or new routes that lack historical data, BusMezzo can be used to make more realistic simulation of vehicle journeys. It is able to include (dynamic) crowding, bunching and the assignment of vehicles to cycles so that spill-over delay from an earlier trip can be modelled.

7.2. Deterministic passenger assignment

The deterministic passenger assignment, which is based on the shortest path at any minute in time, poses a bias. In reality, not all passengers will follow their shortest path. Path choice can be grounded by taste, gender, age, sex or underlying preferences. To cope with this issue, choice modelling can be used that stochastically assigns passengers to alternative paths, according to the taste parameters for each of the alternative paths. A detailed description on how a stochastic assignment could look like in a transit network, can be found in Cats, Koutsopoulos, et al. (2011).

The assignment is done based on the shortest path, existing of in-vehicle edges of simulated trips. The critical assumption is that the future state of the network is known by the passengers. In real-life, there are scenarios thinkable in which passengers make decisions due to the uncertainty of the future state. For example, a passenger does not know yet if a transfer is manageable. There exists a longer route without the need of transferring. A second reality should be modelled in which the passengers can make informed decisions about the predicted future state of the network, instead of the actual unknown future state. The same study looked into the availability of information about estimated future states Cats, Koutsopoulos, et al. (2011) and its relation to the path choice passengers make.

Within this thesis, it was only possible to estimate the egress waiting time, which is more representative for the morning rush hour, since people typically want to make it in time for their appointments. For a better estimate in the afternoon rush hour, the tap-in time would be better representative, since people

typically travel back home whenever work is done. The best data environment is when both tap-in and tap-out data is available, which can yield the exact transfer time within a study area (Bagherian et al. 2016).

The uniformity of passenger arrival is often assumed for short-headway networks. For long-headway services, in literature it is common to assume passenger arrivals are dependent on vehicle departures. There is a caveat, since it is not true that there is no demand when supply lacks. For this reason, in this thesis it is assumed that all passengers that do not transfer arrive uniformly, no matter the chosen frequency. For passengers that do transfer, there is an exception. These passengers can not choose when they arrive at a stop but are dependent of the previous trip. To cope with this uniformity, an additional data source is requested that gives information on the amount of transferring passengers for each stop. These passengers will be assigned onto the network at a specific time. When this is in place, it would support the models representativeness and shows the essence of the time-space simulation, because it is able to handle journeys with transfers.

7.3. Time elapsing parameters

The shortest path algorithm includes weights that make a distinction between time elapsing. A waiting time parameter of 2 is assumed. This means that the time spent waiting for a transit vehicle is experienced twice as long as the time spending in the vehicle. This value highly influences the results, since within the passenger assignment algorithm, this is the weight set to a waiting edge, also used for transferring. This leads to passengers choosing another path, because waiting time gets more expensive. In practice, it is common to use value 2, while in literature, value 1.5 is more common (Brinckerhoff 2013). Since the arbitrary pick of the value can lead to other results, it may cause bias. A recommendation to future research is a sensitivity analysis for specific waiting time parameters.

A transfer penalty of 15 minutes is awarded to every time a passenger changes vehicle. This prevents passengers from making illogical detours to skip the waiting time. Although theoretically this is possible, most passengers will change at the first opportunity lines meet to reduce the ticket fare and the chance of missing the transfer. Ultimately, a separate transfer parameter is set. In the current network-time graph however, it is hard to make a distinction between waiting edges and transfer edges. If the travel time is calculated by an utility function, waiting time and transfer time can have separate parameters. An example study that uses this is Cats, Koutsopoulos, et al. (2011). The advantage is that the duration of a transfer can now have a linear impact on travel time. A 15 minute transfer will then feel triple as long as a 5 minute transfer. In the thesis methodology, there is a marginally decreasing impact for a longer transfer, because the transfer time is determined by an equivalent of the waiting time, accompanied by the transfer penalty that is awarded at the start of the transfer.

7.4. Limited strategies

The methodology can be used for a variety of strategies. While it is possible to simulate the amount of control points and different control point percentiles over the course of a trip, only strategies are included that have a uniform percentile in place for all control points. This does not match with reality, in which for every control point, a unique segment runtime percentile is chosen to determine the strictness of the schedule. An example of a real-life schedule determination can be found in Figure 5.1.

Additionally, no headway-based holding strategies are tested. The current research is scoped to schedule-based holding. However, it is possible within the existing methodology. Due to time limitations and the lack of stakeholder motivation for implementation, headway-based holding is not (yet) tested in this thesis.

For the case study in this thesis, the implementation of headway-based holding is faced with difficulties in the short-term. London (United Kingdom) and Stockholm (Sweden) are two of the cities where headway-based holding is widely implemented (Cats 2012). Although it is technically possible in terms of driver-assistance systems that indicate the ideal holding time or speed that ensures equal intervals, the implementation of headway-based holding is faced with practical implications in The Netherlands. During the focus group, the operator explained these difficulties. Reason being that headway-based holding conflicts with collective labour agreements. In The Netherlands,

agreements are made on the fixation of shift brakes. Headway-based holding will make a scheduled break uncertain, because cycle times are always different due to the amount of holding. Furthermore, to efficiently assign crew to headway-based holding lines, drivers will have to stay on a specific line. This causes less variability in driven routes, which conflicts with the labour agreements, too. In The Netherlands, strikes are common that actively compete for better labour circumstances. This makes it unfavourable to incorporate headway-based holding. A final observation risen by the transit authority was that if the rest of a network is low frequent and drives according to schedule, the passenger arrival rates of transferring passengers will be clustered around the arrival of other vehicles at transfer stations. In that case, spacing departures does not make sense, even at higher frequencies.

Still, the impacts of headway-based holding are to be included within this thesis, in case a proper solution can be found for the above mentioned challenges in the future.

7.5. Elasticity

Elasticity has not played a role directly in this thesis. Elasticity underestimates the effects of a chosen strategy. The amount of passengers is dependent on the experienced travel time (Cats and West 2020; Nello-Deakin 2022; Van Heijningen 2025). Hence, when a strategy outperforms the current strategy, the origin destination data shows too little passengers. Vice versa, when a strategy performs worse, less passengers can be expected. In current setting, it is hard to estimate the performance of an elastic network, since the indicator is generalized. If more passengers use public transit, this would increase the generalized travel time, while the network performs better. A recommendation to future research is to investigate in indicators that can cope with a grow in passengers. For example, an iterative assignment model can aim for a travel time equilibrium, while also respecting the dependency between demand and supply, excluded in this thesis.

The change in RRR for current demand only covers existing passengers. When travel times change, induced demand in terms of journey frequency, journey length and the amount of travellers can be expected. Conversely, disappearing demand can be expected when travel times get longer (Nello-Deakin 2020; Van Heijningen 2025). In light of this induced demand phenomenon, it is suggested that travel time changes are not to be seen as stand-alone variables. A recommendation to future research is to calculate the elasticity effects associated with a change in demand-oriented indicators. The indicator of RRR in this thesis makes it possible to estimate the change in passengers following a timetable change. This is possible due to the exact nature of the performance indicator. While this can yield raw estimates on the growth or decline of passenger numbers, more sophisticated methods are possible, too.

Alongside the pure change in travel time, another suggested direction to future research is the travel time ratio between transit and car (*verplaatsingstijdfactor*). The ratio expresses the competitive position of transit, compared to the car. A ratio smaller than 1.0 indicates transit being faster than the car. According to Salonen and Toivonen (2013), multiple levels of detail exist for the determination of travel time within a network (in essence measuring the accessibility): simple, intermediate and advanced determination. For the transit model in this thesis, it was aimed to develop an intermediate model that was able to produce origin stop to destination stop travel time. For future research, it is recommended that as much data as possible on both modes is collected for an advanced comparison of door-to-door travel times. For transit, it is recommended to include the first and last mile. For the car, travel times have to be collected that include congestion and parking time (Salonen and Toivonen 2013).

Thus, to further improve the methodology in future research, more elements of a complete journey chain should be added, such as the access and egress walking time to reach the doorstep, instead of a transit stop being the origin or destination. Including the alternatives to transit to the network-time structure is recommended, such as walking and cycling. This will make it possible for passengers to change mode. Then, elasticity can be modelled more realistically, because people will decide to walk or cycle to their destination, once the service levels of transit are lacking.

Specific to the case study, the city of Groningen is an interesting play field when it comes to elasticity and modal shift, since the municipality aims to be less car-dependent, while the transit authority aims

in improving the quality of the transit network. Although these interest sometimes conflict, this is an interesting setting for more sophisticated elasticity studies that are able to estimate modal shift more precisely. Recent related articles that could be combined into a combined approach are Cats and West (2020), who present a framework in which passengers are dynamically loaded onto the network. Iteratively, the passengers make a decision based on their experienced travel time. The iteration is repeated, until an equilibrium is found. This study can be supported by (Van Heijningen 2025; Nello-Deakin 2022), both providing evidence in disappearing traffic, occurring after interventions that made the car less attractive.

Supporting this line of research, Borsje (2026) performed a multi-regression analysis on all the service types in the concession Groningen-Drenthe. The dependent variable was the amount of passengers. The explanatory variables used for the regression were ticket price, fuel price (for explaining the attractiveness of personal combustion vehicles), amount of holidays, temperature, precipitation, and the amount of scheduled vehicle runs. For the Q-link network specifically, he found that the amount of scheduled trips positively influenced the amount of passengers. This proves that supply follows demand and when total journey times are decreased, induced demand can be expected.

The punctuality paradox in Chapter 6.3.1 can be seen as a form of elasticity. Not from the demand side (like the ridership spirals), but from the supply side. The province explained that if speed or reliability lack, schedules become infeasible. Cycle times or transfer connections cannot be sustained, because more slack, holding and layover time need to be implemented. This leads to lines being cut, or the frequency to be lowered. On the other hand, when speed and reliability are improved by for example an infrastructure upgrade, the opposite is possible. Tighter schedules and longer routes enable for higher frequent service with more destinations for passengers. This ultimately improves the passenger experience and can attract new passengers.

7.6. Equity

As Chapter 3.4.1 pointed out, an utilitarianism standpoint was chosen as equity principle. Although this gives insight in the utility of the greater good, minorities are easily ignored. This section discusses how other equity principles can be adopted in the methodology to minimize exclusion. This to cope with different types of exclusion. Each of these types are elaborated on how they are related to this thesis, as well as what is suggested to decrease exclusion in future research. The relevance to this thesis is that the equity principle shapes the results. A 'dominating' or 'best performing' strategy may be derived from thesis, but this does not mean it is the best solution for passengers that are excluded from the analysis.

7.6.1. Alternative equity principles

To further minimize exclusion, different equity standpoints can be chosen that can be represented by an indicator. Each standpoint has its strengths and weaknesses (Dijkstra 2025).

- **Utilitarianism** is commonly used in transport related research. For example, it is found back in the utility function, which is used to calculate an individuals' disutility associated with transport. An example study that specifically uses this in determining the stochastic routes in a transit network is Cats, Koutsopoulos, et al. (2011). Later in the article, the travel time is summed over all individuals for an aggregated indicator. While this means all passengers are taken into account, it is hard to trace back the gains or losses of individual passengers in an aggregated indicator.
- **Egalitarianism** strives for equal access and fairness in transportation. It prioritizes accessibility for disadvantaged groups. To include egalitarianism in the methodology, a Gini-coefficient can be used, which aims to equalize an indicator over the population. Šfiligoj et al. (2025) demonstrate how the Gini-coefficient can be integrated with graph theory. In their access graph, the Gini-coefficient is used to give direction to transit policies aiming to equalize the number of accessible locations for everyone.
- **Libertarianism** believes the market can find the optimum transport policies. However, smaller population cores do not have a lot of demand and it is difficult to cover the cost with ticket income, especially in off-peak hours. A long-term political wish is present that aims to create sustainable

transportation. Only optimizing in the economical short-term situation risks ending up in a negative passenger spiral as described in Chapter 6.3.2

- **Sufficientarianism** focuses on ensuring that everyone has a minimum acceptable level of transport access. The RRR indicator makes it possible to derive the gains or losses associated at the node level. The worst-off passengers can be determined. A minimum threshold can be set, which every individual has to exceed. For example, all passengers must be able to take a bus within 30 minutes, or need to arrive at their destination within one hour. For future research, it is recommended that the exact location of the winners and losers is determined, so that geographical exclusion can be minimized.

The passenger representative group in the focus group underlined the importance of listening to minorities. The equity principle is closely tied with elasticity, because if a minor group is not to be seen by an aggregating indicator, it does not mean the group is actually a minority. It can also be that the group is a minority, because of a bad service level. For example, there is invisible demand for a certain transfer that is now used to a minimal extend. If the transfer is better synchronized, induced demand can occur, because a larger group can adapt their travel choices. Then, the group that used to be a minority, grows into a majority. The same counts for direct long-distance passengers. If all the holding time adds up to the in-vehicle time, the resulting attractiveness of these type of journeys is low. Consequently, this results in a minority of passengers travelling on the journey type.

7.6.2. Exclusion types

Luz and Portugal (2021) and Bruno et al. (2024) show how the types of exclusion as a consequence of the chosen equity principle can be conceptualized. Only the concepts from the transportation related social exclusion framework that are directly related to the thesis methodology are discussed.

- **Geographical exclusion** occurs when residence is too far from or not connected to the transportation system (Church et al. 2000). This thesis does not take into account geographical exclusion, since the existing demand is taken as a given. Hence, people that are now not connected, do not have any influence on the indicator. If geographical exclusion is to be measured, it is recommended to use a multi-modal agent-based simulation. As an example, BusMezzo can incorporate walking time that connects demand centroids with the transit system (Cats, Koutsopoulos, et al. 2011). Examples of other agent-based models are the Land-Use and Transport Interaction (LUTI) model (Willumsen et al. 2001), the Urban Dynamics Educational Simulator (UDES) model (Correia 2018) and The Multi-Agent Transport Simulation (MATSim) model (Ballo et al. 2024).
- **Physical exclusion** happens when transport systems or the built environment may impose physical and cognitive barriers to individuals Luz and Portugal (2021). Within this thesis, this is relevant for the assumption of deterministic walking times for a successful transfer. 2 minutes are assumed sufficient for transferring between train and bus and vice versa. However, elderly or people with disabilities may not be mobile enough to reach the second vehicle within 2 minutes. This could overestimate the network presentations, since in reality some people will not be able to make transfers in time. If more time is assumed, there can be an underestimation of network presentations. It is suggested to stochastically model the walking times needed during a transfer, in order to cope with these exclusivity limitations.
- **Time-based exclusion** contains excluded passengers that cannot use the transport system due to a lack of speed or temporal availability (Church et al. 2000). In the case presented by this thesis, time-based exclusion appears with the small fraction of passengers that is a long-distance through passenger and experiences holding for the greater good.
- **Fear-based exclusion** prevents people from using the transport systems due to concerns about safety (Church et al. 2000). The schedule tightness influences the probability of successfully transferring. When this happens at off-peak hours, transfer stations can be abandoned and less lit. In these periods, passengers that do not feel safe do need to be considered in final timetable choices. To incorporate this into the existing methodology, the transfer penalty or the wait/transfer edges can be equipped with a larger weight for different hours of the day or specific stops.

- **Informational exclusion** happens when a lack of knowledge about the transport system prevents people from using it. Not all types of passengers can be informed about real-time information. Especially with tighter schedules or with headway-based holding, it is suggested to equip all stops where passengers alight with dynamic passenger information systems.

The operator and passenger representative organization pointed out that with more heterogeneity in the schedule, people may be excluded that plan their journey ahead or are heavily dependent on a planned transfer. If these feel like the transfer may be missed due to vehicle unpunctuality, they could be spooked in using transit.

Concluding equity and exclusion, they shape the results but can be coped with in future research. The analysis in this thesis had an utilitarianism standpoint, in which all passenger trips are considered equally important, no matter the length of the trip. An accessibility analysis that strives for egalitarianism or sufficientarianism could provide a fairer way of assessing the quality of a timetable. Furthermore, since the current demand is taken as a given, people that are geographically excluded are not part of the analysis. Other people that the analysis in this thesis misses are passengers that are not able to transfer as fast as others, experience fear when transfers are missed at night, or experience difficulties in understanding service that is less tied to a punctual timetable. These exclusion types were acknowledged in the focus group meeting.

7.7. Practical implications of the RRR indicator

This section zooms in on the implications of integrating the RRR indicator in network development or as a concession requirement. Overall, the RRR indicator is able to better describe the total journey quality of passengers. When only vehicle punctuality, or line regularity is taken into account, the transfer time is neglected that passengers have to connect to studied lines. Although this is beneficial for the passenger, it can have implications for other actors. The implications for all the actors are discussed one by one, after which RRR as a concession requirement is discussed.

- **Passenger:** first and foremost, the passenger benefits the most from the new indicator. The general experienced travel time is decreased, with less waiting time, transfer time and in-vehicle time. Remarkably, the indicator is still generalized, which means that some individuals will experience more travel time for the greater good. For example, the recommended tight schedule can lead to the most passengers being better off, while some individuals will be worse off.
- **Transit operator:** the transit operator is influenced most by a tighter schedule, as a result of the RRR indicator. First, with crew assignment. For example, if driver swaps happen along the route, crew assignment becomes a major challenge at the busy transfer stations. This can conflict with labour agreements. Alternative crew assignment schedules may have to be developed that prioritize driver swaps at the line terminals. This is closely related to the decision of the scheduled cycle time. If there is no difference across strategies in the scheduled cycle time, driver swaps at the terminals will be less problematic. When a new schedule is installed, it is recommended to monitor the changes in cycle time, before actually tightening the cycle time, too. Second, the fleet assignment is influenced too, since other assignment blocks have to be made that allows for using less vehicles or a higher frequency.

This is closely related to the expected income of the transit operator. The unused holding time that used to be in the schedule will be moved to the end of the cycle in the pilot phase. Whenever the cycle time is deemed excessive, it can be cut by the authority in order to reduce the cost associated with the TTH and spend it elsewhere. Consequently, the operator will receive less income than it was used to, if no agreements are made on the allocation of TTH. Therefore, it is recommended that all stakeholder wishes and expectations are clear before installing an alternative holding strategy.

- **Transit authority:** for the authority, a tighter schedule can result in lower costs for executing a timetable, while the passengers are better off. The savings can be used to boost the frequency on the line, or elsewhere in the network. A higher frequency means more passengers. More passengers mean more income, which contributes to the ridership spiral as explained in Chapter 6.3.2.

- **Province and municipality:** as a result of an improved RRR level, a region becomes more accessible by transit. This contributes to the sustainability goals by the province and municipalities. Furthermore, the dependence on the car is reduced by the increased attractiveness of transit. A well performing network induces a focus to transit related policies, such as more available budget that can be used for more TTHs and dedicated transit infrastructure. This as well contributes to the positive ridership spiral.

7.7.1. Incorporating a demand-oriented indicator as concession requirement

Finally, the focus group discussed the opportunity to include a demand-oriented indicator in the concession agreements. Currently, only vehicle punctuality and the transfer between vehicles is monitored. The indicators function as a general reliability performance indicator of the network. Although there is a need to better describe the needs of passengers with a combined reliability and speed indicator, it poses challenges for including it in the concession agreements.

First, a demand-oriented indicator for reliability and speed would only be possible when it can be calculated without the need of large assumptions. For example, a full origin-destination matrix should be available with both tap-in and tap-out time of all vehicle trips used. At the moment, this is not possible due to privacy agreements on tracking passengers. Second, the authority noted that the response of an operator should be forecast, to prevent negative implications that were not foreseen by replacing an existing indicator. Finally, the question was raised on how responsible an operator is, due to the large dependency on infrastructure. Concluding, a demand-oriented combined reliability and speed indicator is better suited as indicator for future development on a network-wide level, rather than a punishable concession rule.

Although the stakeholders agreed that a full passenger-oriented concession requirement is not yet implementable, the transit authority wants to head into that direction. This preference was discussed during the focus group. One idea that came to light was the quantification of successful transfers, in order to penalize for missed transfers. Technically, such an indicator is possible if vehicle-based AVL data of connecting trips is analysed. If the design of such an indicator is further researched, the focus group agreed that ultimately, the indicator should be weight according to the amount of passengers using the transfer. This ensures that the operator will focus on guaranteeing important transfers first, before improving the transfers that are barely used. Ultimately, the total transfer time is leading, prioritizing transfers that are used often and face the risk of missing a low-frequent service.

7.8. Research generalizability

Generalizability tells to what extent the methodology and results of the research are valid for cases other than Groningen. This section discusses how generalizable this thesis is. The methodology as presented in Chapter 3 is made general so that it can be replicated for other case studies too.

Although the results are specific to the network its supply and demand in the case study, general lessons can be learned. If other studies in future research share similarities in terms of network characteristics. It is recommended to look into the passenger arrival distribution at major transfer stations, since they can significantly influence the results. This thesis finds that when a large group of passengers is connected to the cordon with low frequency lines, it is important to take into account the transfer time, as it is a substantial part of the full journey.

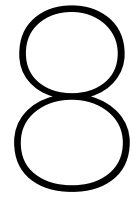
Similar transit networks can expect similar results. This thesis studied two branched diametric bus lines in a mid-sized Dutch city with one central station facilitating transfers between bus and train. Furthermore, multiple connections to other bus lines are possible at minor transfer point along the studied lines. Cities that lie in a metropolitan tissue might see other results, because the network typology is fundamentally different with multiple routes possible, instead of just one central transfer hub. Nevertheless, in The Netherlands, most peripheral cities outside of the *Randstad* consist of one central train station to which the majority of bus network connects to. For these occasions, the results are more directly generalizable.

If the transit network topology is different to the one in the case study, the results will be different. Still,

this thesis provides a methodological framework that is suitable for any network type. In methodological terms, each transit network can be represented as a network-time graph. In former studies, stand-alone networks are studied, without the inclusion of transfers to a cordon area. Hence, a lot of transfer time is neglected upon entering the study area, which this thesis aims to include. One remark is that the growth of the case study area and the associated cordon will increase computational time. All types of public transport can be incorporated. It is not suitable for modelling personal vehicles only, since the time component is not that relevant. A driver can depart any time and would lead to a massive graph. In that case, a two-dimensional graph is more suitable.

The availability of data played an important role in this study. Still, if one of the datasets used in this study is missing, the methodology can still be used. For the successful drawing of a network-time graph, stochastic vehicle runs were estimated. If the data on vehicle runs is missing, a probability density function can be used. Furthermore, a deterministic schedule that has to be optimized can also be drawn in a network-time graph, with. Furthermore, specific insights into transfer stations can be provided by the network-time graph, as visualized in Figure 3.6. It can easily be seen whether a transfer will be feasible or not. For determining the number passengers that use a path, an historical dataset is used for tap-out data only. As well as other data types, an educated guess can be made on the amount of passengers between each stop.

The methodology used in this thesis proves useful for including transfer interactions to other lines in a transit network. The tooling developed is able to quantify the generalized travel time for different timetable design alternatives. This thesis dives into the subject of holding patterns, but the tooling is suitable for analysing other transit networks elements too, such as line (re)routing, frequency and infrastructure upgrades.



Conclusion

8.1. Answers to the research questions

The main research question was:

What is the influence of including transfers in assessing the quality of the timetable as experienced by passengers?

To be able to answer the main question, the sub research questions were formulated as follows:

A) *How can the speed and reliability of a transit timetable be measured if transfers are included?*

This thesis explored opportunities in taking into account transfers for evaluating a timetable. It proposed a methodology that is able to determine the speed and reliability of direct, through and transferring passengers. The methodology used for this study is a simulation tool that reproduces days of service. The simulated vehicle trips, which are based on historical automated vehicle location (AVL) data, are drawn in a novel network-time graph. After, the passengers are assigned according to their shortest path. The demand levels are based on historical smart card data in twofold. First, the smart card data resulted in an origin-destination dataset from within the case study area. Second, the smart card data revealed the number of transfers to other operators, which this thesis used to determine the transfer time before or after having visited the studied network. Dynamic Programming aids in efficiently calculating the shortest paths for every origin-destination pair, for every minute.

A demand-oriented indicator is introduced that combines the reliability and speed as experienced by the passenger, ultimately describing the performance of the network. To include both travel time and reliability into one metric, a new network-wide metric was introduced. RRR aims to capture the *Rider-based Rapidness and Reliability*, by selecting the 80th percentile of passenger-specific journey time over multiple simulation runs. This way, passengers do not only take into account their average travel time, but also buffer some additional time into their journey to cope with expected delays.

B) *What is the effect of including transfers on the performance of holding strategies?*

The influence of including transfers is that the traditional assumption of a uniform arrival distribution is partly replaced by clustered arrivals at transfer stations. Hence, the line regularity becomes less important for the total experienced travel time within a transfer network. As a consequence, the inclusion of transfers had a negative influence on the performance of holding strategies. During this thesis, three different demand data sources were used to determine the total journey time. First, without transfers. Second, with a large group of assumed transfers. Third, real data on the amount of transfers was used. This thesis found that a large amount of transferring passengers favours holding strategies that consist of less scheduled slack. As the transfer real data of the case study consisted of a variation of the amount of transfers over the course of a day, the best slack allocation was dependent of the hour of day.

Results show that schedule-based holding is able to improve the network performance. The schedule tightness influences the effectiveness of holding control interventions. Depending on the chosen

tightness, the total passenger travel time in the network can vary by a maximum of 3.5 per cent. Note that this value is specific to the case study used and is dependent on multiple factors, such as the time of day and the location of control points. In general, schedule-based holding yields better results with a higher percentile to determine the departure time at control points, if transfers are not taken into account. The difference in hourly results seems dependent on the trip time variation. Hence, during peak hours, holding results in better generalized travel time, compared to off-peak hours.

Everything held equal, but now with transfer included, it yields different outcomes. Results show that the now included transfer time makes holding less favourable, compared to when a uniform arrival pattern was assumed. This suggests that the time at which passengers arrive or depart at a transfer stop, influences how well a timetable is measured. An important remark, is that this is dependent on the amount of transferring passengers, the frequency of the connecting services and to what extent the timetable is synchronized.

C) What are the perspectives of stakeholders on a timetable quality indicator that assesses passenger journeys including transfers?

The verification of the methodology and the validation of the results is aided by a stakeholder focus group session. In the session, the relation between the stakeholders main objectives became clear. A narrative was agreed upon: for high-frequent lines, a better indicator is needed than solely vehicle punctuality that is more adequately able to describe the reliability and speed as experienced by passengers.

Stakeholders acknowledge that transfer time is important and ultimately there is an indicator that can cope with complete journeys. Such an indicator is difficult to implement. It was agreed on that it is too complex for a concession requirement. It is better suited as a guide towards timetable development than a strict concession regulation. To move into the passenger-oriented direction, there was a consensus that transfers should be more carefully monitored. For this, more attention should be paid to the probability of successful transfers, as well as guaranteeing transfers that are heavily used.

8.2. Scientific recommendations

The main recommendation of this thesis is to not neglect passengers' transfer time, as it influences the results of a timetable related research question. Specifically, incorporating all transfers that connect a study area with a cordon is recommended, because it can yield different results for the performance of holding control strategies. The root cause of the different outcomes between including and excluding transfers is the assumed passenger arrival pattern at transfer stops. While most holding-related articles assume an uniform arrival distribution, which is totally fair for intermediate stops, it poses a fundamental problem for transfer stops. There, passenger arrivals are known to be centred on specific times. Simply, because the transferring passengers arrive or depart according to the set schedule of the connecting service. This also makes for the poor generalizability of the thesis results, since every network is different in terms of the contribution of transfers. For every network, the amount of passengers that make use of specific transfer connections is different.

While incorporating transfers from a cordon area is proven to be possible by using the network-time graph, the conducted simulation is basic. When holding is to be research specifically, the current approach is not recommended, because a congested assignment is hard within the methodology. Crowding and vehicle capacity constraints are not taken into account. Hence, vehicles and passengers behave independent, while in reality, they do not. The behaviour of passengers and the progression of transit vehicles interact with each other. If a vehicle exceeds capacity, passengers experience denied boarding and have to take the next vehicle. The other way around, when vehicles experience less passenger friction at stops, it can move faster through the network. A recommendation to future research is to combine a more sophisticated simulation tool, like BusMezzo, while connecting the case study with a cordon area.

For the assignment of passengers to the lines that connect to the cordon, this thesis manipulated the origin-destination demand data for journeys to begin or end at the cordon. In general, further research is needed on how to integrate transfers into agent-based models, which also enables the calculation of elasticity effects. While this thesis aims to model passengers as agents, it can still be improved

in terms of tracking passengers. With an iterative approach, the tracking of passengers would make it possible to form proper capacity, bunching and elasticity constraints. When these improvements are made, a transit system can be analysed over time, instead of only quantifying the immediate as conducted in this thesis. As a consequence, the elasticity effects can be estimated, predicting a growth or decline in passenger demand. This thesis proposes a novel graph representation of a transit network. The network-time graph is based on simulated or historical vehicle runs. The graph can be used to determine the shortest path at any moment in time. For the efficient calculation of these shortest paths, dynamic programming is used. This method helps quantifying the total journey time of passengers in a transit network. This thesis demonstrates the functionalities of the network-time graph with a real-life network to assess different holding control strategies.

The methodology used in this thesis proves useful for including transfer interactions to other lines in a transit network. The tooling developed is able to quantify the generalized travel time for different timetable design alternatives. This thesis dives into the subject of holding patterns, but the tooling is suitable for analysing other transit networks elements too, such as line (re)routing, frequency and infrastructure upgrades.

8.3. Practical recommendations

It is recommended that policy makers move towards making more data-driven decisions that aim at increasing passenger satisfaction. Common practice is that the supply-based indicators are a leading principle in timetable related decisions. However, these indicators mostly benefit the operator, with less attention to the passengers. It also has downsides for the transit authority, municipality and province, as they strive for making public transport more attractive to both the existing, and the new passenger.

Although this thesis proposes an indicator to estimate the unobserved passenger travel time, more practical research is needed on the actual design of a demand-oriented indicator. For low-frequent systems, supply-based metrics may still be sufficient, as passengers are bound to a schedule more than in high-frequent networks. For high-frequency systems, it is recommended that more attention is paid to line regularity or network reliability. Furthermore, the full trip should be included in the indicator, because passengers move through the network independent of the type of transit vehicles. The indicator must be able to cover both the reliability and speed. It is recommended that such an indicator is used first to steer policy, and should be necessarily directly installed as a strict rule that guards the performance of a transit operator. An additional advantage of a combined demand-based indicator, is that it can quantify the coordination of transfers and the overall quality of a network-wide timetable.

Specific to the case study in this thesis, the inclusion of transfers into a demand-oriented indicator has led to recommendations in favour of less slack time into the schedule. The holding strategies were assessed first without transfers, after which they were included. Data on transfers was used to determine the transfer time, which replaced the uniform arrival distribution of passengers at transfer stations. With the help of the case study, this thesis provides empirical evidence that transfer time should not be neglected in making timetable-related decisions. If it is neglected, there is a risk that a holding strategy is chosen which on paper yields the best passenger experience, while in reality, passengers experience unobserved transfer time.

References

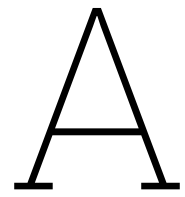
- Ait-Ali, A. (2024). "Exploring Measures to Monitor Passenger-Centric Punctuality in Public Transport". en. In: DOI: 10.13140/RG.2.2.20565.33769. URL: <https://www.researchgate.net/doi/10.13140/RG.2.2.20565.33769>.
- Bagherian, M., O. Cats, N. V. Oort, and M. D. Hickman (2016). "Measuring Passenger Travel Time Reliability Using Smart Card Data". In: URL: <https://api.semanticscholar.org/CorpusID:43136057>.
- Ballo, L., A. Sallard, L. Meyer de Freitas, and K. W. Axhausen (2024). "Is "small" infrastructure the next factory for accessibility?: Evaluating the regional accessibility effects of a cycling-centric transport policy in Zurich". en. In: DOI: 10.3929/ETHZ-B-000688987. URL: <http://hdl.handle.net/20.500.11850/688987>.
- Barabino, B., M. Di Francesco, and S. Mozzoni (2015). "Rethinking bus punctuality by integrating Automatic Vehicle Location data and passenger patterns". In: *Transportation Research Part A: Policy and Practice* 75, pp. 84–95. ISSN: 0965-8564. DOI: 10.1016/j.tra.2015.03.012. URL: <http://dx.doi.org/10.1016/j.tra.2015.03.012>.
- Borsje, R. (2026). *Hoe belangrijk zijn brandstofprijzen voor het OV?* URL: <https://dova.nu/nieuws/hoe-belangrijk-zijn-brandstofprijzen-voor-het-ov>.
- Brinckerhoff, P. (2013). *Transit Capacity and Quality of Service Manual, Third Edition*. Transportation Research Board. ISBN: 9780309283441. DOI: 10.17226/24766. URL: <http://dx.doi.org/10.17226/24766>.
- Bruno, M., M. Kouwenberg, and N. van Oort (2024). "Addressing transport related social exclusion through transportation policy: A novel evaluation method applied to the Amsterdam Transport Region". In: *Transportation Research Interdisciplinary Perspectives* 26, p. 101177. ISSN: 2590-1982. DOI: 10.1016/j.trip.2024.101177. URL: <http://dx.doi.org/10.1016/j.trip.2024.101177>.
- Burghout, W. (2004). "A note on the number of replication runs in stochastic traffic simulation models". In.
- Cats, O. (2012). *A Field Test for Service Regularity Improvement "Tänk på Avståndet" (RETT2 final report)*.
- Cats, O., W. Burghout, T. Toledo, and H. N. Koutsopoulos (2010). "Mesoscopic Modeling of Bus Public Transportation". In: *Transportation Research Record: Journal of the Transportation Research Board* 2188.1, pp. 9–18. ISSN: 2169-4052. DOI: 10.3141/2188-02. URL: <http://dx.doi.org/10.3141/2188-02>.
- Cats, O., H. N. Koutsopoulos, W. Burghout, and T. Toledo (2011). "Effect of Real-Time Transit Information on Dynamic Path Choice of Passengers". In: *Transportation Research Record: Journal of the Transportation Research Board* 2217.1, pp. 46–54. ISSN: 2169-4052. DOI: 10.3141/2217-06. URL: <http://dx.doi.org/10.3141/2217-06>.
- Cats, O., A. N. Larijani, H. N. Koutsopoulos, and W. Burghout (2011). "Impacts of Holding Control Strategies on Transit Performance: Bus Simulation Model Analysis". In: *Transportation Research Record: Journal of the Transportation Research Board* 2216.1, pp. 51–58. ISSN: 2169-4052. DOI: 10.3141/2216-06. URL: <http://dx.doi.org/10.3141/2216-06>.
- Cats, O., F. M. Rufi, and H. N. Koutsopoulos (2014). "Optimizing the number and location of time point stops". In: *Public Transport* 6.3, pp. 215–235. ISSN: 1613-7159. DOI: 10.1007/s12469-014-0092-1. URL: <http://dx.doi.org/10.1007/s12469-014-0092-1>.
- Cats, O. and J. West (2020). "Learning and Adaptation in Dynamic Transit Assignment Models for Congested Networks". In: *Transportation Research Record: Journal of the Transportation Research Board* 2674.1, pp. 113–124. ISSN: 2169-4052. DOI: 10.1177/0361198119900138. URL: <http://dx.doi.org/10.1177/0361198119900138>.
- Cats, O., J. West, and J. Eliasson (2016). "A dynamic stochastic model for evaluating congestion and crowding effects in transit systems". In: *Transportation Research Part B: Methodological* 89, pp. 43–

57. ISSN: 0191-2615. DOI: 10.1016/j.trb.2016.04.001. URL: <http://dx.doi.org/10.1016/j.trb.2016.04.001>.
- Ceder, A. (, Y. Hadas, M. McIvor, and A. Ang (2013). "Transfer Synchronization of Public Transport Networks". In: *Transportation Research Record: Journal of the Transportation Research Board* 2350.1, pp. 9–16. ISSN: 2169-4052. DOI: 10.3141/2350-02. URL: <http://dx.doi.org/10.3141/2350-02>.
- Church, A., M. Frost, and K. Sullivan (2000). "Transport and social exclusion in London". In: *Transport Policy* 7.3, pp. 195–205. ISSN: 0967-070X. DOI: 10.1016/S0967-070X(00)00024-x. URL: [http://dx.doi.org/10.1016/S0967-070X\(00\)00024-x](http://dx.doi.org/10.1016/S0967-070X(00)00024-x).
- Correia, G. H. d. A. (2018). "The Urban Dynamics Educational Simulator UDES tutorial for a tool to teach agent-based and Land Use and Transport LUT Interactions (model version 1.4)". In.
- Cramer, A., J. Cucarese, M. Tran, A. Lu, and A. Reddy (2009). "Performance Measurements on Mass Transit: Case Study of New York City Transit Authority". In: *Transportation Research Record: Journal of the Transportation Research Board* 2111.1, pp. 125–138. ISSN: 2169-4052. DOI: 10.3141/2111-15. URL: <http://dx.doi.org/10.3141/2111-15>.
- Delgado, F., J. C. Muñoz, R. Giesen, and A. Cipriano (2009). "Real-Time Control of Buses in a Transit Corridor Based on Vehicle Holding and Boarding Limits". In: *Transportation Research Record: Journal of the Transportation Research Board* 2090.1, pp. 59–67. ISSN: 2169-4052. DOI: 10.3141/2090-07. URL: <http://dx.doi.org/10.3141/2090-07>.
- Denti, E. and L. Burroni (2023). "Delay Indices for Train Punctuality". In: *Information* 14.5, p. 269. ISSN: 2078-2489. DOI: 10.3390/info14050269. URL: <http://dx.doi.org/10.3390/info14050269>.
- Desaulniers, G. and M. D. Hickman (2007). "Chapter 2 Public Transit". In: *Transportation*. Elsevier, pp. 69–127. ISBN: 9780444513465. DOI: 10.1016/S0927-0507(06)14002-5. URL: [http://dx.doi.org/10.1016/S0927-0507\(06\)14002-5](http://dx.doi.org/10.1016/S0927-0507(06)14002-5).
- Diab, E. I., M. G. Badami, and A. M. El-Genaidy (2015). "Bus Transit Service Reliability and Improvement Strategies: Integrating the Perspectives of Passengers and Transit Agencies in North America". In: *Transport Reviews* 35.3, pp. 292–328. ISSN: 1464-5327. DOI: 10.1080/01441647.2015.1005034. URL: <http://dx.doi.org/10.1080/01441647.2015.1005034>.
- Dijkstra, E. (2025). "Fairness in road pricing policies". MA thesis. TU Delft.
- Eberlein, X. J., N. H. Wilson, C. Barnhart, and D. Bernstein (1998). "The real-time deadheading problem in transit operations control". In: *Transportation Research Part B: Methodological* 32.2, pp. 77–100. ISSN: 0191-2615. DOI: 10.1016/S0191-2615(97)00013-1. URL: [http://dx.doi.org/10.1016/S0191-2615\(97\)00013-1](http://dx.doi.org/10.1016/S0191-2615(97)00013-1).
- Friedrich, M., M. Müller-Hannemann, R. Rückert, A. Schiewe, and A. Schöbel (2018). "Robustness as a Third Dimension for Evaluating Public Transport Plans". en. In: vol. 65. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 4:1–4:17. DOI: 10.4230/OASICS.ATMOS.2018.4. URL: <https://drops.dagstuhl.de/entities/document/10.4230/OASICS.ATMOS.2018.4>.
- Furth, P. G. and T. H. J. Muller (2007). "Service Reliability and Optimal Running Time Schedules". In: *Transportation Research Record: Journal of the Transportation Research Board* 2034.1, pp. 55–61. ISSN: 2169-4052. DOI: 10.3141/2034-07. URL: <http://dx.doi.org/10.3141/2034-07>.
- Furth, P. G. and T. H. J. Muller (2008). "Optimality conditions for public transport schedules with time-point holding". In: *Public Transport* 1.2, pp. 87–102. ISSN: 1613-7159. DOI: 10.1007/s12469-008-0002-5. URL: <http://dx.doi.org/10.1007/s12469-008-0002-5>.
- Furth, P. G. and T. H. J. Muller (2009). "Optimality conditions for public transport schedules with time-point holding". In: *Public Transport* 1, pp. 87–102. URL: <https://api.semanticscholar.org/CorpusID:38888487>.
- Garcia-Martinez, A., R. Cascajo, S. R. Jara-Diaz, S. Chowdhury, and A. Monzon (2018). "Transfer penalties in multimodal public transport networks". In: *Transportation Research Part A: Policy and Practice* 114, pp. 52–66. ISSN: 0965-8564. DOI: 10.1016/j.tra.2018.01.016. URL: <http://dx.doi.org/10.1016/j.tra.2018.01.016>.
- Gavriilidou, A. and O. Cats (2018). "Reconciling transfer synchronization and service regularity: real-time control strategies using passenger data". In: *Transportmetrica A: Transport Science* 15.2, pp. 215–243. ISSN: 2324-9943. DOI: 10.1080/23249935.2018.1458757. URL: <http://dx.doi.org/10.1080/23249935.2018.1458757>.
- Ghofrani, F., Q. He, R. M. Goverde, and X. Liu (2018). "Recent applications of big data analytics in railway transportation systems: A survey". In: *Transportation Research Part C: Emerging Technologies*

- 90, pp. 226–246. ISSN: 0968-090X. DOI: <https://doi.org/10.1016/j.trc.2018.03.010>. URL: <https://www.sciencedirect.com/science/article/pii/S0968090X18303395>.
- Google (2024). *GTFS Realtime Overview*. URL: <https://developers.google.com/transit/gtfs-realtime>.
- Hadas, Y. and A. (Ceder (2010). “Optimal coordination of public-transit vehicles using operational tactics examined by simulation”. In: *Transportation Research Part C: Emerging Technologies* 18.6, pp. 879–895. ISSN: 0968-090X. DOI: 10.1016/j.trc.2010.04.002. URL: <http://dx.doi.org/10.1016/j.trc.2010.04.002>.
- Hendren, P., J. Antos, Y. Carney, and R. Harcum (2015). “Transit Travel Time Reliability: Shifting the Focus from Vehicles to Customers”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2535.1, pp. 35–44. ISSN: 2169-4052. DOI: 10.3141/2535-04. URL: <http://dx.doi.org/10.3141/2535-04>.
- Hernández, D., J. C. Muñoz, R. Giesen, and F. Delgado (2015). “Analysis of real-time control strategies in a corridor with multiple bus services”. In: *Transportation Research Part B: Methodological* 78, pp. 83–105. ISSN: 0191-2615. DOI: 10.1016/j.trb.2015.04.011. URL: <http://dx.doi.org/10.1016/j.trb.2015.04.011>.
- Ibarra-Rojas, O., F. Delgado, R. Giesen, and J. Muñoz (2015). “Planning, operation, and control of bus transport systems: A literature review”. In: *Transportation Research Part B: Methodological* 77, pp. 38–75. ISSN: 0191-2615. DOI: 10.1016/j.trb.2015.03.002. URL: <http://dx.doi.org/10.1016/j.trb.2015.03.002>.
- Imran, A. M., N. Van Oort, O. Cats, S. Hoogendoorn, and P. van der Pot (2019). “Combining Speed Adjustment and Holding Control for Regularity-based Transit Operations”. In: *2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*. IEEE, pp. 1–8. DOI: 10.1109/mtits.2019.8883368. URL: <http://dx.doi.org/10.1109/MTITS.2019.8883368>.
- Joborn, M. and Z. Ranjbar (2022). “Understanding causes of unpunctual trains: Delay contribution and critical disturbances”. In: *Journal of Rail Transport Planning and Management* 23, p. 100339. ISSN: 2210-9706. DOI: 10.1016/j.jrtpm.2022.100339. URL: <http://dx.doi.org/10.1016/j.jrtpm.2022.100339>.
- Laskaris, G., O. Cats, E. Jenelius, M. Rinaldi, and F. Viti (2021). “A holding control strategy for diverging bus lines”. In: *Transportation Research Part C: Emerging Technologies* 126, p. 103087. ISSN: 0968-090X. DOI: 10.1016/j.trc.2021.103087. URL: <http://dx.doi.org/10.1016/j.trc.2021.103087>.
- Lee, K., Y. Jiang, A. (Ceder, J. Dauwels, R. Su, and O. A. Nielsen (2022). “Path-oriented synchronized transit scheduling using time-dependent data”. In: *Transportation Research Part C: Emerging Technologies* 136, p. 103505. ISSN: 0968-090X. DOI: 10.1016/j.trc.2021.103505. URL: <http://dx.doi.org/10.1016/j.trc.2021.103505>.
- Lee, Y., L.-S. Lu, M.-L. Wu, and D.-Y. Lin (2017). “Balance of efficiency and robustness in passenger railway timetables”. In: *Transportation Research Part B: Methodological* 97, pp. 142–156. ISSN: 0191-2615. DOI: 10.1016/j.trb.2016.12.004. URL: <http://dx.doi.org/10.1016/j.trb.2016.12.004>.
- Luz, G. and L. Portugal (2021). “Understanding transport-related social exclusion through the lens of capabilities approach”. In: *Transport Reviews* 42.4, pp. 503–525. ISSN: 1464-5327. DOI: 10.1080/01441647.2021.2005183. URL: <http://dx.doi.org/10.1080/01441647.2021.2005183>.
- Mazloumi, E., M. Mesbah, A. Ceder, S. Moridpour, and G. Currie (2012). “Efficient Transit Schedule Design of timing points: A comparison of Ant Colony and Genetic Algorithms”. In: *Transportation Research Part B: Methodological* 46.1, pp. 217–234. ISSN: 0191-2615. DOI: 10.1016/j.trb.2011.09.010. URL: <http://dx.doi.org/10.1016/j.trb.2011.09.010>.
- Mozzoni, S., M. D. Francesco, G. Maternini, and B. Barabino (2022). “Transfer’s monitoring in bus transit services by Automatic Vehicle Location data”. In: *Transportation Research Procedia* 60, pp. 402–409. ISSN: 2352-1465. DOI: 10.1016/j.trpro.2021.12.052. URL: <http://dx.doi.org/10.1016/j.trpro.2021.12.052>.
- Nello-Deakin, S. (2020). “Environmental determinants of cycling: Not seeing the forest for the trees?” In: *Journal of Transport Geography* 85, p. 102704. ISSN: 0966-6923. DOI: 10.1016/j.jtrangeo.2020.102704. URL: <http://dx.doi.org/10.1016/j.jtrangeo.2020.102704>.

- Nello-Deakin, S. (2022). "Exploring traffic evaporation: Findings from tactical urbanism interventions in Barcelona". In: *Case Studies on Transport Policy* 10.4, pp. 2430–2442. ISSN: 2213-624X. DOI: 10.1016/j.cstp.2022.11.003. URL: <http://dx.doi.org/10.1016/j.cstp.2022.11.003>.
- Nesheli, M. M. and A. (Ceder (2017). "Real-Time Public Transport Operations: Library of Control Strategies". In: *Transportation Research Record: Journal of the Transportation Research Board* 2647.1, pp. 26–32. ISSN: 2169-4052. DOI: 10.3141/2647-04. URL: <http://dx.doi.org/10.3141/2647-04>.
- Nguyen, S. and S. Pallottino (1988). "Equilibrium traffic assignment for large scale transit networks". In: *European Journal of Operational Research* 37.2, pp. 176–186. ISSN: 0377-2217. DOI: [https://doi.org/10.1016/0377-2217\(88\)90327-X](https://doi.org/10.1016/0377-2217(88)90327-X). URL: <https://www.sciencedirect.com/science/article/pii/037722178890327X>.
- Parbo, J., O. A. Nielsen, and C. G. Prato (2016). "Passenger Perspectives in Railway Timetabling: A Literature Review". In: *Transport Reviews* 36.4, pp. 500–526. ISSN: 1464-5327. DOI: 10.1080/01441647.2015.1113574. URL: <http://dx.doi.org/10.1080/01441647.2015.1113574>.
- Peek, G.-J. and M. Van Hagen (2002). "Creating Synergy In and Around Stations: Three Strategies for Adding Value". In: *Transportation Research Record: Journal of the Transportation Research Board* 1793.1, pp. 1–6. ISSN: 2169-4052. DOI: 10.3141/1793-01. URL: <http://dx.doi.org/10.3141/1793-01>.
- Rietveld, P. (2005). "Six reasons why supply-oriented indicators systematically overestimate service quality in public transport". In: *Transport Reviews* 25.3, pp. 319–328. ISSN: 1464-5327. DOI: 10.1080/0144164042000335814. URL: <http://dx.doi.org/10.1080/0144164042000335814>.
- Rudnicki, A. (1997). "Measures of Regularity and Punctuality in Public Transport Operation". In: *IFAC Proceedings Volumes* 30.8, pp. 661–666. ISSN: 1474-6670. DOI: 10.1016/S1474-6670(17)43896-1. URL: [http://dx.doi.org/10.1016/S1474-6670\(17\)43896-1](http://dx.doi.org/10.1016/S1474-6670(17)43896-1).
- Salonen, M. and T. Toivonen (2013). "Modelling travel time in urban networks: comparable measures for private car and public transport". In: *Journal of Transport Geography* 31, pp. 143–153. ISSN: 0966-6923. DOI: 10.1016/j.jtrangeo.2013.06.011. URL: <http://dx.doi.org/10.1016/j.jtrangeo.2013.06.011>.
- Schmöcker, J.-D., W. Sun, A. Fonzone, and R. Liu (2016). "Bus bunching along a corridor served by two lines". In: *Transportation Research Part B: Methodological* 93, pp. 300–317. ISSN: 0191-2615. DOI: 10.1016/j.trb.2016.07.005. URL: <http://dx.doi.org/10.1016/j.trb.2016.07.005>.
- Šfiligoj, T., A. Peperko, and O. Cats (2025). *Access graph: a novel graph representation of public transport networks for accessibility analysis*. DOI: 10.48550/ARXIV.2507.08361. URL: <https://arxiv.org/abs/2507.08361>.
- Uniman, D. L., J. Attanucci, R. G. Mishalani, and N. H. M. Wilson (2010). "Service Reliability Measurement Using Automated Fare Card Data: Application to the London Underground". In: *Transportation Research Record: Journal of the Transportation Research Board* 2143.1, pp. 92–99. ISSN: 2169-4052. DOI: 10.3141/2143-12. URL: <http://dx.doi.org/10.3141/2143-12>.
- Van der Bijl, R. and N. Van Oort (2024). *Betere Bus: BRT en andere vormen van hoogwaardige bus*. Aquire.
- Van der Werff, E., N. Van Oort, O. Cats, and S. Hoogendoorn (2019). "Robust Control for Regulating Frequent Bus Service: Supporting the Implementation of Headway-Based Holding Strategies". In: *Transportation Research Record: Journal of the Transportation Research Board* 2673.9, pp. 654–665. ISSN: 2169-4052. DOI: 10.1177/0361198119845893. URL: <http://dx.doi.org/10.1177/0361198119845893>.
- Van Heijningen, S. (2025). "Modelling changing travel behaviour in response to road capacity reductions". MA thesis. TU Delft.
- Van Oort, N. (2011). *Service reliability and urban public transport design*. Netherlands TRAIL Research School.
- Van Oort, N. (2016). "Incorporating enhanced service reliability of public transport in cost-benefit analyses". In: *Public Transport* 8.1, pp. 143–160. ISSN: 1613-7159. DOI: 10.1007/s12469-016-0121-3. URL: <http://dx.doi.org/10.1007/s12469-016-0121-3>.
- Effect of driving-time determination and holding points on reliability* (2009). Proceedings of the 11th International Conference on Advanced Systems for Public Transport, pp. 1–15.

- Van Oort, N., N. Wilson, and R. Van Nes (2010). "Reliability Improvement in Short Headway Transit Services: Schedule- and Headway-Based Holding Strategies". In: *Transportation Research Record: Journal of the Transportation Research Board* 2143, pp. 67–76. DOI: 10.3141/2143-09.
- Vuchic, V. R. (2005). *Urban transit*. Nashville, TN: John Wiley & Sons.
- Willumsen, L. G. et al. (2001). *Modelling transport*. Wiley-Blackwell.
- Wolters, G.-J. (2016). "Passenger punctuality: an analysis of the method of calculation and describing models". MA thesis. TU Delft.
- Wu, W., R. Liu, and W. Jin (2017). "Modelling bus bunching and holding control with vehicle overtaking and distributed passenger boarding behaviour". In: *Transportation Research Part B: Methodological* 104, pp. 175–197. ISSN: 0191-2615. DOI: 10.1016/j.trb.2017.06.019. URL: <http://dx.doi.org/10.1016/j.trb.2017.06.019>.
- Wu, W., R. Liu, and W. Jin (2018). "Integrating Bus Holding Control Strategies and Schedule Recovery: Simulation-Based Comparison and Recommendation". In: *Journal of Advanced Transportation* 2018.1, p. 9407801. DOI: <https://doi.org/10.1155/2018/9407801>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1155/2018/9407801>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1155/2018/9407801>.
- Yap, M., O. Cats, J. Törnquist Krasemann, N. van Oort, and S. Hoogendoorn (2022). "Quantification and control of disruption propagation in multi-level public transport networks". In: *International Journal of Transportation Science and Technology* 11.1, pp. 83–106. ISSN: 2046-0430. DOI: 10.1016/j.ijtst.2021.02.002. URL: <http://dx.doi.org/10.1016/j.ijtst.2021.02.002>.
- Zhao, J., M. Frumin, N. Wilson, and Z. Zhao (2013). "Unified estimator for excess journey time under heterogeneous passenger incidence behavior using smartcard data". In: *Transportation Research Part C: Emerging Technologies* 34, pp. 70–88. ISSN: 0968-090X. DOI: 10.1016/j.trc.2013.05.009. URL: <http://dx.doi.org/10.1016/j.trc.2013.05.009>.



Pseudo code

Algorithm 1	Bus network simulation and passenger assignment
1.1	<pre> for each origin destination pair do . for each minute do . Calculate demand per minute . Store demand in array . end for </pre>
1.2	<pre> end for Initialize plot to store holding strategy results for each holding strategy do . Set control points with associated percentile . Initialize an empty NetworkX directional graph for storing simulated bus trips . for each bus route do . Set departure headway from terminal . for each departure do . Create terminal node at current time . Define first control point with associated departure time, based on the strategy percentile . for each trip leg do . Update current time . Draw from leg distribution . Create destination node and connecting edge in NetworkX graph . if trip leg ends at control point do . while current time is lower than holding criterium do . Hold vehicle until the next minute . Update current time . end while . Set new control point with associated departure time . end if . end for . end for . end for </pre>
1.3	<pre> end for Initialize empty memory dictionary to store shortest paths for each origin-destination demand minute do . Define NetworkX graph starting and ending node . if the shortest path is already stored in memory then . Extract already calculated route with last line used . else . Initiate an empty list of candidate actions that arrive at the ending node . for each NetworkX neighboring node do . if candidate action is taking the bus do . Add trip time to candidate path . else . Add wait time (times the wait parameter) to another candidate path . end if . end for . Save shortest candidate path as shortest path in memory . end if . Multiply shortest path length with demand . end for . Calculate the RRR value as the 80th percentile of travel time . Bin every hour for hourly strategy performance . Add performance line to plot end for </pre>

Figure A.1: Pseudo code of the algorithm used

B

Network maps

This appendix shows multiple network maps on different scales so it is possible to see all transfer connections within the case study area and the cordon.

Qliner, Q-link en trein

Vanaf juli 2025

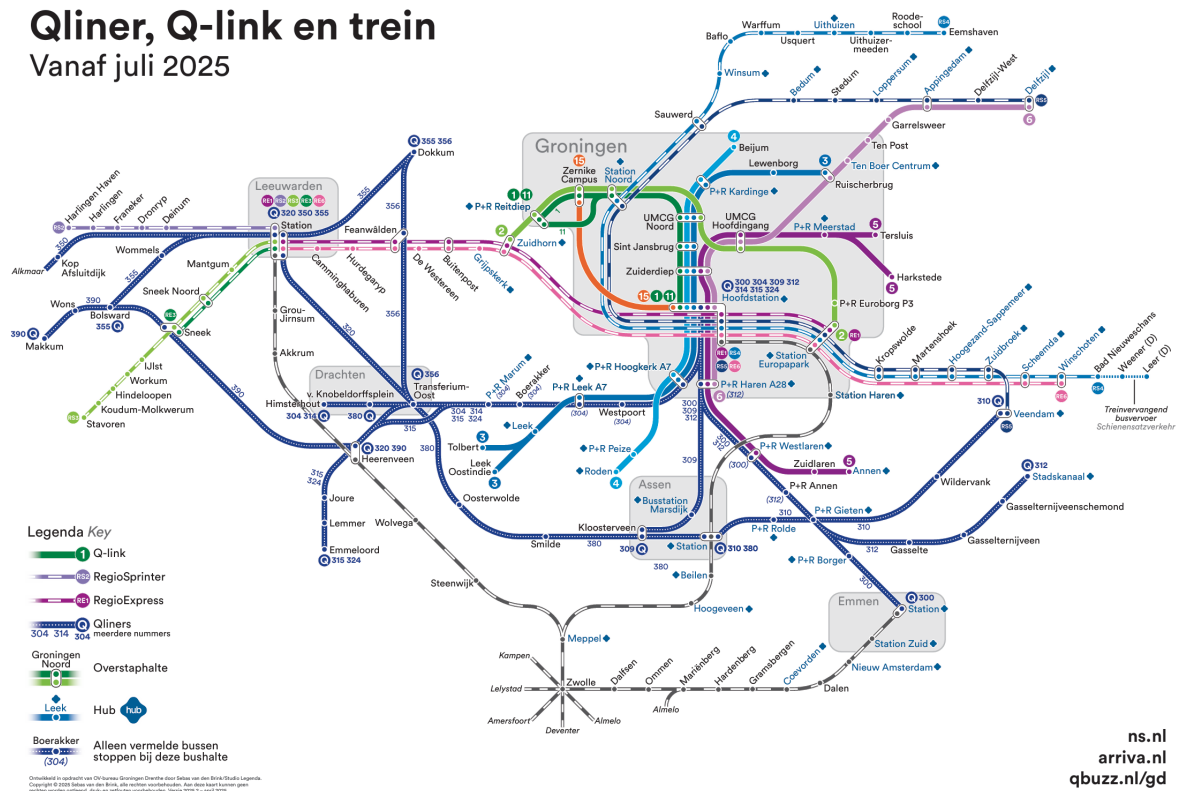
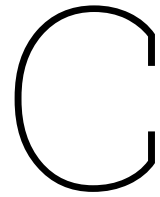


Figure B.1: Metro-style Q-link network and its transfer opportunities to train and long-distance bus services (source: Van den Brink, 2025)





Figure B.3: Metro-style network map from the city of Groningen (source: Van den Brink, 2025)



AI statement

For this thesis, AI is used for three tasks. First, it helped with mathematical conceptualization in earlier versions of this thesis. The simulation design process was led by available data, beginning with a simple simulation which is expanded step by step. In this process, AI is used to resolve errors during the coding process, enabling quicker debugging. Furthermore, in case where the language of Python lacked, AI was able to translate ideas into code. The prompts that were given took the form as the lines in Appendix A. This means that AI was given small tasks at the time, in order to guide the coding process. Furthermore, this enabled for a complete understanding of what was happening in every step in the simulation. *Chat GPT* was used for part of the kick-off document. After, it is replaced by *Mistral Le Chat*, since this is the only AI tool allowed by the OV-bureau.

D

Indicator significance

		strategy				
Arrival hour		SB-multi-85 (base)	SB-multi-65	SB-multi-50	SB-multi-35	SB-multi-0
6	s.d.	195.6	209.5	385.0	208.7	216.6
	mean	261082.8	261165.8	265124.7	263938.8	264997.6
	p-value	-	0.1023	0.0000	0.0000	0.0000
7	s.d.	515.5	528.8	768.1	601.9	593.6
	mean	996471.9	991846.9	992826.5	996876.2	983555.3
	p-value	-	0.0000	0.0000	0.0000	0.0000
8	s.d.	505.5	1688.0	1453.3	1446.2	1458.6
	mean	1465025.3	2129232.4	2118891.2	2134475.7	2126293.5
	p-value	-	0.0000	0.0000	0.0000	0.0000
9	s.d.	1125.1	674.4	1289.0	1191.0	912.9
	mean	1704491.7	1687725.5	1700663.8	1698744.2	1702350.3
	p-value	-	0.0000	0.1234	0.0456	0.3456
10	s.d.	613.5	1024.9	1215.8	587.8	1095.0
	mean	1295432.9	1280352.1	1270014.4	1269808.4	1273373.0
	p-value	-	0.0003	0.0000	0.0000	0.0123
11	s.d.	1003.2	364.5	821.8	413.8	511.9
	mean	1118263.3	1105807.6	1102940.4	1106497.2	1117097.0
	p-value	-	0.0000	0.0000	0.2345	0.0456
12	s.d.	771.6	594.9	626.7	428.8	694.6
	mean	1459511.8	1437632.3	1437741.1	1439446.9	1442314.7
	p-value	-	0.0000	0.0000	0.0012	0.0003
13	s.d.	515.5	422.3	609.3	622.8	604.5
	mean	1477851.3	1478761.9	1486237.6	1484396.6	1488276.1
	p-value	-	0.2345	0.0000	0.0000	0.0000
14	s.d.	505.5	517.1	419.4	567.9	670.5
	mean	1465025.3	1453036.1	1459616.2	1477678.0	1478067.6
	p-value	-	0.0001	0.0123	0.0000	0.0000
15	s.d.	1064.8	760.1	818.1	899.6	1301.1
	mean	1874502.5	1865055.3	1880245.8	1898939.1	1878802.0
	p-value	-	0.0002	0.0001	0.0000	0.2345
16	s.d.	972.6	550.3	653.1	867.8	906.6
	mean	1875184.8	1874884.0	1881744.1	1879057.3	1878692.3
	p-value	-	0.4567	0.0000	0.1234	0.0001
17	s.d.	1758.1	557.3	1264.8	974.1	916.4
	mean	1798624.9	1751736.2	1757940.9	1753951.8	1739686.8
	p-value	-	0.0000	0.0000	0.0000	0.0000
18	s.d.	377.8	464.2	407.5	390.4	358.9
	mean	1180926.4	1191077.1	1192203.1	1188400.3	1190423.6
	p-value	-	0.0001	0.0000	0.0012	0.0000
19	s.d.	277.4	254.5	553.4	180.6	378.5
	mean	792029.6	801091.9	795473.6	800132.8	806491.1
	p-value	-	0.0003	0.0456	0.0001	0.0000
20	s.d.	365.9	209.0	537.7	440.5	426.2
	mean	718642.5	725195.3	727390.6	728187.8	722617.9
	p-value	-	0.0002	0.0000	0.0000	0.0001
21	s.d.	299.4	359.5	301.3	354.8	461.3
	mean	619213.6	628078.8	616493.6	624214.3	623341.1
	p-value	-	0.0000	0.0000	0.0000	0.0002
22	s.d.	381.7	330.4	362.1	273.7	470.9
	mean	585451.2	601233.3	600251.6	592948.0	592538.8
	p-value	-	0.0000	0.0000	0.0000	0.0000
23	s.d.	259.3	296.7	314.2	259.3	328.2
	mean	460935.8	457487.4	461977.2	458220.2	453898.9
	p-value	-	0.0123	0.0003	0.0456	0.0000
24	s.d.	263.1	190.8	245.9	206.7	200.9
	mean	281043.2	281868.8	283165.0	285641.4	283195.2
	p-value	-	0.0456	0.0000	0.0000	0.0001
25	s.d.	143.1	180.4	169.0	334.2	354.3
	mean	244537.9	245637.1	244655.4	246971.7	245782.2
	p-value	-	0.0001	0.0000	0.0000	0.0001

Table D.1: Significance of the RRR indicator with the transfer assumption ($m=50$, $b=10$)

		strategy				
Arrival hour		SB-multi-85 (base)	SB-multi-65	SB-multi-50	SB-multi-35	SB-multi-0
6	s.d.	282.9	286.8	258.9	283.8	239.6
	mean	164074.0	161510.7	165718.3	162873.3	165922.5
	p-value	-	0.0000	0.0000	0.0000	0.0000
7	s.d.	941.1	445.7	602.1	778.0	614.9
	mean	660910.1	656697.4	653901.7	658332.0	661417.3
	p-value	-	0.0000	0.0000	0.0000	0.0164
8	s.d.	971.1	1223.5	1323.5	950.4	1646.0
	mean	1423424.3	1416751.3	1435496.2	143728.1	1434024.4
	p-value	-	0.0000	0.0000	0.0000	0.0000
9	s.d.	1382.6	955.9	830.9	1152.5	826.7
	mean	1125617.8	1124505.1	1124127.5	1134158.6	1140737.6
	p-value	-	0.0006	0.0000	0.0000	0.0000
10	s.d.	641.7	817.1	756.4	547.8	965.8
	mean	790252.8	789699.8	780350.6	782591.9	785301.5
	p-value	-	0.0050	0.0000	0.0000	0.0000
11	s.d.	658.0	494.8	440.2	490.6	657.4
	mean	674303.8	664462.3	66926.5	666902.1	667108.4
	p-value	-	0.0000	0.0000	0.0000	0.0000
12	s.d.	531.0	675.7	614.3	634.1	517.3
	mean	912998.6	912298.4	908759.0	914257.1	917962.2
	p-value	-	0.0000	0.0000	0.0000	0.0000
13	s.d.	587.9	433.0	955.4	537.7	377.1
	mean	947655.8	955254.5	953035.3	950999.7	954420.9
	p-value	-	0.0000	0.0000	0.0000	0.0000
14	s.d.	710.6	640.9	613.3	897.2	542.4
	mean	981400.6	977305.9	985302.8	975117.1	996272.3
	p-value	-	0.0000	0.0000	0.0000	0.0000
15	s.d.	899.5	1024.7	539.9	1082.8	1125.8
	mean	1231123.4	1226710.0	1217789.0	1229521.6	1231853.8
	p-value	-	0.0000	0.0000	0.0000	0.0074
16	s.d.	477.5	755.6	602.3	427.1	974.0
	mean	1225506.7	1240249.9	1227575.1	1232452.2	1239073.6
	p-value	-	0.0000	0.0000	0.0000	0.0000
17	s.d.	1413.8	813.9	812.2	1180.9	667.4
	mean	1164763.6	1170325.9	1183096.1	1183511.9	1187119.6
	p-value	-	0.0000	0.0000	0.0000	0.0000
18	s.d.	523.7	534.9	466.2	512.2	543.6
	mean	774262.1	7780712.0	779420.6	785631.8	785372.2
	p-value	-	0.0000	0.0000	0.0000	0.0000
19	s.d.	286.4	254.0	406.1	324.8	394.9
	mean	493312.3	502332.7	505099.3	505437.4	507693.2
	p-value	-	0.0000	0.0000	0.0000	0.0000
20	s.d.	403.9	542.0	378.3	401.4	447.3
	mean	420925.5	427169.3	434046.4	432882.0	433309.8
	p-value	-	0.0000	0.0000	0.0000	0.0000
21	s.d.	278.6	382.0	386.5	396.8	448.7
	mean	385579.3	383602.3	434046.4	389397.3	391960.4
	p-value	-	0.0000	0.0000	0.0000	0.0000
22	s.d.	357.3	176.9	378.3	412.1	491.6
	mean	354133.7	359360.8	359991.6	358735.3	359202.2
	p-value	-	0.0000	0.0000	0.0000	0.0000
23	s.d.	208.9	279.6	312.5	316.6	286.5
	mean	253244.9	249658.8	254519.2	253537.6	254092.7
	p-value	-	0.0000	0.0000	0.0000	0.0000
24	s.d.	178.0	126.5	206.6	205.4	152.2
	mean	145935.2	148327.0	147849.9	148262.4	149160.1
	p-value	-	0.0000	0.0000	0.0000	0.0000
25	s.d.	221.3	214.5	162.7	79.7	162.5
	mean	136419.2	138052.0	137346.2	136684.2	138938.7
	p-value	-	0.0000	0.0000	0.0000	0.0000

Table D.2: Significance of the RRR indicator with transfer data ($m=50$, $b=10$)

E

Bus-bus transfer numbers

	Transfer station	Current line	Future line	#pax
1	Hoofdstation	Q-liner 309	Q-link 3/4 (eastbound)	102
2	Hoofdstation	Q-liner 300	Q-link 3/4 (eastbound)	98
3	Hoofdstation	Q-link 3/4 (westbound)	Q-liner 309	93
4	Hoofdstation	Q-link 3/4 (westbound)	Q-liner 300	84
5	P+R Hoogkerk A7	Q-link 3/4 (eastbound)	Bus 10	71
6	Ruischerbrug	Q-link 6 (westbound)	Q-link 3/4 (westbound)	69
7	Verzetsstrijderslaan	Q-link 15 (southbound)	Q-link 3/4 (westbound)	67
8	UMCG Noord	Q-link 1 (southbound)	Q-link 3/4 (eastbound)	60
9	P+R Hoogkerk A7	Q-link 3/4 (eastbound)	Bus 18	58
10	Hoofdstation	Bus 50	Q-link 3/4 (eastbound)	52
11	Verzetstrijderslaan	Q-link 3/4 (eastbound)	Q-link 15 (northbound)	51
12	P+R Hoogkerk A7	Bus 10	Q-link 3/4 (westbound)	50
13	Hoofdstation	Bus 9 (northbound)	Q-link 3/4 (eastbound)	49
14	Hoofdstation	Q-link 5 (northbound)	Q-link 3/4 (eastbound)	47
15	Ruischerbrug	Q-link 3/4 (eastbound)	Q-link 6 (eastbound)	45
16	Hoofdstation	Q-link 3/4 (westbound)	Bus 50	45
17	Hoofdstation	Q-link 5 (northbound)	Q-link 3/4 (westbound)	43
18	UMCG Noord	Q-link 3/4 (westbound)	Q-link 2 (northbound)	43
19	UMCG Noord	Q-link 3/4 (westbound)	Q-link 1 (northbound)	43
20	Hoofdstation	Q-liner 314	Q-link 3/4 (eastbound)	41
21	Hoofdstation	Q-liner 315	Q-link 3/4 (eastbound)	41
22	UMCG Noord	Q-link 2 (southbound)	Q-link 3/4 (eastbound)	40
23	Hoofdstation	Bus 51	Q-link 3/4 (eastbound)	40
24	Hoofdstation	Q-link 3/4 (westbound)	Q-liner 312	39
25	Hoofdstation	Q-liner 312	Q-link 3/4 (eastbound)	39
26	Hoofdstation	Q-link 3/4 (westbound)	Bus 10 (southbound)	37
27	Hoofdstation	Q-link 3/4 (westbound)	Q-liner 314	37
28	Hoofdstation	Q-link 3/4 (westbound)	Q-link 5 (southbound)	37
29	Hoofdstation	Q-link 3/4 (eastbound)	Q-link 5 (southbound)	34
30	Hoofdstation	Q-link 3/4 (westbound)	Q-liner 315	32
31	Hoofdstation	Q-link 3/4 (westbound)	Bus 9 (southbound)	30
32	Hoofdstation	Q-link 3/4 (westbound)	Q-liner 304	30
33	Hoofdstation	Bus 19 (northbound)	Q-link 3/4 (eastbound)	28
34	Hoofdstation	Q-link 3/4 (westbound)	Bus 51	27
35	P+R Hoogkerk A7	Bus 18	Q-link 3/4 (westbound)	27
36	Hoofdstation	Bus 10 (northbound)	Q-link 3/4 (westbound)	25
37	Hoofdstation	Q-liner 309	Q-link 3/4 (westbound)	24
38	UMCG Noord	Q-link 3/4 (westbound)	Bus 171 (northbound)	24
39	UMCG Noord	Q-link 3/4 (westbound)	Bus 171 (southbound)	23
40	UMCG Noord	Q-link 3/4 (westbound)	Bus 2 (southbound)	22
41	Hoofdstation	Q-liner 300	Q-link 3/4 (westbound)	22
42	Hoofdstation	Bus 10 (northbound)	Q-link 3/4 (eastbound)	22
43	P+R Kardinge	Bus 65	Q-link 3/4 (eastbound)	21
44	Centrum Leek	Bus 83	Q-link 3/4 (eastbound)	21
45	Hoofdstation	Bus 178	Q-link 3/4 (westbound)	20
46	Hoofdstation	Q-link 3/4 (eastbound)	Q-liner 309	20
47	Hoofdstation	Q-link 3/4 (westbound)	Bus 7	17
48	Centrum Leek	Q-link 3/4 (westbound)	Bus 83	17
49	Hoofdstation	Q-link 3/4 (eastbound)	Bus 10 (southbound)	17
50	UMCG Noord	Q-link 2 (northbound)	Q-link 3/4 (eastbound)	17

Table E.1: Top 50 most used bus on bus transfers (daily average amount of passengers)