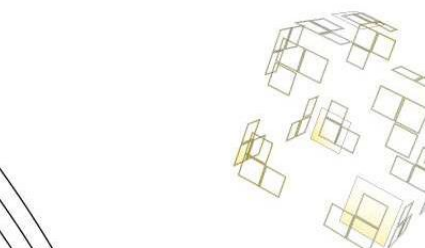


FACULTY OF ENGINEERING TECHNOLOGY (CTW)
TRANSPORTATION ENGINEERING & MANAGEMENT

HOV-scanner



Adding PT route choice and optimizing the
processing time



μ CONSULT

Ramon Peters

UNIVERSITEIT TWENTE.

HOV-scanner

Adding PT route choice and optimizing the processing time

Master thesis

Date: 2015-08-22

Author: ing. Ramon H. Peters
Contact: ramon.peters@gmail.com

Supervisor University of Twente: Prof. Dr. Ir. Eric van Berkum
Dr. Tom Thomas
Supervisor MuConsult: Dr. Edward Rosbergen

Preface

A few years ago I was involved in developing medical equipment and although I enjoyed my job I was considering to pick up a study in the field of traffic and transportation for a while. Three years ago I made a decision to quit my job as a mechanical engineer and I switched towards the study Transportation Engineering and Management. This thesis is the final assignment for the master program Civil Engineering and Management, specialization Transportation Engineering and Management at the University of Twente. It was an interesting and pleasant study and I am looking forward to start in this sector. Perhaps I can combine my interest on both mechanical and transportation engineering in the future.

This report describes a research on the implementation of parallel public transit routes in the HOV-scanner and it describes the approach used to reduce processing time. This research is conducted at MuConsult which is the developer of the HOV-scanner. The HOV-scanner is a tool used to analyze the effect of changes in a public transit network.

MuConsult offered me a lot of support during this period. They offered the opportunity to use their HOV-scanner, data, expertise and a pleasant study environment. I would like to thank MuConsult for that. Especially, I would like to thank Edward Rosbergen, Peter van Bekkum, Rinus Haaijer and Frans Blanker for their feedback and guidance. Furthermore, I would like to thank my supervisors from the UT Tom Thomas and Eric van Berkum for their engagement in our constructive discussion and their critical, but helpful, notes. I would like to thank Lissy La Paix Puello for helping out with Biogeme. I would like to thank the OV-bureau Groningen Drenthe for providing data and their interest. At last but not least I would like to thank my family and friends for their support and interest. Especially I would like to name Kevin, Inge, Marja, Sjors and Ronald because they provided a lot of feedback.

Summary

Motivation and research objective

The HOV-scanner can be used to indicate the effect on public transport (PT) share of changes in PT services, for example, a new PT line. The HOV-scanner is developed and used by MuConsult. The HOV-scanner in its current layout is facing some limitations:

1. In case of parallel PT connections all demand is assigned to the most attractive connection. It is likely to assume that the assignment is more nuanced in real life.
2. The HOV-scanner is used as a first indicator for the feasibility of a new PT connection, but using it for this purpose is complex due to the layout of the HOV-scanner. This results in an operation that is too labor-intensive.

In this study both topics are studied and the HOV-scanner is improved accordingly. As a result the research objective is: "To implement PT route choice into the HOV-scanner to simulate parallel connections more realistic. Furthermore, the work load should be reduced to make the HOV-scanner more suitable as a first indication regarding the feasibility of a new PT system." The objective of this research results in two research questions:

1. How should the HOV-scanner be designed to be able to model parallel PT connections?
2. How can the HOV-scanner be improved to reduce the work load?

The HOV-scanner

The HOV-scanner is a tool that uses origin/destination (OD) matrices and utility functions to calculate a modal split. The HOV-scanner is programmed in MATLAB. It is designed according to the four-stage transportation model. The total trip distribution is calculated and aggregated to match the desired zone size. Then the modal split is modeled for the reference situation by using a Multinomial Logit choice model. Errors are calibrated if any occurred. The changes in the PT network in combination with the sensitivity toward changes results in a new utility for PT. This new PT utility is used to forecast the distribution for the new situation. The HOV-scanner is modeled according to the schedule based approach.

Research approach

As part of the research the current HOV-scanner is analyzed and an introducing literature study has been performed. Thereafter a route set generation method for PT and a choice model are composed and elaborated.

The route set generation method should improve processing time and include all attractive PT routes within the study area. A constrain set is proposed with which the PT route set must comply. Programs and methods are analyzed and the constrained enumeration method is selected as being the most suitable. The constrained enumeration method and the constraints are implemented in a MATLAB program. The program is applied on the "Duin en Bollenstreek" area and resulted in a processing time reduction of a factor five. This is a satisfying improvement over the manual method especially since the HOV-scanner is extended from a single route from towards a route set.

In extension of the original HOV-scanner PT route choice is included and therefore the choice model is altered. A nested Logit model is proposed whereby the different PT routes are nested in the PT branch. To forecast the effect of a new PT scenario a four step approach is conducted:

1. Set up and apply the choice model for the reference situation.
2. Calibrate the outcome of the choice model for each OD relation and mode
3. Calculate the shift in utility for the PT system based upon information that is retrieved in the first step.
4. Forecast the use of the PT system for the scenario.

To model the PT utility it is proposed that in vehicle time, waiting time, transfer time, access an egress time, number of transfers, frequency and PT modes should be included. During the implementation on the "Duin en Bollenstreek" data the frequency and PT mode deviation is neglected due to modeling errors. Furthermore the generalized cost for PT is used to improve the model. The generalized cost for PT consists of the selected parameters and a weight ratio that is retrieved from literature.

Validation

The network changes implemented Marne are used for the validation. The check-in and check-out data of two comparable months in 2014 and 2015 are used. The validation is focused on three different aspects of the HOV-scanner:

1. The route set generated by the program.
2. The forecasted PT share per OD.
3. The route share distribution per OD.

The generated route set revealed two limitations: the detour criterion and the inclusion of unrealistic routes. For the study area a detour criterion of 30% seems to be too narrow and used routes were deleted. Excluding routes is considered to be more harmful than including too many routes. To overcome the deletion of used routes a detour threshold of 15 minutes is proposed. The possibility to walk between stops and transfer lines is added to make the route set more realistic. Nevertheless it allows for unwanted effects to occur. For example, a transfer is made from a direct connection towards another line which resulted in a longer travel time. This is assumed not to be realistic and must therefore be avoided.

The change in PT share can be predicted adequate if the difference in travel time between the reference and the scenario situation is large enough. If there are a reasonable number of trips as well the model can also forecast the direction and the magnitude of the effect. The result of the new model is similar to the results of the original model.

In general the direction of the route share distribution modeled is in line with the route share distribution from the data. Nevertheless exceptions were found; they can be related to small differences in travel time and other influences that are excluded by the model. The magnitude of the trip share distribution on the other hand is modeled much too low, this results in an incorrect prediction. The model fit can be improved by increasing the parameter for the travel time. Based on this validation case it seems to be beneficial to use separate parameters for mode and route choice.

Conclusion

The main objective for this research was to implement PT route choice with parallel connections in the HOV-scanner. The additional goal was to reduce the work load so that the program would be more suitable as a first indication regarding the feasibility of changes in PT services.

During this research parallel PT connections are included in the HOV-scanner nevertheless, the current setup is facing some limitations. The changes in model parameters and parameter values resulted in an improved model fit. Nevertheless on OD level there still are substantial differences in mode shares. The original HOV-scanner is used to forecast the shift in PT share. The validation showed that the shift in mode shares were forecasted quite well by the new HOV-scanner. A requirement to achieve good results is to have enough trips and the difference in travel time between the reference and the scenario should be large enough. The accuracy of the forecast in shift in mode share is identical to the original HOV-scanner. In addition to the original HOV-scanner PT route choice is included to estimate the shares for individual lines. The proposed model was not able to predict route shares correct. The validation case revealed that route shares were distributed far too evenly over the routes.

The work load is reduced by automating the route set generation process. This resulted in a significant time reduction. The used constrained enumeration method itself is working correct. However the validation case showed that the used constraints need some adjustment.

Recommendations:

- Study the effect of overlap inclusion approaches.
- Study the effect of frequencies on trip distribution within a route set.
- Conduct a research regarding the effect of more accurate input data.
- Adjust the route set program to avoid the exclusion of used routes.
- Exclude counter intuitive routes from the route set by improving the program.
- Conduct research on the topic of route choice parameter values to find out if the route choice can be modeled accurate.
- Analyze the results carefully if the model is applied on small differences in travel time or on a limited number of trips, because the result can be inaccurate.
- To make a comparison between the results of the HOV-scanner and a more complex model.

Table of contents

Preface.....	1
Summary	2
Motivation and research objective.....	2
The HOV-scanner.....	2
Research approach	2
Validation	3
Conclusion.....	3
Recommendations:	3
Table of contents.....	4
1. Introduction	6
1.1 Motivation	6
1.2 Introduction on the HOV-scanner	6
1.3 Reading guideline.....	6
2. Problem definition	7
3. Research objective and research question.....	8
4. The HOV-scanner and study area	8
4.1 Choice model and trip components.....	8
4.2 Software flow chart.....	9
4.3 Data requirement.....	10
4.4 Study area	10
5. Literature overview.....	11
5.1 Traffic modeling and the HOV-scanner	11
5.2 Generating the route set.....	13
5.2.1 Route set methods.....	13
5.2.2 Overlap	15
5.3 Mode and route choice.....	16
5.4 Mode and trip components and their weights	17
6. Research approach.....	20
6.1 The research layout.....	20
6.2 Assessment framework	20
6.3 Route set method.....	20
6.3.1 Assessment framework for the route set method	21
6.3.2 Available methods and programs	22
6.3.3 Assessment	22
6.3.4 Conclusion	23
6.4 Choice modeling method.....	23
6.4.1 Data requirement	24
6.4.2 Assessment framework	24
6.4.3 Approach.....	24
6.4.4 Assessment and conclusion	28
7. Route set program	29
7.1 The trip layout.....	29
7.2 The program design	29
7.2.1 Input	29
7.2.2 The structure.....	30
7.2.3 Data processing	30
7.2.4 Adjustable parameters.....	31
7.2.5 Output	31
7.3 Results.....	32
7.4 Discussion	33
7.4.1 The approach.....	33
7.4.2 The results	33
7.5 Conclusion.....	34
8. The choice modeling.....	35
8.1 Assessment framework	35
8.2 Results.....	35
8.2.1 Model parameter estimation	35

8.2.2	Model layout.....	37
8.2.3	Model fit on the input data	37
8.2.4	The PT trip forecast	38
8.3	Discussion	40
8.3.1	Model fit	40
8.3.2	PT trip forecast.....	40
8.4	Conclusion.....	41
9.	Validation.....	43
9.1	Study area.....	43
9.2	Route set generation	43
9.2.1	Results	44
9.2.2	Conclusion	44
9.3	Forecast.....	44
9.3.1	Approach.....	44
9.3.2	Assessment framework	45
9.3.3	Results	45
9.3.4	Conclusion	46
9.4	Trip distribution	46
9.4.1	Approach.....	47
9.4.2	Assessment framework	47
9.4.3	Results	47
9.4.4	Conclusion	48
10.	Answering the research questions.....	49
11.	Conclusions, discussion and recommendations	50
11.1	Conclusion	50
11.2	Discussion	51
11.2.1	NRM/LMS data	51
11.2.2	Trip parameters.....	51
11.2.3	Validation	51
11.3	Recommendations.....	52
	References	53
	Appendix 1 – ‘Duin- en Bollenstreek’.....	56
	Appendix 2 – Weight of trip parameters	58
	Appendix 3 – Program flow chart	59
	Appendix 4 – Program output.....	61
	Appendix 5 – Model process flow chart.....	63
	Appendix 6 – The error in the model fit for PT.....	64
	Appendix 7 – Trip forecast	65
	Appendix 8 – Elasticity.....	67

1. Introduction

1.1 Motivation

The HOV¹-scanner is developed and used by MuConsult. The HOV-scanner can be used to indicate the effect of changes in public transport (PT) services, for example, a new PT line. The effect can be the shift in attractiveness of the PT and the correlated change in PT usage (MuConsult, 2014). However, in its current layout the HOV-scanner is facing some limitations:

1. In case of parallel PT connections all demand is assigned to the most attractive connection. It is likely to assume that the assignment is more nuanced in real life.
2. The HOV-scanner is used as a first indicator for the feasibility of a new PT connection, but using it for this purpose is complex due to the layout of the HOV-scanner. This results in an operation that is too labor-intensive.

The HOV-scanner is a useful tool for a first indication regarding the feasibility of a new PT connection, but it can be enhanced by addressing the limitations. To achieve this MuConsult would like to have improvements on the accuracy and ease of use of the HOV-scanner. During this master thesis both topics are studied and the HOV-scanner will be improved accordingly.

1.2 Introduction on the HOV-scanner

The HOV-scanner is a tool that uses origin/destination (OD) matrices and utility functions to calculate a modal split. The HOV-scanner is programmed in MATLAB². For additional information on the HOV-scanner see chapter four.

The HOV-scanner is designed according to the four-stage transportation model. The total trip distribution is calculated and aggregated to match the desired zoning size. Then the modal split is calculated for the reference situation and eventual errors are calibrated. The changes in the PT network in combination with the sensitivity toward changes results in a new utility for PT. This new PT utility is used to forecast the distribution for the new situation. The HOV-scanner is modeled according to the schedule based approach (see paragraph 5.4). This is done to model the actual origin to destination travel time as realistic as possible.

1.3 Reading guideline

The problem definition is provided in chapter two. The research objective is highlighted in chapter three. Detailed information on the current HOV-scanner can be found in chapter four. The literature research in chapter five provides background information on transportation modeling and the implementation of route choice. The research approach is elaborated in chapter six, route set generation and choice modeling approach are explained in more detail in this chapter. The construction of the route set generation program and the first details are given in chapter seven. The first application of the choice model is the topic of chapter eight. Chapter discusses the validation process. The research questions are answered in chapter ten and the overall conclusion is given in chapter eleven.

¹ HOV can be translated as high quality public transit (e.g. BRT)

² MATLAB 2014B, The MathWorks Inc., Natick, MA

2. Problem definition

MuConsult is not completely satisfied with the current performance of the HOV-scanner. The intended use of the HOV-scanner is a first indication regarding the feasibility of a new PT connection. To improve the HOV-scanner the following elements need to change:

- It is required to make the HOV-scanner less labor intensive.
- A more realistic distribution of trips among PT lines in case of parallel PT routes.

The time consuming parts of the HOV-scanner

Currently it takes a couple of weeks to analyze a new situation (E. Rosbergen, personal communication, July 7, 2014). Process the program itself takes roughly 3 to 4 minutes; on the total processing time this is negligible. MuConsult made clear that processing the PT data and creating the files for the new PT situation is the most time consuming part of the process. Processing the operation schedules and calculating the travel times is done manually. The process is significantly more time consuming if the amount of zones increases and if there are multiple routes.

Modeling parallel PT routes

The previous section mentions that travel times are calculated. During this processes the quickest route is selected and used as an input for the generalized cost function of the HOV-model. Therefore only this alternative is used even if there are attractive alternatives. Nevertheless since the OD matrix is based on LMS/NRM data the trips generated by the less attractive PT mode are included in the input data.

For example: a trip from the University of Twente to Hengelo could be done by taking the bus or train with similar travel times (Figure 1). In this case the train trip is 1 minute shorter than the bus trip (13 versus 14 minutes). The HOV-scanner will use the travel time of the train for the GC matrix, because the model can only handle one PT route. It will use the fastest route for the GC matrix. The OD matrix on the other hand contains information of both the bus and the train trip. So the trips are included but for modeling purpose only the shortest (in time) route is used. This is a limitation of the current HOV-scanner. A more realistic assignment is required for more accurate results with the HOV-scanner.

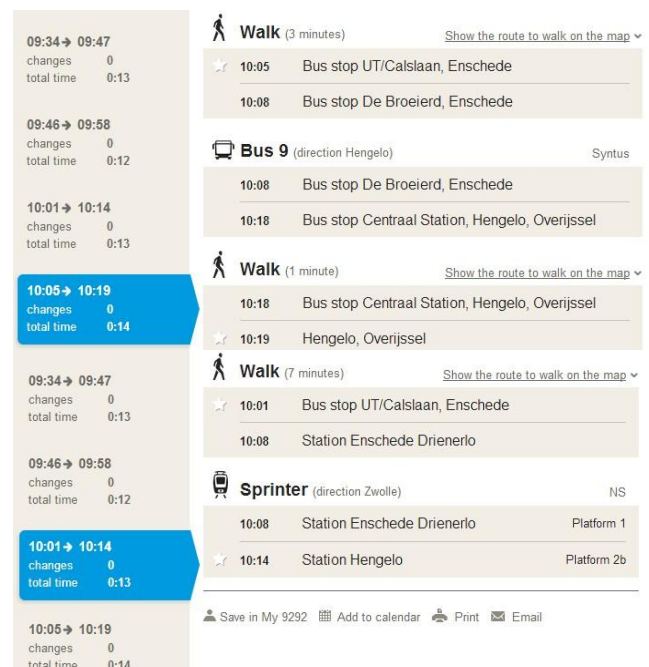


Figure 1: Travel time bus/Train (source: 9292.nl)

3. Research objective and research question

The main objective for this research is: “To implement PT route choice into the HOV-scanner to simulate parallel connections more realistic. Furthermore the work load should be reduced to make the HOV-scanner more suitable as a first indication regarding the feasibility of a new PT system.”

The objective of this research results in two research questions and four sub questions. The sub questions are used to answer the main questions.

1. How should the HOV-scanner be designed to be able to model parallel PT connections?
 - 1.1 Is the HOV-scanner in line with the current knowledge on transportation modeling?
 - 1.2 What is a correct way to implement parallel PT connections in a model?
 - 1.3 How accurate is the prediction value of the modal split for the new HOV-scanner?
2. How can the HOV-scanner be improved to reduce the work load?
 - 2.1 How can the required data for the HOV-scanner be processed more efficient and effective so that the required lead time can be met?

4. The HOV-scanner and study area

After the introduction on the HOV-scanner in section 1.3 a more detailed explanation can be found in this section. An elaborate description is provided on the HOV-scanner, discussing the following elements: choice model, trip components, software flow chart and data requirement. This should explain how the HOV-scanner design and the effect of design changes. The HOV-scanner is applied on the “Duin- en Bollenstreek” in a previous study (MuConsult, 2012). This study and the available data is used and therefore additional information “Duin- en Bollenstreek” is provided.

To avoid confusion there is referred to the first version of the HOV-scanner as original HOV-scanner and the new version as new HOV-scanner when required.

4.1 Choice model and trip components

The original HOV-scanner includes a mode choice model that distributes the total number of trips over three modes; car, bicycle and PT. Discrete choice modeling is applied by using a multinomial Logit (MNL) model. The mode distribution for the future scenario is forecasted by making use of the shift in PT utility and the sensitivity towards these changes. The network changes are extracted by comparing the reference network with the scenario.

A MNL model provides a probability for an alternative based on the utility of all alternatives. As an input for the utility function the travel time per mode is used. For the car and the bicycle the total travel time is used and for PT this travel time is divided in trip components (see formula below). All trip components are specified in minutes. Whereby the transfer time is the time in between arrival and departure of two vehicles and the waiting time is the time used before the arrival of the first vehicle. A schedule based approach is used to retrieve travel times.

$$V_{PT} = \beta_{ivt} * IVT_{PT} + \beta_{transfer} * transfers_{PT} + \beta_{wait} * waiting_{PT}$$
$$V_{car} = \beta_{car} * traveltime_{car}$$
$$V_{bicycle} = \beta_{bicycle} * traveltime_{bicycle}$$

For the study area of the “Duin- en Bollenstreek” the trip parameters are chosen as followed. The ratio towards PT IVT is noted between brackets:

β_{car}	= -0,0562 (1,6)
$\beta_{bicycle}$	= -0,0348 (1,0)
$\beta_{PT \text{ access and egress}}$	= -0,0406 (1,2)
$\beta_{PT \text{ wait}}$	= -0,0435 (1,3)
$\beta_{PT \text{ IVT}}$	= -0,0342 (1)
$\beta_{transfer}$	= -0,0388 (1,1)

4.2 Software flow chart

The HOV-scanner software flow chart figure 2 is used to explain how the HOV-scanner functions. The figure describes the data processing from input towards forecasting. The input data is marked orange, the outputs are yellow, the operations are blue and the result is green. The steps noted between the brackets relate to the four-stage transportation model. For reference purposes each box is numbered. The process can be separated in reference data processing and calculating the effect of the scenario. For a process overview see appendix 5.

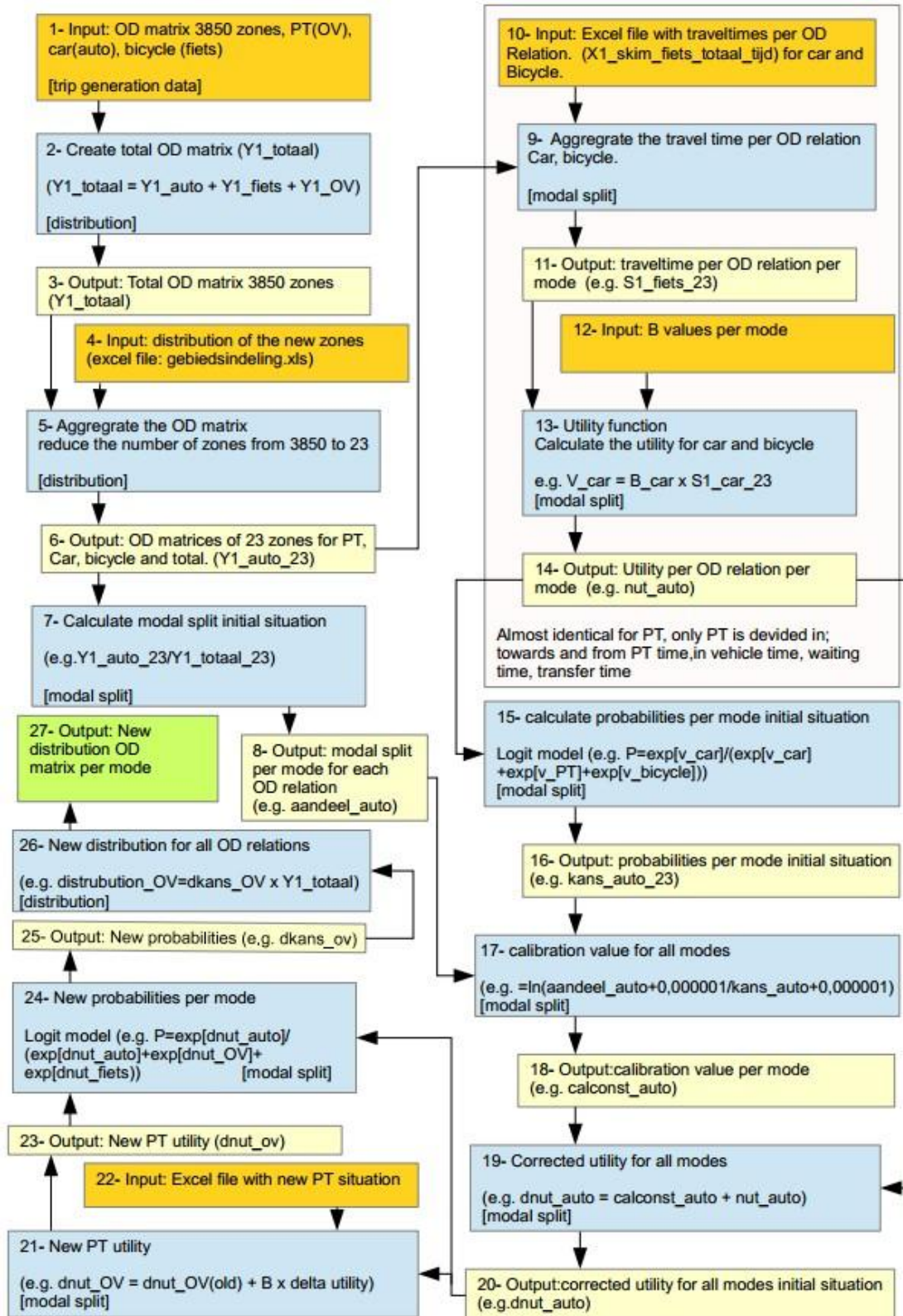


Figure 2: HOV-scanner software flow chart

Reference

First the total number of trips per OD for the reference situation is calculated by summing the PT, bicycle and car trips (2). The LMS/RMS data makes use of small zones. These zones are aggregated to reduce processing time (5). This results in OD trip data for all modes on the new zoning level (6). This data is used to calculate the modal split per OD for the reference situation. In line with the trip data the skim matrices are aggregated as well (9). The skim matrices contain the travel time per mode for all OD's. The skim matrices and the β 's are the input for the utility function (13). The utility function calculates the utility per mode for all OD's. The utility is used as an input for the probability function (15) that results in a probability per mode for all OD's. The probability from the previous step is calibrated (17/19) to match the modal split of the input data (8). This results in a corrected utility per mode and OD for the reference situation (20).

Scenario

The utility of the reference situation is used to calculate the utility in the scenario situation (21). This is done by summing the reference utility with the delta utility. The delta utility is the difference between the utility reference and the scenario. The travel time by PT is used as an input (22) and is manually calculated. The new utility is used to calculate the probability per mode for the scenario (24). The probabilities per mode are used to calculate the scenario mode distribution (26).

4.3 Data requirement

To make use of the HOV-scanner there are five types of input data that are required:

- **OD matrices car, bicycle and PT**
 - The data comes from Rijkswaterstaat's (RWS) LMS/NRM model and is used for the reference situation.
- **Aggregation data for the new zoning table**
 - Aggregated zones are selected by the user and used as an input for both reference and scenario.
- **Travel time per OD relation**
 - The data comes from Rijkswaterstaat's (RWS) LMS/NRM model and PT schedules. For the car and the bicycle this is the total time. The PT time is divided in: access and egress time, in vehicle time (IVT), waiting time and transfer time.
- **The new PT situation**
 - The data file of the PT scenario. This file contains the travel time for PT divided in the different trip segments. This file is created by MuConsult.
- **β values per mode**
 - Values are calculated/estimated by MuConsult

4.4 Study area

The "Duin- en Bollenstreek" is an area that is located at the border between the provinces "Noord-Holland" and "Zuid-Holland". In order to test and calibrate the model this data is used. MuConsult conducted a study to investigate the feasibility of a changed the PT system. These changes should result in a PT system that matches the spatial planning better and improve the connection with Schiphol. An overview of the area and the transit lines can be seen in appendix 1. The overview contains both the reference situation and the scenario. The scenario network contained new PT lines and schedules and is referred to as HOV-basic. Within the study area there are approximately forty different PT lines and forty-seven different transit stops analyzed. In total there are 420 OD combinations and 761558 trips. The car share is 69%, the PT share is 17% and the bike share is 14%. The following information is available for this area:

- PT schedules of the reference situation and possible scenarios (source: 2011.04.01 Dienstregelingen na bespreking.xlsx from MuConsult)
- OD trip matrices and zoning details for car bicycle and PT (source: NRM data that is enriched)
- Travel times for car, bicycle and PT
- The result of a previous study (MuConsult, 2012)

5. Literature overview

To analyze whether the HOV-scanner reflects contemporary knowledge of transportation modeling and to investigate how parallel PT connections are modeled in other studies a literature review is conducted. In general route set generation and choice modeling are considered to be important, because they can make the model outcome more reliable. Therefore they both are investigated. Furthermore, the effect of trip factors is investigated whereby the overlap seems to be of importance for both route set size and route choice. Overlap affects the route set size and can lead to over representation of the overlapping routes if overlap is ignored.

5.1 Traffic modeling and the HOV-scanner

The classic four-stage transportation model (Figure 3) is a result from practice in the 1960s and is more or less unaltered since that time (De Dios Ortuzar & Willumsen, 2011, p. 20). The model distinguishes the steps trip generation, distribution, modal split and assignment. The four-stage transportation model can be used to evaluate a future transportation situation. This model is used for the HOV-scanner.

The trip generation step deals with the question: how many trips? The amount of trips is a forecast based on the base year information and an estimation of future developments. The **trip generation** step is not executed by the HOV-scanner because the number of trips in the reference situation is known and the number of trip in the scenario is expected to be identical. For large changes in PT service, new techniques or changes in urban development this can be a limitation. Because these kind of changes can lead to a shift in the overall trip demand. The **distribution** step deals with the question: from where to where? This step results in an origin destination (OD) matrix for the estimated situation. The distribution data per mode is input information for the HOV-scanner scenario. The **modal split** step deals with the question: which mode? During this step the trips of OD matrix are assigned to the different modes of transportation (i.e. car, bicycle or PT). Calculating this for the scenario situation is one of the main concerns of the HOV-scanner. The **assignment** step deals with the question: Which route? In this step the trips per mode will be assigned to the network. This results in a network load per mode. The assignment step is not executed because there is only one PT route used. For the new HOV-scanner the assignment step must be included because there are multiple PT routes available. This implies that the model should be extended towards mode and route choice. This can be done in separate steps or in a combined step.

A shortcoming of the classic four stage model is the inaccuracy when the effect of congestion should be modeled (De Cea, Fernandez, Dekock, Soto, & Friesz, 2003). Due to differences in link cost during the assignment step and the trip distribution step the four-stage model could mispredict the traffic assignment. To improve the model different combined step models and multimodal models were introduced. Early studies of combined step models are performed by Bruynooghe (1969), Florian, Nguyen, and Ferland (1975), Evans (1976) and Florian and Nguyen (1978). Modeling errors due to congestion does not seem to be applicable for the HOV-scanner. The basic distribution data is generated by an external model (LMS/NRM coming from RWS) and a different model will not change this distribution. Therefore it will not alter the reference situation. In general the PT share in the Netherlands is only 5% (Bakker, Zwaneveld, Berveling, Planbureau, & voor Mobiliteitsbeleid, 2009). Only the link costs for the PT network will change in the scenario. It is assumed that a change in the small PT share has no significant effect on the car share. Therefore congestion effects can be ignored for the HOV-scanner. To summarize the four stage model is suitable to use for the HOV-scanner and therefore it will be used. Nevertheless the correct implementation of mode and route choice is not yet determined.

Network

The original HOV-scanner is using the transit network in combination with one access and egress mode, presumably walking in most cases. This can be a restriction of the model. In the Netherlands bike and ride

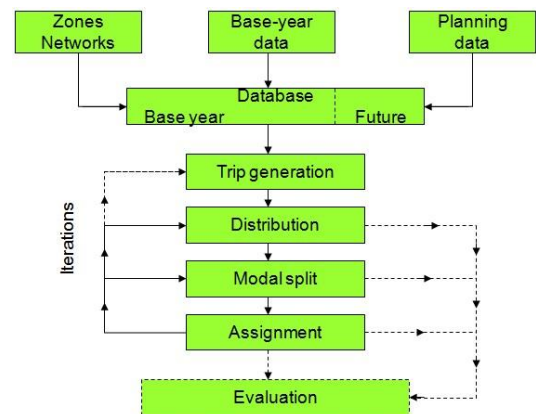


Figure 3: Four-stage transport model

(bicycle and PT) and park and ride (car and PT) are common multimodal combinations. For train access and egress the bicycle is responsible for respectively a 42% and 14% share and the car is responsible for respectively a 14% and 8% share (Van der Heuvel, 2013). The supernetwork approach that is introduced by Sheffi (1985) can deal with these kind of multimodal networks. In the supernetwork approach the networks of the different modes are stacked on top of each other and connected by transfer links. See Figure 4 for an example. The supernetwork approach is suitable for modeling the PT network. This could change catchment areas as well as total travel times and therefore the effect of the PT connection. The main limitation would be processing the required data which will increase the total processing time that is needed to assess a new situation. This is not in line with the aim to reduce the work load and processing time and therefore this option will not be used. The bicycle is responsible for a large share of the train access (42%) and can be added without effecting the processing time of the HOV-scanner too much. Therefore HOV-scanner will be extended with the bicycle as additional access and egress mode.

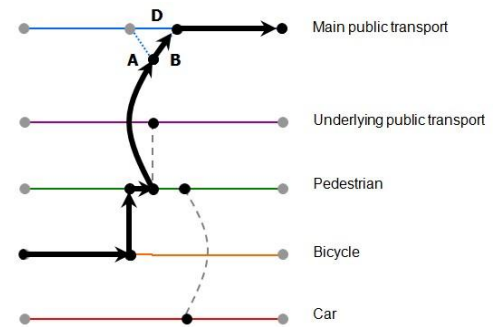


Figure 4: Supernetwork (source: lecture T. Brands 2013-09-12)

For both access and egress modes a catchment area should be selected. The catchment area of a station depends on the combination of the PT mode, station type, transportation modes and user preferences. This results in station and scenario specific catchment areas. Implementing this would increase processing time and it requires more data. To avoid this a fixed catchment area will be used that differs per access and egress mode. According to Yang, Yan, Xiong, and Liu (2013) an acceptable walking limit is 1 kilometer (12 minutes). This is comparable to the results of Debrezion, Pels, and Rietveld (2009) they found that until 1,1 kilometer walking was the most likely mode and between 1,1 and 3,6 to 4,2 kilometer the bicycle is the most likely mode to be chosen. Blij (2010) researched HOV catchment areas and an average catchment area of 800 meters was found for walking and 2350 meters for cycling. To conclude a maximum walking distance of 1 kilometer seems to be realistic. The maximum bicycle distance shows more deviation. For the bicycle distance the average value of 3 kilometer will be used.

Choice modeling

In the mode choice step the mode choice is modeled. The HOV-scanner is modeling a choice between car, bicycle and PT. The HOV scanner uses a Logit model for this choice modeling. The choice distribution in this model is based upon the differences in utility (U_{jq}). The utility (U_{jq}) cannot be measured and is therefore represented by the measurable V_{jq} and the random part ϵ_{jq} ($U_{jq} = V_{jq} + \epsilon_{jq}$). V_{jq} is a generalized cost (GC) function of measurable attributes. For example the disutility of the bus could be composed as following:

$$V_{bus} = \beta_1 * IVT_{bus} + \beta_2 * transfers_{bus} + \beta_3 * waiting_{bus}$$

Where the terms are defined as:

- β_{1-3} = the weights for each of the elements
- IVT_{bus} = the in vehicle time during the trip
- $transfers_{bus}$ = the number of transfers in the trip
- $waiting_{bus}$ = the waiting time during the trip

To predict the chance of an alternative to be chosen the utilities are transformed in a probability between zero and one:

$$P_{bus} = \frac{E^{V_{bus}}}{E^{V_{bus}} + E^{V_{auto}} + E^{V_{fiets}}}$$

Besides the Logit model there are other applicable modeling frameworks. The fuzzy logic technique is such a method. It was introduced in 1965 by Lotfi Zadeh but it is getting an increasing attention in the recent years. It can be used by categorizing upon linguistic based variables whereby each member has a certain degree of membership to the variable. An example of suchlike variable is the medium travel time. The medium travel

time has a certain value and a distribution around this value (figure 5). This implies that a trip with 34 minutes of travel time could be a member of the medium and long travel time group. These linguistic based variables are used as an input for the If-Then rules. They are used to capture the relation between the input and the output. For example IF the travel time is high and the fare is high THEN the utility is very low. An advantage of the fuzzy based techniques is that they can be used if the input is categorized such as the travel time above. A disadvantage could be the processing time required to implement the If-Then rules and to calibrate the model (Kumar et al., 2013). Due to the fact that the input values are exact travel times the fuzzy logic technique will presumably not improve the HOV-scanner.

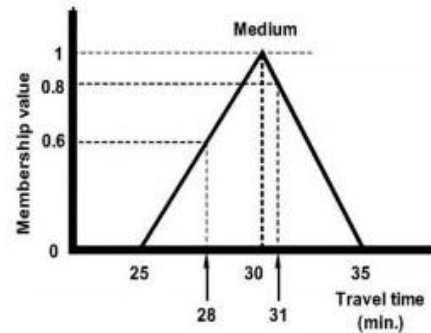


Figure 5: Distribution source: Kumar, Sarkar, and Madhu (2013)

The input data used by the HOV-scanner is constructed with a Logit model. Using an identical model could reduce the error due to model misspecification. According to M. C. Bliemer, Rose, and Hensher (2009) model misspecification can result efficiency loss of 5 to 30%. To avoid this a Logit structure should be used. Furthermore is the fuzzy logic technique not expected to yield a more accurate modeling results. Therefore a Logit model will be used.

5.2 Generating the route set

If the PT route should be extended from one towards more routes it is important to know how many routes and which routes should be included. The methods that can be used to generate these route sets will be elaborated in this section. Overlap within routes is considered to be a main issue regarding both route set size and route choice. Therefore details regarding overlap can found in this section as well.

5.2.1 Route set methods

The considered route set is a set of routes that are considered by a traveler to get from his origin to his destination. The considered route set is a subset of the entire route set. The entire route set contains all existing routes. This set can be narrowed down to, the logical, the feasible, the known and then the considered route set (Lanser, 2005). Most of the time the considered route set (subjective choice set) is much smaller than the objective choice set (Fiorenzo-Catalano, 2007). According to Fiorenzo-Catalano (2007) the objective choice set is in proportion to the network whereas the subjective choice set is limited by the traveler his ability to consider alternatives. For route set modeling it is important to know how much the route set that is generated by the methods differs from the considered route set. Differences in numbers between the generated and the considered route set can be 40 to 15 and 15 to 4 according to Bovy and Stern (1990). In line with this result Lanser (2005) found differences of 63 to 2. Van der Waard (1988) on the other hand was able to generate a route set that had an identical size as the considered route set. Nevertheless in general the considered route set is smaller than the generated route set. For the HOV-scanner this implies that it is necessary to check if the generated route set match the considered route set. It can be beneficial to add constraints during route set modeling to let the generated route set better match the considered route set. This will be elaborated in the research approach.

It is important that the selected method is suitable for a PT network. The number of links within a PT network is in general much more limited than in a car network. This could result in more severe errors in the route set generation for PT when essential links are ignored. This increases the urge to select a method that will find all relevant routes. Furthermore it is required that constraints regarding the routes can be implemented to let the generated route set better match the considered route set.

The route set generation methods can be distinguished in two groups. The group of methods that change network attributes and methods that change other attributes.

Change network attributes

Members of the group that change network attributes are: k shortest path method, the constrained k shortest path (CKSP), link elimination method, link penalty method, k dissimilar paths method and the Monte Carlo approaches. All of these methods apply a shortest path search then they change the network and they apply a new shortest path search. Network changes can be changing or deleting links within the network.

The k shortest path method removes a link of the shortest path each round so that a new shortest path can be constructed. According to Ramming (2002) the KSP result matches with the observed routes

(80% coverage). It is an exact way to find the k-shortest path. Nevertheless the number of routes found depends on the number of k and is therefore not necessarily close to the considered route set size.

The constrained k-shortest path (CKSP) method introduced by Catalano and Van der Zijpp (2001) is an elaborated version of the KSP method. Different constraints (e.g. detour criterion) can be added to let the generated route set match better with the considered route set.

The link elimination method is introduced by Bellman and Kalaba (1960). For each step some or all links of the shortest path are deleted. In comparison to the KSP method the link elimination method requires much less computation power. On the downside the quality of the route set is much lower.

The link penalty method is introduced by Johnson, Joy, Clarke, and Jacobi (1992). This method changes link attributes of the shortest path and then searches the new shortest path. Identical to the link elimination method the computation burden is much lower than with the KSP method. Nevertheless, if the penalty level or strategy that is not correctly chosen could affect the result.

The k dissimilar paths method is introduced by Akgün, Erkut, and Batta (2000). Each round network attributes are changed and routes above a minimum dissimilarity threshold are selected. Resulting in a route set with a certain dissimilarity between the routes.

The Monte Carlo approach is introduced by Sheffi, Mahmassani, and Powell (1982). This method changes all network attributes randomly each iteration. The method is terminated at the maximum number of iterations.

A variant of the Monte Carlo approach is the accelerated Monte Carlo approach of M. Bliemer, Versteegt, and Castenmiller (2004). This approach is similar to the previous one only the variation is increased during the process. This results in more paths for the same amount of iterations.

The Monte Carlo labeling combination approach is introduced by Catalano and Van der Zijpp (2001). Like the title suggest this approach combines the Monte Carlo and the labeling approach. This results in a Monte Carlo approach that can be applied on different criteria (labels). In line with the original approach this approach can be accelerated by increasing the amount of variation.

Change other criteria

Members of the group that change other criteria are: the constrained enumeration, gateway method, essentially least cost path and the labeling approach. In this group the network attributes are unaltered but other changes are made to find different routes.

The gateway method is introduced by Lombard and Church (1993). The routes are searched via gateway points or links in the network. The gateways are set on beforehand and the operation terminates if all gateways are passed. Resulting in a set of routes via preselected locations.

The essentially least-cost paths method is introduced by Hunt and Kornhauser (1996). This is a two step approach. First, a repetitive application of the shortest path method is used. This results in possible route set segments with a predefined cost threshold. In the next step the routes from origin to destination are constructed using the detour segments as constraints.

The constrained enumeration method is introduced by Prato and Bekhor (2006). This is a branch and bound technique that eliminates links that do not apply to one of the criteria. In a branch and bound technique a tree of routes is generated that connects the origin to the destination (figure 6). Every time a node is reached this is an opportunity to switch to other links and thereby the route tree is constructed.

The labeling approach is introduced by Ben-Akiva, Bergman, Daly, and Ramaswamy (1984). Different labels can be applied e.g. shortest in time, shortest in distance, least transfer or specific modes of transportation. For each of these labels the shortest path is

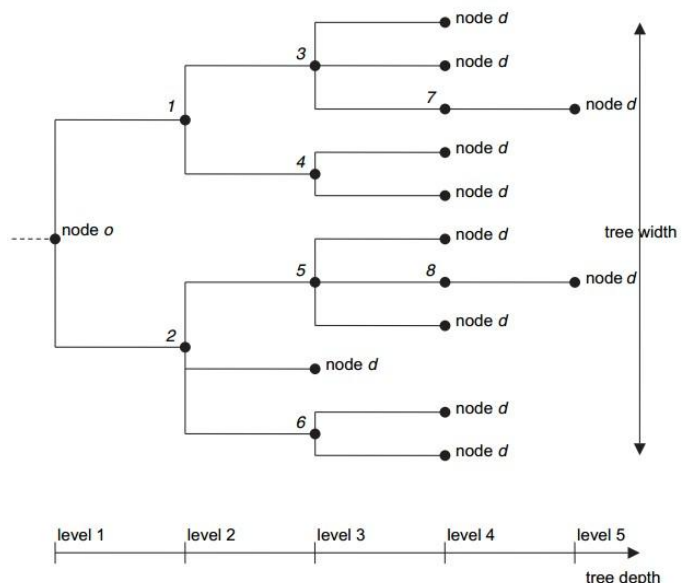


Figure 6: Constrained enumeration (source: Prato and Bekhor (2006))

selected. This can result in a varied route set that include routes with different user preferences. Therefore the total number of routes is related to the number of labels. Nevertheless the total number of routes is not assumed to be identical for all OD's in the study area. A short trip for example could have only one suitable route.

Conclusion

Changing network attributes could lead to disconnecting or deletion of essential links. Guo and Wilson (2011) have acknowledged this effect for the link elimination method and the K shortest path method. Nevertheless it seems to be relevant for all methods that change network attributes. Due to this limitation methods that change network attributes are not the most preferable solution for modeling a PT route set. Of the methods that change other criteria the gateway method is not able to handle route constraints and will therefore be ignored. Taking this into account the most suitable route set methods for creating PT-route sets are: the essentially least-cost paths, the constrained enumeration and the labeling approach. In the research approach a more detailed comparison of these methods will be provided.

5.2.2 Overlap

Overlap is considered to be a problem regarding route choice modeling (Fiorenzo-Catalano, 2007; Lanser, 2005; Ramming, 2002). A route is overlapping if a part of the route is identical to another route e.g. the left figure of figure 7. Due to overlap the trip shares per route can be over or under estimated and this can lead to unrealistic route shares. Figure 7 for example predicts a share for the non-overlapping route from 1/3 to 1/2 with a gradual transition for most of the choice models. For more details on the choice models see (6.4). The MNL model will always predict a 1/3 share per route regardless of the amount of overlap.

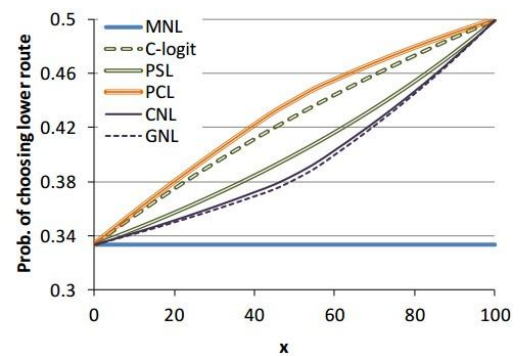
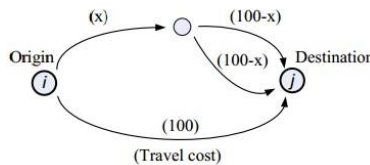


Figure 7: Route overlap (source: Kitthamkesorn (2013))

There are two methods to account for overlap: a choice model that can incorporate overlap in route choice or exclude the overlapping routes from the route set.

If there should be accounted for overlap the question arises when two PT routes are considered to be overlapping and how much overlap should be permitted. Overlap could be considered in case of:

- The same geographical route
- The same PT line
- The same vehicle
- The same time span
- Combinations of the indicators above

Prato, Bekhor, and Pronello (2012) mark a route as similar if they share 3 or more landmarks. This could be a correct implementation but it is a laborious method because land marks should be defined and categorized. Therefore this option will not be used. Lanser (2005) determines overlap by modes and transport services between pairs of nodes. Expressed in transfer free legs between two modes. Time and distance are the most suitable units to measure this overlap according to Lanser (2005). Largely overlapping routes should be removed from the choice set. An overlap boundary of 80% is used by Bovy and Fiorenzo-Catalano (2007). In a succeeding study (Bovy, 2009) the same boundary is mentioned nevertheless the effect of this boundary is not further elaborated.

Most of the overlapping studies are focused on car networks (e.g. Ramming (2002)). Overlap in transit networks is studied less thoroughly according to Lanser (2005). Because overlap in PT is studied less thoroughly the correct implementation in the HOV-scanner is more difficult. Solving overlap in the choice model seems to be a more refined solution than deleting them from the choice set. On the other hand is removing the overlap in the choice set more convenient and therefore it better matches the intended use of

the HOV-scanner. Due the fact that it is unclear how the two methods perform in comparison to each other it is not possible to select the best. Therefore the most convenient method will be used for the HOV-scanner: deleting the overlapping routes from the route set. Overlap will be defined as using the same lines between stations measured in time. This is similar to the approach of Lanser (2005) and intuitively it seems to be the correct definition. For the overlap boundary 80% will be used as suggested by Bovy and Fiorenzo-Catalano (2007). It is uncertain what the best boundary value is but there is not much information available on this topic. Therefore it is desirable to have additional information regarding the deleting overlap in a route set and especially on the overlap boundary. Additional information seems to be required on this topic.

5.3 Mode and route choice

In section 5.1 it is mentioned that the HOV-scanner is using a Logit model. To be more exact, the HOV-scanner uses a Multinomial Logit (MNL) model. Due to the inclusion of PT route choice the choice model is changing from a mode choice model towards a mode and route choice model. Therefore another Logit modeling structure could be required. In this paragraph different models will be explained and compared. With the comparison it is possible to select a suitable choice model.

In the original HOV-scanner MNL is used. This is the most commonly used Logit model (McFadden, 1973). It can be used to derive the probabilities for three or more alternatives if the alternatives apply to the independence of irrelevant alternatives (IIA) property. A clear and well known example of the IIA property is red/blue bus problem. This problem describes the over prediction of the bus alternative because they are assumed to be independent in the model. The red/blue bus problem is applicable for the HOV-scanner because the PT routes are considered to be correlated. They are correlated because the different routes offer comparable modes of transportation and a comparable service. This makes MNL unsuitable if route and mode choice will be combined. Nevertheless if mode and route choice are modeled separately the MNL model could be used.

Nested Logit (NL) is able to overcome some of the restrictions of the MNL. The first exhaustive analysis of NL is done by Williams (1977). NL can be used when alternatives are not independent (e.g. red/blue bus problem, figure 8). NL can be used if the variances in the error term ϵ are not equal. Moreover NL can be used in case of taste variation among individuals. NL can be used to include PT route choice into the mode choice model. In the NL model the logsum is used to capture the correlation between the alternatives within the nest. The logsum parameter can differ from zero and one. If the logsum parameter is one there is no correlation between the alternatives in the nest and the model is identical to a MNL model. If the parameter is zero they are perfectly correlated, and therefore perceived as one.

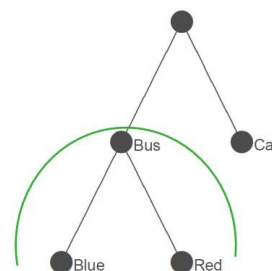


Figure 8: Nested logit

Cross-Nested Logit was introduced by Vovsha and Bekhor (1998). It is comparable to NL but in addition members of a nest can be a member of more than one nest.

Regarding current practice a shift can be seen from Multinomial Logit (MNL) and Nested logit (NL) toward Mixed Logit³ (ML) (Vij, Carrel, & Walker, 2013), (M. C. Bliemer & Rose, 2011), (Hensher & Greene, 2003). This is caused by the limitations of MNL and NL models and due to the possibilities of new software packages according to (Hensher & Greene, 2003). An advantage of ML over MNL is that it can be used to allow for random taste variation, correlation in unobserved factors and it has unrestricted substitution patterns. ML is used for example to combine stated preference (SP) and revealed preference (RP) data. Nevertheless it is a complex model to use. A disadvantage could be the demand for better quality data compared to MNL (Hensher & Greene, 2003). For the HOV-scanner the inclusion of taste variation among groups of travelers could be of use. The preferences from a student with a student PT card for example could differ from a sales man on a business trip.

To deal with unknown or unused routes the Implicit Availability/Perception Logit (IAP) is introduced by Cascetta, Papola, Russo, and Vitetta (1999). IAP is comparable to C-Logit only then for unknown routes instead of overlap. According to Ramming (2002) it is a convenient way to incorporate the awareness of paths into route choice modeling where no explicit choice set generation step is needed. A choice set

³ Mixed logit (MXL) is also referred to in as random parameter logit (RPL), mixed multinomial logit (MMNL), Logit Kernel, error components logit and hybrid logit.

generation step is already included and the effect of unknown or unused routes seems to be less dominant in PT this method will not be used.

To summarize from the preliminary research it seems that there are three choice models that could be used for the HOV-scanner: Multinomial Logit, Nested Logit and Mixed Logit. In the research approach (6.4) a more detailed comparison will be provided.

There are two different approaches introduced to deal with overlapping: C-Logit and Path-Size Logit. C-Logit is introduced by Cascetta, Nuzzolo, Russo, and Vitetta (1996). There is accounted for overlap by introducing a “commonality factor” (CF). This factor is negative to reduce the shares in case of overlapping routes. This can be seen in the formula. For the CF different forms are proposed by the authors.

$$P(i|C_n) = \frac{e^{V_{in}+CF_{in}}}{\sum_{j \in C_n} (e^{V_{jn}+CF_{jn}})}$$

Path-Size Logit is introduced by Ben-Akiva, Bowman, Ramming, and Walker (1998). Comparable to the C-logit approach an additional term is introduced to deal with overlapping:

$$P(i|C_n) = \frac{e^{V_{in}+\ln PS_{in}}}{\sum_{j \in C_n} (e^{V_{jn}+\ln PS_{jn}})} = \frac{PS_{in} e^{V_{in}}}{\sum_{j \in C_n} (PS_{jn} e^{V_{jn}})}$$

Non overlapping paths have a value of one and if more overlap occurs this value drops to 0. If the overlapping paths differ in length additional care must be taken according to Ramming (2002). Nevertheless the Path-Size logit approach outperforms the C-Logit approach (Ramming, 2002). Both models should maintain the simplicity of a Logit model but generate more realistic route shares in situations with overlapping routes. For the HOV-scanner these approaches are not required because during the route set generation overlap will be handled and overlapping routes will be deleted.

5.4 Mode and trip components and their weights

There are several factors that determine the mode and the route choice. For example social economic factors, such as income and family composition. Mode and trip components, such as speed, frequency and comfort but also factors like habit, trip purpose, and destination. In this section the effect of the mode and trip specific components will be elaborated, because these are the components that will be changed when the effect of a new PT scenario is calculated. Most of these changes will result in a changed travel time. There are two approaches distinguished to calculate travel time, they will be elaborated in this section as well. The aim is to determine important factors and the effect of these factors. Furthermore, this chapter contributes towards answering research question 1.1 by providing a reference for trip components and weights used in the original HOV-scanner.

Frequency versus schedule based

For PT modeling there are two approaches used to calculate the total travel time in a PT network. The frequency based approach and the schedule based approach. An example of the frequency based approach is the Zenith algorithm and an example of the schedule based approach is the Diachronic network (Nuzzolo & Russo, 1996). Frequency based models require a low level of input detail. The frequency is used to determine waiting and transfer times for PT trips that contain multiple transit lines. Schedule based approaches on the other hand calculate the transfer and waiting times based on the PT schedules. Both the approaches have their own advantages and drawbacks. The frequency based approach is faster and more easy to use. The disadvantage is lack of realism and accuracy of the model, especially at low frequencies. Schedule based models on the other hand require a high level of input detail and processing power but they have the potential to produce a more realistic outcome. This can be clarified with an example of a combined PT trip that uses the bus as an access mode for the train. If the bus could improve the timetable travel time with 2 minutes the passengers will arrive 2 minutes earlier on the train station. If a passenger still needs to use the same train this will not result in an improved travel time. On the other hand if the passenger can catch an earlier train the improved timetable travel time results presumably in an improved travel time.

The schedule based approach is preferred by MuConsult. Nevertheless, both options were considered. The use of the HOV-scanner is not limited to new connections, but it is also used for analyzing the effect of improved PT schedules. The effect of these kind of network optimizations is better modeled in a schedule based approach. Therefore a schedule based approach will be used.

Mode and trip components

The components that influence the shift from car towards PT are investigated by Qin et al. (2013) and Satiennam, Jaensirisak, Satiennam, and Detsamrong (2015). Both studies highlight the importance of travel time and costs whereby Qin et al. (2013) explicitly mentions fuel cost and emphasizes the effect of bus comfort. There are other studies that have investigated which trip components have the most impact on the utility of a trip. According to Lanser (2005) the most important factors for the utility of a multi modal trip are: IVT, frequency related attributes and mode specific constants (e.g. type of vehicle and feeder mode). According to Lanser (2005) the effect of the rail way station and the costs are small. Zhang, Guan, Qin, and Xue (2013) studied the utility of a bus trip in Jinan China. According to this study the effect of changes in travel time is larger than the effect of changes in the ticket price. Yang et al. (2013) studied the rail transit access mode. In this study was found that the distance to the station is the most important factor that determines the accessibility of the station. Where a more accessible station could result in more PT usage. According to Diana (2008) reliability, flexibility, comfort, rapidity and space are the most important factors that determine trip utility. Where a higher frequency results in more reliability and flexibility in case of disturbances. All of the studies above mention the importance of travel time. Lanser (2005) mentions a small effect of the costs. Nevertheless, it is generally accepted that there are three different components that represent the utility of a trip: travel time, travel costs and the amount of effort (Horowitz & Thompson, 1994; Lanser, 2005; van Hagen, 2011).

It must be determined which components will be used for the HOV-scanner. The more components used the more complex and time consuming the model will be. Complex models can give the impression that they result in a more accurate output. Moreover they are not necessarily more accurate. Therefore it is important to find the sweet spot where the model is accurate enough without requiring too much processing time. A good starting point is to search for the simplest model that explains the data well. The original HOV-scanner uses travel time and effort for PT trips and only travel time for the other modes of transportation. Ignoring the travel costs for both car and PT could lead to inaccurate results. Furthermore, is the effect of frequency ignored despite of the fact that frequency is considered to be important. The used PT mode is ignored whereby a PT mode could be an indication for the comfort and the reliability. To improve modeling accuracy it could be beneficial to include both travel costs and frequency. It is unclear if the effect of different PT modes should be included to capture preferences for a specific mode. This will be elaborated in the next section.

The travel costs are ignored in the HOV-scanner due to the difficulty to obtain car travel cost and the overall difficulty of the implementation of the costs. Especially for short car trips the trip price is more difficult due to the fact that parking costs (if applicable) can be an impressive share of the overall trip cost and are difficult to estimate. Implementing the costs conflict with the intended processing time of the HOV-scanner. The frequency could have an effect on mode and route choice, a more frequent service is considered to be a more attractive alternative. For the original version of the HOV-scanner the frequency could be implemented straight forward due to the use of only one PT route. The new HOV-scanner will use a PT route set instead of a single PT route. Defining frequency for route sets is more difficult, because the sum of the frequencies within the route set is not necessarily the total frequency of the route set. A worst case scenario would be a route set with multiple routes that have an identical first line. This results in a perceived frequency of the first line instead of the sum of the routes in the route set.

To summarize it is preferable to include costs, time, frequencies and used modes into the utility function. Nevertheless cost will be ignored as a result of processing time. For the effect of PT modes and frequency further research is required. This will be elaborated in the next section.

The weight of mode and trip parameters

In this section the mode and trip parameters as mentioned in the previous section will be compared. The mode and trip parameters are: used mode, travel time and frequency. The different trip parameters all have their own effect on the disutility of the trip. For example waiting for a bus is perceived to be less attractive than riding in the vehicle. Therefore, the effect of travel time will be distinguished into different components of the trip. The relations found in this section are used by the HOV-scanner to determine the effect of a PT scenario. For finding the effect of trip parameters different studies have been compared. A table with the results can be found in appendix 2.

Is it possible to distinguish travel time weight for different modes of transportation? The study of

Arentze and Molin (2013) reveals small differences: the train weights 0,8, the bicycle 1 and the car 1,1 in comparison to the bus and tram. Abrantes and Wardman (2011) found different results for the car bus ratio: 1,2. For both studies the time spent in the bus is less attractive than the time in the train. According to Abrantes and Wardman (2011) the time in the car is more attractive than the other modes, but according to Arentze and Molin (2013) the time spent in car is less attractive than the other modes. The difference can be explained by the fact that the attractiveness depends on more factors than the mode alone. For example, trip length and trip characteristics influences the attractiveness of different modes. The tram bonus is researched by Bunschoten (2012): 3,3 minutes for a tram trip compared to a bus trip. On a 20 minute trip this would translate to a tram/bus IVT ratio of 0,85. A bus rapid transit (BRT) time is expected to be in between of the train and the bus. Whereby the actual value depends more on the implementation of the service than on the mode (Currie & Delbosc, 2013). To summarize the effects of modes, there are differences per mode revealed but they are small and complex to describe. This makes it difficult to relate the car time towards PT.

In general the out of vehicle time (OVT) is weighted more than the IVT. The values differ between studies and between modes. On average one transfer is perceived as nine minutes of IVT i.e. the disutility of one transfer is identical to the disutility of nine minutes IVT. Nevertheless there is a large deviation between different modes. Within the study of Fosgerau, Hjorth, and Vincent Lyk-Jensen (2007) differences in transfer time from 1,6 (metro) to 12,2 (bus) were found. For both access and egress and waiting time significant differences in weight were found (0,9 to 2,2). For long trips the weight of access and egress seems to be below IVT value. Presumably due to the smaller effect of access and egress time on the entire trip.

The frequency of a PT line and PT route have an effect on both the attractiveness of the PT alternative (e.g. Goeverden and Van Den Heuvel (1993)) and the trip distribution over the route set. Furthermore the frequency does also effect the transfer penalty by effecting disutility and insecurity (Iseki & Taylor, 2009; Vuchic, 2005). A transfer on a low frequency service increases the chance of large changes in the arrival time when the transfer is missed, this increases trip insecurity. There are a lot of studies on the effect of frequency. Fosgerau, Hjorth, and Vincent Lyk-Jensen (2007) assume a larger effect for headways below twelve minutes than for headways above twelve minutes. Identical Anderson, Nielsen, and Prato (2014) founds weight of 1,3 to 2,8 for headways up to six minutes and 0,6 to 0,7 for headways above six minutes in comparison to bus IVT. According to Anderson et al. (2014) this is caused by the timing of their arrivals. This seems to be more a weighted waiting time effect than only a specific frequency effect. Furthermore it is conflicting with the findings of Goeverden and Van Den Heuvel (1993), in their study a larger headway is expected to have more impact. Intuitively the changes of very low frequencies have more effect than the changes of high frequencies. Schakenbos (2014) found that changing the frequency of the connecting mode from four to three results in increased generalized travel time of about five minutes and changing it from four to two results in an increased generalized travel time of about nine minutes. This seems to be more the effect of added insecurity of an increased travel time due to a missed transfer. Identical to this Szeto, Solayappan, and Jiang (2011) found that travelers with a high risk aversion level are more sensitive to frequency changes than travelers with a lower risk aversion. Based upon this information it does not seem to be possible to relate a frequency effect directly to an IVT.

For this thesis the findings above lead to the following implications that can be used to validate or compose the trip utility parameter weights. It can be beneficial to distinguish the weight of travel time for modes of transportation. Whereby the train is preferred above the bus The BRT is in-between the train and the bus and the position of the car is unclear. The disutility of experiencing a transfer can be weighted identical to 9 minutes of IVT. The weight of waiting and access and egress can differ a lot but is in general experienced to be less pleasant than IVT therefore it will be weighted 1,5 times higher than IVT time. The effect frequency cannot be weighted directly towards IVT based on literature findings. This will be discussed with more detail in the next chapter.

The original HOV-scanner scanner is slightly more conservative on by weighting access and egress 1,2 and waiting 1,3. The biggest difference is however the implementation of transfers. A value of 1,1 is used for the transfer time in minutes. Measured as the time between arrival and departure. This is a combined waiting and transferring effect and therefore 1,1 seems to be too low.

6. Research approach

Details regarding the research approach can be found in this chapter. In the first section the approach is elaborated. An assessment framework can be found in the second section. The third section elaborates the method used for route set generation and the fourth section covers the choice modeling method. The third and the fourth section will contribute in answering the question 1.2: What is a correct way to implement parallel PT connections in a model?

6.1 The research layout

To achieve the research objective and to answer the research questions a research approach is composed. As part of the research the current HOV-scanner is analyzed and an introducing literature study has been performed. To proceed the following topics have been covered:

1. Route set construction method
2. Choice modeling method
3. Routes set program process
4. Choice modeling process
5. Validation
6. Answering the research questions
7. Conclusions and recommendations

First it is investigated how route set can be constructed. The next step is selecting the most suitable choice modeling method. This results in a suitable method to implement parallel PT in the HOV-scanner. Once the method is selected a route set generation program can be constructed and evaluated. The aim of the program is to generate the route set and the necessary information while the workload is reduced compared to the original HOV-scanner. Once the program is finished the constructed route sets can be used in the choice modeling process. In the choice modeling process the new HOV-scanner will be used for the first time and different model configurations will be compared. The most promising model will be validated. By Performing these steps makes it possible to answer the research questions and to give an overall conclusion and recommendations.

6.2 Assessment framework

The assessment framework consists of criteria for the new HOV-scanner. Derived from these criteria specific criteria for elements of the HOV-scanner will be composed. The element specific derived criteria are elaborated in the corresponding section. The new HOV-scanner must meet the following requirements:

- The choice modeling must include both mode and route choice. In extension of the original HOV-scanner the PT trips should be distributed among the different routes.
- The generated route set must be realistic. A realistic route set is the basis for correct distribution modeling and trip forecasting.
- The workload must be reduced in comparison to the original HOV-scanner without becoming less accurate than the original HOV-scanner. This requirement conflicts with the aim to extend the HOV-scanner.
- The output of the HOV-scanner must be suitable to use as a first indication to determine the feasibility of a new PT connection. Therefore it must be accurate enough to fulfill this purpose, but a high accuracy like a full scale model is not required.

6.3 Route set method

The current version of the HOV-scanner uses the quickest PT route as an input for the PT GC matrix. However for many PT trips there are parallel routes available. This could be other lines, other PT modes, other transfer stations, another access/egress station or a combination of these factors. Including parallel routes can make the model more realistic and therefore it contributes towards research question 1 and 1.2.

Not every route from origin to destination is considered by a traveler. Some alternative routes are too long, not known or not used due to other reasons. The set of routes that is considered is called the choice set. For extending the HOV-scanner from a single PT connection towards multiple PT connections first the set of routes (choice set) must be composed. There are different methods and programs to generate the route set. In this paragraph the most suitable method will be selected.

6.3.1 Assessment framework for the route set method

Before the methods and programs can be compared an assessment framework should be composed. This framework is an extension of the framework proposed in paragraph 6.2 and specific for route set generation. The framework describes the requirements for the route set and the routes within the route set. A suitable method can handle these constraints.

According to (Fiorenzo-Catalano, 2007) a single route should apply to the acyclic criterion, the detour criterion and the hierarchic criterion. A route within the route set must should apply to the overlapping criterion, the comparability criterion, the detour-max criterion, the detour-min criterion, the choice set size, the spatial variability and the preferential variability. Nevertheless not all of these criteria seem to be applicable for a PT network. The different criteria will be discussed below.

A hierarchical order could be the use of local road, secondary road and highway in this order. For a car trip it is reasonable to assume that the hierarchy from local road to highway is used from low to high and back without strange transitions in between. For PT different modes can be distinguished e.g. local bus, BRT and train. This criterion does not seem to be suitable for PT trips. A realistic scenario would be walkable links in between stops that improves the connection. Another example would be the use of a local train in between two inter-city rail services. Both of the examples do not comply with the hierarchic criterion. Therefore this criterion will not be used.

An acyclic route is a route that does not contain a loop. The most obvious loop in PT is using the same transit line in the opposite direction after a transfer. But also small loops could occur in an automated root search program whereby the same station is passed twice. It is presumable that travelers avoid loops in their trips and therefore the acyclic criterion will be included.

A detour is the extra time or distance above the shortest route. The detour criterion seems to be a useful option to narrow all the available routes down to a choice set that contains realistic options. After all if a route is much longer than shortest route the much longer alternative is presumably not considered as an option. However the threshold value for the detour is uncertain, guidelines cannot be found in the consulted literature. Because preferred value cannot be found a value will be selected. 30% will be used, this seems to be reasonable for shorter routes but could be on the high side for longer routes. The detour criterion it seems to be an effective and realistic criterion to narrow down the PT route set. Therefore it will be included.

Overlap is considered to be a problem in route modeling as described in paragraph 5.2.2. Therefore the overlapping criterion will be included. As proposed in 5.2.2 an overlap boundary 80% will be used. This implies that is a route is for more than 80% identical to another route it will be perceived as identical and one will of them be deleted. A route segment is defined as identical if the line and the travel time between two stops are identical.

The comparability criterion makes sure that the differences between two routes within the set are not unrealistically large. This criterion is comparable to the detour criterion but can also be applied to other units (e.g. cost and distance). For PT the number of transfers is an example of comparability. If the number of transfers is high it is likely that an alternative is ignored even if the travel time is low due to the fact that a transfer is including additional uncertainty to the travel process. The comparability criterion will be implemented by including a transfer threshold with a maximum of three transfers.

The detour-max criterion is similar to the detour criterion and does not offer additional benefits. The detour-min criterion can, for example, be used to eliminate routes on highway with small on ramp and of ramp detours. A comparable application for PT does not seem to be present. Therefore this criterion seems to be more suitable for road use. These two criteria will not be implemented.

It is important that the choice set size is limited to avoid unrealistic large choice sets and to avoid differences between the generated set and the perceived set. For example Lanser (2005) reports differences from 63 generated routes to 2 reported subjective routes. In case of the new HOV-scanner the number of alternative routes generated in the PT network is expected to be much lower than the number of alternatives in a car or bicycle network. Therefore the route set size is not considered to be an issue and will be ignored. Nevertheless it is necessary to check during implementation if these route sets are avoided.

Both spatial and preferential variation of routes are the result of differences in perception and preferences of travelers. For car routes these criteria could be suitable because the variation in routes can be large. For PT routes on the other hand the variation is much lower due to the more restricted network. If the routes within a route set comply with the criteria mentioned above than they are already different within the provided boundary. Therefore the spatial and preferential criteria is expected not to be useful for a PT

network and it will not be included.

The criteria mentioned above are composed for a car network, but some criteria are also applicable for a PT network. For a PT network there are other criteria relevant that are not be considered for a car network. The specific criteria for PT networks will be discussed below.

The PT network must be accessed and egressed because PT stops are seldom the origin or the destination of a trip. There are a number of combinations possible regarding access and egress modes for the PT system. For example the supernetwork approach of Sheffi (1985) discussed in paragraph 5.1. The car is only responsible for access and egress share of 14% and 8% of all train trips (Van der Heuvel, 2013). Due to the low share and the intention to reduce the work load it seems to be reasonable to ignore the car as access and egress mode and only use walking and cycling. Therefore the chosen method and program must be able to use these two modes.

A PT network is contradiction to car a network restricted by the schedules of the service provider. It is mentioned in the literature section (5.4) that a schedule based approach will be used. This is not a direct route set criterion but the selected method and program must be able to handle the schedule based PT information. Therefore it is included here.

In car networks the car trip is usually not intermitted by another mode. For a PT trip on the other hand it can be beneficial to include a small walking trip to connect two different PT stops. For the application of walkable links see Florian and Constantin (2012) and Lanser (2005). To incorporate this effect the used method and program must be able to handle walkable links.

The selected criteria should contribute towards a realistic route set. Summarizing the above the route set and the including routes must be composed in compliance with the following criteria:

- Acyclic: Routes that contain loops should be avoided.
- Detour: Trips longer than 30% of the shortest route will be excluded.
- Overlap: Routes with overlap above a threshold value are considered as one.
- Comparability: Differences between routes should not be unrealistic large. The maximum number of transfers is implemented as a comparability criteria.
- Schedule based approach: The selected method should be able to handle schedule based data.
- Access and egress by cycling or walking: The selected method should be able to handle access and egress by both cycling and walking.
- Walkable link inclusion: The selected method should be able to handle walkable links between two stops.

6.3.2 Available methods and programs

In the literature section (5.2) a set of route set generation methods are described. The methods mentioned are: shortest path method variants, different Monte Carlo approaches, a labeling approach, a branch and bound technique and combinations of these methods. Besides the available methods there are programs and companies that perhaps could provide the desired route set. The investigated programs and companies are 9292OV, Google Maps, Omnitrans or Arcgis.

6.3.3 Assessment

The most straightforward solution would be the use of existing software or a service to get the desired route set. 9292OV is an obvious company to contact for travel time and trip parameters since they are familiar with processing PT schedules and they are able to provide detailed trip information for PT users. 9292OV has been contacted with the question if they can provide GC matrices that contain travel time for new non-existing PT lines. Unfortunately 9292OV was unable to deliver the required information regardless of the provided information.

The trip planner used in Google maps uses GTFS files to calculate the travel time for public transit. GTFS stands for General transit feed specification. It is a common format to exchange PT schedule information. Google offers users the possibility to test their new GTFS file. The idea was that this option could be used to extract the travel times for new (nonexistent) PT connections. Unfortunately this was not possible. Google only allows the users to perform a consistency test for the data. The next tests are done by Google.

In OmniTRANS version 6.1 and above it is possible to import GTFS files and to process schedule based travel time information. It is unclear if schedule based information without importing GTFS data can be

processed in versions below 6.1. Omnitrans is using the Zenith method to distribute trips among PT lines. Zenith is a frequency based approach and therefore it is probably not possible to distribute trips schedule based by default. Furthermore there is no OmniTRANS license present within MuConsult. The amount of work that could be done with OmniTRANS is not enough to justify the license cost. Therefore this is not the preferred solution for MuConsult. As a result the use OmniTRANS will not be investigated any further.

ArcGIS has included a route analyzer that can find the shortest path. The shortest path can be set to shortest in time or shortest in distance. Furthermore there are additional packages available to process GTFS files in ArcGIS (developed by M. Morang⁴). Unfortunately only the shortest path is given. A route set cannot be generated by this program. Therefore this is not a suitable option.

None of the proposed programs or services comply with the proposed framework. Therefore they cannot be used.

Next the methods will be assessed. The most suitable method can be selected to construct a route set generation program. In the literature section (5.2.1) it is mentioned that the methods that change network attributes are in general not preferable for a PT network, but that the labeling approach, the essentially least-cost paths or the constrained enumeration method are preferred.

The labeling approach has some advantages over the other methods. It is possible to select a predefined number of routes that each have their specific characteristics. A disadvantage is that detour, overlap and comparability cannot be incorporated. Furthermore is the predefined number of routes less suitable if there is much difference in network composition and route length. Therefore it does not comply with all the criteria mentioned in the assessment framework.

The constrained enumeration method is a branch and bound method and therefore computational intensive Prato and Bekhor (2006) were able to find differences of 1.5 to 10 hours compared to a labeling approach. Nevertheless computation time is not considered to be an issue since the number of PT routes is in general much lower than the number of routes in a car network. On the other hand a main advantage is that all possible routes are considered and that due to the implementation of constraints it does comply with all the proposed criteria.

The essentially least-cost paths method is regarding computation effectiveness comparable with the constrained enumeration method. Furthermore can the required constraints be implemented. A disadvantage of the approach is that empirical test data cannot be found in literature.

6.3.4 Conclusion

None of the programs proposed in this chapter comply with the framework. Therefore these programs cannot be used for the HOV-scanner. This means that a custom made program is required. To create the program a suitable route set method must be selected. Of all the proposed methods the constrained enumeration method and the essentially least-cost paths seems to be the suitable. Of these two methods the constrained enumeration method is preferred because good results were obtained in other studies (i.e. Prato and Bekhor (2006)) The constrained enumeration method searches for all possible routes and constraints can be set to comply with the proposed framework. The computational burden is a disadvantage of this approach. However this is not considered to become an issue due to the fact that the number of routes in a PT network is much lower than in a car network.

The route set and the including routes must be composed in compliance with the following criteria: acyclic, detour (30%), overlapping (80%), comparability (maximum of three transfers), schedule based approach, access and egress by cycling or walking and walkable link inclusion. The constrained enumeration method is able fulfill these criteria.

6.4 Choice modeling method

The HOV-scanner uses a reference situation and a future (scenario) network to forecast the PT usage. In this section the choice modeling method for the new HOV-scanner will be selected and explained. In most cases only limited information on PT trips is available. The NRM/LMS data is used as input data, only the total number of PT trips (figure 9, all PT trips) is known but the distribution among different routes is unknown. Nevertheless the aim of the new HOV-scanner is to forecast the trips distribution over the different routes.

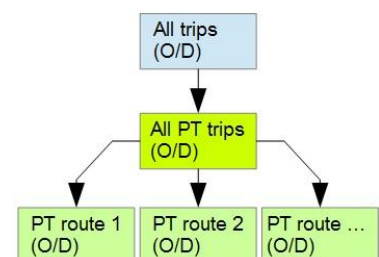


Figure 9: PT trips

⁴ <http://transit.melindamorang.com/>

The main concern in this case is how a correct choice model can be estimated on such limited input information. A correct choice model is important because it contributes towards more accurate results. And thereby it contributes towards research question 1 and 2.1. To construct a correct choice model it is important that:

- A model is composed that is able to distribute the PT trips over the PT routes, using the available input data.
- The changes in PT service can be translated correctly into a new mode choice distribution to forecast in PT usage.

6.4.1 Data requirement

The data availability can differ from case to case. For report it is assumed that the following data is available:

- Trips per OD relation for all modes. Source: preprocessed NRM/LMS data.
- Routes and route characteristics. Source: route matrix (output of route generation process).
- GC matrices for the car and the bicycle, whereby the GC is expressed in minutes of travel time.

6.4.2 Assessment framework

To compose and evaluate the proposed method a framework should be composed. This framework is an extension of the framework proposed in paragraph 6.2 with specific additions for choice modeling. The model used to model mode shares must be able to generate mode shares that are as close to the input data and at least just as good as the original version of the HOV-scanner. The parameter values of the model should be realistic and in line with values found in literature.

The choice modeling must include both mode and route choice. In extension of the original HOV-scanner the PT trips will be distributed among the different routes. It is important that the distribution of the trips along the routes is modeled realistic. However, if the actual distribution for the reference scenario is unknown it is difficult to verify this. Therefore the used approach must match the literature and it should be checked if the outcome is in line with the expectation.

The workload must be reduced in comparison to the original HOV-scanner without becoming less accurate than the original HOV-scanner. For the choice modeling method this implies that it is important to have a method that does not require more set-up time than the original version and it should be convenient to process multiple scenarios. Summarizing the above the following criteria will be used:

- The model must be able to reproduce mode shares in a correct way and the model parameters should be realistic.
- PT route choice must be included with realistic route choice modeling.
- Processing new cases or scenarios may not require too much processing time. This implies that the method can easily be transferred to a new situation.

6.4.3 Approach

To forecast the effect of a new PT scenario a four step approach is conducted:

1. Set up and apply the choice model for the reference situation.
2. Calibrate the outcome of the choice model for each OD relation and mode, for the reference situation.
3. Calculate the shift in utility for the PT system based upon information that is retrieved in the first step.
4. Forecast the use of the PT system for the scenario.

The first step is used to set up the choice model for the reference situation. With this model the mode shares will be reproduced and the effect of the trip components will be revealed. The effect of the trip components is captured by the height of the β 's for the individual trip components. Small errors in mode shares will be corrected in the second step. The last two steps are used to forecast the effect of a scenario. With the effect of the trip components found in the first step the shift in utility for the PT mode is calculated. This is the difference between the reference and the scenario situation. The shift in utility is used to calculate the shift in mode shares and to estimate the route shares. In the fourth step the shifted mode and route shares will be used to redistribute the trips. This results in a forecast of the effect of a scenario.

Step 1: Set up and apply the choice model for the reference situation

To get an accurate fit of the model it is preferable to know how the input data of the NRM/LMS is modeled. Therefore this will be elaborated first. Furthermore this could be used to explain differences between the NRM/LMS data and the output of this first step. Once the details on the NRM/LMS model are known the most suitable choice model will be selected to estimate route shares. In the last section of this step the preferred use of trip components will be elaborated. This should result in a model specification that fits to the NRM/LMS data.

NRM/LMS model

The NRM/LMS data is distributed by using a (Nested) Logit model. It is able to distinguish the following modes: car driver, car passenger, BTM (bus, tram and metro), train, bicycle and walking (figure 10). Whereby the car trips are modeled the most accurate and the other trips are modeled with less accuracy. Therefore the data on car trips is more accurate than the information on PT trips. The train trips are modeled by the station choice model. This model consists of: the access and egress model, the station choice submodel and the train choice model. For the access and the egress model the costs and travel time add up to a generalized travel time. For access and egress travel time the free flow travel time by car is used. The station choice submodel consists of: the generalized access and egress times and the number of connections per origin and destination station. The train choice model is a function of: travel time in train, cost of the train trip, value of time (VOT), waiting time and a transfer penalty (5 minutes). The NS station matrices are the input for the NRM/LMS. It should be noted that the consistency between the synthetic matrices and the NS-station matrices is not checked in detail for all stations. Only the top 10 access stations have been checked. For the train costs $1,5 + 0,1 \cdot \text{distance in kilometers}$ is used. The VOT and the cost are distinguished per mode, group and motive. Three periods are defined for train and BTM: morning rush, evening rush and the remaining day (Tavasszy et al., 2012a). For BTM modeling the skim matrices of the NEA are used (Tavasszy et al., 2012b). Time and cost are important factors. For BTM access and egress a fixed value depending on the urbanization rate is used (Hofman & van Grol). An issue could be the accuracy of the PT trip modeling in the NRM model. This effect will be analyzed in chapter eight where the choice model will be implemented.

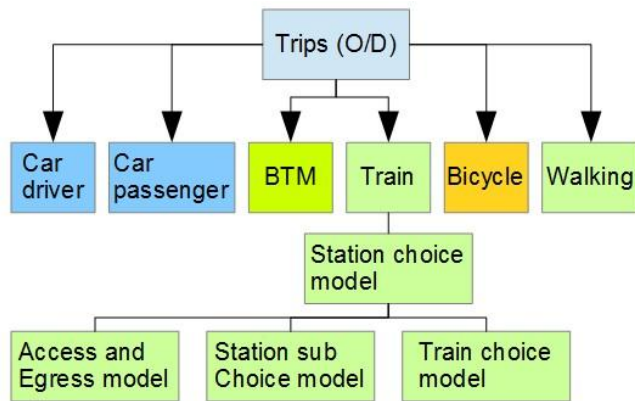


Figure 10: NRM model

Choice model

It is likely that a similar choice model to the one that is used to generate the NRM/MLMS data leads to the best reproduction of the input data. Therefore a Logit model will be used. Different Logit models could be used. In the literature section the following models were suggested: Multinomial Logit (MNL), Nested Logit (NL) and Mixed Logit (ML). Tavasszy et al. (2012a) mentions that ML is not used for the NRM/LMS model. Furthermore data of good quality is required for a ML model according to Hensher and Greene (2003). PT trips are not modeled with that much accuracy in the NRM/LMS model. This presumably affects the quality of the data. Therefore ML is not the most suitable Logit model. In a NL model the PT alternatives

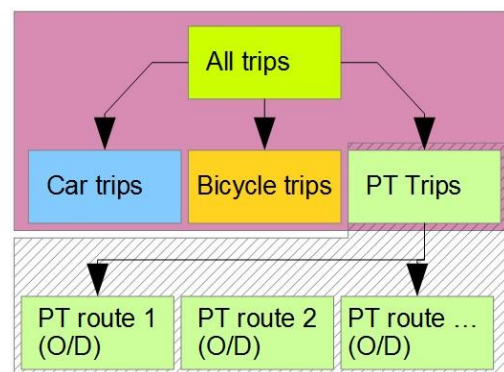


Figure 11: All trips

can be part of one nest (figure 11). This seems to be realistic because all PT alternatives are expected to be strongly correlated. This results in a model with the mode choice: car, bicycle and PT and within the PT nest different routes can be chosen. Because of this effect the NL model is preferred over the MNL model. If the NL model does not provide good results a MNL model will be used. This MNL model is identical to the original HOV-scanner and therefore it is expected to work. For the MNL model the fastest PT route will be used which is identical to the original HOV-scanner. To summarize the NL and MNL models will be

investigated, because they match the design of both the input data and structure of the choice model. Whereby the NL model is preferred. Biogeme will be used to calibrate the model. This is an open source freeware program for the estimation of discrete choice models.

Trip components

In the literature section (5.4) there are suitable trip components suggested to fit the model. For PT trips the trip components are: IVT, number of transfers, access and egress time, waiting time, frequency and mode differences. These components will be used by preference to create the choice model. However it could be possible that this approach does not result in a good model fit. A worst case scenario would be that only mode preferences (car, bicycle and PT) and a total travel time effect can be modeled. This could be caused by the differences in model structure and the used data, despite the fact that an identical choice model is used. If only the total travel time can be modeled a method should be introduced to refine this information. Otherwise the effect of change in trip components (e.g. the number of transfers) cannot be retrieved correctly.

To overcome this limitation the sensitivity for total travel time must be transferred into the sensitivity for individual trip components. This will be done by using the generalized cost for PT instead of the total travel time (see formula below). The PT travel time equivalent is composed out of the trip components multiplied with their weight. The weight for the travel time components are extracted from literature.

$$GC_{PT} = ivt_{PT} + 1.5 * acc_egr_{PT} + 1.5 * wait_{PT} + 9 * transf_{PT}$$

$GC_{PT} =$	Generalized cost for PT
$IVT_{pt} =$	in vehicle time for PT
$acc_egr_{pt} =$	access and egress time for PT
$wait_{pt} =$	waiting time for PT
$transf_{pt} =$	number of transfers for PT

The β_{GC_PT} can then be transferred to β per trip component by multiplying β_{GC_PT} with the weights used to calculate GC_{PT} (e.g. $\beta_{trans_PT} = \beta_{GC_PT} * 9$)

The result

To summarize the first step: A NL or MNL will be used to model the choice process. Whereby the NL model is preferred because it is expected to better match the actual choice, due to the implementation of mode and route choice. The trip components suggested in the literature section will be used for the choice model.

However if the effect of trip components that result from model calibration are inconsistent with the values found in literature the travel time by PT will be replaced by the GC_{PT} . The result of this step is the sensitivity towards change in trip components and a model that can estimate mode shares.

Step 2: Calibration of the reproduced data

In this step the utility per mode for each OD relation will be adjusted to match the mode shares provided by the input data, i.e., differences in distribution that result from the first step will be corrected. This is done to correctly reproduce the mode shares provided by the input data. This is done based on the assumption that the trip distribution provided by the input data (LMS/NRM data) is correct. It is revealed earlier in this section that the PT trips in the NRM model are not always modeled correct. Nevertheless when more accurate information is unavailable this trips distribution will be used.

The method below describes the car calibration. For the other modes the process is identical. This calibration step results in correction factors that can be used to calibrate the trip distribution by effecting the utility per mode. The calibration method is implemented in the original HOV-scanner and is adopted because good results are obtained.

$$\text{Calibration value}_{OD}^{car} = \log \left(\frac{\text{car share}_{OD}}{\text{car probability}_{OD}} \right)$$

For example:

Utility car= -0,5 Utility PT= -2 Utility bike= -2

This results in a probability per mode of: PB car= 0,69 PB PT= 0,15 PB bike= 0,15

The input data has a modal split of: share car= 0,75 share PT= 0,125 share bike= 0,125

$$\text{Calibration value}_{\text{OD}}^{\text{car}} = \log\left(\frac{0,75}{0,69}\right) = 0,081$$

This results in a calibrated car utility of -0,5+0,081= 0,419.

Step 3: Calculate the shift in utility for the PT

For each OD relation a reference route set and a scenario route set is available. Furthermore, there are mode preferences and sensitivity for travel components known as a result of the first step. At step one there is discussed that a MNL or a NL model are expected to provide good results and that NL is preferred over MNL. Therefore, a NL model will be used at this step. The shift in the utility for PT is calculated by multiplying the effect of the scenario with the sensitivity towards trip components.

For example:

Reference trip

IVT= 20 minutes

$$\beta_{\text{ivt}} = -0,04$$

Transfer= 1

$$\beta_{\text{transf}} = -0,36$$

Scenario trip

IVT= 25 minutes

Transfer 0

Utility reference $20 \cdot -0,04 + 1 \cdot -0,36 = -1,16$

Utility scenario $25 \cdot -0,04 + 0 \cdot -0,36 = -1$

The shift in utility for the example is +0.16. This means that the model predicts an increased PT share. In this example only one PT route set is compared with changes in IVT and transfers but the model works identical for multiple trip components and routes.

If there are multiple routes the utility for the PT nest must be calculated. When a MNL model is used in the first step this cannot be done due to the absence of the logsum parameter. To overcome this problem a logsum parameter can be assumed. For the PT nest a strong correlation is expected since the alternatives within the nest are quite similar. Of course there is some variation between the trip length, modes used, and other parameters but in general there will only be small differences between the routes in the nest. A default value for the logsum could not be found in literature. Due to the similarity of the routes a low value of about 0,1 to 0,5 is expected. Different logsum parameters between 0,1 and 0,5 will be applied and compared in chapter eight. The different logsum values will be compared to assess the effect of the logsum and to obtain the most suitable value.

Using multiple routes in a nested model introduces a new issue. In general it is assumed that a higher frequency makes the PT alternative more attractive in comparison to the other modes. Nevertheless, the frequencies of routes within a route set cannot be summed or extracted directly to get an overall frequency of the route set. If the route set for example consists of two routes that both have a frequency of four times an hour, the frequency of the route set could be four if both routes have identical departure times. It could be eight if they are perfectly distributed over the hour. And it could be perceived somewhere in between if they are not perfectly distributed. Furthermore, the literature section did not provide a clear overview of the weight of the frequency in relation to the IVT. Since it is preferable to include the frequency it will be included in the modeling of the data. However if the inclusion leads to unrealistic model parameters it will be excluded.

Step 4: Forecast PT usage for the scenario

The data from the first three steps can be combined to predict the PT trips for the scenario situation. This is done by calculating the new utility. The new utility for the PT is the sum of the reference utility, the correction factor and the shift in utility. This can be calculated for each OD relation. The share per route can be extracted by combining scenario trip utility of the third step and the number of PT trips found in this step.

6.4.4 Assessment and conclusion

The outcome model estimation and model fit will be investigated in chapter eight. Therefore the accuracy of the model cannot be assessed now. Nevertheless, due to the use of a NL or MNL the approach is similar to NRM/LMS model. This contributes towards a good model fit. Therefore, no irregularities are expected for the result of the first step. The aim of the second step is to correct the output of the first step. If the model fits the data well only a small correction is needed. This step cannot affect the outcome of the HOV-scanner in a negative way.

According to the assessment framework the trip distribution among the PT routes should be realistic. If the actual distribution is unknown it is difficult to assess the distribution. Therefore trip distribution seems to be a main concern regarding choice modeling. The trips will be distributed over the route sets using a NL model, fitted trip components and trip component relations found in literature. The value for the logsum is uncertain. This value affects both total PT share as well as the distribution among the routes. Additional research regarding this subject is recommended. It is preferable to have more detailed trip distribution information to verify the assumptions made in the third step and to calibrate the logsum parameter.

By using this approach the processing time is not expected to be an issue. Presumably the most time consuming step of the choice modeling will be the first step. The method itself can quickly process multiple scenarios and it is transferable to other cases. Especially, the possibility of assessing multiple scenarios conveniently is an advantage over the original HOV-scanner.

Correct choice modeling on limited information was the main concern in this chapter. The proposed four step approach is expected to give the desired results without neglecting this concern. Nevertheless, a case study or validation case is required to assess the accuracy of the approach.

7. Route set program

In section 6.3 it is made clear that none of the tested programs complies with the framework nor that they are preferred by MuConsult. Therefore a program has been made in MATLAB to generate the desired route sets. Of the suggested frameworks in section 6.3 the constrained enumeration method is selected as the most suitable method for route set generation. Therefore this method will be used. In this chapter first the trip layout and program design will be discussed. Subsequently the program will be tested and analyzed with the data from a case study.

7.1 The trip layout

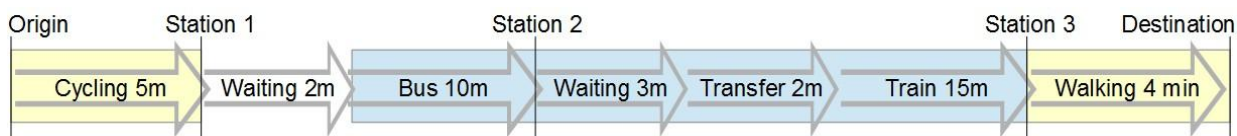


Figure 12: Example PT trip (time in minutes)

In this chapter different terms will be used to describe a trip. To avoid differences in assumptions the terms will be explained here:

- Total trip: This is the trip from the origin to the destination.
- PT trip: This is the trip from boarding the first PT vehicle to leaving the last PT vehicle. Represented as the blue square in figure 12.
- Transfer time: This is the total time spend on transfers. For each transfer there is a prefixed transfer time. The total transfer time for a trip is the amount of transfers in between PT vehicles multiplied with transfer time. In the example (figure 12) there is one transfer of two minutes.
- Waiting time: The total time spend waiting. In the example this is four minutes, two before the first departure and two between the bus and the train.
- Access and egress travel time: The time it takes to get from the origin to the first stop (excluding waiting time) and from the last stop to the destination. These are the yellow squares in the example. Furthermore the time spent on walking in-between PT stops is counted as access and egress travel time as well.

7.2 The program design

The use of input files, data processing, parameter usage and the output is discussed in this section.

7.2.1 Input

The input file with PT schedule information must be as close as possible to default PT schedules to minimize workload. Furthermore, additional required information must be minimized (e.g. detailed access and egress information). The amount of PT lines, PT stops and OD relations must be flexible. This should be done to make the program suitable for multiple scenarios. The input elements are:

- An excel sheet that contains the schedule information.
- A table that contains distances between centroids and stations. Only if the relevant combinations are required in this table.
- A table that contains distances from station to station. Only if the combinations are relevant regarding walkable links between stops.

7.2.2 The structure

The program distinguishes two segments of a total trip:

1. From the origin to all stops within walking or cycling distance of the origin and from the destination to all stations within walking or cycling distance of the destination (figure 13, light blue squares).
2. The PT travel times for all possible combinations of stations in the specific OD relation (figure 13, large yellow rectangle).

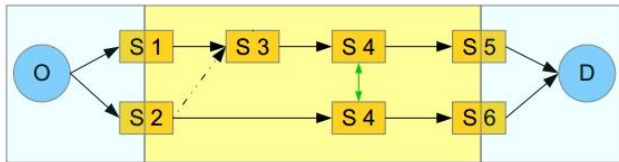


Figure 13: Example program layout

The station and OD information is retrieved from the input files. Therefore the included stations and OD's are restricted to the study area. For the PT trip the program calculates the route sets between all stations that are related to an origin or a destination. The origin and the destinations are the centroids of the corresponding zones. The use of walkable links and transfers are included as displayed in figure 13. For each station to station combination the route set will be saved in the PT matrix. From station 2 to station 5 the program will find two possible routes: One that includes the walkable link from 2 to 3 and one that includes the transfer to the other line at station 4.

Once the PT matrix is finished this matrix will be used to OD matrix that contains the route sets for all OD combinations. In this step the origins and destinations are connected via the suitable stations and routes of the PT matrix. An origin or destination can be linked towards multiple stations in this step.

7.2.3 Data processing

In this section the used data processing method for the program is explained. For more detailed information regarding the way the program processes the information see appendix 3. Note that the information presented in appendix 3 only includes details regarding the code and data processing.

The program starts with reading the input and creating the OD matrix (figure 14, 1) and the PT matrix (2). The origins and destinations are extracted from the input data and only the stations that are related to an origin or a destination are included in the PT matrix.

First the PT matrix is filled with route sets according to the following process. The first station combination in the matrix is set (3) and then the route set is constructed for this station combination (4). Once this is done the route set is filtered (5) and if necessary this process repeats with the next combination. If the entire PT matrix is processed all route sets from station to station are known.

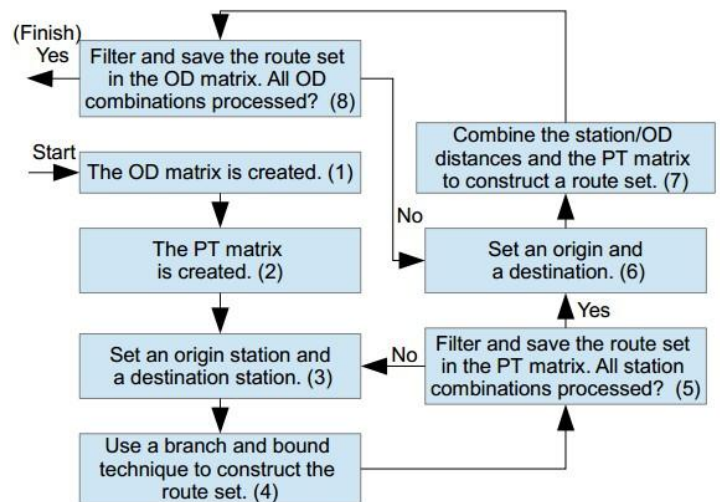


Figure 14: Data flow

After filling the PT matrix the OD matrix is filled. This process starts by selecting the first OD combination (6). The program combines the OD to station distance input information with the PT matrix from the previous step. A route set with all possible combinations is created (7). This route set is filtered and saved (8). If necessary this process is repeated with the next OD combination. If all OD combinations are processed the process is finished. When the program is finished the OD matrix contains the route sets from origins to destinations and the PT matrix contains the route sets from station to station.

7.2.4 Adjustable parameters

The program offers multiple parameters that can be adjusted to suit specific scenarios or needs, the values used for the case study are included:

- Departure time, set to 08:00 to get peak hour travel times.
- Tolerance on the departure time, set to 15 minutes to include both route set with xx:10 and xx:15 departure time.
- Interval on the departure time, set to 7 minutes. For example a departure time of 08:00 with a 30 minute tolerance and a 10 minute interval will result in a departure time of 08:00, 08:10, 08:20 and 08:30. For each of these times the program will search for the first departure time on each line.
- Cycling speed, set to 15 km/h.
- Walking speed, set to 5 km/h.
- Maximum cycling distance, set to 3 kilometers. The value is found in literature.
- Maximum walking distance, set to 1 kilometer. The value is found in literature.
- Transfer time (the time it takes to transfer from one line to the next line), set to 1 minute. One value is used for all transfers. If there are more time consuming transfers this value can be set higher but for the “Duin en Bollenstreek” is assumed to be sufficient.
- The buffer time before the first departure (the amount of minutes the traveler is present at the station before the departure time of the vehicle), set to 2 minutes. For the first stop an additional time is required, to be sure that first vehicle can be boarded.
- Maximum amount of transfers, set to 3. In general having more than three transfers does not seem to be realistic.
- Maximum amount of times the transit loop is repeated, set to 4. This is related to the number of transfers but can be adjusted for test purposes.
- Overlap factor, 0,8 is used. The value is found in literature.
- Detour factor, this is a factor that is used to exclude routes, 1,3 is used. E.g. if the detour factor is 1,3 and the shortest route is 30 minutes. All trips above 39 minutes will be excluded. An exact value could not be found and therefore 1,3 is assumed.

7.2.5 Output

The program determines a route set for each OD relation. For each route within the set trip parameters will be generated. The following trip parameters will be generated:

- In vehicle time (IVT). In addition to the original HOV-scanner it is possible to distinguish different modes of transportation. This is used to filter overlap.
- Access and egress time.
- Waiting time.
- The number of transfers.
- Frequency of the route.
- Route information for each route in the route set. Mainly used for analyzing the route set.

For detailed information regarding output and the layout of the output see appendix 4.

7.3 Results

Applying the route set program on the reference situation of the “Duin- en Bollenstreek” results in the route set distribution displayed in figure 15.

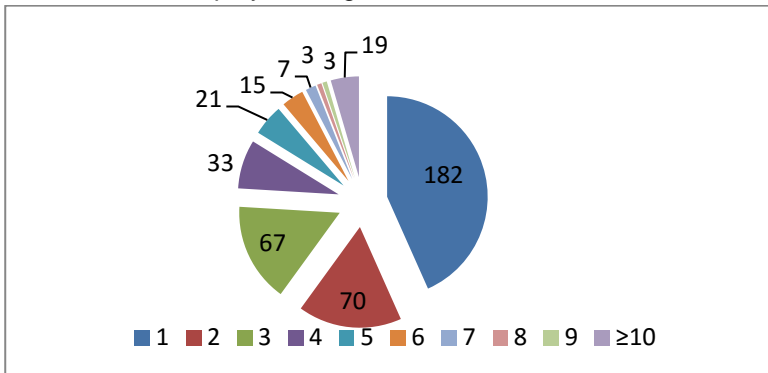


Figure 15: Number of routes in set

There is a total number of 420 route sets. The route set with just one route is still the largest category with 182 route sets. The route set group of two, three and four routes combined is almost the same size. The share of route sets with seven or more routes is low (8%). Having one two or three route options for most of the trips seems to be realistic. To provide more insight in the actual route sets some of them are elaborated in the next section.

From Noordwijkerhout to Amsterdam central station there are nine routes in total. In figure 17 the shortest three and the longest route is displayed. The yellow arrows represent bus lines and the red and green train lines. The time difference between the shortest and the largest route is 18 minutes (76 versus 94). The geographical diversity of the routes is quite large.

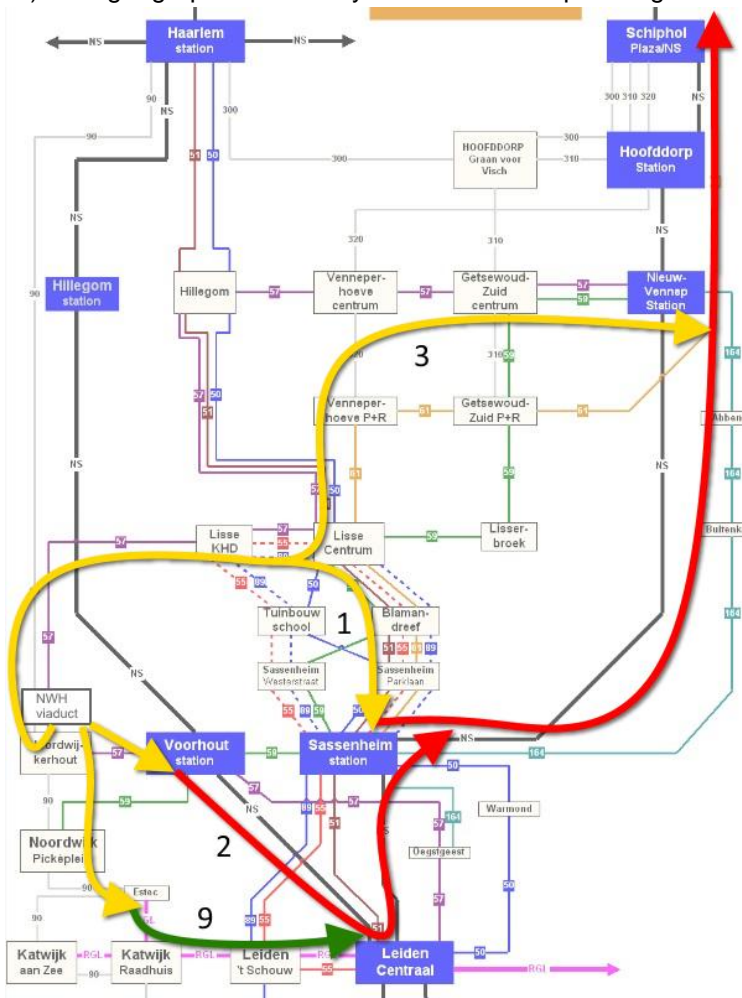


Figure 17: Noordwijkerhout, Amsterdam

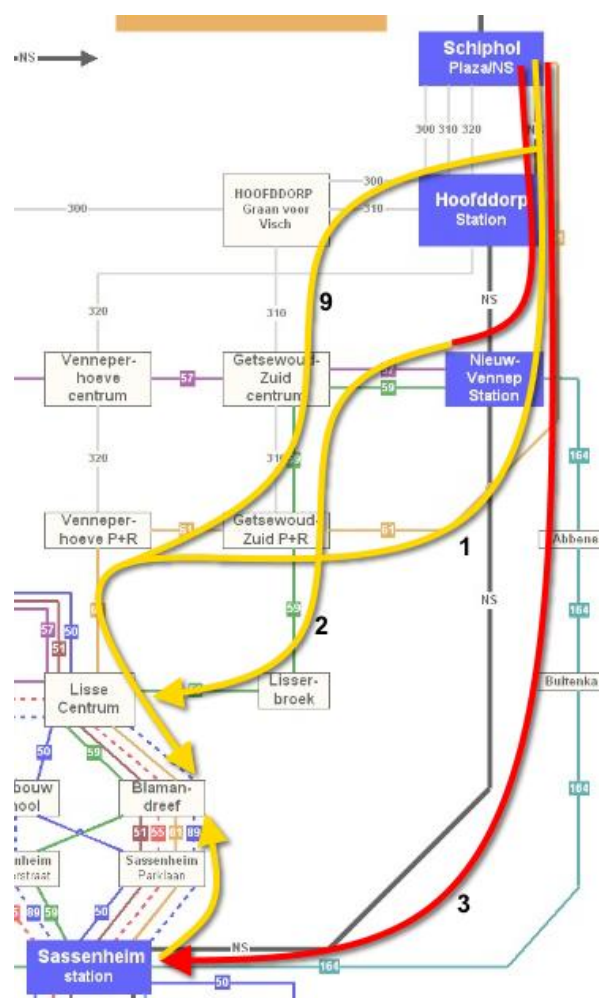


Figure 16: Schiphol, Lisse

From Schiphol to Lisse Gerard Doustraat (close to the Blamandreef). There are 9 possible routes. The shortest one is 49 minutes and the longest one 58 minutes. For the routes see figure 16.

From Hillegom to Lisse there are three direct connections and two of them are considered as choice options (figure 18). The shortest one is 23 minutes and the longer one 25 minutes. They can be considered as two almost identical alternatives. From Hillegom to Lisse Gerard Doustraat there are three alternatives. A direct connection of 33 minutes a connection with one transfer of 33 minutes and a connection with one transfer of 35 minutes.

A two route option with two modes can be seen between Leiden central station and the van Panhuysstraat in Noordwijk (figure 19). The shortest connection is provided via the Rijngouwelijjn. Nevertheless the connection provide by the NS and a bus is only slightly longer (37 versus 38 minutes). A discussion on these route sets can be found in the next section.

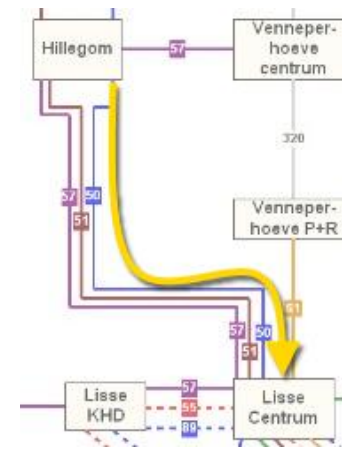


Figure 18: Hillegom, Lisse

7.4 Discussion

The discussion is divided in a discussion regarding the results presented above and regarding the chosen approach.

7.4.1 The approach

The proposed program results in an impressive reduction of processing time for the employee. This despite the fact that there is a route set calculated instead of just one route. In comparison to the manual method the “Duin- en Bollenstreek” route set can be calculated about five times faster. This processing time is mainly affected by the maximum amount of transfers and the number of start times. From a computational point of view there are more effective methods. However, the main goal was to reduce the time an employee was burdened by creating the route set.

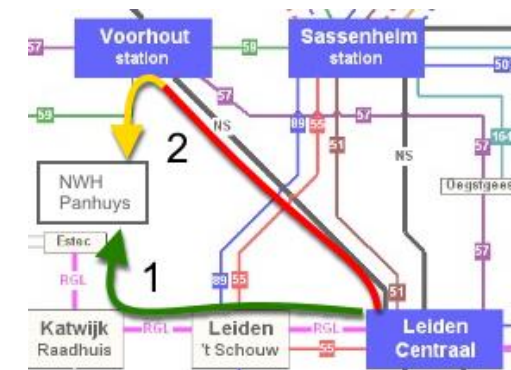


Figure 19: Leiden, Noordwijkerhout

The route modeling within the PT part of a route is extensive since all combinations are considered. The access and egress part of the trip is modeled more simplified. Only walking and cycling are considered as possible options and a strict boundary is used between both modes. The used values are in line with literature but some simplifications have been made. For example, the car and scenario specific access and egress modes are ignored. Scenario specific access and egress is for instance the effect that a bicycle is more frequently used as an access than as an egress mode. This simplification could lead to misperception. The effect of these simplifications are unclear. Nevertheless with the available information it would be difficult to avoid these simplifications. In comparison to the original HOV-scanner some improvements have been made. The original HOV-scanner only allowed for one access and egress mode. Walkable links between stations were ignored.

Another advantage of the new HOV-scanner is the extra check that is included due to the use of the program. In the old situation the user checked his own work. In the new situation the user checks the program output. It is likely that verifying the work other than your own work is more effective.

7.4.2 The results

The route set generated in this chapter applies to the proposed criteria. Nevertheless, the question could arise how realistic these route sets are.

Most of the OD's (76%) consist of one, two or three routes. The share of route sets with seven or more routes is below 8%. This seems to be realistic since there are just a few good route alternatives for most of the PT trips. Furthermore, the number of routes considered by a single user is quite low. In the research approach is mentioned that the generated route set can be much larger than the considered route set. The larger route sets are in general also the route sets that contain the longer routes. In general the shortest routes within these longer route sets are around one hour. Due to the detour criterion of 30% routes until 78 minutes are included in these sets. The combination of the larger distance and longer time span

increases the change of a larger route set. An increase in trip length of almost 20 minutes on an hour trip seems to be quite much. Nevertheless, if route choice modeling is implemented correctly this should not be a problem because it will predict a much lower share for the longer route. Therefore the inclusion of longer routes is not considered to be a problem.

Differences between the route sets that contain short and long routes can be seen in more detail by analyzing the routes they contain. In general the longer route sets allow for more differences in the routes than the shorter route sets. These differences can be geographical or in terms of used modes or a combination of both. The proposed routes are realistic for both the shorter and the longer routes.

7.5 Conclusion

In comparison to the original HOV-scanner a route set instead of a single route is generated. Furthermore the process is automated. To make the route set more realistic the bicycle is implemented as an access and egress mode and the use of walkable links between stations is implemented as well. Due to automation the workload for the employee has been decreased impressively. The computation time has increased, but this is not an issue.

There are some parameters that can be adjusted by the user, for example, transfer time required and walk speed. Other parameters or assumptions cannot be changed, for example, the used criteria and the structure of a trip. When the program is used this should be considered by the user.

The used approach seems to produce realistic route sets. The route sets are in general not that large and within the set the routes differ from geographical locations and by the use of transit lines. The proposed routes by the new HOV-scanner seem plausible and abnormalities within the routes were not found. Nevertheless for the longer trips the detour criterion of 30% seems to allow for large route sets. If route choice is implemented correctly this is not considered to be a problem.

8. The choice modeling

In this chapter the choice modeling approach introduced in paragraph 6.4 will be tested and discussed. Four steps are proposed. The first two are used to fit the model and to capture the effect of the trip parameters. The last two are used to forecast the effect of a scenario. The case of the “Duin- en Bollenstreek” will be used. The chapter starts with an assessment framework and the effect of the proposed framework will be analyzed.

8.1 Assessment framework

To assess the choice modeling method an assessment framework should be composed. This framework is an extension of the framework proposed in paragraph 6.2 and specific details are added for the each step.

The degree in which the model fits the data is an indicator for accuracy of the model and the indicators that are used. Therefore it is important that the input data can be reproduced accurate. To assess the accuracy the fit between the model and the input data can be analyzed. The accuracy will be compared with the original HOV-scanner and by comparing the accuracy of the different models mutual. Furthermore should the parameters that are resulting from the modeling structure be logical and in proportion to each other. A comparison to the literature can be made to evaluate the results.

The trip forecast of a new scenario is the main output for the HOV-scanner. Without a validation it is difficult to assess the output. For the used dataset validation data is not available. Nevertheless two indicators can be used to evaluate the model: The elasticity of the model can be compared towards literature. The output can be judged on the direction and the magnitude of the effect.

8.2 Results

The results are displayed in this section and provided separately for the reproduction and the forecast. First section the model configurations will be evaluated, and the model layout will be elaborated. Then the model fit and the result of model forecast will be presented.

8.2.1 Model parameter estimation

Biogeme is used to estimate model parameters whereby different model structures are fitted on the NRM/LMS data. This should result in the correct parameters to reproduce the mode distribution from the NRM/LMS data. Regarding the choice modeling structure a MNL or NL model is suggested in the research approach. The NL model was used with different trip parameter configurations but none of them led to plausible results. This implies that the β values directions or levels were implausible. This seems to be the result of a difference in modeling structure between the input data and the used model. Presumably because route choice is not implemented for train and bus in the NRM/LMS.

For MNL model the results are better. Nevertheless using too many parameters results in unrealistic parameter values or ratios (table 1). The models 2 and 3 make use of individual travel time parameter for each mode. This can be the result of correlation between the factors whereby Biogeme makes use of the correlation and the additional factors to improve the model fit (R^2); due to the unrealistic parameter values this does not result in a better model.

When the travel times are estimated per mode huge differences can be found for the time β 's between the different modes. According to the literature the travel time weight for different modes of transportation should not differ more than factor 1,5. For model 2 this is 13 and for model 3 this is 30. Therefore these two models are rejected. The total travel time parameter is used to replace the travel time per mode. This is one time parameter that is used for all modes (model 1, 4 and 5). The models that use the total travel time show comparable results for both the time parameter and the R^2 . Model 4 shows travel time to transfer ratio that is in line with literature findings unfortunately the effect of the transfer is positive instead of negative. It is unrealistic that a route is more attractive if it includes more transfers, therefore this model will be rejected. Including the frequency also led to an unrealistic outcome. According to model 5 a higher frequency makes the route less attractive. This is unrealistic because a higher frequency should make the PT alternative more attractive therefore this model will be rejected. In the research approach is described that the effect of the frequency should be determined in this step because the literature section did not provide a clear relation between frequency and travel time the effect of the frequency will be excluded. The effect of the frequency cannot be modeled correctly and will therefore be ignored for this analysis. Nevertheless it seems to be preferable to include the frequency when it is possible.

Summarizing the findings above: adding parameters in addition to the total travel time parameter results in an unrealistic model. Therefore only model 1 can be used. This model is using mode preferences and the effect of the total travel time.

Model	ASC bicycle	ASC PT	β total time	β car time	β bicycle time	β PT time	β transfer	β frequency	R ²
Travel time (1) $V_{pt}=ASC_{pt} + \beta_{tt} * totaltime_{pt}$ $V_{bicycle}=ASC_{bicycle} + \beta_{tt} * totaltime_{bicycle}$ $V_{car}=\beta_{tt} * totaltime_{car}$	-0,661	-0,492	-0,0451						0,245
Travel time per mode (2) $V_{pt}=ASC_{pt} + \beta_{ptt} * totaltime_{pt}$ $V_{bicycle}=ASC_{bicycle} + \beta_{bt} * totaltime_{bicycle}$ $V_{car}=\beta_{ct} * totaltime_{car}$	0,595	-1,86		-0,0582	-0,112	-0,00879			0,293
Travel time per mode and transfer (3) $V_{pt}=ASC_{pt} + \beta_{ptt} * totaltime_{pt} + \beta_{transf.} * nr.oftransfers$ $V_{bicycle}=ASC_{bicycle} + \beta_{bt} * totaltime_{bicycle}$ $V_{car}=\beta_{ct} * totaltime_{car}$	0,603	-1,92		-0,0545	-0,111	-0,00365	-0,162		0,294
Travel time and transfer (4) $V_{pt}=ASC_{pt} + \beta_{tt} * totaltime_{pt} + \beta_{transf.} * nr.oftransfers$ $V_{bicycle}=ASC_{bicycle} + \beta_{tt} * totaltime_{bicycle}$ $V_{car}=\beta_{tt} * totaltime_{car}$	-0,505	-0,479	-0,0561				0,512		0,25
Travel time and frequency Travel time and transfer (5) $V_{pt}=ASC_{pt} + \beta_{tt} * totaltime_{pt} + \beta_{frequency.} * frequencyofroute$ $V_{bicycle}=ASC_{bicycle} + \beta_{tt} * totaltime_{bicycle}$ $V_{car}=\beta_{tt} * totaltime_{car}$	-0,656	-0,360	-0,0456					-0,0319	0,254

Table 1: Reproduction

In the literature section is mentioned that it is preferable to distinguish modes of transportation. However it is not possible to distinguish the travel time weight for PT from the other modes by these models. A clear relation between car and PT travel time sensitivity is not provided in the literature section. Nevertheless it is clear that they are close to each other (factor 1,5). Therefore the total travel time parameter will be used for PT, car and bicycle trips. The PT modes themselves will not be distinguished for the same reason; it is unclear how they relate to the car and the bicycle time. This is in contradiction with the recommendations made in the literature section.

As proposed in the research approach the total travel time for PT will be replaced by the total travel time equivalent to retrieve the sensitivity of individual trip segments. According to the literature section the IVT, waiting and access and egress time is used in combination with the number of transfers. The fit of this model will be elaborated in the next section.

8.2.2 Model layout

In the research approach a four step method is proposed for choice modeling. The results of the model fit in the previous section has some implications on the layout of the model. To keep track of the outcome of these implications the used method will be explained.

The previous section revealed that only a MNL mode choice can be used to fit the model to the data. PT route choice should be implemented and therefore a NL model is preferred. This implies that the MNL mode distribution model should be extended towards a NL model that includes mode and route choice. Furthermore it is proposed in the research approach that the total travel time for PT should be replaced with the generalized cost for PT that incorporates the sensitivity towards different modes.

The first step is used to retrieve model parameters and mode distribution with a MNL model. The model can be found in the previous section (nr. 1). Whereby the travel time used in the utility function for the PT alternative is replaced with the generalized cost for PT:

$$V_{PT} = ASC_{PT} + \beta_{TT} * (ivt_{PT} + 1.5 * acc_egr_{PT} + 1.5 * wait_{PT} + 9 * transf_{PT})$$

$$V_{bicycle} = ASC_{bicycle} + \beta_{TT} * totaltime_{bicycle}$$

$$V_{car} = \beta_{TT} * totaltime_{car}$$

The result of this step is the correct parameters (ASCs and β total travel time). With these parameters the model should be able to model the trip distribution for each OD. Mismatches on OD level will be corrected in the second step, by adjusting the utility per mode. This is done based on the assumption that the distribution found in the input data is more accurate than the model outcome.

The third step is used to capture the shift in utility of the PT system that is used for the forecast. The effect of multiple PT routes is implemented in this step by using a NL model. The parameters from the MNL model will be used for the NL model and the logsum parameter is unknown. As proposed in the research approach the effect of values between 0,1 and 0,5 will be analyzed. The different PT routes will be part of the PT branch and two other branches are used for the car and the bicycle. The utility function for the car and the bicycle are unaltered and the utility for PT is replaced by a utility per route:

$$V_{PT\ ri...rj} = ASC_{PT} + \beta_{TT} (ivt_{PT\ ri...rj} + 1.5 * acc_egr_{PT\ ri...rj} + 1.5 * wait_{PT\ ri...rj} + 9 * transf_{PT\ ri...rj}) * totaltime_{PT\ ri...rj}$$

The shift in PT utility is used to forecast the effect of the scenario. The shift is the difference between the reference PT utility and the scenario PT utility for each OD:

$$\Delta PT\ utility = scenario\ PT\ utility - reference\ PT\ utility$$

In the fourth step the new mode distribution is calculated by combining the corrected reference utility from the second step with the shift in utility from the third step. Finally the PT share can be distributed by using the route utility from the third step.

8.2.3 Model fit on the input data

It is important to know how the model fits the input data and how accurate the distribution that is found in the input data can be reproduced. This is done by analyzing the errors in PT share and the regression. The data displayed in this section is the outcome of the first choice model step.

Trips per OD

Table 2 reveals the summed absolute deviation per OD displayed as a ratio of the actual total number of trips. To explain the used method an example with two OD combination is provided. The first OD has 8 PT trips modeled and 10 according to the input data the second OD has 23 trips modeled and 20 according to the input data. The summed absolute deviation is $|-2|+|3|=5$ and the total number of trips is $10+20=30$. This results in an error of $(5/30)*100=17\%$. The error for each OD can be found in appendix 6. The new model is the model proposed in this study and the original HOV-scanner uses the model as it was proposed by MuConsult in a previous study.

Absolute deviation actual trips ratio (%)	new model	original HOV-scanner
PT	49	69
Car	14	38
Bicycle	68	179
Weighted total	28	63

Table 2: Absolute deviation actual trips ratio

Regression

Figure 20 and 21 display the regression between the trip share of the NRM and the trip share predicted by the model for all ODs. The line plotted in both graphs is the line where the model data is identical to the input data.

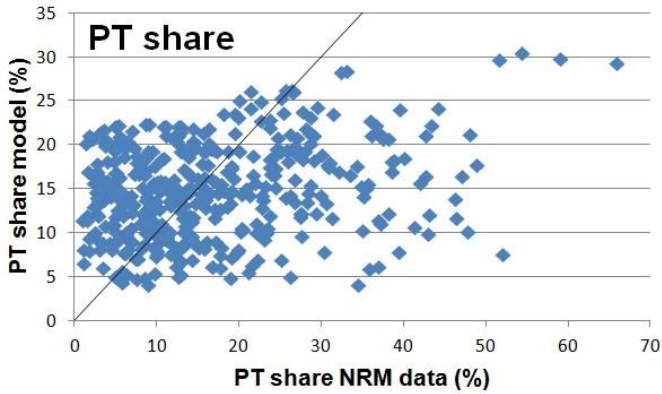


Figure 21: PT share

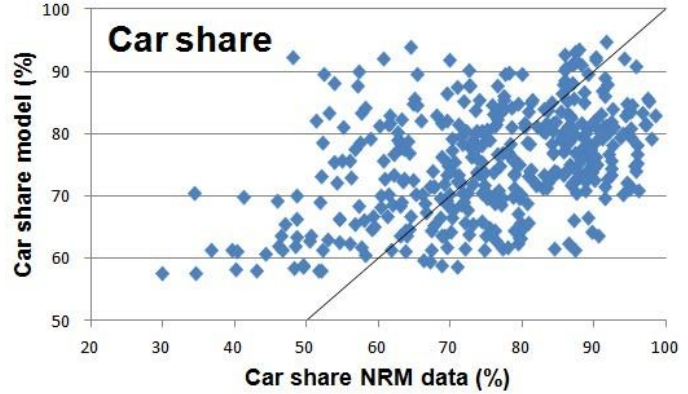


Figure 20: Car share

8.2.4 The PT trip forecast

Forecast of the HOV-scanner will be elaborated by comparing the new scanner with other alternatives:

- New HOV-scanner NL (logsum 0,1; 0,25 and 0,4). Mode and route choice with the parameters found in this chapter.
- New HOV scanner MNL. Mode choice with the parameters found in this chapter.
- Original HOV-scanner. Mode choice with the parameters from the original HOV-scanner.

The effect of the different models is assessed by comparing the outcome for the HOV-basic scenario in the

“Duin- and Bollenstreek”. The elasticity of the models is assessed by increasing the travel time of the reference with 10%. This should provide additional insight in the effect of the models. The overall effect can be seen in table 3. The more detailed effect per OD can be found in appendix 7. A discussion on the models and this data can be found in the next section.

Forecast:	scenario		travel time +10%	
	Number of trips	Growth (%)	Number of trips	Growth (%)
Reference	39842			
New HOV NL LS 0,10	43189	8,4	36747	-7,8
New HOV NL LS 0,25	43055	8,1	36725	-7,8
New HOV NL LS 0,40	43109	8,2	36709	-7,9
New HOV MNL	43368	8,6	36759	-7,7
Original HOV-scanner	43369	8,9	36468	-8,5

Table 3: Growth HOV-basic scenario and growth travel time +10% elasticity

More details on a selected number of OD pairs can be found below. OD combinations that show large differences between different model configurations are selected. For the different models the shift in trips are displayed in table 4. The reference trips for car, bicycle and PT are displayed and per model the change in PT trips for the scenario are displayed.

			bicycle	car	PT	Δ trips	Δ trips	Δ trips	Δ trips	Δ trips
code	Origin	Destination	Reference trips (nr.)	Reference trips (nr.)	Reference trips (nr.)	Log-sum 0,1 (nr.)	Log-sum 0,25 (nr.)	Log-sum 0,4 (nr.)	New HOV MNL (nr.)	Ori HOV MNL (nr.)
2 4	Hoofddorp	Nw Vennep Getsewoud	567	2075	432	206	206	206	206	129
6 3	Schiphol	Nw Vennep station area	27	916	453	-95	-62	-37	-116	-70
7 2	Hillegom	Hoofddorp overig	570	4214	413	430	326	250	505	403
8 9	Lisse other	Lisse Gerard Dousstraat	3057	3774	1014	-16	-101	-181	25	31
11 8	Noordwijkerhout	Lisse overig	138	2031	163	-90	-87	-82	-90	-84
14 13	Voorhout	Sassenheim Noord	845	2076	304	177	150	116	179	104

Table 4: Delta PT trips

The changes in the route set characteristics between the reference and the scenario are displayed in table 5. This is done for the ODs of table 4.

code	Origin	Destination	nr of routes in set reference	nr of routes in set scenario	fastest PT route reference (min)	fastest PT route scenario (min)	bicycle (min)	car (min)
2 4	Hoofddorp	Nw Vennep Getsewoud	1	1	29	20	19	9
6 3	Schiphol	Nw Vennep stationsgebied	2	3	31	38	48	13
7 2	Hillegom	Hoofddorp overig	4	2	58	35	24	10
8 9	Lisse other	Lisse Gerard Dousstraat	2	1	16	15	9	4
11 8	Noordwijkerhout	Lisse overig	1	2	22	44	21	8
14 13	Voorhout	Sassenheim Noord	2	1	31	22	13	5

Table 5: Route set changes

The distribution over the routes within the route set is displayed below (table 6). Only the route sets from the previous table that contain multiple routes in the scenario are displayed.

			total (min)	IVT (min)	transf. (nr.)	wait (min)	acc. egr. (min)	log-sum 0,1 (nr.)	log-sum 0,25 (nr.)	log-sum 0,4 (nr.)
6 3	Schiphol	Nw Vennep station area	38	18	2	8	10	123	135	143
6 3	Schiphol	Nw Vennep station area	39	17	2	10	10	117	127	136
6 3	Schiphol	Nw Vennep station area	42	12	1	19	10	118	129	137
7 2	Hillegom	Hoofddorp	35	21	1	3	10	524	459	412
7 2	Hillegom	Hoofddorp	44	27	2	5	10	318	279	250
11 8	Noordwijkerhout	Lisse other	44	31	1	2	9	45	46	50
11 8	Noordwijkerhout	Lisse other	49	28	2	10	9	28	29	31

Table 6: Trip distribution

8.3 Discussion

The discussion is divided in a section regarding the model fit and a section regarding the resulting trip forecast.

8.3.1 Model fit

In contradiction to the suggestion made in the research approach, only the MNL could be used to fit the model with realistic parameters. Presumably because route choice is not implemented for train and bus in the NRM/LMS. This is a limitation of the used data nevertheless it was the only data available in this situation. Furthermore the generalized cost for PT are used because only the sensitivity towards mode specific preferences in combination with total travel time could be captured. As a result of this approach a logsum for the NL forecast model cannot be estimated, the effect of trip components is not case specific and the effect of the PT route modeling is more vague. For the logsum and the effect of trip components a solution is proposed in the research approach.

The error in the summed absolute deviation per OD of the new HOV-scanner are compared with the original version. The new HOV-scanner results in much less deviation than the original model for the HOV-scanner. In general the results for the car modeling are much better than the results for PT and bicycle modeling. Nevertheless there is still an overall trip reproduction error of 46%.

Figure 20 displays the difference between the PT share of the input data and the PT share modeled. The PT share on some ODs is remarkably high in the input (NRM) data, values above 50% are found. This seems to be unlikely high. The model is much more conservative and in general the shares are below 25%. A comparable trend is revealed for the care share (figure 21). In general the model is predicting a smaller range than the input data from the NRM. The differences between the input (NRM) data and the model could be explained by modeling errors, accuracy of the input data and differences in modeling assumptions between both models. For the NRM/LMS model the focus lies on modeling the car trips and the train trips are modeled with less details. The train model is not fully calibrated and the train costs are overestimated on long distances. This could effect in errors in the PT trips and underestimation of the train on long distances. On the other hand, the train costs are ignored by the HOV-scanner. Overall the mode shares of both car and PT trips are not modeled close to the input data. Nevertheless this could be a limitation of both the input data (NRM) as the used model.

The improvements of the model fit over the previous model of the HOV-scanner can be explained by the method used to fit the data. In the previous study of MuConsult the shares per OD are used to fit the data while this study used the actual number of trips per OD. This resulted in a greater weight for ODs with more trips. Presumably this are also the OD's with more accurate trip data.

8.3.2 PT trip forecast

The implementation of the HOV-basic scenario results in an overall increase of PT trips over the reference scenario of 8,9% according to the original HOV-scanner, 8,1 to 8,4% according to the new NL HOV-scanner and 8,6% according to the new MNL HOV scanner. These values are close and the exact value is unknown. Differences are caused by model structure parameter values and the use of parameters. In addition to the HOV-basic scenario the elasticity of the model is tested by implementing a scenario with a 10% increase travel time over the reference scenario. The outcome of the elasticity is close. The original HOV-scanner predicts a reduction of 8,5%, the new NL HOV-scanner 7,8 to 7,9% and the new MNL HOV-scanner 7,7%. An increased travel time of 10% should result in an overall reduction of about 10%. As a result the original HOV-scanner is expected to provide the best results. However the range of the elasticity's found in literature is wide (appendix 8) and subjected towards a wide range of characteristics. All tested models are with the provided range and therefore they are all plausible.

Besides the total effect, the effect on specific OD combinations is interesting because it reveals additional information regarding modeling effects. To do this there are six combinations selected with strong differences between the different models (table 4 and 5). The first combination has 432 trips in the reference and the original model predicts an increase of 130 trips and the new models around 200. The PT travel time improves from 30 to 20 minutes and the car takes 10 minutes. The strong difference can be explained by the elimination of the transfer in the scenario. This effect is modeled stronger in the new HOV-scanner. Due to the huge improvement the outcome of the new HOV-scanners could be realistic. For the second combination the travel time by PT increases from 31 to 38 with an additional transfer. The effect of an increased route can be seen here. The models with the logsum 0,25 and 0,4 estimate a smaller decrease in the trips. This is an

undesired effect because a larger route set does not imply more departure times. Overall the decrease is larger for the new model which could be a realistic effect. A strong increase in trips is forecasted for all models on the third OD combination. The trips are roughly doubled and as a result the original HOV-scanner is expected to be the most realistic. This model forecast a smaller grow than the new HOV-scanner MNL and NL logsum 0,1. The fourth combination reveals contradicting forecasts. The MNL forecast a growth and the NL forecast a decline in the number of trips. Again this is the result of the changes in route set size. Due to the minimal change in travel time of only one minute a trip difference around zero is expected. As a result the original HOV-scanner, the MNL new HOV-scanner and the NL new HOV-scanner with LS 0,1 provide the best result. The fifth OD combination had a PT travel time that doubles from 22 to 44. As a result the PT share that is already quite low for the area reduces strongly. If the PT share is already quite low I would expect only small changes. This because the share of captives is presumably higher. As a result of this effect the outcome of the original HOV-scanner seems to be the most realistic. The sixth OD combination has an improved travel time from 31 to 22 minutes and one transfer that is eliminated. This combination reveals the difference in the use of parameters. The new HOV-scanners are predicting larger growths due to the elimination of a transfer. The new MNL and NL model with logsum 0,1 could provide realistic results here.

In general the effect of the implementation of the transfer parameter differs clearly between the models. Whereby the parameter used in the new HOV-scanner matches the literature value better. Furthermore is the predicted growth according to the new model in general lower than the growth of the original model, this is the result of differences in the travel time parameters and parameter values. When the NL models are compared the logsum of 0,1 seems to be preferable over 0,25 and 0,4. The models with a logsum value of 0,25 and 0,4 seems to be to sensible toward changes in the route set size. This seems to be unrealistic and therefore the logsum should be low from this point of view. A higher logsum value implies a lower correlation between the alternatives within the nest. If the logsum is too low on the other hand the utility will only be affected by the most attractive alternative. It could be useful to test these models with a route set size that is identical for the reference and the scenario situation.

In addition to the original HOV-scanner the new HOV-scanner should be able to include route choice for the PT. From the selected OD combination there are three with a route set that is larger than one. As stated previous the real distribution is unknown and therefore the modeled distribution can only be judged. It seems to be that the trips are distributed quite evenly, presumably even too evenly. From Hillegom to Hoofddorp for example the difference in travel time is 35 to 44 minutes and a 2/5 to 3/5 distribution is assumed. Based on the difference in travel time a larger difference would be expected. Nevertheless there are other factors that influence PT share and route shares. If frequency and mode preferences were included the result could differ. The frequency was not included because the effect could not be modeled correct. Another important factor could be the match between the preferred arrival time and the actual arrival time. A validation could provide the additional information on both mode and route distribution.

8.4 Conclusion

The choice modeling approach is implemented and tested by using the “Duin- en Bollenstreek” data. The aim is to create a model that fits the input data and can be used to forecast the effect of a scenario. By using a MNL with mode specific preferences and a total travel time parameter the best model fit was obtained. The total travel time for PT is replaced by the generalized cost of PT to capture the effect of trip segments. This resulted in a model with an improved fit in comparison to the model of the original HOV-scanner. Nevertheless there is still a lot of variation between the modeled number of PT trips and the input data. Presumably due to inaccuracy of the train and bus modeling in the input (NRM/LMS) data, and due to different modeling assumptions. Fortunately the fit of the car trips is better. Due to the use of the total travel time parameter it is assumed that the sensitivity towards travel time is captured correctly.

The forecast is more difficult to assess because the effect of the implementation is unknown. The forecast is predicting a slightly smaller growth than the previous study of MuConsult with the original HOV-scanner. The elasticity of the model on the other hand slightly below the average value found in literature. Nevertheless this range is quite wide. Overall the total effect of forecast seems to be correct.

Besides the total number of trips the distribution over the route is important. The trip distribution among routes within a route set seems to be relatively evenly. This effect seems to be more predominant for longer routes because they allow for a larger detour. Nevertheless this cannot be verified due to the absence of more detailed trip distribution information. A validation should provide additional details on the trip distribution.

Due to the used approach the logsum value is unknown. Multiple values have been tested. The height of the logsum parameter is mainly effecting the trips on OD level. The change of total number of trips is just slightly affected by the height of the logsum parameter. The trip distribution on the other hand is not influenced by the logsum value. Overall a lower logsum value seems to be preferable to overcome the effect of differences in route set size.

To summarize: the model fit is improved over the original HOV-scanner. Nevertheless large deviations in the mode shares between the data and the model can be found. The forecast is in line with expectation from the original version nevertheless in cannot be verified. Especially the distribution and the logsum effect is unknown and further research is required. A Validation should provide this additional information.

9. Validation

The “Duin- Bollenstreek” that was used in the previous section did not provide a PT trip distribution for the reference case and the actual outcome for the scenario was unknown. Furthermore it is preferable that the data used for validation differs from the data used to set up the model. Therefore the need for a validation case arose. The validation provides insight on the accuracy of the HOV-scanner and answers research question 1.3: How accurate is the prediction value of the modal split for the new HOV-scanner. During this validation the effect of the new HOV-scanner will be analyzed on three different aspects of the HOV-scanner:

1. The accuracy of the route set.
2. The accuracy of the forecast.
3. The accuracy of the trip distribution.

Furthermore recommendations for improvements on the HOV-scanner will be provided based on the results of this validation.

9.1 Study area

In the Marne a municipality in the north west of Groningen some network changes have been implemented. In 2015 bus line 67 removed and the schedules of the other lines are adjusted and replace line 67. An overview of the area and the lines can be found in figure 22. Removing line 67 resulted in a small reduction of PT trips. This effect will be used to validate the HOV-scanner.



Figure 22: Study area

Average nr. of trips per workday		
Winsum station	Wehe- Den Hoorn	21
Winsum station	Eenum	37
Winsum station	Kloosterburen	18
Winsum station	Hornhuizen	3
Wehe- Den Hoorn	Winsum station	18
Eenum	Winsum station	33
Kloosterburen	Winsum station	24
Hornhuizen	Winsum station	3

Table 7: Number of trips per workday

The trips between Winsum and Eenum, Wehe- Den Hoorn, Kloosterburen and Hornhuizen are analyzed. These combinations are directly affected by removing line 67. Furthermore is Winsum generating and influenced most of the trips. The average number of trips per work day is displayed in table 7.

Check-in and check-out data is available per line on OD level. This data is available for two similar months in 2014 and 2015 and is provided by OV-bureau Groningen Drenthe. For 90% of all trips in Groningen and Drenthe a PT card is used and for the remaining 10% a paper card or another method is used (OV bureau Groningen Drenthe, 2015). Of the 10% group trip data is unknown. Therefore this 10% is left out although it can lead to a small bias. Nevertheless is the used data of high accuracy due to the use of chip data. A disadvantage of this data is that the effect of transfers cannot be validated, because transfer data is not included. Despite the fact that the new approach is matching the literature better on the effect of transfers additional research seems to be required, to select the most suitable solution for the HOV-scanner.

9.2 Route set generation

For the validation it is necessary to check if the generated route set matches the route set provided by the data. Due to the fact that the provided data is check-in and check-out data per line it does not provide routes that are composed of multiple lines. Therefore it is impossible to make a one-to-one comparison between both route sets. Nevertheless it is possible to search for all routes that are provided in the check-in and check-out data and to search for abnormalities in the generated route set. The program is generating routes correct when all used routes are reproduced and when there are no illogical routes generated. It is judged if the generated routes are illogical.

9.2.1 Results

In the check-in and check-out data Winsum is always connected through two lines with the other villages. The number of routes provided by the route set generation program is shown in table 8.

Firstly it is remarkable that Winsum, Kloosterburen and Winsum, Hornhuizen both have just one route option, instead of the two provided by the check-in and check-out data. This is the result of excluding line 68 from the route set for these combinations. Line 68 exceeds the 30% detour criterion for these combinations. Nevertheless line 68 is used and therefore a detour criterion of 30% seems to be too narrow. The total trip length is 21 and 27 minutes for line 67 and the total trip length 33 and 39 minutes for line 68. Therefore the route with line 68 are 12 minutes longer than the route with line 67.

		nr. of routes
Winsum station	Wehe- Den Hoorn	2
Winsum station	Eenrum	2
Winsum station	Kloosterburen	1
Winsum station	Hornhuizen	1
Wehe- Den Hoorn	Winsum station	4
Eenrum	Winsum station	3
Kloosterburen	Winsum station	2
Hornhuizen	Winsum station	4

Table 8: Number of routes, route set program

According to the route set program there are four routes possible from Wehe to Winsum. The two longest routes contain a transfer, despite the fact that the first entered vehicle provides a direct connection. It does not seem to be realistic that a direct connection will be left to make a transfer that does not shorten the overall travel time. The route set from Hornhuizen to Winsum suffers from the same limitation.

9.2.2 Conclusion

For the study area a detour criterion of 30% seems to be too narrow, as a result used routes were deleted from the route set. It is undesirable that used routes are excluded from the route set because that can lead to unrealistic choice modeling. Including too many routes on the other hand is not sensitive for this problem if the route choice is modeled correctly. This will typically result in a very small assignment for the unattractive routes. To overcome the deletion of used routes there are two solutions:

- A threshold value for the minimum detour can be added. The proposed value is 15 minutes.
- The detour value can be increased. The detour value of 35% was necessary to solve it for this specific case.

Overall the detour threshold of 15 minutes seems preferable because this solves identical problem for route sets with shorter routes.

The possibility to walk between stops and transfer lines is added to make the route set more realistic. Nevertheless the opposite can be seen on some combinations. While a direct connection was used a transfer has been made that resulted in a longer trip. This does not seem to be realistic. It would be better if a transfer or a walkable connection only would be included if it improves travel time in comparison to the current vehicle, given the fact that the current vehicle provides a direct connection. There are two possible solutions: Updating the program or deleting the routes manually. To improve processing time it is preferable to update the program.

9.3 Forecast

For the validation it is necessary to check if the forecasted growth per OD is similar to the growth found in the data. In this section the forecast of the new HOV-scanner will be assessed and compared with the original HOV-scanner as well. The preferred logsum value is not revealed in chapter eight and therefore the effect of different values will also be assessed in this section.

9.3.1 Approach

To assess the accuracy, the shift in probability is compared with the shift in trips from the data. The shift in probability and the shift in actual trips both result in a growth factor (see the formula below). The growth factor can be compared on OD level. If the growth factor is 1,06 this implies that the number of trips in 2015 is 1,06 times higher than in 2014.

The parameters retrieved in the “Duin en Bollenstreek” study will be used for the HOV-scanner and

for comparison the original HOV-scanner included as well. The new NL model that includes mode and route choice will be compared with the mode choice model of the original HOV-scanner. To calculate the probability for PT the travel time by car is retrieved with Google maps. The bicycle is ignored because the travel times by bicycle are high: around an hour for most of the combinations. For 2015 there is one route per OD (line 65 or 68) and in 2014 there are two routes (line 67 and 65 or 68). The probability for 2014 is the probability of the PT nest in case of NL or the probability of the fastest alternative in case of MNL.

$$\frac{P_{65 \text{ or } 68}^{2015}}{(P_{67}^{2014} \text{ and } P_{65 \text{ or } 68}^{2014})} = GF_P$$

$$\frac{N_{65 \text{ or } 68}^{2015}}{N_{67}^{2014} + N_{65 \text{ or } 68}^{2014}} = GF_N$$

9.3.2 Assessment framework

The growth factors are compared to judge if the model is predicting correct. The forecast is perfect when the growth factors from the model match the data exact. The larger the difference the worse is the forecast. When the direction of the growth factor differs between the model and the data the direction of the change is predicted wrong.

Of the provided ODs only Winsum, Kloosterburen has both a substantial shift in travel time and a reasonable number of trips. This combination makes this OD the most suitable to examine the effect of the differences in parameters and model structure. The IVT changes from 21 minutes to 30 minutes and from 19 minutes to 31 minutes. On average there are 18 and 24 trips in 2014.

9.3.3 Results

forecast		Growth factor				
parameters "Duin- en Bollenstreek"		data	MNL	NL LS 0,1	NL LS 0,25	NL LS 0,4
(1) Winsum station	Wehe- Den Hoorn	1,06	0,98	0,95	0,89	0,83
(2) Winsum station	Eenrum	0,77	1,02	0,98	0,91	0,85
(3) Winsum station	Kloosterburen	0,87	0,83	0,82	0,79	0,74
(4) Winsum station	Hornhuizen	0,77	0,84	0,84	0,80	0,75
(5) Wehe- Den Hoorn	Winsum station	1,21	1,02	0,97	0,90	0,83
(6) Eenrum	Winsum station	0,83	1,02	0,98	0,91	0,84
(7) Kloosterburen	Winsum station	0,79	0,84	0,84	0,80	0,75
(8) Hornhuizen	Winsum station	1,20	0,84	0,83	0,80	0,74

Table 9: Forecast "Duin en Bollenstreek" parameters

forecast		Growth factor				
parameters original HOV-scanner		data	MNL	NL LS 0,1	NL LS 0,25	NL LS 0,4
(1) Winsum station	Wehe- Den Hoorn	1,06	0,98	0,96	0,92	0,87
(2) Winsum station	Eenrum	0,77	1,02	0,99	0,94	0,89
(3) Winsum station	Kloosterburen	0,87	0,84	0,84	0,82	0,79
(4) Winsum station	Hornhuizen	0,77	0,84	0,84	0,82	0,79
(5) Wehe- Den Hoorn	Winsum station	1,21	1,02	0,98	0,92	0,87
(6) Eenrum	Winsum station	0,83	1,02	0,99	0,94	0,89
(7) Kloosterburen	Winsum station	0,79	0,85	0,85	0,83	0,79
(8) Hornhuizen	Winsum station	1,20	0,84	0,84	0,82	0,78

Table 10: Forecast original HOV-scanner parameters

When the original HOV-scanner is compared to the new HOV-scanner (table 9 and 10) it is noticeable that the results of the MNL mode choice model is similar. This implies that the effect of the both models and parameter values is similar.

In general all models reveal a lower growth factor than the grow factor of the data itself. This implies that the total number of trips in 2015 will presumably be underestimated with most of the models. If a lower travel time parameter value (β) would be used the outcome would be closer to the 2015 value.

When the NL models are compared mutually the models with the lower logsum value seem to provide a more distinct outcome for each OD, where the high logsum value on the other hand seems to level out the outcome more. This is presumably a side effect of the route set that reduces from two to one for all ODs. As a result of this effect the outcome does not seem to be modeled correct for the higher logsum values.

When the change in travel time from 2014 to 2015 is used to explain the growth factor of the data some counter intuitive results can be revealed. From Winsum towards Wehe- Den Hoorn the travel time increases but a growth in trips can be seen (table 9 nr. 1). A similar effect can be seen from Eenrum towards Winsum (nr. 6). Both combinations have a minimal change in travel time (1 minute) and therefore the change in trips is presumably overshadowed by other effects.

The trips from Winsum and Kloosterburen (table 9 and 10, nr. 3 and nr. 7) are expected to be the most suitable to analyze the effect due to the large change in time and a reasonable number of trips. The conclusion of this analysis is that overall both the direction as the magnitude of the growth is forecasted accurate. For the trip towards Winsum the models with the logsum of 0,4 provide accurate results. For the other direction the MNL and NL model with a logsum of 0,1 seem to provide the best results. When the parameters from the “Duin and Bollenstreek” are compared with the original HOV-scanner parameters the “Duin and Bollenstreek” parameters seem to provide slightly better results for the MNL and NL logsum 0,1 model. Nevertheless the difference is negligible.

9.3.4 Conclusion

If the difference in travel time between the reference and the scenario situation is large enough and there are a reasonable number of trips the model can forecast the direction and the magnitude of the growth. The results of the MNL model and de NL model with a logsum of 0,1 are similar.

The result with parameters from the “Duin and Bollenstreek” are compared with the result of the original HOV-scanner parameters. The difference between both parameters is negligible, especially for the MNL model and the NL model with a logsum of 0,1. Based upon this case a logsum value between 0,1 and 0,25 seems to be preferable over a logsum value of 0,4. A logsum value of 0,4 appears to be sensitive for changes in the route set size.

If the difference in travel time is small the forecast of the number of trips will not be accurate. In this case a small change is found to be 1 or 2 minutes difference on 15 minutes of IVT. Other changes seem to be dominant over the change in travel time. This could be for example: the share of captives, the connection with other lines, the preferred arrival time, the availability of substitutes and changes in trip patterns. This is a point of attention when the HOV-scanner is used.

9.4 Trip distribution

As an extension of the original HOV-scanner, route choice is implemented in the new HOV-scanner. In this section the trip distribution from the model is compared to the distribution from in the provided data.

9.4.1 Approach

The model is using the difference in utility and the resulting difference in probability to distribute the trips over the different routes. The modeled distribution of the 2014 data can be compared with the distribution retrieved from the data to validate the model. This can be done by comparing the ratio of line 67 and the other line for both the model and the actual data. The ratios are calculated as followed:

$$\frac{P_{67}}{P_{65 \text{ or } 68}} = ratio_p$$

$$\frac{N_{67}}{N_{65 \text{ or } 68}} = ratio_n$$

A ratio of five implies that the share of trips on line 67 is five times higher than the ratio of the other lines. The models and parameter values that are proposed in the previous section are used.

9.4.2 Assessment framework

The ratios are compared and it is judged if the model is predicting accurately. The ratio is modeled perfectly when both ratios match exactly. The larger the difference between both ratios, the worse is the ratio modeled.

Whether the ratio is above or under the one is an indication for the share of line 67 being smaller or larger than the share of other lines. The actual share and the modeled share should provide similar results.

9.4.3 Results

Ratio 67/(65 or 68)			parameters "Duin- en bollenstreek"	parameters original HOV- scanner
		data	model	model
(1) Winsum station	Wehe- Den Hoorn	0,3	0,9	0,9
(2) Winsum station	Eenrum	0,6	1,0	1,0
(3) Winsum station	Kloosterburen	4,8	1,4	1,5
(4) Winsum station	Hornhuizen	3,0	1,4	1,5
(5) Wehe- Den Hoorn	Winsum station	0,3	1,0	1,0
(6) Eenrum	Winsum station	0,6	1,0	1,0
(7) Kloosterburen	Winsum station	3,5	1,3	1,4
(8) Hornhuizen	Winsum station	5,3	1,4	1,5

Table 11: Route ratio

When table 11 is assessed the first thing that draws the attention is that the distribution over the two routes is modeled far too evenly. The model is prediction maximum ratio of 1,5 the actual ratio can be around 5. This could be corrected by increasing the travel time parameter.

The travel time from Wehe- Den Hoorn to Winsum (5) is 18 minutes for both line 65 as line 67. Therefore the model estimates a 50/50 distribution between both lines. This results in a ratio of one. Nevertheless according to the data the distribution is 0,3. On an identical travel time the proposed model will never provide an asymmetric distribution, this is a limitation of the approach. A possible explanation for the

actual ratio could be that line 65 is driving four times an hour during morning peak and line 67 is driving two times an hour in the same period.

9.4.4 Conclusion

In general the direction of the route share distribution modeled is in line with the route share distribution from the data. The only exception is Wehe- Den Hoorn towards Winsum (5) due to the identical travel times of both routes.

The magnitude of the share distribution on the other hand is modeled much too low. This could be the effect of the parameters being calibrated for mode choice and not on for route choice. Both models reveal identical results. The model fit can be improved by increasing the parameter for the travel time. Currently identical parameters are used for both mode choice and route choice. Based on this validation case it seems to be beneficial to use separate parameters for mode and route choice. The mode distribution parameter values are quite good but for the route distribution parameters values should be higher. With the NL model used as originally intended this is impossible because the utility of the mode is the result of the PT nest utility. Due to the contradicting effect of the trip time parameters on mode and route choice it seems to be beneficial to use separate parameters for both choices. This implies that the modeling structure must be adjusted. To select the correct model structure and parameters for route choice additional research is required.

Besides the travel time other factors are involved at route choice. From Wehe- Den Hoorn to Winsum (5) the frequency seems to influence route distribution. A line with a higher frequency is assumed to attract more trips than the lower frequency alternative if the other conditions are similar. Furthermore the desired arrival time could be of influence, especially for low frequency services. If there are two routes for example that differ twenty minutes in travel time but the slower alternative matches the desired arrival time perfectly and the faster alternative arrives 30 minutes before the desired arrival time. Then the alternative with the slowest travel time has most chance of being used. For instance when a PT line that is mainly used to access a school this could have a huge influence on the distribution. This is an indirect travel time effect that is difficult to capture.

10. Answering the research questions

The research questions helped to achieve the research objective. In this chapter the research questions as proposed in chapter four are answered. The first main research question is: “How should the HOV-scanner be designed to be able to model parallel PT connections?” There are three sub questions that contribute to answering this main research question:

- 1.1 Is the HOV-scanner in line with the current knowledge on transportation modeling?
- 1.2 What is a correct way to implement parallel PT connections in a model?
- 1.3 How accurate is the prediction value of the modal split for the new HOV-scanner?

The first sub question is answered by using the information gathered in the literature study whereby section 5.1, 5.3 and 5.4 are especially important. The used model, the choice model and the modeling assumptions of the HOV-scanner must be in line with the current knowledge on transportation modeling. The HOV-scanner is modeled according to the four-stage model. This model is still used despite the main limitation namely the lack of the availability to implement congestion effects. Fortunately this is not an obstruction for the HOV-scanner. The four-stage model as used in the HOV-scanner is therefore suitable and in line with current knowledge on transport modeling. As a choice model the HOV-scanner makes use of the MNL model. Although MNL is still widely used a shift towards ML can be seen in the recent literature. This implies that the choice model is not state of the art, but is still a good option because it matches the available data and the intended use. Regarding modeling assumptions the HOV-scanner makes use of one route for PT, car and the bicycle trips. To model the utility of a trip the parameter time is used. Whereby the PT trips are distinguished in IVT, access and egress time, waiting time and transfer time. The used trip cost calculation is not in line with the current knowledge of transportation modeling, because the monetary cost are excluded. It is generally accepted that cost, time and effort should be included. Furthermore trip purpose and a population segmentation could be included. This is done in the NRM/LMS model. This can improve accuracy but also increases the data need. The use of only one PT route for PT trips is in line with other models like the NRM/LMS model, but it is considered as an oversimplification of the model because the focus of the HOV-scanner is on the modeling of changes in the PT network.

The second sub question can be answered by using the information gathered in section 5.2, 5.3 and 6.3. To implement parallel PT connections correctly the first concern is the composition of the route set. From all possible routes only the considered routes are included in the route set. This can be done by applying suitable constraints and by selecting an appropriate route generation method. In this study the constrained enumeration method is used successfully. Nevertheless the used constraints require some improvement. The detour criterion of 30% has proven to be too narrow and routes with unrealistic transfer behavior were included during validation. Attention is required regarding the route overlap as well, because overlap can lead to over or under representation of routes. There are choice models that can account for route overlap or overlap can be deleted from the choice set. It is unclear which of these methods provides the best result. The choice model has an effect on the model fit and thereby on the distribution, it is therefore important that the correct choice model is chosen. This implies that the choice model should match the available data and the situation. It is also important to include the correct utility factors in the model. Cost, time and effort are generally accepted. In addition important PT route specific factors are: IVT, number of transfers, frequencies, waiting time, access and egress time and modes used. For this study a NL model is chosen whereby the PT routes are part of one PT nest. The validation case revealed that this model has his limitations. The parameter values used for mode distribution do not match the parameter values used for route distribution. In general the sensitivity towards travel time differences is much larger in route choice than in mode choice. Therefore the used method is not the correct method to implement parallel PT.

The third sub question can be answered by using the information in chapter eight and nine. In chapter eight only an indication is provided. The new HOV-scanner is predicting slightly lower PT shares and the model fit is slightly improved in comparison to the original HOV-scanner. The elasticity of the new HOV-scanner is somewhat below literature values but still in the correct range. The fit of the new HOV-scanner indicates an improvement nevertheless it is unclear if the forecast for the scenario is improved. Chapter nine covers the accuracy of the mode share prediction in detail. When using a MNL model or a NL model with a logsum of 0,1 the prediction value seems to be a little bit better than the value of the original HOV-scanner. Nevertheless the difference is negligible. Overall the prediction value of the model split is quite good as long

as there are enough trips and the difference in travel time between the reference and the scenario is large enough.

The second main research question is: “How can the HOV-scanner be improved to reduce the work load?” Followed by the sub question: “How can the data required for the HOV-scanner be processed more efficient and effective so that the required lead time can be met?” To answer these questions it is important to know which parts of the HOV-scanner can be addressed that consume most of the processing time. Transferring PT schedule information into travel times per OD is labor intensive because it is done manually. Due to the implementation of route sets at the new HOV-scanner this would become an even more extensive process. In chapter seven a program is introduced which automates the process. With this program the processing time for generating route sets is reduced by roughly a factor five. This is a satisfying result.

11. Conclusions, discussion and recommendations

In this chapter the conclusion regarding the objectives will be drawn, the limitations will be pointed out, the results are discussed and the recommendations are provided.

11.1 Conclusion

The main objective for this research was to implement PT route choice with parallel connections in the HOV-scanner. The additional goal was to reduce the work load so that the program would be more suitable as a first indication regarding the feasibility of changes in PT services.

During this research parallel PT is connections are included in the HOV-scanner nevertheless, the current setup is facing some limitations. The original HOV-scanner is based upon a MNL model. This model was not expected to be suitable to handle parallel PT due to the large correlation in the PT alternatives. To overcome this limitation a NL model is proposed. Whereby all PT alternatives are nested in the PT branch. In extension of the original HOV-scanner model specific constants are added. Compared to the original HOV-scanner the transfer is defined in number of transfers instead of time between two PT lines. For the parameter estimation the number of trips per mode are used instead of the share. This resulted in an improved model fit. Nevertheless on OD level there still are substantial differences in mode shares. The original HOV-scanner is used to forecast the shift in PT share. The validation showed that the shift in mode shares were forecasted quite well by the new HOV-scanner. A requirement to achieve good results is to have enough trips and the difference in travel time between the reference and the scenario should be large enough. The accuracy of the forecast in shift in mode share is identical to the original HOV-scanner. In addition to the original HOV-scanner PT route choice is included to estimate the shares for individual lines. The proposed model was not able to predict route shares correct. The validation case revealed that route shares were distributed far too evenly over the routes.

The work load is reduced by automating the route set generation process. This resulted in an impressive time reduction despite the fact the HOV-scanner is extended from a single route from towards a route set. For automating the route set generation process a constrained enumeration method is used. The used constrained enumeration method itself is working correct. However the validation case showed that the used constraints need some adjustment.

11.2 Discussion

11.2.1 NRM/LMS data

The NRM/LMS data that is used for the “Duin- en Bollenstreek” has its limitations and therefore implications on the model fit and the forecast. The NRM/LMS data is modeled data that uses different modeling assumptions and is focused on car trips. The train trips are modeled with less detail. Therefore a part of the error in the model fit is expected to be the result of the used NRM/LMS data. The used NRM/LMS data also affects parameter value calibration because the distribution in the NRM/LMS data is influenced by the modeling assumptions made in that model. Furthermore is the forecasted number of trips based on the number of PT trips retrieved from the NRM/LMS data. However the number of PT trips in the NRM/LMS is expected not to be accurate. The latter limitation was known on forehand, but for this study it was the only available data set.

11.2.2 Trip parameters

The selected parameters in the HOV-scanner are characteristics of the model. The literature section made clear that it is desirable to include the parameter frequencies. Nevertheless the effect of frequencies could not be modeled correct and the literature did not provide a clear relation between IVT and frequencies. Furthermore, the effect of frequencies is more difficult to model once it is included in a route set. When frequencies are included in a route set the perceived frequency will influence the outcome. How the perceived frequency should be included is also not known.

The match between the preferred arrival time and the actual arrival time can be of significant influence on the share distribution over routes. When there are two lines that arrive once an hour both with a 30 minute distribution. The slowest line of the two can be more attractive when the difference in travel time stays below 30 minutes. This effect is very difficult to model when the preferred arrival time is unknown. Nevertheless it can have a strong effect on trip distribution. When the HOV-scanner is used it should be noted that the exclusion of these effects are expected to lead to some degree of inaccuracy.

11.2.3 Validation

The validation revealed that the parameter values used for route choice deviate from the values used for mode choice. The parameter values used for mode choice modeling are too low for route choice modeling. This contradicts with the modeling assumption that the attractiveness of the PT mode is correlated with the attractiveness of the PT route. If the new assumption is that the attractiveness of the PT mode is not correlated with the attractiveness of the PT route a different model structure can be used. For instance a separate MNL model for both mode choice and PT route choice might be suitable. In that case the new HOV-scanner can be used for forecasting mode choice and PT route choice separately with different parameter values.

When the difference in travel time between the reference and the scenario is small no significant change in mode choice will be predicted. In some cases validation revealed a higher growth in PT trips than expected based on the small difference in travel time. This could be due to other effects that overshadow the small difference in travel time which are not included in the model, random variations or other reasons. The HOV-scanner is not suitable for predicting the effect of small changes for mode choice. For PT route choice the same effect could be applicable. This implies that for forecasting the effect of a change in a network, utilizing the HOV-scanner should be done carefully when the change in the network contains small changes.

11.3 Recommendations

The recommendation for further research relates to HOV-scanner specific subjects that are in interest of MuConsult and transportation modeling research in general.

The literature section and the modeling section made clear that there are still some gaps in the knowledge on PT route choice modeling. There are two different approaches to deal with overlap; it can be included in the route set or in the choice model. It is unclear how the result of both approaches relate to each other. Therefore additional research on this topic is recommended.

The effect of frequencies within a route set is unclear and influenced by the additional complexity caused by headway distribution. Therefore additional research regarding the effect of frequencies on the utility of a route set is desirable.

The route set program should be adjusted avoid the exclusion of used routes. This can be done by changing the detour criterion or by including a minimum threshold. For the validation case 35% or 15 minutes is enough. A detour threshold seems to be preferred. Additional research could be required regarding the correct value. Furthermore should counter intuitive routes be excluded from the route set. These are the routes where a transfer is made from direct connection that results in a longer trip.

In this study identical model parameters are assumed for both mode and route choice. The validation case revealed that the parameters for mode and route choice are different. In this study the β total travel time is calibrated for mode choice and experienced to be too low for route choice. Therefore additional research is required concerning the correct parameters for route choice.

The HOV-scanner is not suitable to analyze small differences in travel time between the reference and the scenario. If there are only a limited number of trips the results of the HOV-scanner can be inaccurate. It is recommended to take this limitations into account when the HOV-scanner is used.

It is unclear how accurate the NRM/LMS data is regarding PT trips. It is very well possible that the data on PT trips is inaccurate. This would affect the outcome of the HOV-scanner. Therefore it is recommended to analyze the effect of more accurate data

During this research the effect of the changes to the HOV-scanner are compared with the original version of the HOV-scanner. No comparison with another model has been made. For MuConsult it seems to be beneficial to make a comparison between the results of the HOV-scanner and a more complex model. This could provide more insight in the accuracy of the HOV-scanner.

References

- Abrantes, Pedro AL, & Wardman, Mark R. (2011). Meta-analysis of UK values of travel time: An update. *Transportation Research Part A: Policy and Practice*, 45(1), 1-17.
- Akgün, Vedat, Erkut, Erhan, & Batta, Rajan. (2000). On finding dissimilar paths. *European Journal of Operational Research*, 121(2), 232-246.
- Anderson, Marie Karen, Nielsen, Otto Anker, & Prato, Carlo Giacomo. (2014). Multimodal route choice models of public transport passengers in the Greater Copenhagen Area. *EURO Journal on Transportation and Logistics*, 1-25.
- Arentze, Theo A, & Molin, Eric JE. (2013). Travelers' preferences in multimodal networks: design and results of a comprehensive series of choice experiments. *Transportation Research Part A: Policy and Practice*, 58, 15-28.
- Bakker, Peter, Zwaneveld, Peter, Baveling, Jaco, Planbureau, Centraal, & voor Mobiliteitsbeleid, Kennisinstituut. (2009). *Het belang van openbaar vervoer: de maatschappelijke effecten op een rij*: Centraal Planbureau.
- Bellman, Richard, & Kalaba, Robert. (1960). On k th best policies. *Journal of the Society for Industrial & Applied Mathematics*, 8(4), 582-588.
- Ben-Akiva, M, Bergman, MJ, Daly, Andrew J, & Ramaswamy, Rohit. (1984). *Modeling inter-urban route choice behaviour*. Paper presented at the Proceedings of the 9th International Symposium on Transportation and Traffic Theory, VNU Press, Utrecht.
- Ben-Akiva, M, Bowman, J, Ramming, S, & Walker, Joan. (1998). Behavioral realism in urban transportation planning models. *Transportation Models in the Policy-Making Process: Uses, Misuses and Lessons for the Future*, 4-6.
- Bliemer, MCJ, Versteegt, HH, & Castenmiller, RJ. (2004). *INDY: a new analytical multiclass dynamic traffic assignment model*. Paper presented at the Proceedings of the TRISTAN V conference, Guadeloupe.
- Bliemer, Michiel CJ, & Rose, John M. (2011). Experimental design influences on stated choice outputs: An empirical study in air travel choice. *Transportation Research Part A: Policy and Practice*, 45(1), 63-79.
- Bliemer, Michiel CJ, Rose, John M, & Hensher, David A. (2009). Efficient stated choice experiments for estimating nested logit models. *Transportation Research Part B: Methodological*, 43(1), 19-35.
- Blij, FL. (2010). *van der, JS Veger & IC Slebos, "HOV op loopafstand"*. Paper presented at the Bijdragen Colloquium Vervoersplanologisch Speurwerk.
- Bovy, Piet HL. (2009). On modelling route choice sets in transportation networks: a synthesis. *Transport reviews*, 29(1), 43-68.
- Bovy, Piet HL, & Fiorenzo-Catalano, Stella. (2007). Stochastic route choice set generation: behavioral and probabilistic foundations. *Transportmetrica*, 3(3), 173-189.
- Bovy, Piet HL, & Stern, Eliahu. (1990). *Route choice. Wayfinding in transport networks. Studies in operational regional science*.
- Bruynooghe, M. (1969). Un modèle intégré de distribution, de trafic sur un réseau. *Departement de Recherche Operationnelle et Informatique, Arcueil*.
- Bunschoten, TM. (2012). *To Tram or not To Tram: Exploring the existence of the tram bonus*. TU Delft, Delft University of Technology.
- Cascetta, Ennio, Nuzzolo, Agostino, Russo, Francesco, & Vitetta, Antonino. (1996). *A modified logit route choice model overcoming path overlapping problems: specification and some calibration results for interurban networks*. Paper presented at the Proceedings of the 13th International Symposium on Transportation and Traffic Theory.
- Cascetta, Ennio, Papola, Andrea, Russo, FRANCESCO, & Vitetta, ANTONINO. (1999). *Implicit availability/perception logit models for route choice in transportation networks*. Paper presented at the World Transport Research: Selected Proceedings of the 8th World Conference on Transport Research.
- Catalano, SF, & Van der Zijpp, N. (2001). *A forecasting model for inland navigation based on route enumeration*. Paper presented at the Proceedings of the aet european transport conference, held 10-12 september, 2001, Homerton college, Cambridge, uk-cd-rom.
- Currie, Graham, & Delbosc, Alexa. (2013). Exploring comparative ridership drivers of bus rapid transit and light rail transit routes. *Journal of Public Transportation*, 16(2), 47-65.
- de Beer, P, Tiemersma, D, & Ploeg van der, R. (2011). Modal split onder druk? Gevolgen bezuinigingen openbaar vervoer Amsterdam *DIVV (Bijdrage aan het Colloquium Vervoersplanologisch Speurwerk 2011)*.
- De Cea, J, Fernandez, JE, Dekock, V, Soto, A, & Friesz, TL. (2003). *ESTRAUS: a computer package for solving supply-demand equilibrium problems on multimodal urban transportation networks with multiple user classes*. Paper presented at the annual meeting of the Transportation Research Board, Washington, DC.

- De Dios Ortuzar, Juan, & Willumsen, Luis G. (2011). Modelling Transport.
- Debrezion, Ghebreegzabiher, Pels, Eric, & Rietveld, Piet. (2009). Modelling the joint access mode and railway station choice. *Transportation Research Part E: logistics and transportation review*, 45(1), 270-283.
- Diana, Marco. (2008). Making the “primary utility of travel” concept operational: a measurement model for the assessment of the intrinsic utility of reported trips. *Transportation research part A: policy and practice*, 42(3), 455-474.
- Evans, Suzanne P. (1976). Derivation and analysis of some models for combining trip distribution and assignment. *Transportation Research*, 10(1), 37-57.
- Fiorenzo-Catalano, Maria Stella. (2007). *Choice set generation in multi-modal transportation networks*: Netherlands TRAIL Research School.
- Florian, Michael, & Constantin, Isabelle. (2012). A note on logit choices in strategy transit assignment. *EURO Journal on Transportation and Logistics*, 1(1-2), 29-46.
- Florian, Michael, & Nguyen, Sang. (1978). A combined trip distribution modal split and trip assignment model. *Transportation Research*, 12(4), 241-246.
- Florian, Michael, Nguyen, Sang, & Ferland, Jacques. (1975). On the combined distribution-assignment of traffic. *Transportation Science*, 9(1), 43-53.
- Fosgerau, Mogens, Hjorth, Katrine, & Vincent Lyk-Jensen, Stéphanie. (2007). The Danish value of time study: final report: The Danish Transport Research Institute.
- Goeverden, CD van, & Van Den Heuvel, MG. (1993). De verplaatsingstijdfactor in relatie tot de vervoerwijzekeuze. *Delft University of Technology under the authority of PB IVVS/Ministry of Transport, Public Works and Watermanagement, Delft*.
- Guo, Zhan, & Wilson, Nigel H. M. (2011). Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground. *Transportation Research Part A: Policy and Practice*, 45(2), 91-104. doi: <http://dx.doi.org/10.1016/j.tra.2010.11.002>
- Hensher, David A, & Greene, William H. (2003). The mixed logit model: the state of practice. *Transportation*, 30(2), 133-176.
- Hofman, Frank, & van Grol, Rik. Gedragsveranderingen in LMS/RMS: Rijkswaterstaat.
- Horowitz, Alan J, & Thompson, Nick A. (1994). Evaluation of intermodal passenger transfer facilities.
- Hunt, David T, & Kornhauser, Alain L. (1996). Assigning traffic over essentially-least-cost paths. *Transportation Research Record: Journal of the Transportation Research Board*, 1556(1), 1-7.
- Iseki, Hiroyuki, & Taylor, Brian D. (2009). Not all transfers are created equal: Towards a framework relating transfer connectivity to travel behaviour. *Transport Reviews*, 29(6), 777-800.
- Johnson, PE, Joy, DS, Clarke, DB, & Jacobi, JM. (1992). HIGWAY 3.01, An enhanced highway routing model: Program, description, methodology and revised user's manual, Oak Ridge National Laboratory. *ORNL/TM-12124. Oak Ridge, TN*.
- Kitthamkesorn, Songyot. (2013). *Modeling overlapping and heterogeneous perception variance in stochastic user equilibrium problem with weibit route choice model*. UTAH STATE UNIVERSITY.
- Kumar, Mukesh, Sarkar, Pradip, & Madhu, Errampalli. (2013). Development of fuzzy logic based mode choice model considering various public transport policy options. *International Journal for Traffic and Transport Engineering*, 3(4).
- Lanser, S. (2005). *Modelling travel behaviour in multi-modal networks*: TU Delft, Delft University of Technology.
- Lombard, K, & Church, RL. (1993). The gateway shortest path problem: generating alternative routes for a corridor location problem. *Geographical systems*, 1(1), 25-45.
- McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, 105-142.
- MuConsult. (2003). Elasticiteiten treinmobiliteit. State of the art anno 2003. Amersfoort.
- MuConsult. (2012). Planstudie kwaliteitsimpuls OV-corridor Duin- en Bollenstreek – Schiphol.
- MuConsult. (2014). Producten.
- Nuzzolo, Agostino, & Russo, Francesco. (1996). Stochastic assignment models for transit low frequency services: Some theoretical and operative aspects *Advanced methods in transportation analysis* (pp. 321-339): Springer.
- OV bureau Groningen Drenthe. (2015). Trend monitor 2014.
- Prato, Carlo Giacomo, & Bekhor, Shlomo. (2006). Applying branch-and-bound technique to route choice set generation. *Transportation Research Record: Journal of the Transportation Research Board*, 1985(1), 19-28.
- Prato, Carlo Giacomo, Bekhor, Shlomo, & Pronello, Cristina. (2012). Latent variables and route choice behavior. *Transportation*, 39(2), 299-319.
- Qin, Huanmei, Guan, Hongzhi, Zhang, Zhihu, Tong, Liu, Gong, Liyuan, & Xue, Yunqiang. (2013). Analysis on Bus Choice Behavior of Car Owners based on Intent–Ji'nan as an Example. *Procedia-Social and*

- Behavioral Sciences*, 96, 2373-2382.
- Ramming, Michael Scott. (2002). *Network knowledge and route choice*. Massachusetts Institute of Technology.
- Raveau, Sebastián, Guo, Zhan, Muñoz, Juan Carlos, & Wilson, Nigel HM. (2014). A behavioural comparison of route choice on metro networks: Time, transfers, crowding, topology and socio-demographics. *Transportation Research Part A: Policy and Practice*, 66, 185-195.
- Satiennam, Thaned, Jaensirisak, Sittha, Satiennam, Wichuda, & Detdamrong, Sumet. (2015). Potential for modal shift by passenger car and motorcycle users towards Bus Rapid Transit (BRT) in an Asian developing city. *IATSS Research*.
- Schakenbos, Rik. (2014). Valuation of a transfer in a multimodal public transport trip: a stated preference research into the experienced disutility of a transfer between bus/tram/metro and train within the Netherlands.
- Sheffi, Yosef. (1985). Urban transportation networks: equilibrium analysis with mathematical programming methods.
- Sheffi, Yosef, Mahmassani, Hani, & Powell, Warren B. (1982). A transportation network evacuation model. *Transportation Research Part A: General*, 16(3), 209-218.
- Szeto, WY, Solayappan, Muthu, & Jiang, Yu. (2011). Reliability- Based Transit Assignment for Congested Stochastic Transit Networks. *Computer- Aided Civil and Infrastructure Engineering*, 26(4), 311-326.
- Tavasszy, Lóri, Snelder, Maaïke, Haaijer, Rinus, Meurs, Henk, van Nes, Rob, Verroen, Erik, . . . Jansen, Ben. (2012a). Audit LMS en NRM eindrapport stap 2: TNO.
- Tavasszy, Lóri, Snelder, Maaïke, Haaijer, Rinus, Meurs, Henk, van Nes, Rob, Verroen, Erik, . . . Jansen, Ben. (2012b). Audit LMS en NRM Syntheserapport: TNO.
- Van der Heuvel, Jeroen. (2013). Netherlands Railways company overview: NS stations.
- Van der Waard, J. (1988). *The relative importance of public transport trip time attributes in route choice*. Paper presented at the PTRC Summer Annual Meeting, 16th, 1988, Bath, United Kingdom.
- Van der Waard, J. (1990). *Koncept elasticiteiten handboek*. Rijkswaterstaat: Dienst Verkeerskunde, Rotterdam.
- van Hagen, Mark. (2011). *Waiting experience at train stations*: Eburon Uitgeverij BV.
- Vande Walle, Stefaan, & Steenberghen, Therese. (2006). Space and time related determinants of public transport use in trip chains. *Transportation Research Part A: Policy and Practice*, 40(2), 151-162. doi: <http://dx.doi.org/10.1016/j.tra.2005.05.001>
- Vij, Akshay, Carrel, André, & Walker, Joan L. (2013). Incorporating the influence of latent modal preferences on travel mode choice behavior. *Transportation Research Part A: Policy and Practice*, 54, 164-178.
- Vovsha, Peter, & Bekhor, Shlomo. (1998). Link-nested logit model of route choice: overcoming route overlapping problem. *Transportation Research Record: Journal of the Transportation Research Board*, 1645(1), 133-142.
- Vuchic, Vukan R. (2005). *Urban transit: operations, planning, and economics*.
- Williams, Huw CWL. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A*, 9(3), 285-344.
- Yang, Rongrong, Yan, Hai, Xiong, Wen, & Liu, Tao. (2013). The Study of Pedestrian Accessibility to Rail Transit Stations based on KLP Model. *Procedia-Social and Behavioral Sciences*, 96, 714-722.
- Zhang, Zhihu, Guan, Hongzhi, Qin, Huanmei, & Xue, Yunqiang. (2013). A Traffic Mode Choice Model for the Bus User Groups based on SP and RP Data. *Procedia-Social and Behavioral Sciences*, 96, 382-389.

Appendix 1 – ‘Duin- en Bollenstreek’

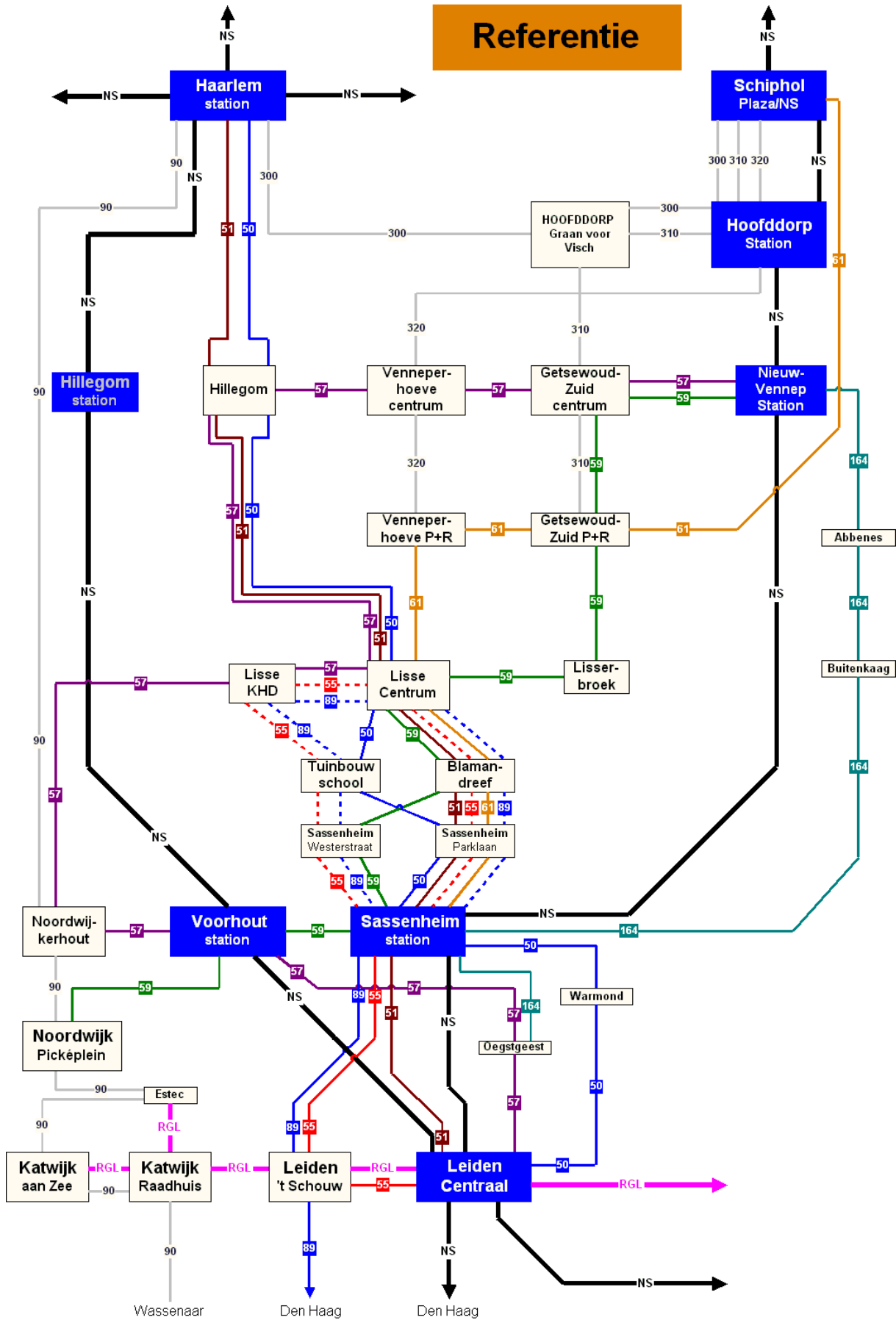


Figure 23: “Duin- en Bollenstreek” reference situation

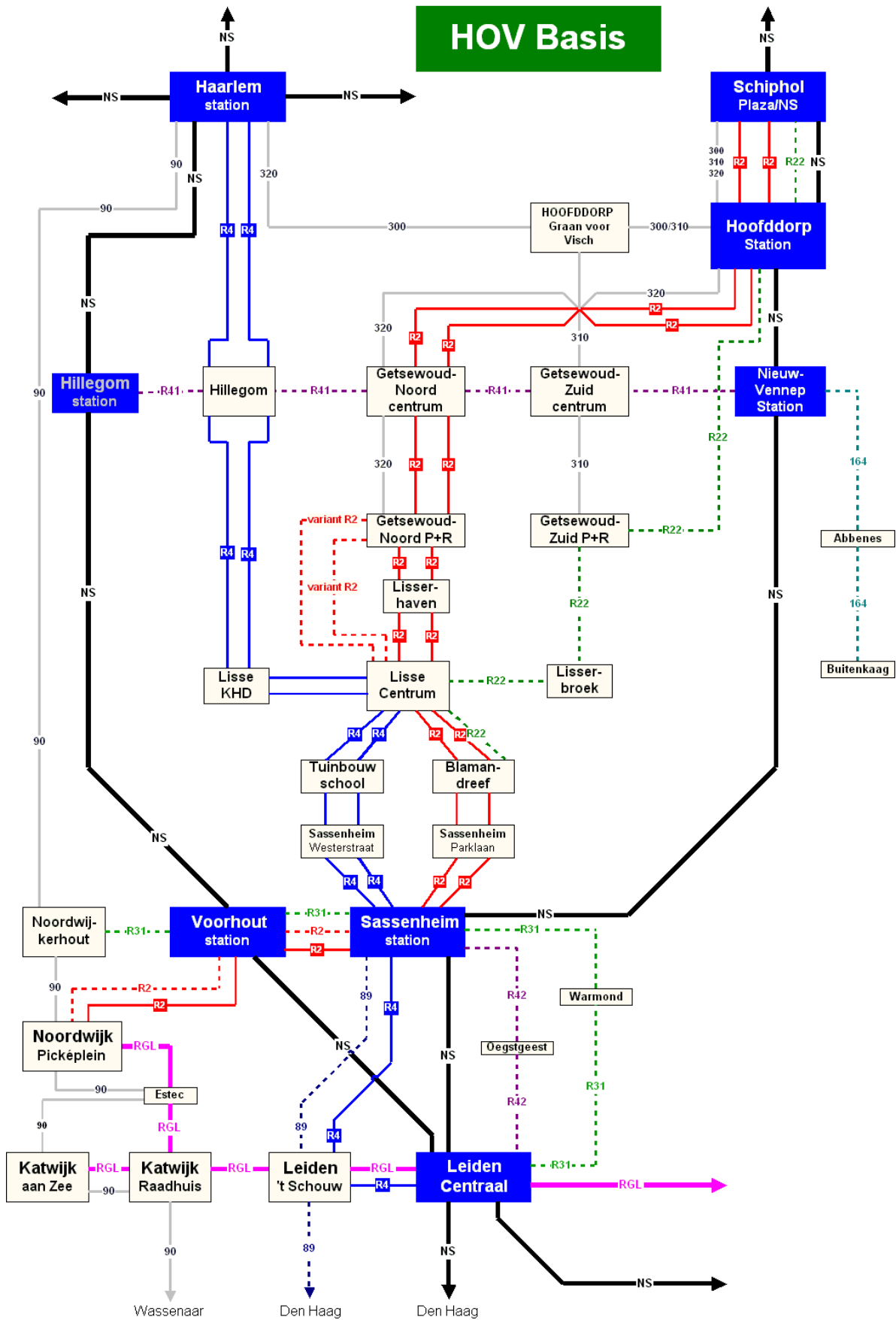


Figure 24: “Duin- en Bollenstreek” HOV basic scenario

Appendix 2 – Weight of trip parameters

Author and remarks	parameter	Weight
	IVT	1,0
Z. Guo and N. H. Wilson (2011)	Entry/exit walking	0,9
	Initial waiting	1,1
	Nr. of transfers	4,9
Abrantes and Wardman (2011)	Walk	1,7
	Out of vehicle time	1,4
	wait	1,7
	Search	1,4
	Late	6,4
	Headway	0,8
	Departure time shift	0,7
(versus car)	Main trip bus	1,2
(versus car)	Main trip train	1,1
(Raveau, Guo, Muñoz, & Wilson, 2014)	Waiting	1,6
Santiago metro morning peak average	Transfer	5,9
	Walking	2,2
London underground morning peak average	Waiting	1,9
	Transfer	5,4
	Walking	1,3
Vande Walle and Steenberghen (2006)	Out of vehicle time	1,5-2,3
Arentze and Molin (2013)	Main trip car	1,1
In comparison to main trip bus / tram	Main trip bicycle	1,0
	Main trip train	0,8
	Access/egress walk	1,4
	Access/egress bike	1,5
	Wait on access	1,0
	Transfer	10
Fosgerau, Hjorth, and Vincent Lyk-Jensen (2007)	Additional driving time due to congestion	1,0
	Headway <12	1*H
	Headway >12	12+0,4*(H-12)
Compared to car IVT	egress walk	1,5
Compared to bus IVT	Access/egress time	1,8
	Transfer	12,2
	Waiting on transfer	2,0
Compared to train IVT	Access/egress time(<=50 km)	1,3
	Access/egress time (>50 km)	0,9
	Transfer	12,9
	Waiting on transfer	1,4
Advised values PT	Access/egress time	1,5
	Transfer	6
	Waiting on transfer	1,5
Original HOV-scanner	Access and egress	1,2
Compared towards PT IVT	Transfer	1,1
	Waiting	1,3
	bicycle	1,0
	car	1,6

Table 12: Weight of trip parameters (IVT normalized to 1)

Appendix 3 – Program flow chart

In this appendix will be explained how the data is processed by the program code. For a flow chart of the program see Figure 25. This flowchart provides detailed information about the method used to process the schedule information into travel times by PT. This is considered to be the core of the program. Other elements are explained in more general terms. The numbers between brackets in this paragraph correspond with the numbers steps of Figure 25.

First of all the program generates a matrix with all unique start and stop PT stations (1). Then the departure time is set (2). For the first OD combination the program searches for the first possible departure time per transit line. The first temporary table will be filled (inter zonal combination will be filled with zeros) (3).

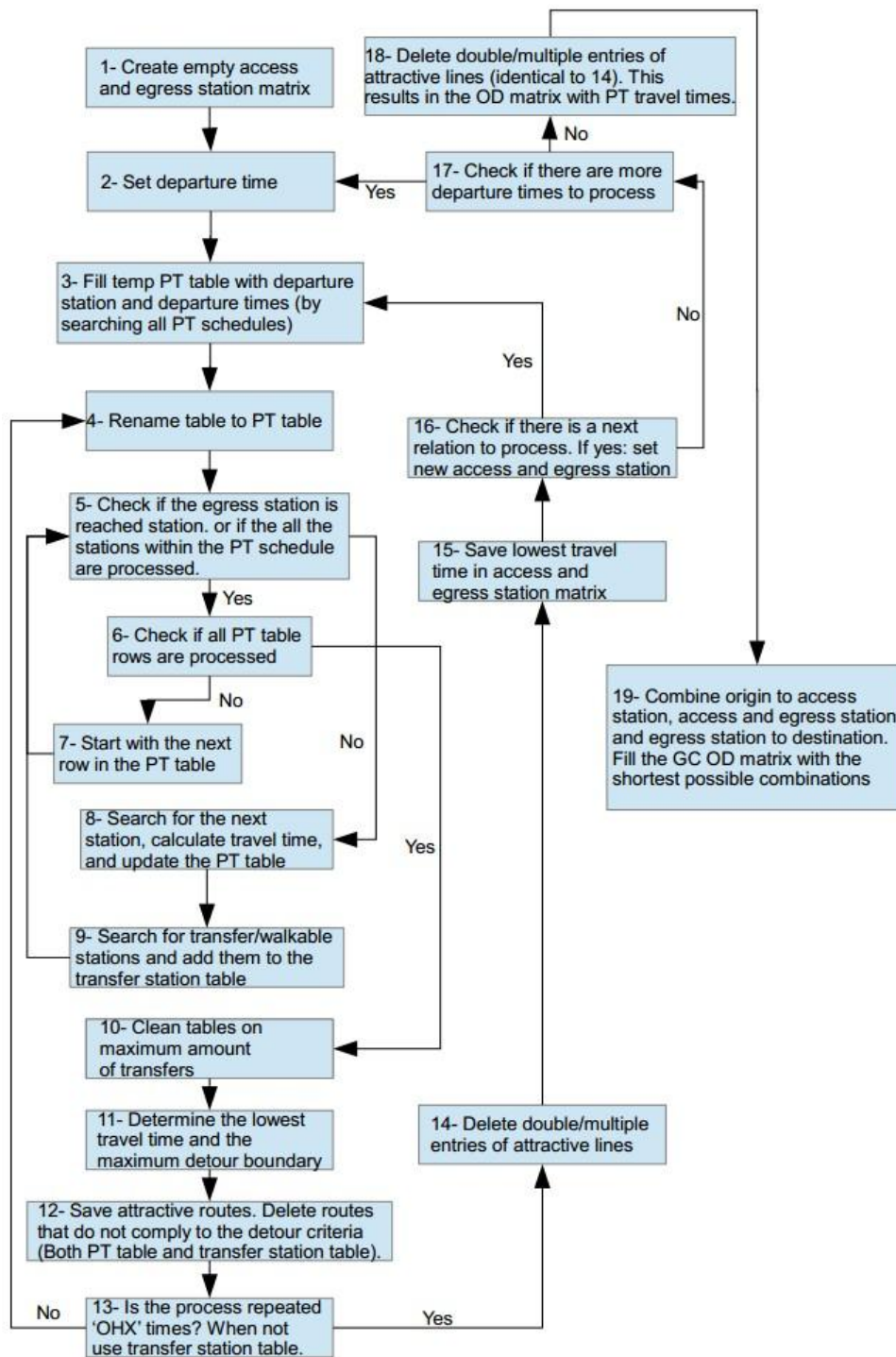


Figure 25: Travel time program


Then the table is renamed (4). The first OD combination is set and there will be checked if the egress station is reached (5). If the current station is not the egress station the program will move to the next station within the current schedule. Once this is done it calculates the distance (in time) to the next stop (8). The result is stored in a table. When the next stop is found the program searches possible transfers and for walkable stops (9). Once this is done the iterative process is repeated to find the next stop. This process ends when the egress station is found and when all stations within the time table are checked (5). Once the entire table is processed (6) the tables will be cleaned by removing routes that contain too much transfers (10). Then the minimum travel time for this OD relation and the maximum detour boundary are updated if applicable (11). Subsequently the attractive routes are saved and the routes that do not comply with the detour criteria are deleted in both the PT table and the transfer station table (12).

When this iterative loop is finished the table that is filled with transfer lines and walkable stops is used as the new input file and the process repeats (13). If the number of loops is reached double entries will be deleted (14) and the attractive routes will be saved in matrix that contains the travel times between all stops (15). Once all OD relations are processed this loop is finished (16). If not all OD relations are finished the next OD combination will be set (16). If there is a wider interval and a tolerance set for the departure time this process repeats with the next departure time (17). If all departure times are processed alternatives that have multiple entries will be deleted so that each alternative only appears once in the table (18). This is identical to step 14. The last step here is to delete routes that do not comply with the overlap filter. The result is a matrix that contains the choice set and GC information from all access to all egress stations.

Once the station distance table is created the origin/destination distance (in time) to the corresponding stations is calculated (19). All combinations of access station, egress station and modes that comply to the detour criterion are included in the OD matrix.

Appendix 4 – Program output


The tables and matrices used by the program are explained in this appendix. For every iteration the following information is stored in a 24 column matrix (figure 26):

 2x24 cell

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	18	4	4	76	76	0.3715	0.3715	0.3715	30.0000	30.0000	19	1	26.0000	0	3.0000	0	13.0000	0	13.0000	0	[88;53;52;76]	4	[4;2;2]	[13;9;4]
2	18	4	4	76	76	0.3674	0.3674	0.3674	24.0000	24.0000	17	1	21.0000	0	2.0000	0	4.0000	0	17.0000	0	[88;53;52;76]	4	[4;4;2]	[13;4;4]

Figure 26: PT time table

1. Sheet number of the schedule, used to locate the station/stop within the file.
2. Row number within the sheet of the current stop, used to locate the station/stop within the file.
3. Row number within the sheet of the current stop, used to locate the station/stop within the file.
4. The code of the current stop. This is a unique code assigned to a stop/station.
5. The code of the current stop. This is a unique code assigned to a stop/station.
6. Departure time at the previous stop
7. Arrival time at the current stop
8. Departure time at the current stop
9. Delta arrival time. Arrival time current stop minus departure time previous stop adding the delta departure value if applicable.
10. Delta departure time. Departure time current stop minus departure time previous stop adding the delta departure value if applicable.
11. The column number in the schedule where the departure time from the previous stop is found.
12. The amount of transfers to reach the current stop
13. Total IVT
14. Delta walk time. The amount of time in minutes spend walking. This is the walking time within the PT section (figure 13, yellow square).
15. Delta waiting time. The amount of time spend waiting in the PT section.
16. Delta in vehicle time vehicle type 1
17. Delta in vehicle time vehicle type 2
18. Delta in vehicle time vehicle type 3
19. Delta in vehicle time vehicle type 4
20. Delta in vehicle time vehicle type 5
21. List with passed stations
22. Frequency of the route
23. List with used modes
24. List with IVT per mode

 32x24 cell

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	18	2	0	53	0	0.3625	0	0	0	17.0000	19	1	13.0000	0	3.0000	0	0	0	13.0000	0	[88;53]	4	4	13
2	19	4	0	53	0	0.3625	0	0	0	17.0000	15	1	13.0000	0	3.0000	0	0	0	13.0000	0	[88;53]	4	4	13

Figure 27: Transfer and walk stops

The table that contains the transfers and walkable stops Figure 27 is almost identical to the PT time table. The difference is that the information on position 3, 5, 7, 8 and 9 is not used (and stored as 0) in this table.

HO4matrixtest						
21x21 cell						
	1	2	3	4	5	6
1	1x21 cell	1x24 cell	1x24 cell	1x24 cell	4x25 cell	1x24 cell
2	1x24 cell	1x21 cell	1x24 cell	1x24 cell	5x25 cell	4x25 cell
3	2x25 cell	1x24 cell	1x21 cell	1x24 cell	1x24 cell	1x24 cell
4	3x25 cell	1x24 cell	1x24 cell	1x21 cell	1x24 cell	1x24 cell
5	3x25 cell	6x25 cell	1x24 cell	1x24 cell	1x21 cell	1x24 cell
6	7x25 cell	5x25 cell	3x24 cell	2x24 cell	2x24 cell	1x21 cell

Figure 28: PT matrix

The final data that contains the route set per OD and the GC information is stored in the PT matrix (figure 28). Each cell in this table is a route set and per route the information displayed in the PT time table (figure 26) is stored.

ODmatrixtotaal						
21x21 cell						
	1	2	3	4	5	6
1	[]	1x32 cell	1x32 cell	1x32 cell	1x32 cell	1x32 cell
2	1x32 cell	[]	1x32 cell	2x32 cell	2x32 cell	4x32 cell
3	1x32 cell	1x32 cell	[]	1x32 cell	1x32 cell	1x32 cell
4	1x32 cell	2x32 cell	1x32 cell	[]	1x32 cell	4x32 cell
5	1x32 cell	2x32 cell	1x32 cell	1x32 cell	[]	5x32 cell
6	1x32 cell	5x32 cell	1x32 cell	5x32 cell	4x32 cell	[]

Figure 29: OD matrix

ODmatrixtotaal{13, 4}								
	25	26	27	28	29	30	31	32
1	46	69	0	0	4.5000	4.5000	65	1
2	46	69	0	0	4.5000	4.5000	65	1
3	46	69	0	0	4.5000	4.5000	80	1
4	46	69	0	0	4.5000	4.5000	80	1

Figure 30: OD matrix stored information

The total trip information is stored in the OD matrix (figure 29). The information in each cell for the first 24 columns identical to the PT time table (figure 26). The additional information regarding the total trip information is stored in column 25 until 32 (figure 30):

- 25. Access station
- 26. Egress station
- 27. Access mode (0=walking 1=bicycle)
- 28. Egress mode (0=walking 1=bicycle)
- 29. Access time
- 30. Egress time
- 31. Total travel time
- 32. Transfer time

Appendix 5 – Model process flow chart

To get an overview of the entire process used to forecast the scenario the flowchart is included (figure 31). In this overview the purple blocks are input data, the green blocks represents data processing or calculations and the blue blocks are data blocks. The flowchart is used to provide more insight about data flows and process steps. Across this section there is referred to blocks of this figure by providing a number between brackets.

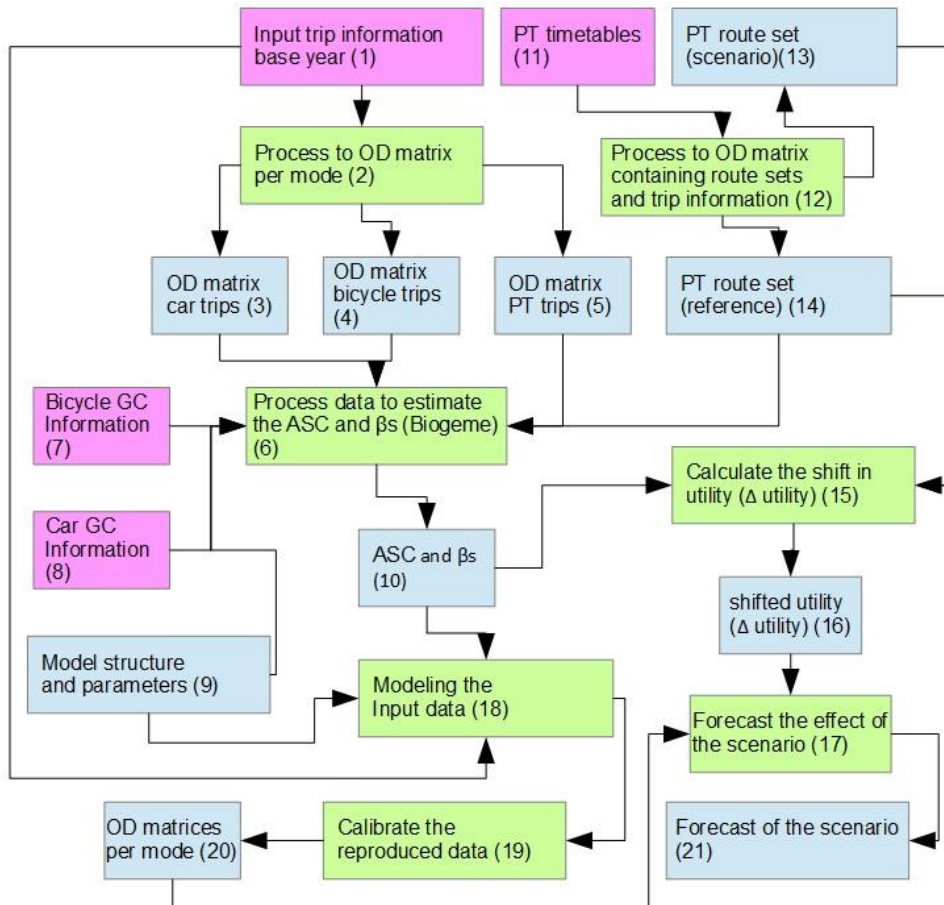


Figure 31: Flowchart

First the PT route sets will be conducted (12) for both the reference (14) and the scenario (13). The base year information processed and OD matrices per mode are constructed (3, 4 and 5). GC information (7, 8 and 14), trip information (3, 4 and 5) and the model structure (9) are used to estimate the model parameters and alternative specific constants with Biogeme (6). This results in parameters that are specific for this case (10). The next step is the reproduction of the input data with the model (18) and a calibration per cell of the reproduced data (19). This results in OD matrices per mode (20). The reference route set (14), the scenario route set (13) and the weights of the parameters (10) are used to calculate the shift in utility (15). Finally the shifted utility (16) and the OD matrices per mode (20) are used to forecast the effect of the scenario (21).

Appendix 6 – The error in the model fit for PT

absolute error (nr.)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
1 Hoofddorp stations.		3499	4	194	-62	456	-137	9	6	7	-5	16	34	5	2	-14	-18	-56	-394	-74	-5	
2 Hoofddorp overig	3837		-41	35	-181	-52	-309	-25	6	-2	-17	9	23	-5	-9	-14	-31	-433	-1159	-234	-68	
3 Nw Vennepe stations	-21	-59		1212	176	-229	222	75	30	2	11	12	16	7	-5	-3	-12	-61	-618	-79	-44	
4 Nw Vennepe Getsew	191	34	1213		504	-47	-13	49	6	0	14	10	12	12	-5	1	-8	-144	-323	-122	-50	
5 Nw Vennepe Vennepe	51	-5	180	414		-53	460	115	10	-1	21	3	5	-3	1	0	3	-58	-150	-15	-18	
6 Schiphol	316	-301	-228	-80	-107		-224	-3	4	13	-5	6	40	-8	9	-29	-42	-106	26	0	-84	
7 Hillegom	-154	-244	242	12	302	-257		753	131	-3	33	48	100	-34	5	-6	0	-179	-494	-179	-145	
8 Lisse overig	-16	-29	145	31	94	13	761		673	20	139	31	81	50	-43	22	2	20	-99	141	-62	
9 Lisse Gerard Dousst	10	11	24	10	3	3	124	669		1	13	8	31	-4	-9	-7	-14	-3	-7	-5	-61	
10 Noordwijk	22	10	4	0	-4	39	33	44	-1		993	13	55	423	8	38	670	18	43	222	177	
11 Noordwijkerhout	-2	-17	15	9	2	-6	28	341	6	835		8	-148	-168	-16	-10	278	-79	-178	-35	-123	
12 Sassenheim Zuid/st	25	10	6	3	-8	41	48	75	7	-11	4		727	-20	69	-5	-36	-9	35	25	-149	
13 Sassenheim Noord	31	17	9	4	-12	40	75	150	29	21	-123	593		-15	-46	3	3	-15	12	159	-141	
14 Voorhout	-21	-30	6	2	-18	-49	-82	32	-13	438	-177	-2	97		8	86	183	-198	-195	-14	-345	
15 Warmond	1	-8	-9	-10	-5	-3	0	-34	-9	4	-24	97	-44	9		55	20	-4	-4	262	9	
16 Oegstgeest	-23	-19	-8	3	-2	-24	-8	22	2	25	-16	32	56	49	54		233	-26	-101	831	-73	
17 Katwijk	-30	-15	-9	-5	-12	-65	1	33	3	704	187	8	49	173	-22	119		31	-146	1105	281	
18 Haarlem	-157	-408	-134	-200	-131	-160	-276	-37	-19	15	-47	-5	2	-105	-2	-24	-29		-4075	-315	-152	
19 Amsterdam	-179	-1077	-719	-346	-265	983	-422	-56	-11	-30	-115	-22	22	-125	12	-132	-147	-1233		-1155	-976	
20 Leiden NS	-95	-294	-190	-142	-64	148	-218	84	-3	208	-37	113	155	7	195	789	1144	-340	-935		294	
21 Den Haag	9	-53	-137	-79	-45	50	-165	-8	-36	114	-64	-18	-39	-213	7	-21	229	-208	-761	662		
error (%)		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1 Hoofddorp stations.		150	1	72	-22	30	-50	10	37	29	-37	67	298	9	9	-23	-16	-7	-14	-13	-2	
2 Hoofddorp overig	161		-9	8	-58	-4	-63	-15	26	-7	-53	48	109	-11	-44	-38	-35	-16	-52	-41	-23	
3 Nw Vennepe stations	-6	-16		455	104	-56	100	131	367	80	499	655	607	55	-29	-23	-32	-35	-70	-25	-25	
4 Nw Vennepe Getsew	75	10	452		576	-19	-3	64	65	-5	284	285	228	274	-28	13	-27	-53	-68	-42	-31	
5 Nw Vennepe Vennepe	34	-3	173	528		-37	282	280	109	-23	342	88	77	-32	62	-11	28	-45	-64	-25	-42	
6 Schiphol	21	-20	-50	-27	-45		-69	-3	30	59	-41	10	216	-7	44	-27	-28	-19	0	0	-17	
7 Hillegom	-57	-59	111	3	115	-74		454	537	-17	25	394	382	-29	146	-48	0	-19	-79	-62	-60	
8 Lisse overig	-12	-13	194	24	87	11	304		66	47	89	21	25	35	-43	371	2	11	-39	238	-31	
9 Lisse Gerard Dousst	53	41	316	143	21	19	567	72		11	105	24	40	-12	-57	-52	-45	-11	-22	-7	-65	
10 Noordwijk	131	42	169	-9	-23	940	331	116	-14		1132	36	103	349	42	1059	37	37	215	123	468	
11 Noordwijkerhout	-16	-44	678	184	8	-35	19	210	38	1366		11	-55	-33	-81	-29	680	-45	-85	-25	-65	
12 Sassenheim Zuid/st	412	58	315	83	-60	876	299	50	18	-25	4		285	-12	73	-7	-38	-34	151	10	-53	
13 Sassenheim Noord	287	64	422	82	-62	219	243	58	42	45	-48	287		-5	-29	4	3	-35	16	120	-49	
14 Voorhout	-26	-43	167	28	-69	-37	-49	16	-31	322	-30	-1	32		18	355	131	-55	-61	-2	-50	
15 Warmond	5	-47	-70	-74	-74	-9	-4	-42	-62	25	-88	118	-26	20		129	50	-49	-13	123	14	
16 Oegstgeest	-29	-45	-50	109	-36	-22	-59	481	89	678	-49	82	200	211	146		93	-56	-47	104	-30	
17 Katwijk	-26	-18	-34	-25	-48	-40	3	85	30	53	356	17	74	168	-34	59		25	-49	70	46	
18 Haarlem	-19	-13	-63	-59	-66	-29	-28	-26	-60	39	-33	-17	5	-42	-20	-54	-21		-72	-44	-36	
19 Amsterdam	-6	-46	-83	-70	-72	17	-77	-37	-34	-51	-80	-23	35	-55	35	-57	-54	-19		-45	-49	
20 Leiden NS	-14	-45	-56	-56	-58	23	-75	133	-8	209	-23	76	140	1	96	100	67	-49	-39		11	
21 Den Haag	3	-15	-74	-49	-56	13	-84	-7	-71	422	-55	-13	-23	-56	26	-9	43	-53	-39	25		

Table 13: Model fit error for PT

Appendix 7 – Trip forecast

Reference	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		
1 Hoofddorp stationsgebied		2331	362	268	279	1511	276	84	16	24	13	23	12	60	18	64	113		
2 Hoofddorp overig	2386		437	432	312	1470	495	164	24	25	31	19	22	44	20	37	87		
3 Nw Vennep stationsgebied	341	367		266	169	406	222	57	8	3	2	2	3	12	18	14	37		
4 Nw Vennep Getsewoud	254	333	269		87	246	384	76	9	4	5	3	5	4	16	8	29		
5 Nw Vennep Vennepelhoev	147	186	104	78		145	163	41	9	5	6	4	6	10	2	4	11		
6 Schiphol	1530	1507	453	293	236		326	75	14	22	13	57	18	109	19	105	153		
7 Hillegom	272	413	219	358	263	347		166	24	15	129	12	26	116	4	12	41		
8 Lisse overig	140	225	75	129	109	117	250		1014	43	156	147	327	144	99	6	73		
9 Lisse Gerard Dousstraat	18	28	8	7	15	16	22	935		10	12	32	77	37	16	13	32		
10 Noordwijk	17	23	2	5	15	4	10	38	9		88	35	53	121	19	4	1817		
11 Noordwijkerhout	15	39	2	5	21	17	144	163	16	61		73	269	501	20	33	41		
12 Sassenheim Zuid/station	6	17	2	4	14	5	16	150	39	43	95		255	165	95	72	94		
13 Sassenheim Noord	11	26	2	5	20	18	31	259	68	47	256	207		276	159	86	111		
14 Voorhout	80	70	4	7	26	134	165	192	41	136	592	151	304		41	24	140		
15 Warmond	16	17	12	14	7	28	4	82	15	15	28	82	170	42		43	40		
16 Oegstgeest	77	42	17	3	6	108	14	5	2	4	33	39	28	23	37		250		
17 Katwijk	114	84	27	20	26	162	39	38	11	1322	53	49	67	103	63	201			
																		Total trips	39842
Forecast logsum 0,1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		
1 Hoofddorp stationsgebied	0	0	-119	0	23	0	-6	42	11	11	8	-2	0	25	2	-2	0		
2 Hoofddorp overig	0	0	16	206	204	3	90	99	19	8	8	-1	1	12	-4	-10	-28		
3 Nw Vennep stationsgebied	64	175	0	30	18	56	36	14	2	3	-1	0	0	1	-12	-5	-9		
4 Nw Vennep Getsewoud	1	215	30	0	0	3	17	9	5	5	-4	0	1	2	-10	-2	0		
5 Nw Vennep Vennepelhoev	14	62	11	0	0	10	9	20	9	10	-2	2	5	6	-1	1	3		
6 Schiphol	0	0	-95	0	13	0	13	22	3	10	8	-4	0	44	4	-3	0		
7 Hillegom	84	430	36	16	14	78	0	-11	-7	13	-71	4	7	-2	0	1	33		
8 Lisse overig	69	135	14	8	50	34	-13	0	-16	13	-79	25	48	22	-22	1	41		
9 Lisse Gerard Dousstraat	6	6	-2	-1	12	3	-4	23	0	2	-4	3	6	6	0	-1	3		
10 Noordwijk	2	6	1	1	9	1	-2	6	5	0	0	22	38	7	-3	0	0		
11 Noordwijkerhout	10	21	-1	-3	-7	11	-102	-90	-3	0	0	21	31	0	23	-16	0		
12 Sassenheim Zuid/station	0	-1	0	1	13	0	3	34	4	26	20	0	7	101	-39	-9	34		
13 Sassenheim Noord	1	0	0	1	22	1	5	52	5	35	16	5	0	185	-56	-9	49		
14 Voorhout	5	7	1	2	13	8	-16	34	20	4	0	77	177	0	19	-9	10		
15 Warmond	1	3	-3	3	5	4	4	0	1	10	24	-13	-20	35	0	5	-4		
16 Oegstgeest	-12	-4	-1	0	0	-2	-1	0	-1	2	-16	-9	-6	-8	-18	0	0		
17 Katwijk	1	3	-9	3	7	1	10	9	0	0	0	7	12	-6	6	30	0		
																		Total trips	43189
																		Growth (%)	8,4

Table 14: Forecast reference and logsum 0,10

Forecast logsum 0,25		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17			
1	Hoofddorp stationsgebied	0	0	-119	0	23	0	-8	30	6	8	4	-2	2	31	1	-7	3			
2	Hoofddorp overig	0	0	16	206	152	41	132	64	12	5	1	2	4	12	-3	-10	-21			
3	Nw Vennep stationsgebied	64	222	0	30	18	69	36	9	2	2	-1	0	0	1	-11	-4	-8			
4	Nw Vennep Getsewoud	14	215	30	0	0	28	17	7	4	4	-3	1	2	2	-9	-1	1			
5	Nw Vennep Venneperhoeve	14	84	11	0	0	12	9	20	9	8	-2	2	6	6	0	1	6			
6	Schiphol	0	0	-62	-3	20	0	49	22	3	8	5	-4	2	54	4	-12	4			
7	Hillegom	85	326	36	16	14	82	0	-19	-8	9	-70	3	8	10	0	-1	20			
8	Lisse overig	45	89	6	-1	50	33	-35	0	-101	15	-68	22	48	34	-18	0	21			
9	Lisse Gerard Dousstraat	4	5	-3	-1	12	3	-6	23	0	2	-5	3	6	8	0	-2	0			
10	Noordwijk	2	4	1	1	8	0	-1	7	3	0	0	22	38	7	0	0	0			
11	Noordwijkerhout	7	15	-1	-2	-6	9	-97	-87	-4	0	0	23	46	0	19	-14	1			
12	Sassenheim Zuid/station	0	-1	0	1	11	0	3	51	4	22	23	0	7	110	-39	-12	23			
13	Sassenheim Noord	1	0	0	2	18	2	6	80	5	30	31	5	0	200	-56	-11	34			
14	Voorhout	5	6	1	2	11	8	2	28	14	4	0	65	150	0	19	-9	16			
15	Warmond	0	3	-1	2	4	6	3	0	1	9	20	-13	-20	41	0	0	-2			
16	Oegstgeest	-17	-6	-2	0	0	-8	-2	-1	-1	1	-14	-10	-7	-8	-17	0	9			
17	Katwijk	5	7	-6	2	8	9	4	5	-1	0	-2	10	17	-1	3	10	0			
																			Total trips	43055	
																				Growth (%)	8,1
Forecast logsum 0,4		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17			
1	Hoofddorp stationsgebied	0	0	-119	0	23	0	2	22	2	4	0	-2	5	38	-1	-14	10			
2	Hoofddorp overig	0	0	16	206	99	101	195	30	5	1	-5	5	8	12	0	-10	-10			
3	Nw Vennep stationsgebied	64	273	0	30	18	87	36	4	2	2	-1	0	1	0	-9	-3	-6			
4	Nw Vennep Getsewoud	35	215	30	0	0	60	17	0	3	2	-3	1	3	3	-7	0	4			
5	Nw Vennep Venneperhoeve	14	108	11	0	0	21	9	20	9	6	-2	3	7	8	0	1	9			
6	Schiphol	0	0	-37	-7	23	0	87	24	4	4	0	-4	6	67	4	-22	13			
7	Hillegom	91	250	36	16	14	92	0	-23	-9	6	-67	2	10	21	0	-3	12			
8	Lisse overig	25	57	-5	-12	50	28	-56	0	-181	15	-54	17	48	56	-16	0	6			
9	Lisse Gerard Dousstraat	2	5	-3	-1	12	3	-8	23	0	2	-6	3	6	11	-1	-4	-2			
10	Noordwijk	2	1	1	1	8	0	0	10	1	0	0	22	38	7	4	-1	0			
11	Noordwijkerhout	3	7	-1	-2	-4	5	-85	-82	-5	0	0	21	68	0	14	-10	3			
12	Sassenheim Zuid/station	0	-1	0	1	10	0	3	68	4	16	23	0	7	113	-39	-18	16			
13	Sassenheim Noord	1	0	1	2	13	5	9	110	5	21	52	5	0	206	-56	-14	24			
14	Voorhout	5	6	1	1	9	8	11	17	8	4	0	50	116	0	19	-8	24			
15	Warmond	-2	3	0	2	3	9	1	0	1	6	15	-13	-20	48	0	-4	0			
16	Oegstgeest	-21	-7	-3	-1	0	-13	-4	-2	-1	1	-12	-11	-8	-9	-16	0	27			
17	Katwijk	12	10	-2	2	9	18	1	3	0	0	-4	16	25	12	1	-24	0			
																			Total trips	43109	
																				Growth (%)	8,2

Table 15: Logsum 0,25 and 0,40

Appendix 8 – Elasticity

In Dutch studies the band of the elasticity toward changes in travel time is quite broad. Differences from -0,4 to -1,6 can be found. This is in line with international studies. On average a value around -1 is assumed.

Source	area	details	elasticity	classification
Van der Waard (1990)	Bus streetcar metro	Short term	-0,88 morning peak -0,85 evening peak	
Van der Waard (1990)	Bus streetcar metro	Long term	-1,2 morning peak -1,21 evening peak	
LMS2011	Bus streetcar metro	Long term IVT	-0,593	Model output
de Beer, Tiemersma, and Ploeg van der (2011)	Bus streetcar metro		-0,9 tot -1,1	recommendation
MuConsult (2003)	train	Trip time	-0,5 tot -0,6	recommendation
Van der Waard (1990)	train	Long term travel time	-1,63 morning peak	Data 1995
Van der Waard (1990)	train	Long term travel time	-1,64 evening peak	Data 1995
LMS7 (LT)	train	Long term IVT	-0,812	
LMS2011	train	Long term IVT	-0,439	Model output

Table 16: Elasticity towards travel time