DELFT UNIVERSITY OF TECHNOLOGY / ROYAL HASKONING DHV

MASTER'S THESIS

Performance analysis on joint mode and destination choice: a tour-based model of Limburg

Author: Yuri Groeneveld (4389263), TU Delft
Faculty of Civil Engineering and Geosciences

Company supervisors: Erik de Romph, Royal Haskoning DHV
Alex Nugteren, Royal Haskoning DHV

Chairman: Adam Pel, TU Delft
Faculty of Civil Engineering and Geosciences

First supervisor: Winnie Daamen, TU Delft
Faculty of Civil Engineering and Geosciences

Second supervisor: Baiba Pudāne, TU Delft
Faculty of Technology, Policy and Management

June 5, 2023

Royal HaskoningDHV
Enhancing Society Together

Preface

This thesis concludes the era of me being a student. After this thesis, I completed my master's degree in Civil Engineering. Transport modelling is not something I wanted do to since I was a child. I didn't even know back then it was done by companies. It was only during my Civil Engineering bachelor's that I became interested in the field of transport modelling. I am glad I was introduced in this topic as I didn't choose the bachelor Civil Engineering because of transport modelling. I had lots of fun during my master's courses, applying small transport models and modelling traffic.

During my studies, I have never developed a large-scale transport model. Royal Haskoning DHV gave me the opportunity to work on such a model: Verkeers Model Limburg. I got to see first-hand what it is like to develop a large-scale transport model. For a course, I only experienced discrete choice modelling on a simple case. But now I got the experience in estimating and implementing a discrete choice model on a large-scale, and everything that is involved. For this opportunity, I would like to thank Erik and Alex. They also helped me a lot during my thesis with my writing skills and implementing the tour-based discrete choice model in Verkeers Model Limburg.

I would also like to thank Winnie. She helped me be critical of my thesis and explain every step I took. She always kept me in the right direction by exposing weak points in my thesis. She also introduced me to discrete choice modelling and other fields of transport modelling during my master's. I would also like to thank Adam and Baiba for the general discussions of my thesis and tips if I asked for help.

My thesis was a tough process for me. In the beginning, I had difficulties in my literature study finding and describing the specific elements I needed. Also writing my thesis was difficult but fortunately, my supervisors helped me with this aspect by reading my thesis over and over again.

I hope you enjoy reading this thesis as much as I do.

Yours sincerely,
Yuri

Management Summary

Transport models were invented in the 1960s to substantiate the decisions for new highway locations. The first generation transport model was a 4-step transport model. The four steps of a 4-step transport model consist of trip generation, trip distribution, modal split, and assignment. The trip generation step models the number of trips each zone produces and attracts. The trip distribution step matches all the origins with destinations to form trips. The third step is the modal split, where a mode is selected based on the trip. The fourth and final step is to assign the trips to the network to determine the load on the network. Over the years, new methods were developed to increase the level of detail and accuracy of these transport models.

Today, more advanced transport models are being used, but there is still a demand for innovations within transport modelling. Royal Haskoning DHV (RHDHV) develops (tour-based) transport models for provinces and municipalities to assess their transport networks and evaluate their policies. As such, RHDHV desires to improve its transport models because there is room for improvement in approximating the ground truth. RHDHV uses the tour-based distribution function model (TBDFM) to model the joint mode & destination choice of individuals in the transport model. But this method to model the mode & destination choice has some limitations. A different method to model the mode & destination choice is the usage of a discrete choice model (DCM). To determine if RHDHV should adopt this new method, this thesis answers the following research question:

How does the performance of modelling the OD matrix with a discrete choice model compare to modelling the OD matrix with the distribution function choice model, in a joint mode & destination choice?

To answer this research question, the following sub-questions are defined:

- How to construct a discrete choice model for a tour?
- Which data sources are available to compose the choice set?
- What are the estimated parameters of the tour-based discrete choice model and its effects?
- How is the performance of a choice model defined?

To answer the sub-questions, a literature study is performed. The literature pointed out that the multinomial logit (MNL), nested logit (NL), and cross nested logit (CNL) are the conventional DCMs used for the mode & destination choice. The MNL is a basic form. A choice is made between all the mode & destination alternatives available as depicted in Figure 1.

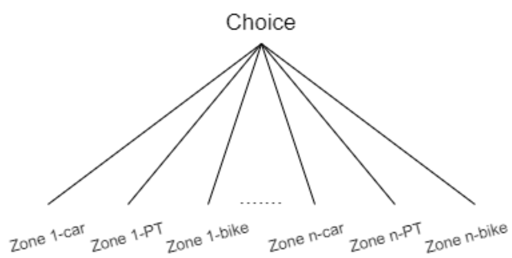

Figure 1: MNl model.

However, alternatives that have similar characteristics may be correlated. Hence a NL can be applied. The alternatives can be correlated by either the destination choice or the mode choice. As a result, a certain hierarchy is implemented as one choice is seen as more dominant. The choice structures of the NL-destination and NL-mode are presented in Figure 2 and Figure 3.

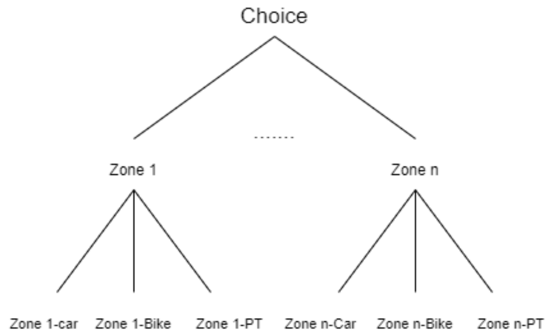

Figure 2: Destination choice nested.

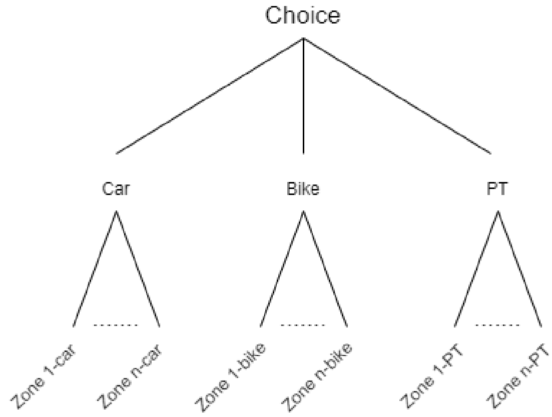

Figure 3: Mode choice nested.

The nesting of the mode alternatives and the destination alternative can also be combined. In a CNL model, each alternative is nested within both the mode nest and the destination nest as presented in Figure 4.

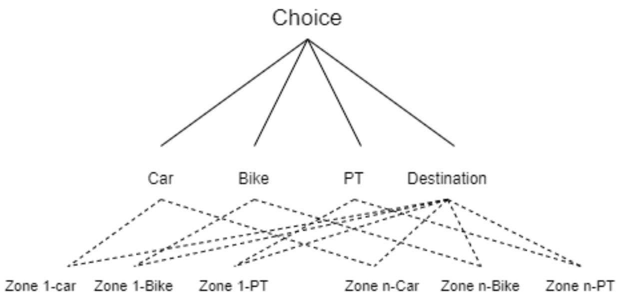

Figure 4: CNL model.

The chosen and non-chosen alternatives are required to estimate any DCM. The non-chosen alternatives of the mode choice are straightforward; there are only three modes to choose from. The destination choice is not straightforward. It is computational challenging to estimate a DCM with all the destinations as an alternative. In the literature, previous studies have tackled this problem by sampling the destinations. Hence, a strategy to sample the destinations was chosen: to sample 20 random zones as the alternatives. From those destinations the car, public transport (PT) and bike alternative were taken to compose the non-chosen alternatives for the choice set. Adding the chosen alternative to the choice results in a total of 63 alternatives.

A methodology was developed to answer the main research question. A data set needs to be composed to be able to estimate a DCM. Hence, data needs to be selected and analysed if fulfils the needs for this thesis. If the data set is finalised, the various DCMs can be estimated. Attributes are added and removed, one by one, to iteratively determine the best possible tour-based discrete choice model (TBDCM). The nesting structures have a corresponding parameter. This parameter will indicate the correlation among the alternatives. This parameter also has a lower bound: 1. Hence, these nest parameters have to be checked whether these are higher than 1. If that is not the case, the corresponding nest structure is rejected and the TBDCM will not be used for the remainder of this thesis. All the valid TBDCMs and the TBDFM will be subjected to a validation, sensitivity analysis and a complete destination set. For the validation an independent choice will be used to validate the predictions of the choice models. The sensitivity analysis will point out the elasticities of the choice models. By subjecting the choice models to a complete destination set, instead of the 63 alternatives in the choice set, it can be analysed how the choice models perform with similar alternatives.

Furthermore, a case study will be executed. RHDHV has a tour-based model: Verkeers Model Limburg (VML), which will be used for this thesis. One of the key performance indicator (KPI)s to determine the performance is the comparison with ground truth observations. These observation were supplied by the province and municipalities of Limburg. It will be used to analyse the performance of modelling an origin-destination (OD) matrix and also act as a data source. The chosen and non-chosen alternatives are required to estimate a DCM. Onderzoek Verplaatsingen in Nederland (OViN) is a revealed preference (RP) survey conducted by the Central Bureau of Statistics (CBS) each year. OViN is a great source for observations, but it lacks the non-chosen alternatives. When analysing the travel times of the OViN observations, it was concluded that the travel times are inaccurate and unsuitable for estimating a DCM. In the OViN data is an over-representation of certain travel times as depicted in Figure 5. Individuals have the tendency to round up or down, their travel time. Hence the over-representation of certain travel times do not reflect reality. A potential solution is to use the travel times of VML. These travel times are modelled; hence they do not contain human inaccuracies. The trips from OViN are between PC4 areas, while the travel times in VML are between VML zones. The size of the VML zones is inconsistent. Only in Limburg does the PC4 zones match with the VML zones. Hence only observations within Limburg are taken into account.

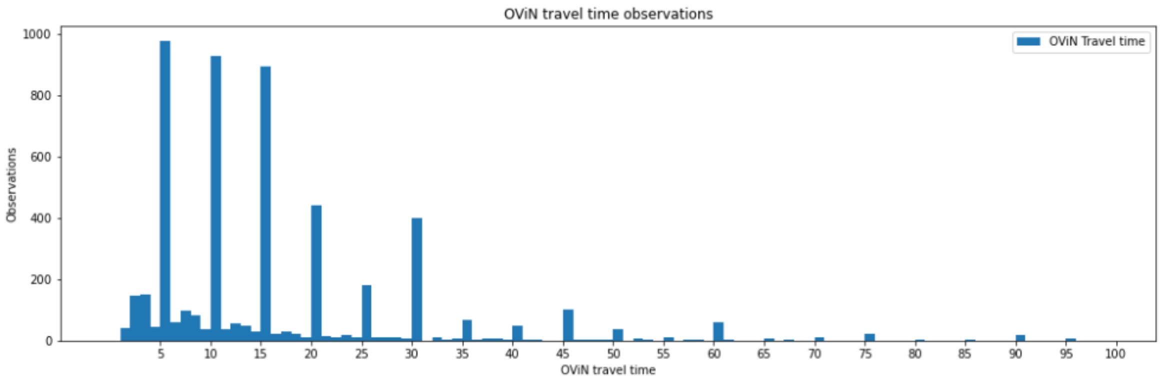

Figure 5: OViN travel times.

A TBDCM is needed to estimate a complete tour, a chain of multiple destinations. Based on the literature, the best method to determine the utility of the tour is to use the DCM for every trip sequentially. The utility of each trip is summed to determine the utility of the whole tour. In this thesis only uni-modal tours are taken into account. The tour length and composition are input from VML.

The literature also presented different methods to determine the performance of a choice model. Based on the literature, in-sample validation, out-of-sample validation, and a case study were executed. To be able to fit the TBDCM in VML, VML was analysed to determine the available attributes for the TBDCM. Only personal characteristics, zonal characteristics, and Level of Service (LoS) attributes that are present in VML can be used as an attribute in the TBDCM. The personal and zonal characteristics present in VML are presented in Table 1 and Table 2. The zonal characteristics relate to the corresponding trip purposes in VML.

Table 1: VML personal characteristics

Characteristics	**Categories**
Age	0-17, 18-34, 35-64, 65+
Employment	None, part-time, full-time
Household-income	<30k, 30-50k, >50k
Car availability	Yes/no

Table 2: Zonal characteristics.

Zonal characteristics	
Work	Shopping
Business	Other
Education	Urbanisation
Bring/get	

When the different types of DCMs were applied and estimated, only the MNL and the nest structure of the NL-mode were estimated significantly. The NL-mode was estimated with a μ for car, PT, and bike 1.04, 1.00, and 1.15, respectively. These low values for μ suggest a weak correlation among the car and bike alternatives and no correlation among PT alternatives. The $\overline{\rho}^2$ of the NL-mode and MNL is for both 0.577. An equal $\overline{\rho}^2$ indicates that the performance of the DCMs is equal. The added nest structure is estimated significantly but does not contribute to predicting the choices better. Hence there is no hierarchy between the mode choice and destination choice. The results are presented in Table 3 and Table 4.

Table 3: Estimated MNL model

Attributes	Car	PT	Bike
ASC	1.13**	-2.83**	0 (ref)
Travel time	-0.185**	-0.0107**	-0.123**
Age 0-17	-	-	3.7**
Age 18-34	-	-	-
Age 35-65	-	-0.598**	-
Age 66+	0.393**	-	-
Work full-time	0.608**	-	-
Work part-time	0.393*	-	-
Work nan	-0.773**	-	-0.609*
Income high	-	-	-
Income middle	-	-	-
Income low	-	-	-
Urban dest 1	-0.55**	0.802**	-
Urban dest 2	-	-	-
Urban dest 3	-	-	0.296**
Urban dest 4	-	-1.18**	-
Urban dest 5	-0.39**	-2.63**	-
Urban orig 1	-	1.19**	-
Urban orig 2	-	-	-
Urban orig 3	-	-	-
Urban orig 4	-	-	0.358**
Urban orig 5	0.251*	-1.00**	0.752**
Log(attraction) B	0.296**	-	0.659**
Log(attraction) D	0.549**	-	0.345**
Log(attraction) E	-	0.236**	0.328**
Log(attraction) S	0.414**	0.292**	0.523**
Log(attraction) W	0.726**	0.653**	0.933**
Log(attraction) O	0.409**	0.279**	0.581**
$\overline{\rho}^2$ = 0.577			

** = p<0.05, ** = p<0.01*

Table 4: Estimated NL-mode model

Attributes	Car	PT	Bike
ASC	1.04**	-3.01**	0 (ref)
Travel time	-0.178**	-0.0104**	-0.113**
Age 0-17	-	-	3.63**
Age 18-34	-	-	-
Age 35-65	-	-0.601**	-
Age 66+	0.391**	-	-
Work full-time	0.579**	-	-
Work part-time	0.263*	-	-
Work nan	-0.805**	-	-0.612*
Income high	-	-	-
Income middle	-	-	-
Income low	-	-	-
Urban dest 1	-0.529**	0.806**	-
Urban dest 2	-	-	-
Urban dest 3	-	-	0.296**
Urban dest 4	-	-1.18**	-
Urban dest 5	-0.378**	-2.63**	-
Urban orig 1	-	1.14**	-
Urban orig 2	-	-	-
Urban orig 3	-	-	-
Urban orig 4	-	-	0.328**
Urban orig 5	-	-0.97**	0.705**
Log(attraction) B	0.284**	-	0.616**
Log(attraction) D	0.537**	-	0.308**
Log(attraction) E	-	0.244**	0.299**
Log(attraction) S	0.396**	0.283**	0.486**
Log(attraction) W	0.698**	0.637**	0.878**
Log(attraction) O	0.392**	0.272**	0.542**
μ mode	1.04**	1.00***	1.15**
$\overline{\rho}^2$ = 0.577			

** = p<0.05, ** = p<0.01, *** = active bound*

The in-sample validation confirmed that the estimated MNL is very similar to the NL-mode. An independent choice was used to perform the out-of-sample validation, which pointed out that the TBDCMs performed better in predicting the chosen alternative. The TBDCMs assigned the chosen alternative with the highest probability of all possible alternatives more often than the TBDFM. Especially predicting the correct mode contributed to an improvement compared to the TBDFM. The out-of-sample validation also pointed out that the TBDCMs and the TBDFM predict a too high mode share for the car, compared to the observed modal split. Between the TBDCMs, the MNL performs slightly better than the NL. The nest structure does not lead to predicting the choices better.

The result of estimating the TBDCM is a list of attributes, as presented in Table 3 and Table 3. A sensitivity analysis was executed to determine the effect of these attributes. First, the elasticities of the travel times are determined. The travel times of one mode are increased by 1% to analyse the modal shift.

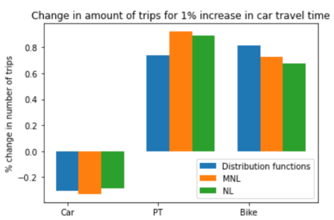

Figure 6: Elasticity of car travel time.

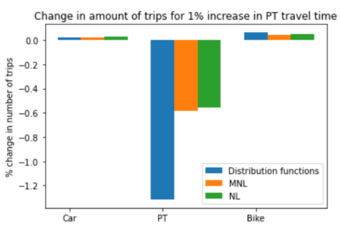

Figure 7: Elasticity of PT travel time.

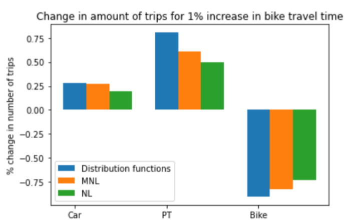

Figure 8: Elasticity of bike travel time.

Based on Figure 6, Figure 7 and Figure 8, pointed out that the TBDFM has a too high elasticity for the PT travel time. The elasticities for the car and the bike are similar for the TBDFM and the TBDCM.

Second, the effect of the attributes on utility of the alternative is analysed. The sensitivity analysis pointed out that the sensitivity of attributes is significantly different for each choice model. The TBDCMs predicts that for the car, the predominant attribute is the travel time. The travel time has the most influence on choosing the car or not. The TBDFM predict that the travel time is not as important and that the attraction of the zone has a bigger influence on the probability. For PT, the TBDCMs predict a small influence on the travel time and a major influence on the attraction of the destination. The TBDFM predict a larger influence of the travel time, but the attraction is also important in the decision. The sensitivity for the bike is very similar for the TBDCMs and the TBDFM.

An important aspect of the performance is the probability distribution among similar alternatives. The sensitivity analyses already pointed out a difference between the TBDCMs and the TBDFM. An analysis is performed with a choice set containing all the destinations as an alternative to analyse the impact of this difference in sensitivity. The trip length distribution of the car and the bike are similar. But, the TBDFM predict more trips on shorter travel times for the car while the TBDCMs predict more trips on shorter travel times for the bike. The trip length distribution for PT is significantly different. The TBDCMs predict a wider range of and longer trips. The TBDFM predict a shorter travel time for most of the trips. The MNL performs similarly to the NL, but there is a slight difference. The MNL performs slightly better; hence only the MNL will be used for the remainder of this thesis.

The TBDCM and the TBDFM are implemented in VML to compare their performance in a large-scale transport model. The MNL predicts a modal split of 70% car, 29% bike, and 1% PT, the TBDFM predict 61% car, 37% bike, and 3% PT, while the observed modal split is 62% car, 33% bike and 5% PT. It can be concluded that the TBDFM are more similar to the observed modal split than the TBDCM are to the observed modal split. The trip length distributions are similar for the car and the bike. Still, in both cases, the TBDFM predict more long distances trips, which better represents the observed trip length distribution. For PT, the TBDCM predicts more long-distance trips, which is more similar to the observed trip length distribution.

VML contains the ground truth for certain links, which is a different method to compare the performance of the distribution function and the TBDCM. To determine if the modelled flow, X_b, is comparable with the ground truth, X_w, a t-test is done, which is depicted in Equation 1.

$$T = ln\left(\frac{(X_b - X_w)^2}{X_w}\right) \tag{1}$$

The modelled flow is *good* if $T < 3.5$, *acceptable* if $3.5 < T < 4.5$ and *bad* if $T > 4.5$. The results of the ground truth observations are presented in Figure 9, Figure 10, Figure 11, Figure 12, Figure 13 and Figure 14.

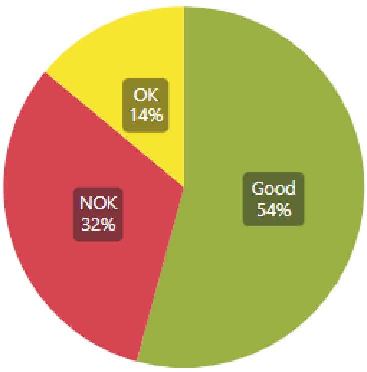

Figure 9: Car ground truth observations TBDCM

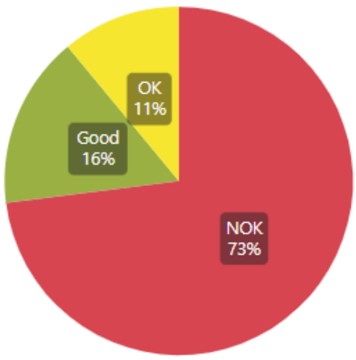

Figure 10: PT ground truth observations TBDCM

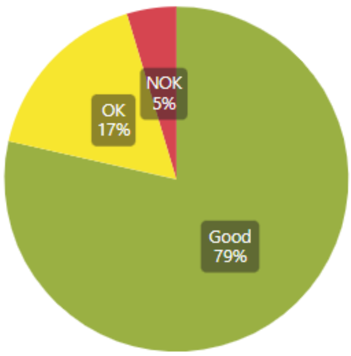

Figure 11: Bike ground truth observations TBDCM

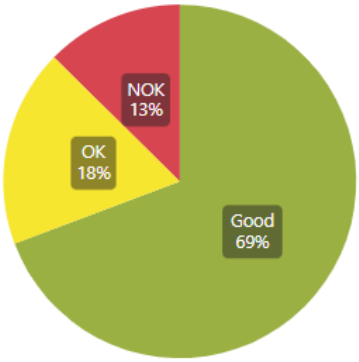

Figure 12: Car ground truth observations TBDFM

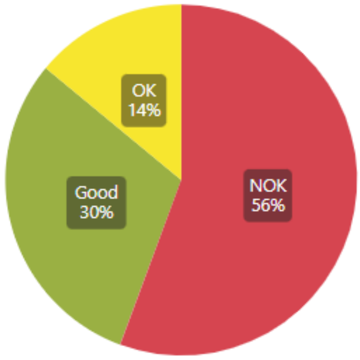

Figure 13: PT ground truth observations TBDFM

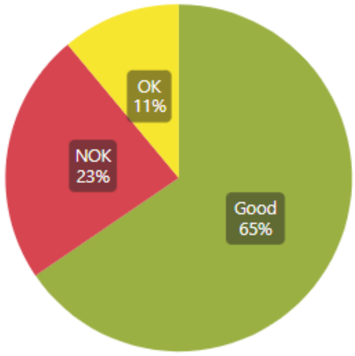

Figure 14: Bike ground truth observations TBDFM

Based on the Figure 9, Figure 10, Figure 11, Figure 12, Figure 13 and Figure 14, it can be concluded that the performance of the TBDFM and TBDCM differs per mode. The TBDFM performs better for the car and PT. Although the performance for PT is bad for both the TBDFM and the TBDCM. The TBDCM performs better for the bike.

Based on all the results, it can be concluded that the TBDCM performs better in predicting the correct chosen mode-destination alternative. Still, the TBDFM performs better in the case study. That the TBDFM perform better in the case study is not surprising. The TBDFM is specifically estimated and tailored for VML. The TBDCM is a generalised choice model and no modification have been made specifically for VML. Thus there is potential in tailoring the TBDCM specifically for VML. Tailoring the TBDCM will possibly increase the result on the ground truth observations.

This thesis has certain limitations. The main limitation is the input data. Because this thesis lacks a data source for the non-chosen alternatives, the input data for the TBDCM was limited to only Limburg observations. Hence, only the travel behaviour of individuals in Limburg is taken into account, and long-distance trips are absent in the choice set. As a result, the TBDCM is estimated on a specific range of individuals and trips. Also, the TBDCM is sequentially applied for every trip in the tour. A TBDCM that estimates the tour at once might perform better for longer tours. But, it does require huge amount of data preparation as each destination combination is one alternative.

Recommendations for future research are to research the cause of the high $\overline{\rho}^2$, elaborate the TBDCM structure, and analyse the potential of the TBDCM. The TBDCM yielded a very high $\overline{\rho}^2$. A possible explanation is the usage of VML for the travel times but this has to be confirmed by further research. The TBDCM was estimated using only observation from Limburg. VML is only one of the transport model RHDHV is developing. The applicability of the TBDCM must be researched to apply the TBDCM in a transport model with a different study area. The TBDCM estimates the trips sequentially to estimate the whole tour. The TBDCM could be improved if it estimates the whole tour simultaneously instead of multiple trips. Also, the potential of the TBDCM could be analysed. The TBDCM is estimated based on trips but VML models tours. Extra data preparation is needed which is not trivial. It should be researched if this extra data preparation is worth the extra detail. Could the TBDCM be better on the ground truth observations if the TBDCM is specifically modified to fit VML? It would create a more fair comparison between the TBDCM and the TBDFM.

Acronyms

List of Figures

LIST OF TABLES

Contents

1. Introduction

In the 1950s, highway construction started to accelerate worldwide. Originally, transport systems were assessed and analysed based on traffic counts. But with the expansion of traffic, in order to be able to assess policies and impacts, new, more sophisticated methods were needed to predict traffic, Kane and Behrens, 2002. Luckily, the development of computational technologies was also advancing. These developments and the first insights into transport models led to a 4-step transport model, which was developed in the 1960s, Hilty et al., 2001. Over the years, transport models have been continuously improved to overcome limitations in their usage. Dynamic methods were developed to improve the static nature of the models, while activity-based methods were developed to improve analysis of travel behaviour.

A (classic) transport model consists of 4 steps: trip generation, trip distribution, modal split and assignment. The trip generation step models the number of trips each zone produces and attracts. The trip distribution step matches all the origins with destinations to form trips. The third step is the modal split, where a mode is selected based on the trip. The fourth and final step is to assign the trips to the network to determine the load on the network.

The trip distribution and modal split steps are related because destination and mode can depend on each other due to the correlation between accessibility and utility. If a destination is more accessible because there are multiple alternatives, it is more likely to be chosen. Therefore, the trip distribution and mode choice are often modelled together, as separated choice models may produce misleading results, Ben-Akiva, 1973.

Provinces and municipalities use transport models to assess the implementation of transport policies or potential changes in the current transport network in order to determine their effect on congestion and the environment. To model these effects more accurately, RHDHV has developed a tour-based model (TBM). The main reasons to apply a tour-based approach are to be able to model trip chain behaviour, main activity versus secondary activity, heterogeneity in activity patterns and consistency in space, time and motives. Differentiating trips is not possible in a trip-based model because trips are seen as stand-alone trips and do not correlate with each other in any way.

In the TBM developed by RHDHV, each person has different characteristics, such as age, gender, income, car availability and household size, leading to differences in travel behaviour. But a difference of one year, or even a month, in age may not change travel behaviour significantly. Furthermore, it is computationally impossible to continuously determine travel behaviour for each combination of characteristics. Therefore, travel behaviour is aggregated. People are categorised based on personal characteristics that influence travel behaviour. The resulting groups are called *personas*, each having a different combination of characteristics. A distribution function is estimated for each persona to model the willingness to travel with each mode. Together, these distribution functions form the TBDFM. The data needed to estimate these functions is collected through a survey, which gathers the stated or revealed preference. This process requires a lot of data because there are many different characteristics, and thus, many personas. The CBS conducts a yearly survey to collect the revealed preference of the population of the Netherlands. Although this data set is quite large, with around 200,000 observations per year, there is still a shortage for some personas due to segmentation.

Within a TBM, there are different approaches to modelling the travel behaviour of each person. There are also approaches that can model each person's choice individually, using an activity-based model (ABM). In an ABM, each person has their own activity schedule and is modelled separately. This means that each individual will choose independently of other individuals with the same characteristics/persona. This method is computationally more challenging as each choice needs to be evaluated for each individual person. For a TBM, all the persons with the same persona are distributed among the choices according to the probability of each alternative.

A transport model aims to model the current and future traffic as accurately as possible. Because the transport model is used as a tool to determine the influence of policies, RHDHV aims to continuously improve its transport model to deliver a better-quality product. With an improved model, provinces and municipalities will be able to better substantiate their decision-making in transport policies on congestion and the environment. A potential improvement for the TBM is to replace the TBDFM with a DCM. In a DCM, personal characteristics are a part of the utility functions. The effect of each individual characteristic is estimated instead of the whole set of characteristics, resulting in a more detailed overview of the influence per characteristic and its significance. Furthermore, the data set is not disaggregated when estimating the model, which should increase the confidence interval of the result.

However, a DCM is not necessarily better than the TBDFM. A comparison must be made between the proposed DCMs and the TBDFM. Therefore, it needs to be determined what makes the choice model perform *better*. Multiple KPIs are defined to evaluate the models' performance, judge the models, and compare them.

1.1. Problem definition

RHDHV has developed a tour-based model with a TBDFM to model the mode & destination choice. However, the approach for the mode & destination choice leaves some room for improvement. As with every model, assumptions are made to focus on certain aspects or simply to make concessions. Yet, these assumptions can lead to an oversimplified reality. Additional methods can be added to address these issues.

Data is required to assess the travel behaviour of the personas representing individuals with a certain combination of characteristics. When there are more personas, the required data is increased as the data is split up into more sub-sets. When there are only a few observations for a persona, the accuracy of the estimated distribution functions of the TBDFM will be too low. Some personas can be combined to increase the number of observations, as they are very similar, but some cannot. Even though the estimated functions are smoothed to reduce the fluctuations and obtain a more realistic result, the TBDFM remain undesirable as the fluctuations remain significant which leads to inaccurate results. Another side of this problem is the limitation in the number of personas. The fewer the personas, the more aggregated the travel behaviour, which leads to less detail in the travel demand.

Moreover, the distribution functions of the TBDFM are estimated after segregating the data; hence the influence of each individual characteristic is not measured. Instead, only the total influence of all the characteristics combined is modelled. This means that the influence and significance of the individual characteristics on travel behaviour will remain unknown.

The TBDFM estimate the probability for each alternative for every persona. These results in OD matrices per mode. When analysing the OD matrix, the TBDFM leave room for improvement because some distribution functions of the TBDFM can have large fluctuations. As a result, trips with specific travel costs might have a unilateral modal split as one mode has a predominant preference. These irregularities trickle down when the trips are assigned to the network. As a result, the link's final network load does not pass the test compared with the traffic counts on the roads. A mismatch with the ground truth undermines the quality of the model, which is undesirable.

To sum up, the main shortcomings of using the TBDFM to model the mode & destination choice are:

- The number of observations per persona will decrease too quickly with more personas. Thus, there is a limit to the level of detail.
- By defining personas, the total effect of the combination of characteristics is modelled instead of the effect of the characteristics separately.
- For some personas the the number of observations is too low to estimate an accurate distribution function.

To overcome the limitations of the current approach, a new approach is needed to model the mode & destination choice. There are many potential approaches for doing this, but the main distinction is between aggregated and disaggregated models (Horni, 2013, Barff et al., 1982). An aggregated model averages the travel behaviour of subgroups of decision makers. By averaging their behaviour, heterogeneity among decision makers is lost. Different disaggregated choice models are described in Horni, 2013. However, the dominant method in transport and planning is DCM according to Horowitz, 1985.

In a DCM, personal characteristics are part of the utility function instead of a means to classify people. An advantage of a DCM is that the model is based on the entire data set and applies to every persona depending on the characteristics present in the model. The model contains every characteristic, which will be estimated with the whole data set. As a result, all choices can be modelled with the same model. Furthermore, in developing a DCM, the significance of certain personal traits or properties will become apparent. Thus, insignificant characteristics can be excluded, while with the TBDFM, the significance of each characteristic is not tested with the used data set.

There are, however, downsides to using a DCM. According to Hörl et al., 2019, DCMs fail to include the context of the trips. The infrastructural characteristics may not be the same for each choice. Thus, the parameters are estimated for a traveller's "average" trip. Furthermore, DCMs do not include dynamics of the travel demand. The parameters are based on a specific travel time, as there is no information on the congestion experienced. The expected travel time may not correspond with the actual travel time. However, it can be deduced based on the theoretical free flow travel time.

The main problem of using a DCM is estimating the destination choice. When estimating, alternatives are evaluated to determine the influence of the attributes. But if every destination alternative is considered, there will be many alternatives, which increases the estimation time. Although this process is executed only once, this is not desirable, and the added value of including all the alternatives instead of a well-considered selection is debatable. In short, the main problem is that it is unknown whether a DCM can overcome the limitations of the TBDFM. Thus arises the question: to what extent is a DCM able to better predict the choices for an entire population?

1.2. Scope

RHDHV wants to determine whether they should move forward with their current approach in the development of transport models or if they should adopt a DCM. In this thesis, both approaches will be compared to investigate the differences between them. To draw a fair comparison, boundaries must be set which are equal for both methods. The main reason is that both approaches need to have the same number of resources. Furthermore, the approach will need to be able to fulfil the needs of RHDHV. The approach must also be compatible with VML and the tour-based model. This creates extra limitations which will need to be dealt with.

The model designed in this thesis will be tested in VML. The implementation in VML poses various limitations. The main challenge is to fit DCM in the current tour-based model. To prevent changing the rest of the model, the DCM needs to work with the same input and produce the same output as the current method. The input is the personas and their tour frequencies, and the output is the destinations and modes of each tour. Because the persona represents a group of people, and not a single person, they must be distributed among the alternatives.

As mentioned earlier, CBS collects transportation data for Dutch citizens through the OViN. OViN is a national survey where respondents are asked to fill in their travel behaviour on a randomly selected day, including all the origins, destinations, modes, travel times, etc., for every trip they made. They are also asked to fill in their personal and household characteristics. To extrapolate the survey for the entire population of the Netherlands, the occurrence of personal characteristics in the survey is compared with the personal characteristics of the population as a whole. As a result, observations are weighed based on this relative occurrence.

A DCM needs data to estimate the model's coefficients. However, data is not available in large quantities. Thus, the model must provide significant results with limited data. OViN will be used for revealed preference. Data from several years can be stacked to increase the number of observations. However, due to developments in the transport networks and policy changes, historical differences within the data should be limited. Thus, a maximum of 3 years of OViN will be used. To have the same data set for the TBDFM and DCM, the same 3 years will be used for estimating the DCMs as for theTBDFM.

RHDHV aims to refine their transport models on a daily basis. This means that a computational limitation arises. For practical and efficiency reasons, a complete run of VML has to be finished in the morning when it starts running at the end of the previous day. In its current state, the model will finish well before the next day's start. This means that the available slack time within the total running time can be used to refine the model.

Finally, the model will also need to be future proof. The transport system and network are continuously developing. Governments are continuously (re)developing areas in their municipalities as there is a constant need for more space. This often leads to a redesign of the current transport network. Furthermore, transport policies can shift towards a different mode due to technological advancements. Topics such as Mobility as a Service (MaaS) and zero emission zones are examples of policies being implemented today. These policies affect both the mode choice and availability. Changes like these should be easy to implement in the TBDCM.

1.3. Research Question

This thesis aims to investigate the differences in modelling the mode & destination choice with TBDFM versus using a DCM. The outcome should determine which approach will be more suitable in future transport models at RHDHV. Thus, the main research question will be:

- How does the performance of modelling the OD matrix with a discrete choice model compare to modelling the OD matrix with TBDFM in the utility functions, in a joint mode & destination choice?

A DCM comes in various shapes and designs, based on different principles. The main aspect is how alternatives are evaluated and compared. Other factors are how the models deal with the correlation between alternatives or heterogeneity among decision makers. A selection of potential models will need to be made that theoretically fit the purpose of this study. Those models need to be designed and estimated. The models can be easily compared to each other on a theoretical level. Complex models will have a longer computational time but may provide more explanatory power. Based on the computational time and the $\overline{\rho^2}$, which indicates the model's fit and the estimated parameters. These aspects combined will influence the practical usage of the model. Also, an alternative model must be compatible with the tour-based model because it will be implemented in VML. The TBDCM should be able to handle the same input and generate the same output as the TBDFM.

The mode choice in this thesis is fixed as the case study contains only 3 alternatives: car, PT and bike. The destination choice is inherently different as all the zones are a possible destination and thus an alternative. In a mode & destination choice, the number of alternatives is $\#m * \#d$, which means that an increase in destination alternatives considerably increases the computational time. For tours, this problem increases exponentially as for every activity that is added to the tour, the alternatives grow #d times. Thus, not every destination should be an alternative, as this is also not the case in real life. Therefore, it should be investigated how to limit the destination alternatives without compromising the model's outcome. Hence, the next sub-question is: *How to construct a discrete choice model for a tour?*

To estimate a DCM data is needed for the chosen and non-chosen alternatives. As mentioned before, OViN will be used for the chosen alternatives. A second data source is needed to complement OViN with the non-chosen alternatives. This source needs to be compatible with OViN to prevent bias in the choice set. Hence data sources need to be identified and analysed to determine the compatibility with OViN. Thus, the next sub-question is: *Which data sources are available to compose the choice set?*

After collection the required data the estimate a DCM, a TBDCM can be constructed and estimated. The results of estimating the TBDCM is a list of attributes and its value. These values can also indicate if the choice structure is valid. Hence, these values are analysed to get a better comprehension of the estimated attributes which leads to: *What are the estimated parameters of the tour-based discrete choice model and its effect?*

The performance of the DCM needs to be evaluated to determine if it exceeds the performance of the TBDFM. There are multiple methods to compare the performances of the models. These methods need to be identified and executed. The VML is a great opportunity to test how the models perform in a complete transport model.

Furthermore, it can be compared to the ground truth data with the means of OViN. Therefore, the final sub-question is: *How is the performance of a choice model defined?*

To sum up, the structure of the thesis is as follows:

- How does the performance of modelling the OD matrix with a discrete choice model compare to modelling the OD matrix with the TBDFM, in a joint mode & destination choice?
  - How to construct a discrete choice model for a tour?
    - Which type of models are used in practice?
    - How relates the mode choice to the destination choice?
    - How are the activities prioritised?
    - How to sample destinations sampled?
  - What data sources are available to compose the choice set?
    - Which characteristics are present in OViN?
    - Which characteristics are present in VML?
    - What are available sources for the non-chosen alternatives?
  - How is the performance of a choice model defined?
    - How are discrete choice models validated?
    - How do the choice models perform on predicting the chosen alternatives?
    - How do the choice models perform on the ground truth observations?
    - Which model is the most suitable for RHDHV?

1.4. Research Contribution

This thesis contributes not only to the academic research but also to society. First, the scientific contributions are described. Second, the societal contributions are explained.

1.4.1. Scientific contribution

This thesis aims to replace the TBDFM with a DCM, in a TBM. To the author's knowledge, there is a significant amount of theoretical research on the methods for choice modelling in transport models. The different methods to model the mode & destination choice are widely researched.

However, there is a gap in comparing a distribution functions choice model with a DCM in the context of a tour-based transport model. Mishra et al., 2013 and Molloy, 2016 compared a gravity model, with distribution functions, with a DCM in a transport model but only for the destination choice. The mode choice and destination choice are often modelled simultaneously, as there is a relation between the two choices, which should not be overlooked.

Furthermore, there is a lack of knowledge on how the results of transport models are best compared. There are theoretical indicators of the choice model's fit, but the effects on the results and the performance in a complete transport model is often overlooked.

Thus, this thesis contributes to the comparison between the TBDFM and the TBDCM, in the context of a joint mode & destination choice within a tour-based transport model. The focus lies on comparing the results of the TBM, hence comparing the performance of the TBDCM and the TBDFM in approximating the ground-truth.

1.4.2. Societal contribution

This thesis also aims at a societal level to contribute to developing transport models within RHDHV and to quantifiably improve their tour-based models and other future models in terms of explanatory power, computational time and robustness. As a result, RHDHV will be able to model traffic more accurately and more reliably. The outcomes of this thesis will help provinces and municipalities in their challenge to transition their transport network to a more sustainable model. Thus, it will indirectly contribute positively to society and the environment.

1.4.3. Report Structure

In Chapter 2, a literature study is performed. The literature study is conducted to gain insight into discrete choice modelling and to answer the questions: *what are the differences between the discrete choice models?*, *how to model a complete tour?* and *how is performance defined?* An overview of studies that have developed a DCM for a tour-based model will be given. The differences and similarities will be pointed out, and the pros and cons will be determined. Also, the methods that can measure an indication of performance are identified.

In Chapter 3, the methodology will be laid out to determine how to answer the main research question. The steps that lead to answering the main research question are described. The required input and output of each step are determined. A case study is required to execute the methodology; hence in Chapter 4, VML is explained, and its characteristics are analysed. But the limitations that come hand-in-hand with VML are also identified. In Chapter 5, the data from OViN is combined with the data from VML to compose a choice set which will be used to estimate the DCM.

In Chapter 6, the different DCMs are estimated. The DCMs, and the TBDFM are analysed. Also, the performance of the choice models is determined. At last, the choice models are implemented in VML to determine the choice models' performance in a large-scale transport model. The results are discussed in Chapter 7. The main research question is answered, and areas of this thesis that can be elaborated are identified for future research.

2. Literature Review

This chapter aims to analyse previous studies on transport modelling and to summarise the methods used. Based on the findings in these studies, the knowledge gaps identified in the previous chapter can be filled. Literature describes many different methods for constructing a DCM and validating DCM models. Therefore, it is key to create a solid overview from which the ideal method can be selected for this study. First, the approaches of DCMs are identified and assessed to determine the best approach for this study. Then, the types of models that fall within the selected approach are assessed, followed by an analysis of which DCMs are used for tour-based models. Models are categorised, and key differences are highlighted and discussed. To investigate if a DCM performs better than the deterrence functions, a DCM needs to be estimated. A practical problem that arises for estimating a destination choice is the sampling of non-chosen destinations. Therefore, methods to sample destinations are discussed. The characteristics that influence the mode & destination choice are collected, and an overview is given. Finally, methods to analyse and validate the DCM are presented. Also, the KPIs for analysing the performance are identified.

2.1. Discrete Choice Model Types

The different model types must be identified to determine which DCM type is suitable for further elaboration. The most dominant type of DCM, according to Hess et al., 2018, is the RUM model. It is based on the principle that people choose based on certain attributes of an alternative and a random factor that accounts for taste heterogeneity. Thus, it combines the (dis)utility of attributes with a random parameter to determine the total utility of an alternative.

The utility of each alternative is calculated based on these attributes and the random parameter. The choice is calculated based on the difference in utility of the alternatives. The first variant of the RUM model was developed by McFadden et al., 1973. By estimating a RUM model, the value of each attribute is given to replicate individuals' choices. Over the years, the model has been further developed, and new RUM models were presented, increasing the complexity and fit of the model.

Although according to Hess et al., 2018, RUM models have been dominant in discrete choice modelling, the main assumption of RUM has been questioned. Some behavioural aspects that have been studied and argued relevant for transport behaviour are not included in the method. With that in mind, it is important to analyse whether a RUM model should be used or, indeed, a different type of discrete choice model should be applied.

One point of criticism is based on Loomes and Sugden, 1982, which states that the utility of an alternative is not only based on the alternative itself but also on the regret of the non-chosen alternatives. Consequently, Chorus et al., 2008 developed a new kind of discrete choice model: random regret minimisation (RRM), as an alternative for RUM. An RRM does not maximise the positive pay-off of the alternatives but minimises the avoidance of negative aspects. It considers the regret that a different alternative would turn out to be the better choice in hindsight. It is argued that an RRM is useful in determining travel demand as it models the negative emotions of missing a bus or getting stuck in traffic. The attributes of the alternative are compared with each other to determine the regret of choosing alternative i over alternative j. As a result, the number of function evaluations is higher. For a small number of alternatives, this increase is negligible, but when there are around 1,600 zones and 3 modes, this results in $4,800*4,799 \approx 23 million$ combinations (van Cranenburgh and Chorus, 2018), which is significantly more than 4,800 evaluations for the RUM.

Another point of critique is the decision-making. A RUM evaluates the probability of every available alternative based on the utility of attributes. Sælensminde, 2006 argued that people choose differently. Not only do they choose based on the most important attributes, a subset of the alternatives, but also, they do not consider all the available alternatives. Some people will always choose the car, no matter what. Therefore, Hess et al., 2018 argued that heuristics could be applied to simplify an individual's decision-making. Arentze and Timmermans, 2004 developed a heuristic approach to model the destination choice. A destination is chosen based on a set of rules regarding destination characteristics, such as travel time and attractiveness. It concluded that the decision tree was able to achieve considerable predictive accuracy.

However, Hess et al., 2018 also pointed out the main benefits of RUM: consistency, flexibility and agility. Within RUM, multiple options exist to increase the complexity of the model. The availability of these options creates a certain flexibility because how people choose between their alternatives is unknown. Furthermore, the RUM is computationally less challenging, which is an important factor for this study. Therefore, for this study, RUM will be selected and explored to determine which specific model should be used.

2.2. Existing Discrete Choice Models In Tour-Based Models

In this section, DCMs that have been developed for tour-based models are presented. Different approaches and assumptions are categorised to outline the differences between the models.

Hasnine and Nurul Habib, 2021 performed a literature overview on the different approaches to model the mode choice in different activity-based modelling frameworks. They identified 7 different categories:

- Simplified main tour mode. To limit the number of possible mode combinations within the tour, an individual is assumed to not change mode during the tour.
- The two-tier nested logit model. This states that the destination choice is on a higher level than the mode choice. The level of service of a destination alternative is influenced by the travel times of the different modes. That factor influences the destination choice, after which the corresponding mode is chosen. The mode & destination choice is evaluated for every trip in the tour and is separately conditioned with the previous trips.
- Simplified main tour mode and conditional trip-level mode. First, the main tour mode is modelled based on the log sums of the alternatives. Next, the destinations and mode choices on the trip level are estimated based on the main tour mode, detours and previous trip mode.
- An activity-based model with exogenous mode choice. A dynamic activity-based model jointly models the activity type, location and duration, based on a 24-hour schedule and activity constraints. Although the mode choice is modelled exogenously, various mode-specific attributes are incorporated into the destination choice.
- Simulation-based tour-based mode choice. The Travel/Activity Scheduler for Household Agents consists of sub-models. It combines the separate trip utilities to obtain the utility of the complete tour. It consists of multiple sub-models in a sequence where the location and mode choices are separate.
- Combinatorial tour-based mode choice. To decrease the number of alternatives, (feasible) combinations of modes and destinations within the tour are specified to simplify the model.
- Dynamic tour-based mode choice. A dynamic discrete choice model takes future expectations and state dependencies into account. By setting a time frame of 24 hours, it can evaluate all the possible combinations backwards. It also updates the choice set during the tour. Because it is computationally challenging, the tours cannot be complex.

From these categories, it can be concluded that there are many different approaches to applying a DCM in a TBM. Multiple characteristics can be identified from the categories above, which are always present in the DCM. Namely: temporal resolution, tour utility, tour mode, choice structure and trip hierarchy.

The temporal resolution determines the activities based on the activity schedule and the activity duration. A limitation of this model, a dynamic discrete choice model (DDCM), is that the number of trips in the tour and the activity schedule are required beforehand. For the current study, the tour itself is an input which contains the number of trips and its purposes, but there is no time frame. As a result, a DDCM may be a strong option, but it would only be suitable for more complex tours. In this thesis, the tours are too simplified to apply a DDCM; hence, it is unsuitable for this case.

Two methods can be used to calculate the tour utility. A DCM can calculate the utility based on the entire tour, or it can calculate the utility for each trip in the tour separately and sum up the results to determine the utility of the tour. An advantage of the former method is that tour characteristics are captured. However, for each tour combination of activities, a different model must be estimated. The advantage of the latter method is that it is very flexible as it can be used for every combination of activities. Because there are many different tour combinations that will serve as input for the TBDCM, the utility of the tour will be calculated by summing up the trip utilities. The drawback is that the attributes are estimated on trips and not tours. The attributes characterises the behaviour on trip while it will be used for tours. If the travel behaviour is different for tours than for trips, it will not be taken into account.

The mode for the tour can either be determined for the entire tour or for each trip. The distinction is between single-mode tours versus multi-mode tours. Some studies assume that people do not change their modes during a tour, while others assume that they do. There is, however, a small percentage that does change mode, which differs per study, while the vast majority do not change their mode. This requires a more in-depth comparison which will be discussed further on in this thesis.

The choice between mode and destination choice can be structured in various ways. The two choices can be seen as equal; both are made simultaneously, while one choice can be seen as "more important". If so, which is the more important choice? There can also be a correlation between destination choices and mode choices. To determine the choice structure, studies will be categorised to determine the most suitable method.

Because a tour will consist of multiple activities, it is possible to implement a hierarchy among activities. It can be assumed that one activity is the most important one, on which people determine their mode choice. In most cases, this trip's purpose is mostly work or education because this is seen as more important than shopping, for instance. The reason is that the activity duration is longer and more often fixed in a person's time schedule. To determine if a trip hierarchy is desired, studies will be analysed.

To get an overview of the distinctions outlined above, a summary is presented in Table 2.1. For each aspect, the options that are available are given, and every study which developed a relevant DCM is analysed and placed in the corresponding box.

Table 2.1: Overview of studies on RUM models.

No trip hierarchy	**Choice structure**	**Single mode**	**Multi modal**
	Nested logit	Newman and Bernardin, 2010	Kitamura, Fujii, et al., 1998
	MNL	Eluru et al., 2009	
Trip hierarchy	**Choice structure**	**Single mode**	**Multi modal**
	Nested logit	Yagi and Mohammadian, 2010	Bradley et al., 2010 Shiftan et al., 2003
	MNL	Bowman and Ben-Akiva, 2001	

2.2.1. Single mode vs multi modal

Eluru et al., 2009 defined purpose, time-of-day and mode combinations to analyse the tours, thus limiting the tour to a single mode. Yagi and Mohammadian, 2010 modelled a mode choice based on the primary activity. For the secondary trip, the mode is fixed; hence a multi-modal tour is not possible. Bowman and Ben-Akiva, 2001 also modelled the mode for the tour instead of the trip. They acknowledged that the mode alternatives could be enhanced with more complex mixes of modes. Newman and Bernardin, 2010 also focused only on the primary mode of the tour.

Kitamura, Fujii, et al., 1998 modelled the mode and destination for every trip conditioned with the previous trip and the next fixed activity. Bradley et al., 2010modelled the tour step by step. Their approach first models a main tour mode, and after the locations of the intermediate stop have been modelled, it re-evaluates the trip mode based on the specific OD pair of the trip and a previously modelled trip to incorporate the possibility of a multi-modal tour. Shiftan et al., 2003 also modelled the primary and secondary modes and destinations in steps. Their approach models the destinations first and then the corresponding mode.

To conclude, the main reason for single-mode tours given in the studies that used this method is that only a small percentage of the tours are multi-modal. To focus their research, previous researchers accepted the lower level of detail in order to simplify the model. The reason for this is that evaluating a new mode for every new destination requires a lot more computational time as the number of alternatives grows significantly. A decision will need to be made in the methodology as to whether the added value of multi-model tours is worth the extra computational time.

2.2.2. Trip hierarchy vs no trip hierarchy

Bowman and Ben-Akiva, 2001 assumed that performing a tour has a motive. The activity that corresponds with this motivation is the primary activity. The model uses the primary activity as a starting point in modelling the destination and mode of the tour. The modes and destinations of the other activities are determined, based on the mode and destination of the main activity. Bradley et al., 2010 stated a difference between trip purposes. School and work trips usually go to a determined destination; hence they can be seen as the primary activity. The secondary destinations are conditioned by the main tour mode. Yagi and Mohammadian, 2010 defined a primary activity for each tour based on different trip purposes. The model is also estimated for each of those main trip purposes. Shiftan et al., 2003 modelled a primary activity first but with all the different purposes incorporated in the same model.

When there is no hierarchy among trip purposes, the sum of each trip is used to determine the total utility of the tour. With this approach, the destinations and modes are simultaneously chosen, which will incorporate no trip hierarchy among the activities.

Whether trip hierarchy will be implemented in the TBDCM will be determined further on in this thesis.

2.2.3. Choice structure

When combining the choice of the mode and the destination, an assumption is needed. Bowman and Ben-Akiva, 2001, Hasnine and Habib, 2019 and Västberg et al., 2020 assumed that there is no correlation between combined mode & destination choices. Each alternative is seen as independent from all the other alternatives. With this assumption, a MNL can be estimated as a critical assumption that the alternatives have no correlation.

However, one can argue that there is, in fact, a correlation between combined mode & destination choices. Alternatives that correlate can be merged with a nest. There are multiple ways to structure the nests. The main distinction lies in which is seen as the dominant choice: the mode choice or the destination choice. This relates to nesting all alternatives that have the same mode or that have the same destination.

Newman and Bernardin, 2010, Bradley et al., 2010, Kitamura, Fujii, et al., 1998, Shiftan et al., 2003 all used a "classic" hierarchy with the destination choice as the upper level and the mode choice as the lower level. They all stated that this is the most common approach as people are more likely to change mode than to change destination. A log sum variable of the accessibility of each destination is used in the upper level to implement the nests. Yagi and Mohammadian, 2010 also followed this line of reasoning but combined the destination nests with mode nests. The upper level is still the destination choice, but it also considered the nests of the mode choice by combining the destination with those nests as an alternative.

However, Newman and Bernardin, 2010 used a "reversed hierarchy" due to geographical and infrastructural differences between Europe and the United States. European cities are denser and their infrastructure is not focused on one specific mode. American cities are relatively more spread out and focused on cars. Because the case in their study is based on an American city, they argued that in this car-dominated community, people will more likely change their destination than their mode. Furthermore, Newman and Bernardin, 2010 argued that for some trip purposes, such as school and work, the hierarchy is reversed as people who have access to a car will almost certainly choose the car and people who do not have a car will choose CNL.

Based on literature, it appears that nesting is a useful tool to cope with the complicated choice between an NL mode and an NL destination. Thus, an NL will be created to model the choice of a combined mode &

destination choice. However, this requires a priority among the two choices. It is noticed that in European cities, the destination choice is more often the higher level, while in American cities, the mode choice is. As the TBDCM in this thesis will be applied in the Netherlands, it is logical to have the destination choice as the upper level. Nonetheless, both nest structures will be estimated.

The nesting within the model can be expanded. Instead of destination nests or mode nests on the higher level, it is possible to have both nests on the upper level. As a result, each alternative has two nest parameters instead of one. This is called a CNL. Ding et al., 2014investigated the different logit models: MNL, NL and CNL, in a joint shopping mode & destination choice. They found that the CNL has a higher predictive power than the NL. Therefore, a CNL will also be developed to have a better understanding of the choice structure and hierarchy of a mode & destination choice.

2.3. Correlation among alternatives

As mentioned in the previous section, different methods exist to implement a choice structure. The main reason to implement a choice structure is that alternatives can correlate with each other. Correlation means that an individual does not view them as separate alternatives but as one alternative because the characteristics of the alternatives are very familiar. In this section, the different choice structures are described in more detail.

2.3.1. Multinomial Logit

When an MNL structure is assumed, all alternatives are independent of each other. Thus, the choice structure will be as depicted in Figure 2.1.

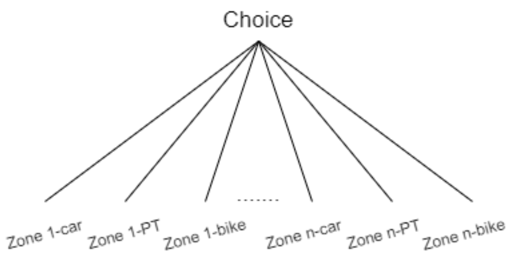

Figure 2.1: MNL choice structure

As the figure shows, each mode & destination combination is independent on all the other alternatives. Attributes of alternatives can still be present in other alternatives to incorporate a form of priority between the destination and mode choice. Newman and Bernardin, 2010, used two DCMs of which the log sum of characteristics of the mode choice in the destination alternatives. The destination will be chosen first, after which the mode is chosen, as seen in Figure 2.2.

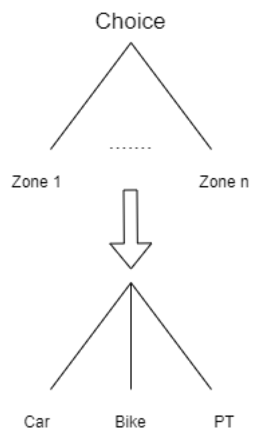

Figure 2.2: Double DCM with a logsum.

However, it can be argued that alternatives do correlate. Mode & destination combinations that have the same destination or mode have similar characteristics. Not only in terms of attributes but also for individuals preferring a certain destination or mode, there can be a correlation between alternatives that lead to that specific destination or destinations that are accessible by a specific mode.

2.3.2. Nested Logit

To quantify the correlation between alternatives with the same mode or destination, a nested choice structure can be implemented. An NL will indicate if alternatives correlate with each other. However, depending on the researcher's goal, the nested choice structure can differ. As mentioned before, there are two different structures for the NL. The destination choice is often seen as "more important" and used to determine the correlation between alternatives to that destination, as shown in Figure 2.3. However, it is possible that all the alternatives with the same mode correlate, which will lead to a structure as depicted in Figure 2.4

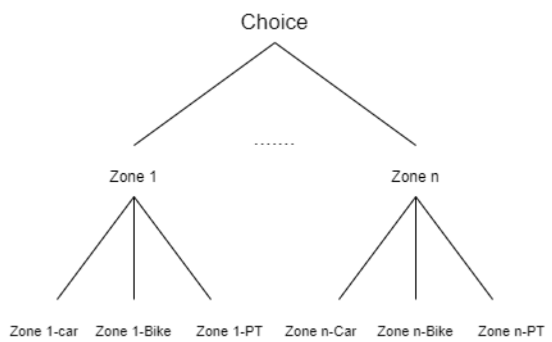

Figure 2.3: Destination choice nested.

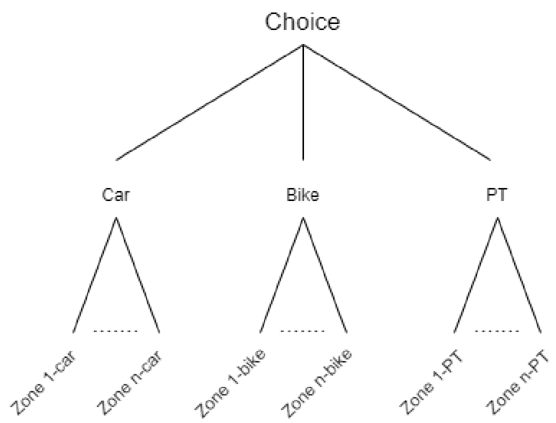

Figure 2.4: Mode choice nested.

2.3.3. Cross Nested Logit

As mentioned before, the CNL is a more sophisticated version of the NL. This means that the same logic holds for both the CNL and the NL. There are, however, small adjustments needed. The main adjustment is that now both mode and destination nests are present. Each alternative now belongs to both nests. The model will estimate a distribution among the nests for the alternatives. The distribution indicates which nest is more dominant. The choice structure will be as follows:

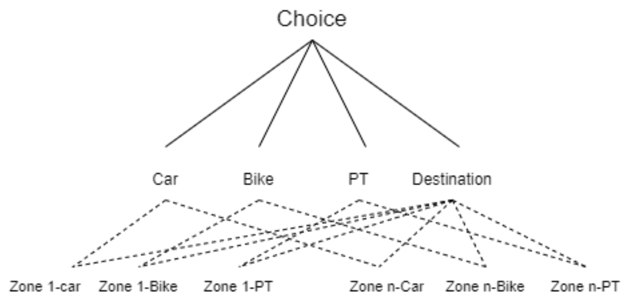

Figure 2.5: CNL choice structure

As a result, two separate functions can calculate the utility of each alternative. Each alternative has an α parameter for each nest to determine the distribution among the two nests. This parameter determines the ratio of utilisation of each utility function. The utility functions of the mode nests have a dedicated α, and the utility functions of the destination nests have a dedicated α. The sum of the α-destination and α-mode equals one. The α parameters also indicate which nest is more dominant.

2.4. Destination Sampling

When estimating a DCM with a destination choice, a practical issue arises in the determination of the choice set, i.e., the chosen and non-chosen alternatives. The mode choice is straightforward as there are a limited number of modes. The destination choice is more complex, because not every destination zone in the model was considered as an alternative. While it is still possible to assume every zone was considered as an alternative, the number of destinations in the choice set will significantly affect the number of calculations, which will increase the computational time. More destinations will increase the computational time when applying the model and estimating the model; evaluating all the non-chosen alternatives also takes time. A solution to limit the computational time is to reduce the number of destinations by sampling the destination alternatives.

McFadden, 1977 proved that it is possible to limit the number of alternatives without compromising the consistency of the estimators. He also provided multiple strategies to construct a subset of the full choice set and its corresponding probability function. The condition is that the chosen alternative is part of the subset, which is equal to or a part of the full choice set.

The sampling of destinations is not necessary, as Shiftan and Ben-Akiva, 2011 applied no sampling and used all 1,244 zones as potential destinations when they discussed this subject. However, they concluded that sampling the choice set of destinations is a reasonable method to obtain a reasonable complexity and running time.

An obvious strategy to sample the destinations is by randomly selecting zones. Jonnalagadda et al., 2001 and Shiftan, 1998 used random sampling to reduce the number of alternatives. The Florida model used a subset of 40 zones, while the Idaho model used a similar number. Furthermore, Eluru et al., 2009 randomly sampled 30 locations from the full set of 1,099 zones in San Francisco to compose a subset. Although random sampling of destinations is a straightforward method, it is proven to be a reliable solution to limit the number of alternatives.

Curtis-Ham et al., 2021, in a study on estimating a DCM on crime location choice, compared nine different sampling strategies, which can be categorised into three categories. One category is distance importance sampling (DIS), which samples the destination location based on the distance. The second category is simple random sampling (SRS) which uses sampling based on randomly drawn destinations. The third is strategic importance sampling (SIS) which combines the first two categories by sampling the nearest zone plus randomly sampled zones.

In an empirical study to investigate the average choice set based on the location of a crime, they used four different distance boundaries to create different sample sizes. The first strategy, DIS1, uses all the zones within 5 km of the activity plus 10 randomly selected zone of any distance. DIS2 is DIS1 with 20 randomly selected zone between 5-10 km. DIS3 is DIS2 plus 15 zones between 10-50 km and DIS4 is DIS3 plus 10 zones from 50-100 km. Due to the difference in zone sizes, the choice set sample size can differ greatly.

For simple importance sampling, they used all the zones within 5 km of the activity as a basis. SIS1 is the basis, with 30 randomly sampled zones, SIS2 is with 55 zones and SIS3 with 100 zones.

For simple random sampling, they used only randomly sampled zones. But because the sample size differs significantly between activities, SRS1 sampled as few as the smallest sample size of the previous strategies, and SRS2 sampled as many as the largest sample size. This strategy aims to make a fair comparison between the strategies and random sampling.

They also researched the sample size in comparison to the full choice set and found different studies with diverse sample sizes ranging from 1% to 12.5%. In their study, the strategies lead to a range of the average sample size from 5 to 15%.

When estimating the parameters of their model, it appeared that increasing the distance, DIS2, DIS3 and DIS4 did not lead to significantly different parameters and standard errors with respect to the full model. Nor did increasing the sample size with SIS1, SIS2 and SIS3. The random sampling, SRS1 and SRS2, did lead to a deviation in the coefficients of the model, but for SRS2, the difference is marginal. From this research, it can be concluded that for crime locations, random sampling will lead to a small difference in the coefficients when estimating the model.

Yagi and Mohammadian, 2010 used a different strategy, which combined the distance aspect and characteristics of the zones. They created categories based on the relative distance and relative attractiveness compared to all the other zones. Thus, the distance to each zone and its attractiveness were collected and categorised. The distance was categorised in 0-20th , 60th-20th and 60th-100th percentile while the attractiveness was categorised in 0th-50th and 50th-100th percentile. The categorisation creates a total of six bins. From each bin, the same number of zones are sampled to create the choice set. This approach may result in both an over- and an underrepresentation by zones. If there are only a few zones in specific categories, they will appear in almost every choice set, which may not have been the case if the zones were sampled randomly.

2.5. Utility Attributes

As mentioned earlier, the utility functions consist of attributes and a random parameter. What these attributes are, is unknown. There are many personal characteristics that can be implemented in the DCM, but they may not produce a statistically significant and usable result. To get an overview of potential attributes, previous studies have been reviewed. From these previous studies, the (significant) attributes in the utility functions are categorised and summarised in Table 2.2. In this thesis, three categories are formed: LoS, personal and zonal attributes.

Table 2.2: Used attributes used in literature.

Attibutes	Bowman and Ben-Akiva, 2001	Newman and Bernardin, 2010	Eluru et al., 2009	Kitamura, Fujii, et al., 1998
LOS				
Travel time	✓	✓	✓	✓
In-/out-vehicle time	✓			
$Distance^2$	✓			
Distance	✓			
Waiting time				
Socioeconomic				
Gender			✓	✓
Age	Cat.		Cat.	Cat.
Drivers license				✓
Vehicle availability	✓	✓	✓	✓
HH income	✓	✓		✓
Destination characteristics &				
Area			Log	
Land-use			%	
Population				✓
Employment	Cat.	Total	Log	Total
Intrazonal		✓		✓
Attibutes	Västberg et al., 2020	Yagi and Mohammadian, 2010	Bradley et al., 2010	Shiftan et al., 2003
LOS				
Travel time	✓	✓	✓	✓
In-/out-vehicle time			✓	✓
$Distance^2$				
Distance				
Waiting time	✓	✓		✓
Socioeconomic				
Gender		✓		
Age			Cat.	Cat.
Drivers license				
Vehicle availability			✓	✓
House-hold income		✓	✓	
Destination characteristics &				
Area		Log		
Land-use		%/dummy		Dummy
Population				✓
Employment	Total	Density	Density	
Intrazonal	✓			

There are a few things to notice in this table. First, literature is not unanimous on the attributes. Travel time and employment are present in some form in all papers apart from one. For the other attributes, the occurrence varies; some studies include the variable while others do not.

Second, there is a difference in categories but also in usage. Some attributes are used as continuous variables, while others are used categorically. There are even differences within a paper.

This literature review shows that travel time and employment play an important role as all the studies have used these attributes. But for the other attributes, there is no consensus. This means that all these variables have an expected influence on the mode & destination choice.

2.6. Performance Measurement

When a DCM is estimated with statistically significant attributes, the model needs be analysed. The performance needs to be quantified to be able to determine the differences between variations of the TBDCM and the deterrence functions. The performance analysis consists of two parts: validation and predictive power.

2.6.1. Discrete choice model validation

When estimating a DCM, the most common method to determine whether one model performs better than the other is to compare the $\overline{\rho}^2$. To determine this value, ρ^2 needs to be determined first. It is calculated by comparing the likelihood of the null model $\mathscr{L}(0)$, the model where all the attributes are set to zero, and the likelihood of the estimated model $\mathscr{L}(\beta)$. This quantifies the effect of the attributes on the probability of the correct choices.

$$\rho^2 = \frac{\mathscr{L}(0) - \mathscr{L}(\beta)}{\mathscr{L}(0)} \tag{2.1}$$

After determining ρ^2, $\overline{\rho}^2$ can be calculated. The "score" needs to be adjusted based on the sample size N and the number of independent attributes p. With this correction, it is possible to compare models with a different sample size and, more importantly, a different number of independent attributes because it is easier for a model to score a higher ρ^2 with more attributes.

$$\overline{\rho}^2 = 1 - \frac{(1-\rho^2)(N-1)}{N-p-1} \tag{2.2}$$

However, Parady et al., 2021 warned for over-fitting.This may occur when estimating the model if there is too much focus on the fit of the estimation data. The model may have a relatively high $\overline{\rho}^2$, but that does not mean it performs the same on an independent data set. They warned that a lack of validation increases the risk that models may perform "well" on estimation data but poorly on similar data sets which were not used to estimate the model. They summarised different concepts on validation based on Justice et al., 1999.

There are two main methods to validate a DCM model: in-sample and out-of-sample testing. For in-sample testing, the consistency of the model is determined. For this method, the data that is used for estimation is also used for testing the model. By splitting the data set, multiple smaller data sets are created. The model is estimated on each of these sets. The variation of the $\overline{\rho}^2$ of each set will indicate the consistency of the model

For out-of-sample testing, a data set that is not used for estimating the model is used to validate the model. It requires a data set which is independent of the data set which was used for estimation. This can be achieved by collecting a different data set from a different country, time period or source. This approach measures the transferability to indicate whether the model can also be used for different cases.

2.6.2. OD-MATRIX VALIDATION

A different method to validate a DCM is to generate a complete OD matrix. Vajjarapu et al., 2020 validated the OD matrix based on three aspects: link flows, trip length distribution and modal split. To validate the link flows, observed data were collected from several locations. Based on the observed data, it was then determined whether the modelled link flows were within acceptable limits. The same holds for the trip length distribution and the modal split.

REPRODUCIBILITY

Different methods exist to split the data for validation purposes, as Parady et al., 2021 showed. The simplest method is the holdout validation (HOV). For every observation, it is randomly determined whether it will be used for the estimation data set or the validation data set. After estimation, the model can be validated using the validation set. For each observation, it is determined if the observed alternative is also the predicted alternative. The method counts how many correct predictions the model predicts, which will lead to a certain score. This will produce an unbiased holdout data set.

$$Q[y_n, \overline{y}_n] = \begin{cases} 1 & \text{if } y_n = \overline{y}_n \\ 0 & \text{if } y_n \neq \overline{y}_n \end{cases} \tag{2.3}$$

Where Q represents the estimated model and y_n the observed outcome & $\overline{y}_n$ the predicted outcome for observation n. Thus the HOV for data set N holds:

$$\text{HOV} = \frac{1}{N} \sum_{n=1}^{N} Q[y_n, \overline{y}_n] \tag{2.4}$$

This approach can be executed numerous times to minimise the influence of the randomness, known as the cross-validation (CV). To apply the CV, the HOV will be executed I times. The results of each HOV are recorded and used to determine the result of the CV. The result will indicate how consistently the model performs based on different subsets of the data.

$$\text{CV} = \frac{1}{I} \sum_{i=1}^{I} HOV_i \tag{2.5}$$

Bootstrapping methods were developed to tackle some limitations of CV methods. The main limitation of CV is that it can produce high variances between instances, especially for small sample sizes. However, the sample size for the current case is quite big; thus, a more extensive approach is not needed.

TRANSFERABILITY

To determine transferability, a different data set is needed. This means that an aspect of the data set needs to be different. This can be achieved by collecting the data using a different method, from a different population or from a different time period. As OViN is executed every year, a previous year can be used to determine the temporal transferability.

The models try to predict the correct outcome based on the data set from a different year. The results will indicate which model performs better on an independent data set and therefore has a higher transferability, Parady et al., 2021

2.6.3. PREDICTING THE CORRECT ALTERNATIVE

When applying these methods, one issue is to determine whether a DCM correctly predicts the choice. For example, suppose there are three alternatives, A, B and C, and the model predicts a [34%, 33%, 33%] probability distribution. In that case, it may present the chosen alternative A as the highest probability. However, a model that predicts a [80%, 10%, 10%] produces a significantly better prediction. de Luca and Cantarella, 2016 discussed this problem and concluded that a threshold value should be set with a minimum of 50%. For this study, however, there are many more alternatives than three, which automatically leads to a higher probability distribution. A decision on the method to be used will be made further on in this thesis.

2.7. Conclusion

To construct a DCM for a TBM, numerous decisions must be made. There are a few choices which relate to the tour itself. The first choice is whether there will be a trip hierarchy. This is because the mode & destination choice needs to be determined based on the main activity of the tour or on the whole tour. Second, it needs to be decided whether the tour will be uni-modal or multi-modal. Can the individuals change their mode during the tour, and if so, will it be conditioned that the vehicles (car and bike) need to return home when used for the first trip leg? Lastly, there is the choice between the mode and destination nesting. Is one choice superior to the other?

Moreover, the computational time for estimating a DCM can be very high if all destinations are viewed as an alternative. Therefore, destinations need to be sampled to be able to estimate a model within an acceptable time frame. There are numerous ways to sample the destinations. After estimating the model, it must be validated, and the performance must be measured to substantiate which model performs better. Also, the criteria for when the model correctly predicts a choice need to be defined.

3. Methodology

This chapter describes the methodology that is used to answer the main research question: *How does the performance of a discrete choice model compare to the TBDFM in a joint mode & destination choice?* An overview of the methodology is presented in Figure 3.1.

The core of this thesis is estimating a TBDCM. A choice set is required to estimate any DCM, which contains the chosen and non-chosen alternatives & characteristics. A RP survey, OViN, will be used to acquire the chosen alternatives. A second data source is needed as OViN does not provide the non-chosen alternatives and the social economic geographical (SEG) data for the zones. These databases must be compatible because this data source is combined with OViN. These databases are analysed and combined to finalise the choice set for estimating the TBDCM and validate the TBDCM.

A literature review is performed to determine the structure of the TBDCM and the possible attributes of the utility functions. Multiple types of TBDCMs will be estimated to analyse which type performs the best. Input data is needed to develop and estimate a TBDCM. The choice set which was developed in the previous step will be used to estimate the TBDCMs. The nest structure has to yield significant results before the TBDCM can be used. Hence the estimated TBDCMs are filtered based on the results of the nest parameters. All the TBDCMs that yield significant results will be used for the remainder of this thesis.

The result of estimating the TBDCM is a list of significant attributes and their values. The TBDCMs will be analysed to understand the meaning and effects of these values. First, the TBDCMs are subjected to internal validation. The choice set, which was used to estimate the TBDCM, is used to validate the TBDCM. Second, an independent choice set is used to validate the choices the TBDCM predict to validate the choice model externally. the TBDFM are also subjected to external validation to compare the performance of the TBDFM with the performance of the TBDCM. A sensitivity analysis is performed to analyse the effect of the attributes of the TBDCMs and the TBDFM. The effect on the probability of an alternative of each attribute is analysed to determine how the choice model distributes the probabilities among similar alternatives. These validation methods will also determine if a TBDCM fails to predict correctly the independent choices. If that is the case, the TBDCM will not be used for the remainder of this thesis.

An analysis is done with a complete set of destinations to analyse the performance of the remaining TBDCMs and the TBDFM. The TBDCMs are estimated by sampling the destinations. But it is important to analyse the probability distribution among all alternatives. Based on the performance of the complete destination set, it can be determined which TBDCM is the best performing. Implementing a TBDCM costs a lot of effort; hence only one TBDCM will be implemented. Based on the literature, KPIs were identified to compare the case study results for both the TBDCM and the TBDFM. Based on all the results, a conclusion can be drawn on the overall performance of the TBDCM and the TBDFM.

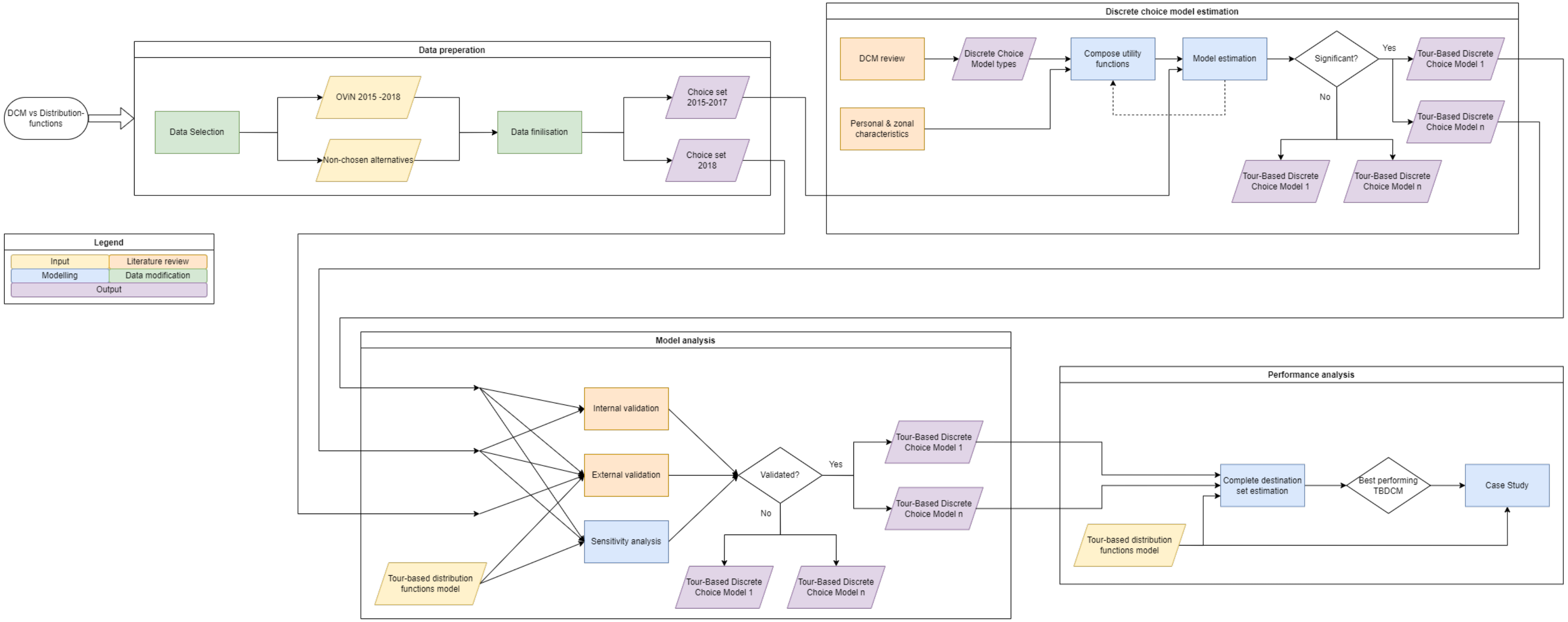

Figure 3.1: Approach

In the following sections, the steps are explained in more detail.

3.1. Data Preparation

A choice set is needed to estimate any DCM. This choice set consists of the chosen alternatives, non-chosen alternatives, personal characteristics and zonal characteristics. The data is collected from different sources. To combine the different data sets, the data needs to be analysed to determine the compatibility. Mismatches in the data sets need to be tackled and equalised. Choice sets for 2015 - 2017 and 2018 will be constructed to perform the discrete choice model estimation and model analysis, respectively

3.1.1. Data Selection

The choice set that is used will need to fulfil the requirements of estimating a DCM. The main requirement is that it contains the choices of individuals. This can be either stated preference (SP) survey, where people indicate what they would choose for a certain scenario, or RP survey, where people indicate what they have chosen. The main advantage of SP is that it can be used for theoretical scenarios. When the goal is to model a non-existing scenario/mode, a scenario will need to be proposed. In this case, it won't be possible to use an RP as no data are available. Moreover, when conducting an SP survey, data is collected for the chosen alternative and for the non-chosen alternative as both are combined in the survey.

However, the main issue with an SP survey is theoretical choices that people make. People may choose differently when faced with the same choice in real life. They indicate that they would choose the most sustainable alternative but, in fact, will choose the cheapest alternative when faced with the consequences. With the RP, this problem does not arise because people indicate their chosen alternatives. The downside of an RP survey is that it only contains the chosen alternatives. Thus, a second data set needs to be collected, which contains the non-chosen alternatives. Because for this study, the mode & destination choice contains existing alternatives and focuses more on the difference between a DCM and the TBDFM, an RP survey will be used.

To draw a fair comparison between the TBDCM and the TBDFM the available data for estimating the models should be equal. OViN 2015 - 2017 was used to estimate the TBDFM; hence this data set will also be used to estimate the TBDCM. Because OViN does not only contain the choices of the individuals but also a large set of personal characteristics useful for estimating the model, OViN will be used for the chosen alternatives. To complement the chosen alternatives, the non-chosen alternatives are needed. The main challenge is that the chosen and non-chosen alternatives have comparable data. If one data set overestimates the LoS, especially the travel times, there will be a bias when estimating the model.

OViN also collects the personal characteristics of the respondents. Hence only the zonal characteristics – facilities within the zone – are missing. Not all facilities attract the same type of trips and number of trips. Thus, the zonal characteristics must be categorised and scaled according to the facility's size. Larger facilities will attract more trips than smaller facilities.

Panel Effect

For OViN, respondents fill in their travel behaviour for a specific day. Because this is a random day, some people will have travelled more than others, and will have filled in more trips. If a DCM is estimated with these trips, the preference of people who have filled in more trips than average has a bigger influence than people who have filled in less than average. This panel effect needs to be reduced to a minimum to prevent bias in the model.

The panel effect can be reduced by taking one random trip from every questionnaire. However, this will reduce the number of observations. Estimating the model with and without the reduced number of observations will give an insight into the effect of reducing the panel effect.

3.1.2. Data Finalisation

Data is collected from multiple sources and will therefore need to be analysed. There will be differences between the multiple data sets, which must be equalised. In order to do this, these differences will first have to be identified. Furthermore, a case study will be performed in the performance analysis. The characteristics of the case study must match the characteristics of the TBDCM. The data preparation results in a choice set for the discrete choice model estimation and the model analysis.

The results of the data finalisation is a choice set which will be used to estimate the TBDCM and a choice set which will be used for the out-of-sample validation.

3.2. Discrete Choice Model Estimation

When constructing a DCM, the first aspect is determining what type of DCM will be estimated. A literature review was performed to define the nesting structure of the TBDCM. Based on this literature review, it is determined that a MNL, NL, and CNL will be estimated. The NL will be estimated for both nesting the modes and nesting the destinations. Furthermore, based on the literature, an overview is made of important personal and zonal characteristics that influence the mode & destination choice.

There is no guarantee that these characteristics will also be significant in this study. Adding or removing a characteristic to or from the utility functions will impact all the present characteristics. Hence estimating the different DCM types will be done iteratively, where each characteristic is added and removed one by one. Each type of DCM will be estimated and determined if the correlation of the alternatives is significant. This DCMs will be subjected to the model analysis.

3.2.1. Mode Changes

When an individual determines the mode(s) for the tour, they can use one mode for the whole tour or use multiple modes. They can for example leave the car or bike behind, knowing it will be picked up later in the day and brought home again. However, as discussed in Section 2.2, only a small percentage of the tours are multi-modal. A multi-modal tour is much more challenging computationally, compared to a uni-modal tour. Because computational time is a limiting factor for this study, the added value of a multi-modal tour is not worth the extra computational time. Therefore, the assumption is made that individuals cannot change their mode during the tour.

3.2.2. Activity Hierarchy

A tour consists of multiple activities in a specific sequence, for example, home-work-shopping-home. These activities need to have a destination in order to be performed. These activities are quite different. The duration of a *Work* trip or *education* trip is significantly longer than a *shopping* trip or *bring/get* trip.

As previously mentioned in Chapter 2, researchers have assumed that tours are based on the main activity in the tour. This means that first, the destination of the main activity is determined, which is only based on the trip from and back to the home location. This means that the destinations of the secondary activities are determined based on the destination of the main activity.

However, the complete tour should be considered. If the total utility of a tour is higher, the probability of that tour should be higher, regardless of the activities in the tour or a possible difference in priority. When evaluating the tour in steps, it may be that the tour with the highest utility does not have the highest probability. Therefore, there will be no activity hierarchy in the TBDCM, and the mode & destination combinations will be determined based on the complete tour.

3.2.3. Destination Sampling

When estimating a DCM, the non-chosen alternatives need to be known. For the mode choice, this is simple. If the car had been chosen, the individual could have also used the bike or PT. But for the destination, this is not straightforward. All the alternatives are initially all the zones in the model. The individual could have chosen any other zone. The number of alternatives in a mode & destination choice $= \#m * \#d$ alternatives. This requires a substantial amount of computation time, while in reality, not every destination is considered by the individual.

To reduce the number of alternatives, the destinations can be sampled. There are four strategies to sample the destinations, as mentioned in Section 2.4: 1) based on the distance to the chosen zone, 2) random sampling, 3) a combination of distance and random sampling, and 4) a combination of distance and zonal characteristics. How destinations are chosen differs per activity motive. For work trips, the destination choice is more or less fixed. In day-to-day travel behaviour, the work location might vary because of multiple work locations, but people do not change jobs very often; thus, even those multiple locations remain fixed. For other motives, such as shopping, the location could change daily as people may go shopping at locations that are most convenient for them at a certain point in time. Therefore, a large variety of zones should be considered as alternatives. They should however also represent the surrounding area in the best way possible. To conclude, apart from the chosen alternative 20 zones have been randomly sampled to construct the choice set. This results in a full choice set of 63 alternatives.

3.2.4. Attributes

Once the different structures have been laid out, the next step is to compose the utility functions. Based on models in literature, an overview of the occurrence of different attributes is presented in Table 2.2. This overview will be the starting point for the potential characteristics that will be used. However, these characteristics are limited by the survey that is used to collect the data and the environment which will be used to test the TBDCM. Together, these sources will determine which characteristics can be used to compose the utility functions.

3.2.5. Model Development

When the choice structure is decided, a DCM can be constructed. Attributes will be added to the utility functions, and the model will be estimated based on the choice set provided by the previous step. Based on the $\overline{\rho}^2$, the quality of the model can be monitored when adding new attributes. The goal is to compose a DCM with all the main attributes included and with the highest $\overline{\rho}^2$ possible. It is, unfortunately, impossible to know when the model is on its limit.

When a DCM is estimated, the values of the attributes are determined with a corresponding significance parameter p. This parameter is an indication of the significance of the estimated attribute. To ensure only statistically significant variables are implemented in the TBDCM, a threshold value of 0.05 will be used for p. If an attribute has a $p > 0.05$, the value is estimated with such a certainty that the attribute can be used in the TBDCM. Adding and removing attributes from the model will affect the significance of all other attributes. Thus, the attributes will be added and removed, one by one, to monitor the effect on all other attributes. This iterative process will be continued until there are no more attributes that can be estimated significantly and increase the $\overline{\rho}^2$.

Although the quality of DCM alternatives will be monitored, it is unknown how they will perform during the model analysis and performance analysis. To keep a certain flexibility in the research, three different DCMs will be constructed to minimise an early focus on solely one type of model.

3.2.6. Biogeme

To estimate the parameters in the model, a software package is used called Biogeme developed by Bierlaire, 2003. It is designed for estimating random utility models using the maximum likelihood estimation. The maximum likelihood estimation estimates the parameters such that the observed data has the highest predictive results possible.

3.3. Model Analysis

When estimating the model, Biogeme produces a report on the results. It provides the $\overline{\rho}^2$, BIC and AIC, which each have their own characteristics. These will indicate the fit of the model on the provided choice set. As mentioned before, it is important to validate the model with different methods.

Based on the literature, different methods were identified, which analyses different aspects of the model. For DCMs, there are two main validation principles apart from the model's statistics: internal & external validation.

3.3.1. Internal Validation

For internal validation, a choice set is needed, which consists of choices of the same individuals. Not only the individuals should be the same, but also the alternatives and the temporal characteristics. The only data set that fulfils these requirements is OViN 2015 - 2017.

The models will be subjected to a HOV. The data set will be split up between an estimation data set and a validation set. This method will indicate how strongly the model can reproduce the choices made by the same individuals. But to limit randomness and coincidental predictions, a CV will be applied. The HOV will be executed five times, with each a separate score. The variance of these results will indicate the dependency of the data and the consistency of the model.

3.3.2. External validation

A choice set with different criteria is required to perform an external validation method. Thus, a different choice is needed where either the individuals, time frame or location is different. To fulfil this requirement, OViN 2018 will be used. This OViN data set is collected from the same population with the same methods, yet the time frame is different. Furthermore, it may be that the individuals making the choices are different.

The DCM will be provided with the choices included in the data set and it will evaluate the probability of each alternative. Based on the reviewed literature, it is concluded that a threshold is needed to ensure that the model convincingly predicts the correct alternative. However, as the models have to deal with 63 alternatives, the probability will be distributed such that the alternative with the highest probability will be viewed as the predicted choice. Furthermore, the position of the chosen alternative will be monitored to analyse the overall predictive power of the model.

3.3.3. Sensitivity analysis

It is important for a choice model to predict the correct alternative. But it is also important for the choice model to distribute the individuals among multiple alternatives accordingly. These statements can contradict each other because the chosen alternative should have the highest probability, but similar alternatives should have similar probabilities. Therefore, a sensitivity analysis will be performed to analyse the sensitivity of similar alternatives.

A sensitivity analysis will determine the effect on the utility, and thus also the probability, by changing one attribute of an alternative. This change can be either a difference between two destinations or a change in personal characteristics. The choice will consist of two alternatives that are identical except for the specific attribute that has been changed. Therefore, the probability of a choice with two alternatives is determined where one of the alternatives is different for one characteristic. This isolates that characteristic, and the result in probability will be directly caused by changing this characteristic. This comparison will be made for every category of characteristics to determine each effect of those characteristics.

Moreover, it is essential to investigate the elasticities of the LoS attributes. If certain policies are applied that influence the travel time or price of certain modes, a shift in the modal split is expected. The change in the modal split must be in accordance with the change in the LoS values. Hence a sensitivity analysis is used to measure these effects.

3.4. Performance Analysis

In this section, the methods to measure the performance are discussed. These methods are focused on quantifying the effect of calculating the utility and its effect on the probabilities of the alternatives. This means that first, a sensitivity analysis will be performed to determine the influence of attributes on the probability of the alternatives. The TBDCM is estimated with a sample of possible destinations, and the validation is performed with the same sampling strategy. During this process, it is crucial to determine the probabilities of every destination if every zone is seen as an alternative. As a final step, the models will be subjected to a complete OD matrix generation. A case study will show how the models perform when a mode & destination choice is modelled for an entire population.

3.4.1. Complete destination set

The TBDCM is estimated by sampling other destinations as a possible alternative. But when applying the model to the complete set of destinations, the distribution of the probability might change. Therefore, the models will be subjected to choices on the complete set of destinations. This also gives an indication of the trip length distribution as long-distance trips will be directly compared to short-distance trips.

3.4.2. Case study

The final step is to analyse the performance of the models when applying them to an entire population. To accommodate this analysis, a large-scale transport model is needed. In the literature reviewed, multiple KPIs were found for analysing the modelled OD matrix. To compare the characteristics of the matrix, observed data is needed. Therefore, OViN 2018 will be used as it contains data from the entire Netherlands and is independent of the data used to estimate the models. The first KPI that will be used is the modal split. This will indicate the basic preference of each of the modes. It is important to analyse if the TBDCM shows a significant preference for one of the modes.

The second KPI is the trip length distribution for each mode. It is important to analyse the deterrence of the distance for a trip. Also, this will be different per mode as the behaviour for the bike is different than for PT.

The final KPI is ground truth observations. For this analysis, local traffic counts are needed to assess the flows. Based on the error, the results of the assignment can be assessed.

4. VERKEERS MODEL LIMBURG

As mentioned in Section 3.4, a case study is used to execute the performance analysis, which will need to fulfil certain requirements to be deemed useful. The main requirement is that the case study covers an entire country. To compare the results of the TBDCM and the TBDFM, a large-scale transport model is needed which covers that entire country. A large-scale transport model can run all four steps of a 4-step model, including trip generation, trip distribution, modal split and assignment.

RHDHV has a large-scale transport model which fulfils all the requirements: VML. This transport model executes all four steps and defines Limburg, a province of the Netherlands, as its study area, with the rest of the Netherlands, Belgium, France and Germany as its influence and periphery area. The core of VML is the TBM. Each persona has a distribution of different tour types, with different numbers of activities that are performed. The goal of VML is to model an average workday in Limburg. For each tour, the probability combination of destinations and mode is calculated by the TBDFM; hence the destination choice and mode choice are jointly determined. All the probabilities of each destination combination and mode will add to the OD matrix. This method ensures that the result is the average of each tour, hence the average of a complete workday.

As the TBDCM will be implemented in VML, the TBDCM will need to be compatible. VML consists of different modules that have specific input and output, which need to match with the TBDCM. This concerns mainly the personal and zonal characteristics. VML is based on OViN which means that all of the characteristics in VML are present in OViN but not all. Furthermore, the trip purposes of OViN are categorised to limit the potential tour types, a combination of trip purposes.

The TBDCM will be implemented in VML to run it in a large-scale transport model. The implementation in VML brings both advantages and disadvantages. The advantages are that a large population performing daily activities can be generated. A disadvantage is that the TBDCM will be limited by VML because only the characteristics that are present in VML can be used by the TBDCM. Because OViN data is also used for the TBDFM, VML is compatible with OViN. The same categories for characteristics are used for VML which can be found in OViN.

The TBDCM will replace the distribution function module in VML; hence the input and output of the TBDCM should match with VML. This chapter aims to analyse VML and identify key characteristics of the transport model. First, the personal characteristics are analysed. Individuals are modelled by VMLs population generation. The same characteristics should be present in the TBDCM. Second, the different trip purposes are defined, which are used to compose the different tours. Third, the zoning is analysed. Not only how VML divided its region into different zones but also the zonal characteristics. Fourth, the transport networks of VML are analysed. The quality of the travel times is important for the TBDCM. Lastly, the time periods of VML are analysed and the traffic is distributed in time periods to limit the computational time.

4.1. PERSONAL CHARACTERISTICS

Within VML, people are categorised based on a number of personal characteristics. This is done because a small change in a personal characteristic will translate into a small change in travel behaviour. Furthermore, if multiple people can be combined into the same category, this saves computational time as the distribution function will be the same. Thus, the categories of the characteristics are bounded so that the model still captures the essence of travel behaviour differences between different groups and reduces the computational time. For example, in the Netherlands, the minimum age for driving independently is 18. Thus, travel behaviour can change significantly when a person turns 18. Therefore, there is a boundary between the ages of 17 and 18. The categories for all characteristics are as follows:

Table 4.1: VML personal characteristics

Characteristics	**Categories**
Age	0-17, 18-34, 35-64, 65+
Employment	None, part-time, full-time
Household-income	<30k, 30-50k, >50k
Car availability	Yes/no

Based on Table 4.1, it is expected that VML contains 72 personas. However, some combinations of personal characteristics do not make sense, individuals below 17 year old are not employed. Apart from impossible combinations of personal characteristics, there are also personas which had too few observations to estimate a distribution function. Hence, some personas with multiple overlapping characteristics have been fused together. All in all, 42 personas remain which are presented in Table 4.2.

Table 4.2: Personas in VML

Persona	Car availability	Age	Household income	Employment
P1	Car	Age 0-17	Income 0-30k	Employment none/part-time/full-time
P2	Car	Age 0-17	Income 30-50k	Employment none/part-time/full-time
P3	Car	Age 0-17	Income 50k+	Employment none/part-time/full-time
P4	Car	Age 18-34	Income 0-30k	Employment none
P5	Car	Age 18-34	Income 0-30k	Employment part-time
P6	Car	Age 18-34	Income 0-30k	Employment full-time
P7	Car	Age 18-34	Income 30-50k	Employment none
P8	Car	Age 18-34	Income 30-50k	Employment part-time
P9	Car	Age 18-34	Income 30-50k	Employment full-time
P10	Car	Age 18-34	Income 50k+	Employment none
P11	Car	Age 18-34	Income 50k+	Employment part-time
P12	Car	Age 18-34	Income 50k+	Employment full-time
P13	Car	Age 35-64	Income 0-30k	Employment none
P14	Car	Age 35-64	Income 0-30k	Employment part-time
P15	Car	Age 35-64	Income 0-30k	Employment full-time
P16	Car	Age 35-64	Income 30-50k	Employment none
P17	Car	Age 35-64	Income 30-50k	Employment part-time
P18	Car	Age 35-64	Income 30-50k	Employment full-time
P19	Car	Age 35-64	Income 50k+	Employment none
P20	Car	Age 35-64	Income 50k+	Employment part-time
P21	Car	Age 35-64	Income 50k+	Employment full-time
P22	Car	Age 65+	Income 0-30k	Employment none/part-time/full-time
P23	Car	Age 65+	Income 30-50k	Employment none/part-time/full-time
P24	Car	Age 65+	Income 50k+	Employment none/part-time/full-time
P25	No car	Age 0-17	Income 0-30k	Employment none/part-time/full-time
P26	No car	Age 0-17	Income 30-50k	Employment none/part-time/full-time
P27	No car	Age 0-17	Income 50k+	Employment none/part-time/full-time
P28	No car	Age 18-34	Income 0-30k	Employment none
P29	No car	Age 18-34	Income 0-30k	Employment part-time/full-time
P30	No car	Age 18-34	Income 30-50k	Employment none
P31	No car	Age 18-34	Income 30-50k	Employment part-time/full-time
P32	No car	Age 18-34	Income 50k+	Employment none
P33	No car	Age 18-34	Income 50k+	Employment part-time/full-time
P34	No car	Age 35-64	Income 0-30k	Employment none
P35	No car	Age 35-64	Income 0-30k	Employment part-time/full-time
P36	No car	Age 35-64	Income 30-50k	Employment none
P37	No car	Age 35-64	Income 30-50k	Employment part-time/full-time
P38	No car	Age 35-64	Income 50k+	Employment none
P39	No car	Age 35-64	Income 50k+	Employment part-time/full-time
P40	No car	Age 65+	Income 0-30k	Employment none/part-time/full-time
P41	No car	Age 65+	Income 30-50k	Employment none/part-time/full-time
P42	No car	Age 65+	Income 50k+	Employment none/part-time/full-time

In the trip generation of VML, a population set is generated based on these characteristics. Thus, each individual falls into a category for each personal characteristic. Individuals with the same category for each characteristic fall into the same persona. Each persona has one distribution function that calculates the probability for each tour.

4.2. Trip Purposes

Individuals perform a wide range of activity types daily. The behaviour is different per activity. Certain facilities, such as schools, attract specific activity types. Each activity is performed with a certain purpose. Thus RHDHV has differentiated between activities, and each activity is assigned a purpose. The differentiation also allows to generate a daily travel pattern which contains multiple trip purposes. A tour type consists of multiple trip purposes, but many different tour types are possible with numerous different trip purposes.

For each persona, the travel behaviour is estimated beforehand. The travel behaviour is described by a selection of the tour types, including the frequency of these tour types. The frequency indicates how often a specific tour type is made per day. Because the goal of VML is to model an average day, this frequency is also the average for a single day.

Thus, the number of different trip purposes leads to more tour types, which in turn leads to more computational time. Seven trip purposes are defined to limit the computational time while having varying travel patterns: home, work, shopping, education, business, bring/get and other. Because OViN will be used to estimate the TBDCM, the trip purposes in OViN have to match the trip purposes in VML. However, in OViN, more trip purposes are defined. This means the trip purposes of OViN must be re-categorised to match VML. This is done as follows:

Table 4.3: Trip purposes.

OViN	**VML**
Home	Home
Work	Work
Business	Business
Shipment	Business
Pick-up/drop-off person	Bring/get
Pick-up/drop cargo	-
Education	Education
Shop	Shop
Visit/staying over	Other
Touring/trail	Other
Leisure	Other
Leisure other	Other
Service/personal care	Shop
Other purpose	Other

A combination of these trip purposes will compose a tour type. Every tour type will start and end at the home location. This means that there are no trips where people stay the night or leave the country for a period of time, either on holiday or on a business trip. Based on the personal characteristics, a tour frequency for 15 specified tour types is estimated by RHDHV as input. These tour types differ per persona.

4.3. Zoning

Individuals need to depart from a specific location and arrive at a specific location. For a large-scale transport model, it is impossible to determine these locations at the address level. Therefore, zones are defined which contain a certain area. Every trip departs from or arrives at the centroid of the zone instead of the specific location. The study area of VML is Limburg which means that for all the trips in Limburg, a more precise centroid is needed compared to trips in Groningen. Therefore, the zone in Limburg are defined small but zone will increase in size when they are further away from Limburg. For practical reasons, RHDHV has based this zoning on the PC4 zones. The zoning can be viewed in Figure 4.1 and Figure 4.2.

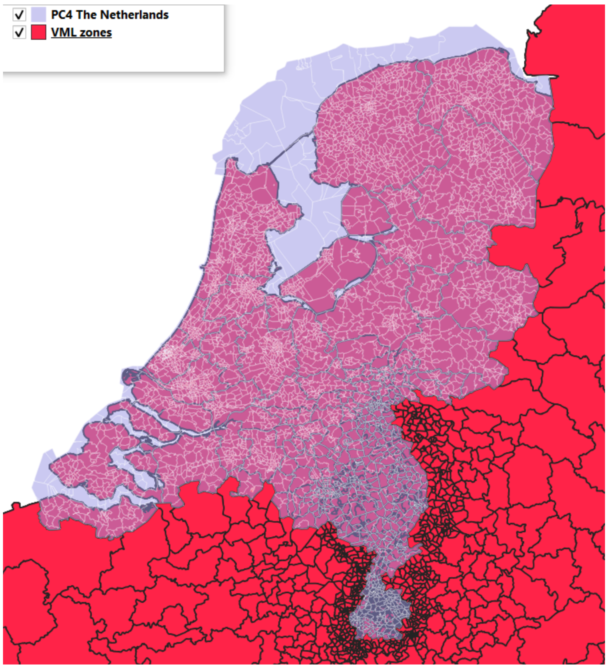

Figure 4.1: Zoning of The Netherlands.

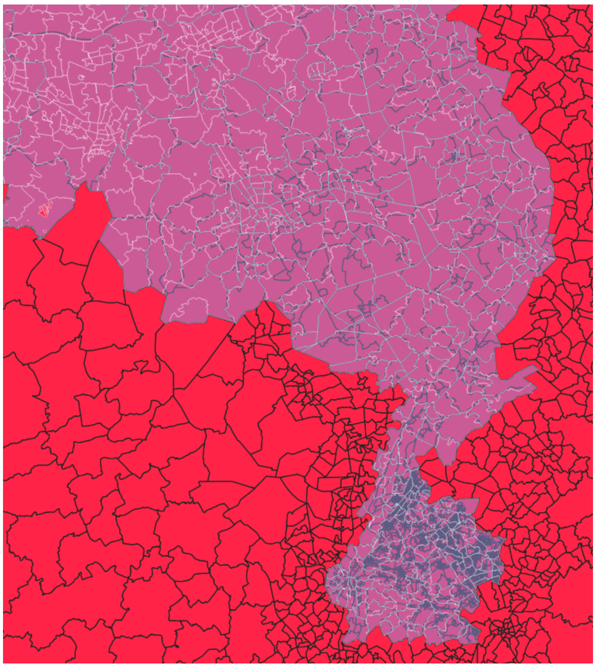

Figure 4.2: Zoning of Limburg.

Zones with more facilities attract more trips. More individuals travel to a large hospital than to a small hospital. Hence each trip purpose should have a corresponding attraction value for each zone. These attraction values are based on multiple characteristics of the given zone. Thus, an overview is presented of which zonal characteristic is used for which attraction value in VML. Apart from these attraction values, a specific zonal characteristic is the urbanisation degree. The urbanisation degree is not based on the number of facilities in the zone but on the number of addresses of the zone. The density of addresses within the zone indicates the urbanisation degree. VML has categorised all the zones into five urbanisation categories, with 1 being high-density zones and 5 being rural areas.

Table 4.4: Zonal characteristics.

Category	**Zonal characteristics**
Work	Employment: agriculture, service, industry, retail, care, hospitality, education, government, other
Business	Employment: care, government, services, other
Education	Capacity: school, university
Bring/get	Care m^2, capacity: childcare, primary school
Shopping	m^2: construction, living, supermarket, other
Other	Inhabitants & m^2: care, leisure, hospitality
Urbanisation	Density of addresses

4.4. Transport Networks

When modelling trips between zones, the travel times between all the zones are needed, in the form of a skim matrix. To model this skim matrix, data is needed on the transport networks for each mode. For the bike and car, this data is the roads, and for PT, it is the timetables of the bus, tram, metro and train.

VML has an extensive network for bike, car and PT. The bike and car networks are based on Open Street Map (OSM). OSM contains all the different roads and a classification system to identify the road types. There is also information on bicycle paths to differentiate between the availability of the car and the bike. Apart from the network itself, VML also has a junction delay model for the car and the bike. The junction delay model adds travel time to routes with junctions to compensate for the time lost while waiting for a traffic signal or traffic that has the right-of-way.

The network for PT is based on General Transit Feed Specification (GTFS). GTFS contains all the timetables of the PT suppliers. This data can be processed to model the travel time between all the zones in VML. Because it contains the timetables, detailed information, such as the in-vehicle time and transfer time, can be derived from the routes.

4.5. Time Periods

VML models the traffic in Limburg for an average workday. Hence the traffic needs to be distributed over the day. Certain trips are usually performed at a specific time of the day. Work trips are usually performed in the morning, leisure trips during daytime and home trips during the evening. VML is not an ABM; hence the time period of the trip is needed as an input.

Three time periods are defined in VML to distinguish between certain parts of the day: morning peak, evening peak and daytime. Before the trips are assigned to a time period, a 24h-ODmatrix is modelled. The trips are divided among the time periods based on the trip purpose and tour length.

5. Data Preprocessing

Data is needed to estimate a DCM. This data needs to contain choice sets which includes both the chosen and the non-chosen alternatives. OViN only contains the chosen alternatives; hence the non-chosen alternatives must be collected from a different data source. The quality of the choice set will influence the estimated TBDCM quality. In this chapter, the OViN data will be analysed to determine the quality of this data set. Errors within OViN must be identified and removed to increase the quality.

Because a different data source is needed for the non-chosen alternatives, there can be a difference between this data set and OViN. This different data sources needs to be compatible with OViN. Hence, the same personal characteristics need to be present and the travel time should be comparable. If one of the two data sets significantly under- or overestimates the travel time, the quality of the TBDCM will decrease. Therefore, a comparison will be made to ensure a certain similarity between the two to prevent bias.

As a final step, the two data sets will be merged into one choice set. Because the data sets come from different sources, adaptions must be made to create compatibility between the two data sets.

5.1. Data Analysis

This section analyses the data from OViN. For this study, three years of OViN data will be used: 2015, 2016 and 2017. These three years together create a choice set of 110,000 observations performed with either the car, PT or the bike and with a specific trip purpose. Because the home location of the individuals is given in VML, trips that go back home at the end of a tour or the end of the day are not needed to estimate the TBDCM. Hence these trips are not taken into account.

The main reason for choosing these three years is that the TBDFM are also estimated using these years of OViN. This will give both models the same number of observations for the estimation process. As CBS continuously develops their survey, it is difficult to stack multiple years. The main challenge are historical differences. The mobility behaviour of people and the network change over time. Thus, using outdated data will decrease the quality of the TBDCM. Furthermore, if the choices within the choice set are too different, this will create noise and decrease the quality.

The data from OViN is compared to the data from VML with a focus on travel time. It is important to compare the travel times from OViN with those of VML. If the TBDCM is estimated with the travel times from OViN, and they are not comparable with those of VML, it would provide incorrect results. Travel time attributes are estimated based on the deterrence found in OViN. But if the deterrence of travel time is different in VML, the TBDCM will predict the wrong choices. Hence a comparison is needed between OViN and VML.

5.1.1. Analysis of OViN data

This section analyses the data of OViN. Because this data set will be used to estimate the TBDCM, it is important to assess the quality beforehand. A survey will always contain errors; hence, these errors must be found and removed from the data set. First, the trips of OViN will be filtered. Different errors in the trips are identified and removed. Removing these errors will increase the quality of the TBDCM. When the errors of the data set are filtered out, the data set will be analysed to determine the quality.

Trip filtering

OViN is a survey, meaning mistakes can be made or strange observations can be recorded. People can make mistakes when submitting their travel behaviour, leading to incorrect submissions. The most important mistakes that must be filtered out are the trips themselves. A mistake in the destination will lead to a significant mismatch in the distance travelled. Especially if this mistake happens with a bike trip, the average speed of the trip can be unrealistically high or low. The same thing happens if the travel time is submitted incorrectly. There can also be mistakes in personal characteristics.

OViN is a large survey with many questions and observations. Different kinds of trips are included. However, some trips are part of a person's travel behaviour but will not make sense for estimating the model, such as tours where a person would go for a walk or bike around a park or forest. Such trips have the same origin as the destination and a significant travel time. While this is a common type of trip, the model would not understand it. The main indicator to identify these observations is the given travel time and travel distance as the crow flies between origin and destination. The resulting ratio indicates that the person travelled a logical path. Because there are observations which have a high travel distance, but the origin and destination in the same PC4 zone, these trips have to be filtered out as they are not modelled by VML.

Lastly, mistakes in the survey have to be identified to prevent estimating the model on invalid data. When people have filled in that they have chosen the car but do not have access to a car or do not have a driver's license, it would not be possible for them to have the car as the chosen mode. Thus, for car trips it needs to to be checked whether the person has a driver's license and has access to a car. If not, the trip will be left out of the choice set.

Travel Time

The observations of OViN are reported manually, which poses a number of challenges. The first aspect is that travel time is an estimated value. People have to recall and estimate their travel times. Because this is not a precise method, people do not precisely report their actual travel time but make a rough guess in minutes. Furthermore, longer trips are rounded up or down to a multiplication of five. This manual recording of travel times causes noise in the data set, leading to an inaccurate TBDCM. How often a trip is rounded up or down can be seen in Figure 5.1.

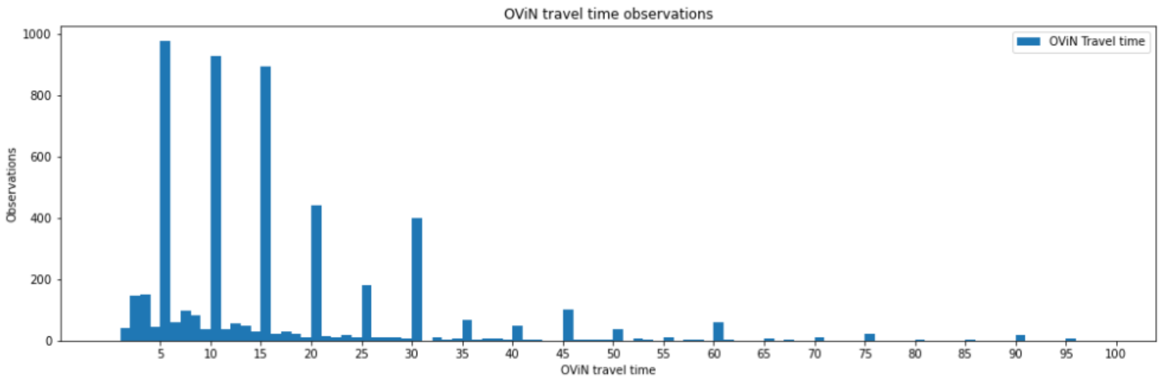

Figure 5.1: OViN travel times.

Based on Figure 5.1, it can be concluded that reported travel times are very inaccurate. Trips with a travel time with a multiplication of five are over-represented. This over-representation will significantly affect the estimation of the TBDCM.

Because OViN contains only the travel time, there are also unavoidable inaccuracies. The main inaccuracy is whether people have chosen the shortest route to their destinations. People may have different preferences in route choice, which causes them to take a different route than expected. Also, there may be accidents and other events that cause an incidental delay. Furthermore, people can miss their bus or train if they depart from home too late or if an incident occurs en route. Moreover, a bus or a train can also experience delays which in turn delays the individual.

5.1.2. Social Economic Geographical data

The corresponding zonal data is needed to estimate the zonal attributes of the utility functions. The SEG data is needed per PC4 area as this matches the zoning in OViN. The SEG must contains the m2 of activity types (shops, education, office, etc.), employment types and student places. This data is spread over different sources. Fortunately, RHDHV has processed this data already for VML. However, the study area of VML is Limburg, which means that the zone sizes are small in Limburg, but become coarser as they are further away from Limburg. It is more convenient to aggregate the SEG data to larger zones because the sum of the zones would suffice for the new zone. It is inconvenient to segregate a bigger zone's SEG data into multiple smaller zones, because it is difficult to retrace the specific location of all the shops and offices to determine in which part of the zone it is. Thus, this problem must be solved because the SEG are essential for estimating the TBDCM.

5.1.3. Non-chosen alternatives

OViN is a great source of RP, but when estimating a DCM the non-chosen alternatives are also required. These are not present in OViN, so they must be collected from a different source. There are two methods available to gather the non-chosen alternatives. The first one is with the means of open-source data. OSM is an open-source map which can be used to generate the network for cars and bikes. This network is very detailed and contains plenty of data on the network links. However, OSM does not contain the PT network. The alternative, GTFS, is also open source and contains all the timetables of every transport provider in the country. However, a huge disadvantage here is the time needed to process all this data and create one super network for every mode. Because time is limited, this method will not be used.

The second option is using Google API. The advantage of Google API is that the travel time of trips can be easily retrieved. It also contains GTFS for detailed PT trips. A big disadvantage is the costs. 62 requests are needed to complete the choice set for each observation. One request of Google API costs 0.005. The complete data set contains 110,000 observations. The costs are too high to create a data set with sufficient observations.

Another method is to use VML. The advantage of this method is that it does not have the main problem that is present in the open data solutions, as the transport networks for every mode are already created and pre-processed. Thus VML already has a skim matrix per mode. However, that skim matrix is only for the zones in VML. As mentioned before, the zoning of VML and PC4 areas does not match. The study area of VML is Limburg; hence the zones in Limburg are relatively small and zones outside of Limburg are relatively large. To match the zoning, the skim matrix needs to be adjusted. The aggregation of a skim matrix is cumbersome because there is no indication of the travel times to the new centroids of the new zones. Travel time may be lower because the destination is more accessible than before or higher because it is less accessible.

A comparison is made to check if the travel times of VML correspond with the travel times of OViN. A comparison can only be made for the area of Limburg, where the zoning of VML matches the zoning of OViN. For each trip in OViN, the corresponding travel time of VML is compared. Ideally, the travel times are similar, leading to a perfect bundle on $y = x$. A perfect relation is impossible, but the travel times of VML should at least approximate those of OViN. The comparison is presented in Figure 5.2, Figure 5.3 and Figure 5.4.

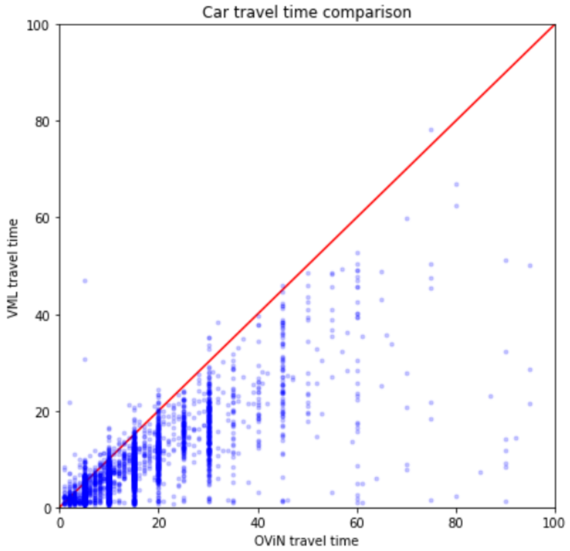

Figure 5.2: Car travel times

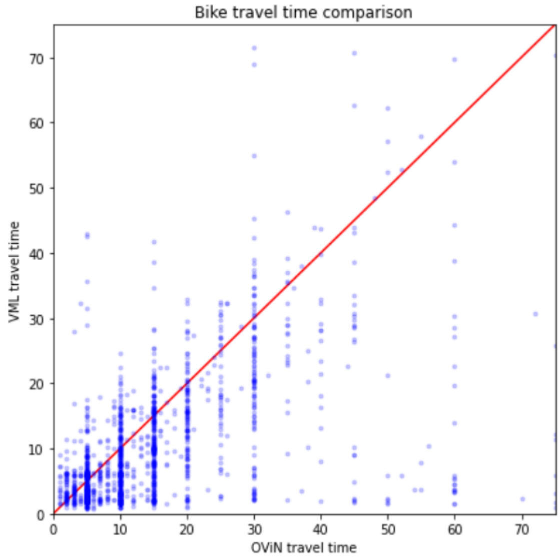

Figure 5.3: Bike travel times

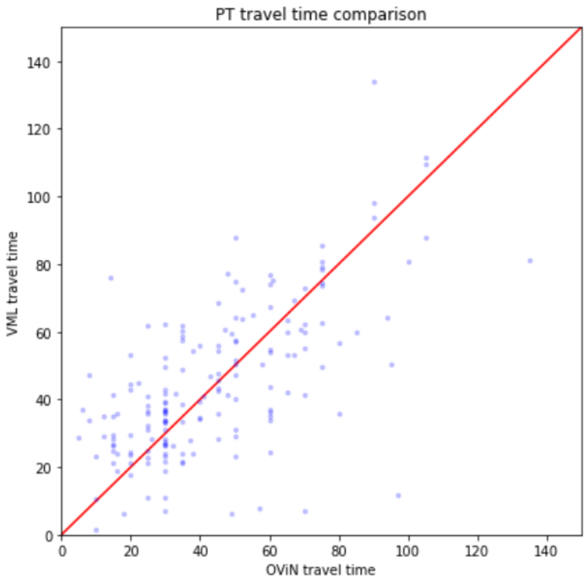

Figure 5.4: PT travel times

Based on Figure 5.2, Figure 5.3 and Figure 5.4, it can be concluded that the travel times of OViN and VML are not similar. Especially for the bike, there is a significant spread between the two databases. Many data points deviate from $y = x$ with a tendency that VML has lower travel time than OViN. The deviation can also be seen for the car. The spread is less, but almost all data points are below $y = x$. There is some spread for PT, but VML does not systematically have lower travel times. Overall, it is also quite clear that people tend to round their travel time to a multiplication of 5, as these observations are over-represented.

To determine if the deviation between the travel times of VML and OViN is too significant, and thus prevents the use of the travel times of VML, a two-sided t-test will be performed. This t-test will determine if the two distributions are significantly alike. If so, the two data sets can be used together. For every trip submitted in OViN, the corresponding travel time for VML is used for comparison. The trips are split up per mode to analyse whether a mismatch is only for a single mode or for the whole data set. A t-test will require a null hypothesis (H0) and an alternative hypothesis (H1) to test if H0 holds. It will produce a t-value and p-value. Based on the significance level that is set for the test, the boundaries for these values are set. For this test, a significance level of 0.05 is set. Thus, the H0 will be accepted, and H1 rejected if $-1.96 < t_{value} < 1.96$ & $p < 0.05$. If $1.96 < |t_{value}|$ & $p < 0.05$ H0 will be rejected and H1 will be accepted.

To test if the travel times of VML and OViN can be used together, the t-test will be as follows:
H0: The travel times of VML and OViN are similar
H1: The travel times of VML and OViN are different
$p = 0.05$
Accept H0 if: $-1.96 < t_{value} < 1.96$ & $p < 0.05$
Reject H1 if: $1.96 < t_{value}$ & $p < 0.05$

This test is done per mode to analyse if one of the modes may be used while the other modes are not suitable. The results of the t-tests are presented in Table 5.1.

Table 5.1: T-tests

T-test	Car t-test	PT t-test	Bike t-test
T-value	-21.49	3.43	-7.97
P-value	0.00	0.00	0.00

Based on the results, H0 can be rejected, and H1 is accepted for each mode. The travel times of OViN and VML are significantly different. This means that the VML travel times cannot be used as the non-chosen alternatives to estimate the TBDCM.

A solution to this problem is to replace the travel times of OViN with those of VML. The observations and choice are used from OViN, but the travel time is adjusted. The consistency of both travel times is compared to determine if this leads to more consistent travel times. For each trip in OViN, the distance is plotted with the travel time for both OViN and VML. It will show if the database is consistent with travelling a certain distance. Similar distances are not always travelled with the same travel time, but there should be a certain consistency with the data set. The comparison is made with each mode separately as each mode is expected to have a different relationship between the travel distance and travel time. The results are presented in Figure 5.5, Figure 5.6 and Figure 5.7.

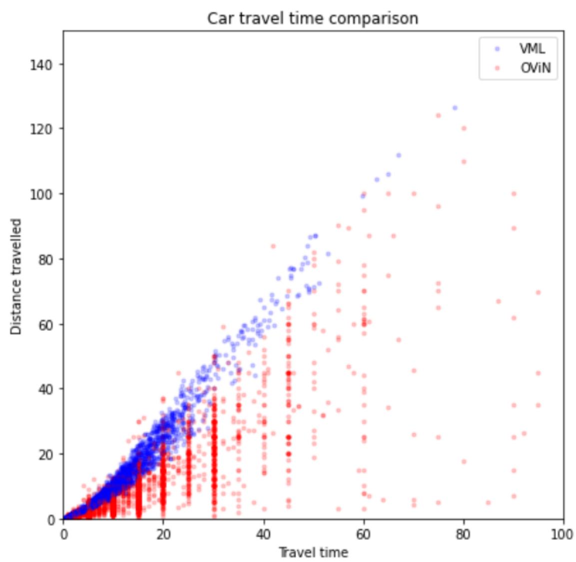

Figure 5.5: Car travel time consistency

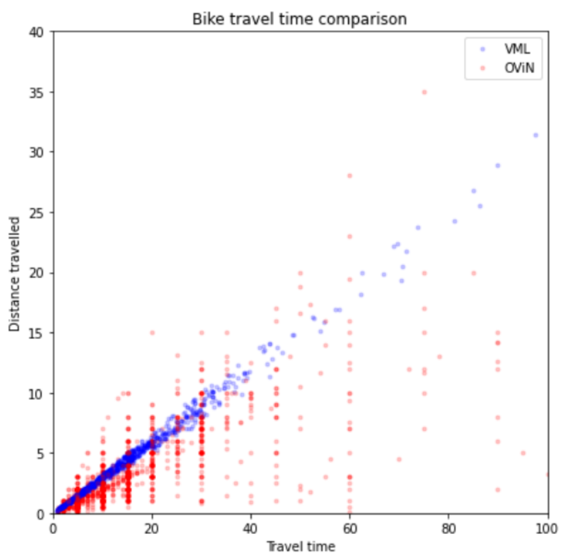

Figure 5.6: Bike travel time consistency

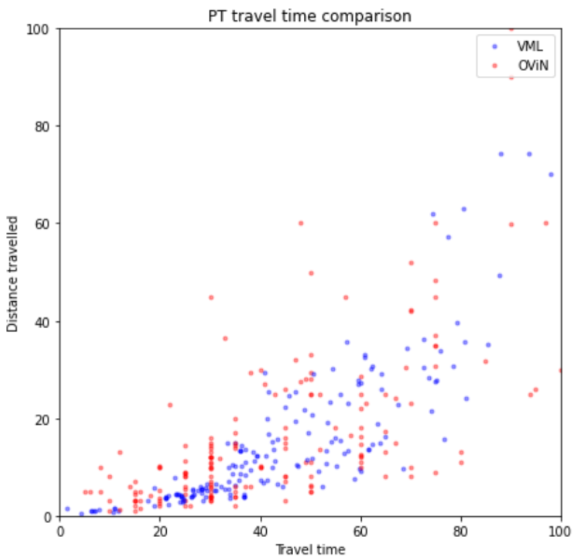

Figure 5.7: PT travel time consistency

From Figure 5.5, Figure 5.6 and Figure 5.7, it can be concluded that overall, VML is much more consistent for a certain travel distance than OViN. Especially for car and bike, the travel times are significantly more consistent. For PT, a slight change is expected as consistent travel times are highly dependent on a dense transport network, which is not the case for PT.

To conclude, the travel times of VML will be used for the non-chosen and chosen alternatives. For each observation in OViN, the recorded travel time will be replaced by the travel time from VML. The main reason for this switch is the difference in quality between OViN and VML. The travel times of OViN are unrepresentative since it is a survey. The travel times of VML are modelled. Thus, travel times are more detailed, rational and precise, leading to a higher distribution quality among the observations.

A disadvantage of replacing the travel times is that the zoning of VML needs to be adjusted to match the PC4 zoning. However, this had to be done either way to use the zonal characteristics of VML. How this modification will be executed will be discussed in the next section.

5.2. Data Preparation

This section prepares the choice set for estimating the TBDCM. As mentioned in the previous section, the travel times of VML will be used for the chosen and non-chosen alternatives. To use the zonal characteristics and travel times of VML, the zoning of VML must be changed to PC4 zoning.

To change the zoning of VML to PC4 zoning, there are two procedures to be defined: segregating and aggregating zones. The zones of VML are either larger than, equal to or smaller than PC4 zones. If the zones are larger than PC4, each zone needs to be segregated into smaller zones to create precise PC4 zones. If the zone is equal, no action is required. If the zone is smaller than the PC4 zone, the remaining VML zones that will create the larger PC4 zone need to be identified and aggregated to create the PC4 zone.

5.2.1. Segregating Zones

Segregating a VML zone into multiple PC4 zones is challenging. It is important to note that the zoning of VML is based on the PC4 zones. Thus, a larger VML zone always consists of multiple PC4 zones. Zone in VML also become larger if they are further away from Limburg, increasing the difficulty of segregating the zone. The zoning of VML and PC4 was presented in Figure 4.1 and Figure 4.2.

It is difficult to segregate the zonal characteristics and the skin matrix. To segregate the zonal characteristics, the precise location of each building that generated that attraction must be retrieved to correctly distribute the attraction of the VML zone to the multiple PC4 zones. This will require substantial effort. A simplistic approach is to equally distribute the zonal characteristics among the new zones. This approach can be quickly achieved, but it assumes the PC4 zones are equally attractive. An equal distribution may hold if a VML zone is split into two or three PC4 zones but not when a VML zone is an entire province.

Segregating the skim matrix is also troublesome. Travel times have to be created from and to all the existing zones. Furthermore, new intrazonal travel times and travel times from and to the new PC4 zones must be defined. This is dependent on the transport networks present in and near the VML zone but also on the shape of the PC4 zones. These results cannot be achieved with only the information that is available.

Is segregating a VML zone into multiple PC4 zones too problematic? The area where the VML zones are larger than the PC4 zones cannot be used for estimating the TBDCM. As a result, if a trip starts or ends in this area, it will be removed from the data set, resulting in a loss of observations. The exact number of observations lost will be determined further on in this thesis. Reducing the are from The Netherlands to only Limburg may create a bias in the data set. If there is a difference in travel behaviour between individuals from Limburg and the rest of The Netherlands it will not be taken into account.

5.2.2. AGGREGATING ZONES

To aggregate multiple VML zones to a single PC4 zone is simpler. As mentioned, the zoning in VML is based on the PC4 areas. Thus, multiple VML zones will fit perfectly in a larger PC4 zone. To aggregate the zonal characteristics, the characteristics of the VML zones that create a PC4 zone need to be summed.

Aggregating the skim matrix is a more difficult task. Multiple travel times have to be combined into one travel time for every possible destination in the model. An observed trip in OViN is from one PC4 area to another PC4 area, or to be more precise, from one address to another address. Thus, the individual also travelled from one VML zone to one PC4 area. It is however unknown which of the VML zones was used. Thus, assuming that all VML zones that were combined to create a PC4 zone are used evenly, the travel time from the new PC4 zone to all the other PC4 areas is the average of the travel times from the VML zones. The same holds for determining the intrazonal travel time. The new intrazonal travel time is the average of the travel times of the VML zones that create the new PC4 area. This method is visualised in Figure 5.8 and Figure 5.9.

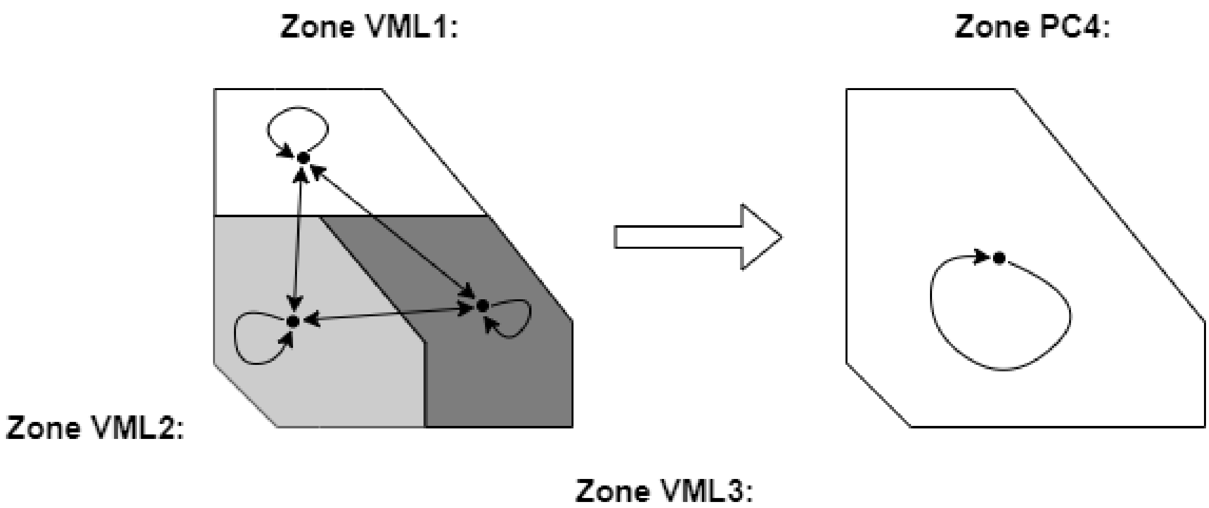

Figure 5.8: Skim aggregation intrazonal.

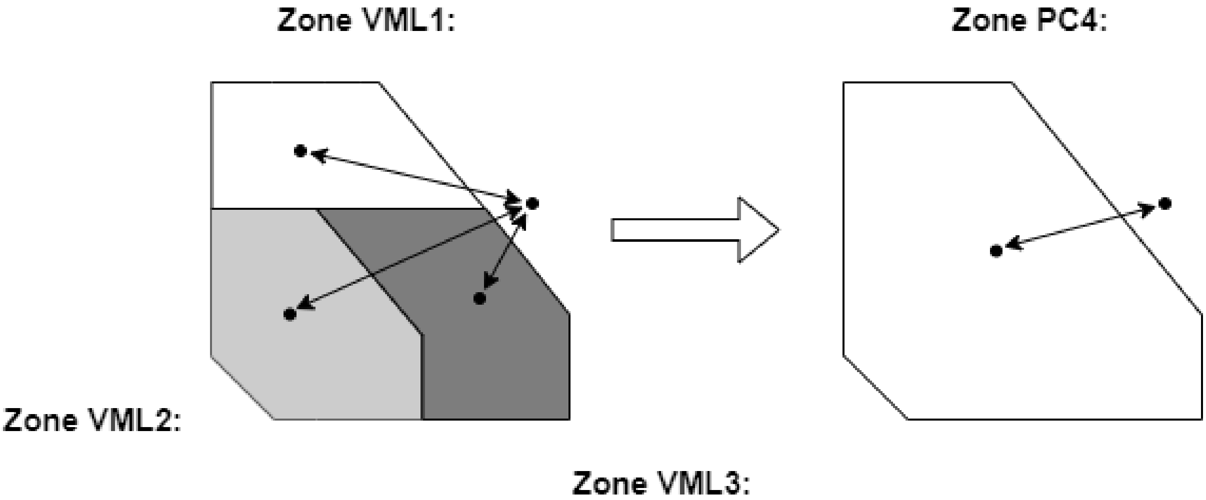

Figure 5.9: Skim aggregation zonal.

5.2.3. Finalising The Choice Set for Case Limburg

An observation that chooses a certain PC4 zone also chooses one of the VML zones within that PC4 zone. Thus, combining the VML zones into a PC4 zone will keep the choice valid in OViN. As a result, applying these methods will not affect the number of observations in the choice set.

However, as mentioned before, not segregating the VML zones will exclude a significant area of the model. Wherever the VML zones are larger than the PC4 zones, they will not be included in the choice set. The study area of VML is Limburg which means that in Limburg, the zones are smaller than PC4 zones, but outside Limburg, the zones are larger than PC4 zones. Thus, only trips within Limburg can be used for estimating the TBDCM. Of the original 110,000 observations, 5,500 observations start and end in a Limburg PC4 area. Compared to the original 110,000 observations, this is a massive reduction but still sufficient to estimate a DCM.

Another consequence is behaviour segregation. By taking only the data from people who travel in Limburg, the model estimates the coefficients for this specific behaviour; people may have different preferences in other provinces. The difference in the coefficient depends on how different the travel behaviour is.

5.3. Descriptive Statistics

In this section, statistics of the finalised choice set are visualised. The aim is to gain insight in influential characteristics. These insights will be used in the estimation process for the TBDCM. These characteristics can be either personal or zonal characteristics. First, the personal characteristics are analysed by comparing the modal split of a certain group with the modal split of the complete data set

5.3.1. Personal Characteristics

In OViN multiple characteristics are collected. Not all personal characteristics are going to be used in the TBDCM. Only, age, household and employment can be used, as presented in Table 4.1. The modal split for the complete data set and the model split for every sub category is evaluated to determine if these personal characteristics may have an influence on the mode choice.

If there is a significant difference between the modal split of the complete data set and the sub category it is not necessarily means that the specific personal characteristics will influence the mode choice. There could be an underlying relation which is not visible in the visualisation.

Moreover, the increase in a specific mode cause a decrease in the other modes. Hence it is difficult to determine whether an increase in a specific mode is related to the preference for that mode or a dislike for the other modes.

The modal split for the complete data set and the age categories are presented in Figure 5.10.

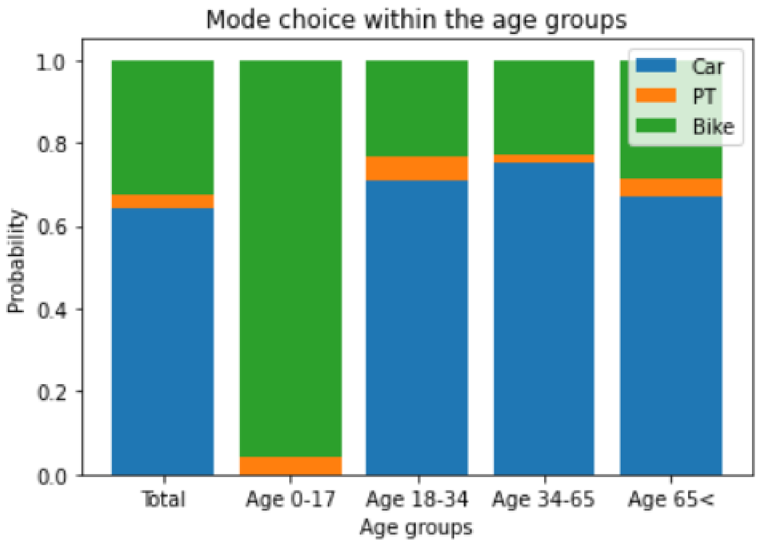

Figure 5.10: Modal split for age categories.

Based on Figure 5.10 it can be concluded that the age group 0-17 has a massive impact in the mode choice. This impact is mainly caused by the unavailability of the car. Individuals under 18 years old are not allowed to drive the car. Hence, the modal split for the car is zero. Also, the mode share for PT in the age group 34-65 is significantly lower than the share of PT in the other age groups.

The modal split for the complete data set and the employment categories are presented in Figure 5.11.

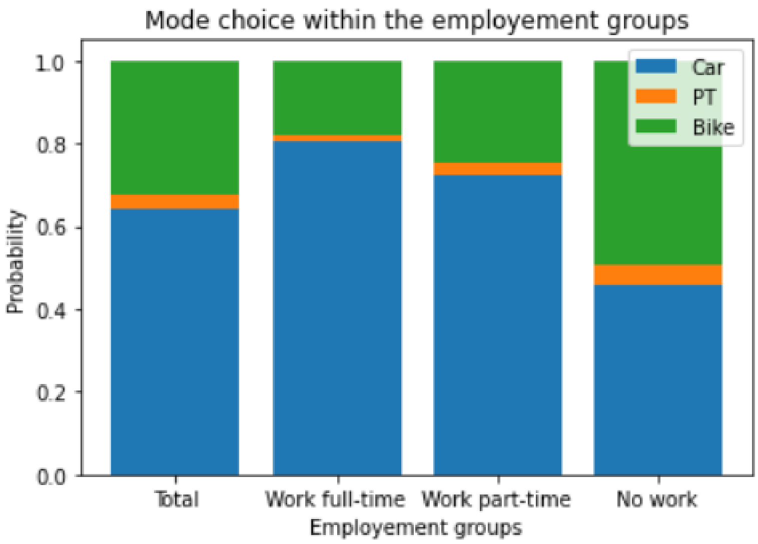

Figure 5.11: Modal split for employment categories.

Based on Figure 5.11 it can be concluded that being employed influences the mode choice significantly. There is a significant difference between the mode share of unemployed and employed individuals. As a result, the unemployed group have a higher mode share for the bike. Hence, the estimation process should determine whether employed individuals prefer the bike of unemployed individuals prefer the bike.

The modal split for the complete data set and the income categories are presented in Figure 5.12.

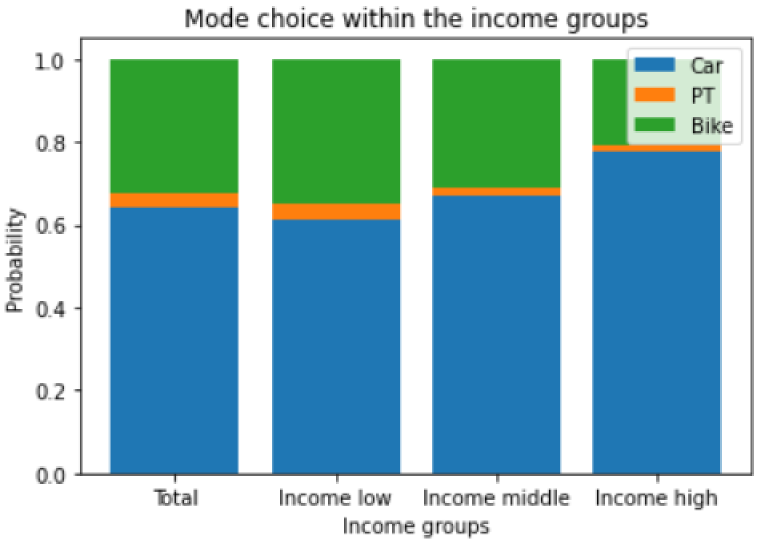

Figure 5.12: Modal split for income categories.

Based on Figure 5.12 it can be concluded that income has the lowest influence on the mode choice of the personal characteristics. Only for the high income group there is a slightly higher mode share for the care and lower for the bike. Income low and middle are similar to the modal split of the complete data set.

5.3.2. Zonal Characteristics

The amount of trips to each zone is plotted with its attraction value to determine if the zonal characteristics might have an influence in the destination choice. If there is a correlation between the number of trips to a zone and the attraction, it may be the case that there is a causation. The trips to the same zones are depicted in Figure 5.13, Figure 5.14, Figure 5.15, Figure 5.16, Figure 5.17 and Figure 5.18.

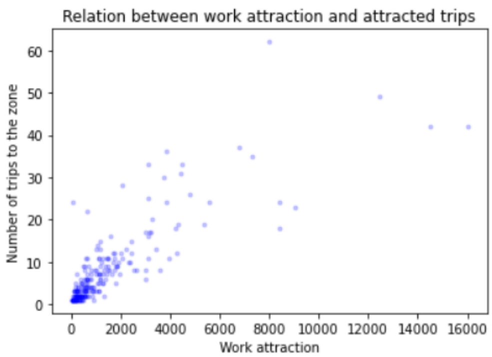

Figure 5.13: Work trips to the same zone.

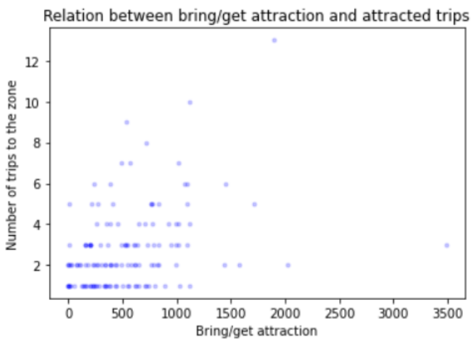

Figure 5.14: Business trips to the same zone.

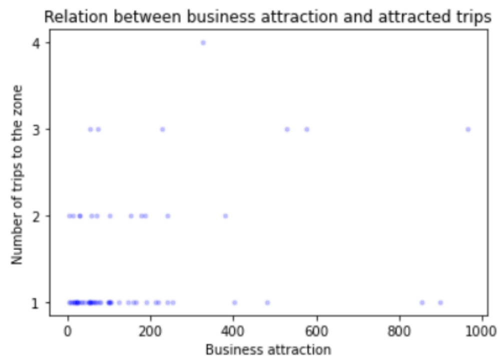

Figure 5.15: Education trips to the same zone.

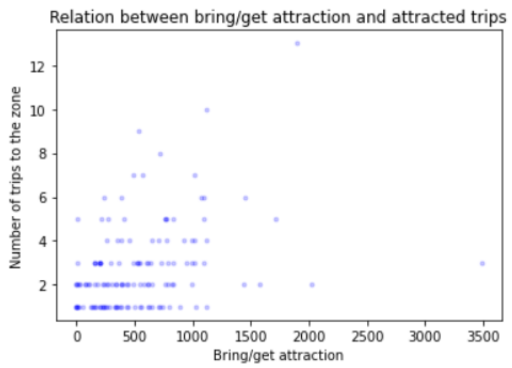

Figure 5.16: Bring/get trips to the same zone.

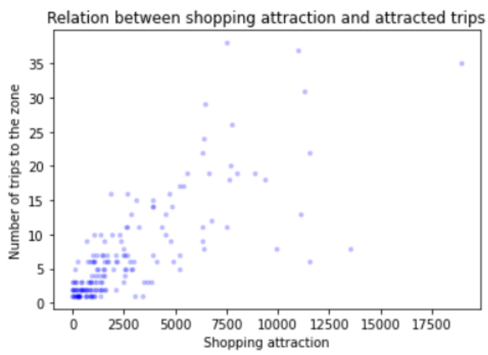

Figure 5.17: Shopping trips to the same zone.

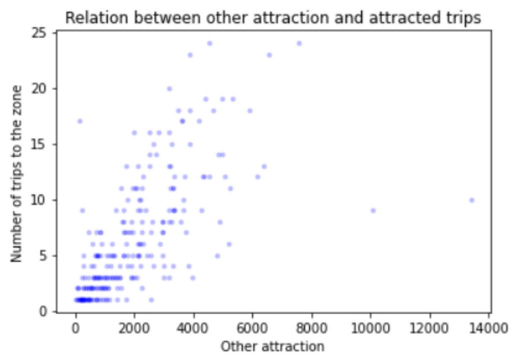

Figure 5.18: Other trips to the same zone.

Based on Figure 5.13, Figure 5.14, Figure 5.15, Figure 5.16, Figure 5.17 and Figure 5.18, it can be concluded that there are differences between the different zonal characteristics. For work, shopping and other, there is a clear relation between the number of trips to a zone and its attraction. If the attraction increases, the number of trips to that zone also increases. For these characteristics there are a few outliers but generally the relation holds for all the zones.

The correlation between the number of trips and the attraction, is for education and bring/get less visible. For bring/get there are relative fewer observations. Hence a correlation is difficult to determine. For education has a variety of number of trips for zones with a relative low attraction.

For business it is very difficult to determine if there is a correlation between the number of trips and the attraction because of the few observations present in the data set. Most of the zones have only one trip to that zone. Because there is a lack of business observations, it is not possible to determine if there may be a correlation or not.

5.4. Conclusion

OViN is a large data set which contains only the chosen alternatives. Only one source was found to complement OViN with the non-chosen alternatives: VML. Concession had to be made to finalise the data set. Only the area of Limburg could be used. Also the travel times of VML had to be aggregated to match the zoning of OViN.

Because the travel times of VML were more consistent relating to the distanced travelled and the occurrence of certain travel times, the travel times of OViN were replaced with the travel times of VML.

6. Results

In this chapter, the results are presented. This thesis aims to determine how the DCM and the TBDFM perform on a joint mode & destination choice. Therefore, the results consist of three sections: TBDCM estimation, model analysis and performance analysis.

In the first section, a TBDCM is estimated. Chapter 2 describes how to define a DCM for a TBM and identifies multiple choice structures that could describe the choice between the mode & destination choices. With the data set from Chapter 5, first, a standard MNL model is estimated. To elaborate the MNL, nest structures are added to the MNL to test if there is a correlation among certain alternatives. This results in an NL and a CNL model. The strength of the correlation is indicate by the nest parameter: μ. The condition $\mu > 1$ must be met for the nest structure to be valid. If not, the nest structure is invalid, and the NL is rejected.

Secondly, if a TBDCM is estimated, the results have to be analysed and validated. The estimated TBDCMs are subjected to two validation methods. For comparison, the TBDFM are also subjected to the same method for validating the models. First, the models are validated based on the data set that was used to estimate the TBDCMs. The in-sample validation shows how consistently the TBDCMs perform on the estimation data set. Secondly, the TBDCMs and the TBDFM are subjected to an out-of-sample validation. The models must predict the choices of a different data set than the one used for estimating the TBDCM. This will indicate how the models perform on different choices than the choices that were used to estimate the models. Lastly, the sensitivity of the parameters of the TBDCMs, and the TBDFM are analysed. The effect of the parameters is determined by comparing them with each other and determining their influence on the probability of an alternative.

Lastly, the TBDCMs, and the TBDFM are subjected to a performance analysis. First, the models must distribute the probability on a choice set with all the Limburg zones as an alternative. The destination alternatives were sampled when estimating the TBDCMs. Now all the destinations are seen as potential destinations. Lastly, the TBDFM and the best performing TBDCM are implemented in VML. A complete run of VML is executed. The performance of the models on a large population and a larger choice set can be compared. Based on the modal split, trip length frequency and ground-truth observations, the performance of the TBDFM and the TBDCM is assessed.

6.1. Estimating Tour-Based Discrete Choice Model

In Chapter 2, the basis of the TBDCM was defined. In the literature reviewed, the most common types of DCM are the MNL and NL. To ensure a diverse selection of potential TBDCMs, the MNL, NL mode, NL destination and CNL are estimated. The tour would be constructed by applying the TBDCM for each activity separately. Connecting these activities would create the probability of each complete tour. As a result, the trip purpose is a parameter in the TBDCM. The trip purpose variable requires separate attraction attributes for each trip purpose. To specify the influence of each characteristic, the attributes are mode specific. This means that each characteristic is independent of each mode. Thus, a characteristic may have three values, for each mode, instead of generalising the characteristic for multiple modes.

To estimate a DCM a general utility function needs to be defined which contains all the personal & zonal characteristics. Every alternative should contain an alternative specific constant (ASC) to describe the taste variations. For a mode & destination choice this would be one ASC for every mode & destination combination. However, in the destination choice a ASC is rarely used. Hence the ASC is only specified for the different modes. The utility function for each alternative, for each mode m and destination d combination, is presented in Equation 6.1.

$$V_{m,d} = ASC_m + \sum_i \beta_{m,d} * x_i \tag{6.1}$$

Potential attributes that influence the mode & destination choice are personal, zonal and LoS characteristics. As previously mentioned, an overview of all the different characteristics that were found in literature is presented in Table 2.2. Characteristics that are not present in this overview are not excluded from this thesis as the goal of the literature study was not to collect every characteristic used in a mode & destination choice. Hence all the characteristics present in VML are estimated to determine if they have a significant influence. The personal characteristics of VML are shown in Table 4.1, and the zonal characteristics are presented in Table 4.1, and the zonal characteristics are presented in Table 4.4.

6.1.1. Multinomial logit model

Attributes are specifically assigned to a mode to isolate the effect it has. Hence, each attribute can occur for each mode. The significant estimated values are presented in Table 6.1. It is possible that some attributes could not be estimated significantly as not all attributes influence the mode & destination choice. If that is the case, the value is not present in this overview. The meaning of the estimated attributes will be discussed below.

Table 6.1: Estimated MNL model

Attributes	**Car**	**PT**	**Bike**
ASC-mode	1.13**	-2.83**	0 (ref)
Travel time	-0.185**	-0.0107**	-0.123**
Age 0-17	-	-	3.7**
Age 18-34	-	-	-
Age 35-65	-	-0.598**	-
Age 66+	0.393**	-	-
Work full-time	0.608**	-	-
Work part-time	0.393*	-	-
Work nan	-0.773**	-	-0.609*
Income high	-	-	-
Income middle	-	-	-
Income low	-	-	-
Urban destination 1	-0.55**	0.802**	-
Urban destination 2	-	-	-
Urban destination 3	-	-	0.296**
Urban destination 4	-	-1.18**	-
Urban destination 5	-0.39**	-2.63**	-
Urban origin 1	-	1.19**	-
Urban origin 2	-	-	-
Urban origin 3	-	-	-
Urban origin 4	-	-	0.358**
Urban origin 5	0.251*	-1.00**	0.752**
Log(attraction) B	0.296**	-	0.659**
Log(attraction) D	0.549**	-	0.345**
Log(attraction) E	-	0.236**	0.328**
Log(attraction) S	0.414**	0.292**	0.523**
Log(attraction) W	0.726**	0.653**	0.933**
Log(attraction) O	0.409**	0.279**	0.581**
$\overline{\rho}^2 = 0.577$			

** = p<0.05, ** = p<0.01*

The most striking value is the $\overline{\rho}^2$, 0.577 is very high. It is however difficult to retrace why the $\overline{\rho}^2$ is high as the estimation process is execute by an external software package. A possible explanation is the replacement of OViN travel times by VML travel times. Further research should indicate if this is the case.

There are significant differences between the attributes per mode. The ASCs are based on the bike. This means the ASC for the bike is 0, and the ASC for the car and PT are estimated and act as a reference to the bike. The ASC is a collection of all non-observed variables. The car has the highest ASC, which indicates it is the preferable mode.

As expected, the travel time attribute for PT is the highest. PT is more often used for long-distance trips; hence the deterrence for an extra minute of travel time is lower than for the car or the bike. However, the value for the car is lower than for the bike. The lower value indicates that the bike is preferred to the car for longer trips. However, the car is faster, the travel time is usually quicker than the bike for the same OD pair. This effect can be seen in Figure 6.1, the median travel time in Limburg in VML is for car: 29.22, PT: 84.36 and bike: 136.12 minutes. Hence if the travel time for the car is equal to that to the bike, the bike is the preferred alternative because the travel time for the car indicates it is an inconvenient alternative.

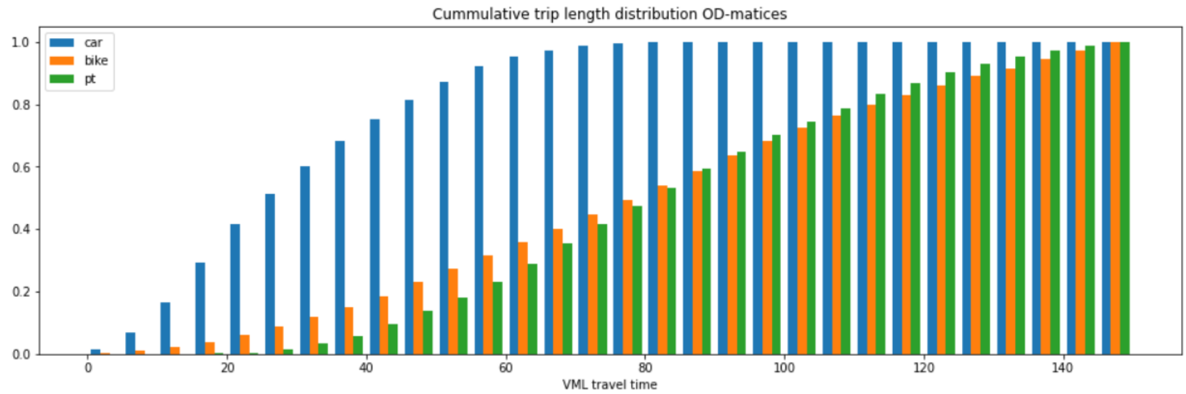

Figure 6.1: Cumulative OD-pair length distribution in VML

Individuals under the age of 18 will have a strong preference for the bike. This makes sense as people under 18 are not allowed to drive. The TBDCM takes this into account by limiting the car's availability. The strong preference for the bike is for comparison with PT. No attribute was estimated with a significance limit for the 18-34 age group. There is no estimated attribute, mainly because the 18-34 age group is a sample of people with varying travel behaviour. It comprises students, who have varying schedules, and employees, who have more fixed schedules. These people all have different travel behaviour and preference, which leads to the model being unable to estimate an attribute. The 35-65 age group is indifferent on PT whereas the 66+ age group prefers the car.

People who are employed have a clear preference for the car. This preference is even stronger for people who work full-time. The car gives them the flexibility to go to work and work at locations which are difficult to access with PT. Some companies offer employees a lease car; hence employed individuals have more often access to a car. Employed individuals would use the car more often than unemployed individuals. It is a reasonable assumption that individuals would acquire a car only if they are going to use it regularly.

The absence of income attributes substantiates the condition for using a car. If only money was the factor in car use, high-income individuals would use the car more often. However, the high income attribute has no significant influence on the mode choice and does certainly not have a stronger influence than the employment attributes. Thus, a high income alone does not influence using the car.

When people travel towards highly urbanised zones, PT is preferred, while the car is undesired. A reasonable assumption is that high urban zones get easily congested with cars while PT is more developed in those zones. The effect is almost mirrored for lower urban zones. PT is undesired when travelling towards rural areas as the PT network is less developed in those zones. The same holds when people are starting their trip. If the trip starts in a highly urbanised zone, PT is preferred, and if it starts in a rural area, PT is undesired. If the trip starts in a rural area, the car is preferred but the bike even more.

When estimating the attributes of the attraction values, it was noticed that the logarithmic relation performed better than the linear relation. It flattens the desirability of very large zones instead of limitlessly increasing the desirability while the attraction value increases. It is also more common in the reviewed literature to implement this relation as logarithmic, as shown in Table 2.1. Moreover, it is also more suitable when applying the

TBDCM in VML because of the zoning. In VML, the zones become larger as they are further away from Limburg. Thus the SEG also increase, which includes the attraction values. If a linear relation was applied in the TBDCM, the utility for numerous zones would be extremely high due to an extremely high attraction value. The logarithmic relation eliminates this phenomenon, as an excessive high attraction does not correspond with an excessively high value.

For the attraction attributes, only a few could not be estimated. For PT, the attractions for business and bring/get trips were insignificant. The insignificant attributes are presumably as business trips are not likely done with PT. The same holds for bring/get. For the car, education is not estimated, which is most likely for the same reason. It is a reasonable assumption that individuals who go to school or university have no access to a car yet. Hence it is not possible to estimate a value for the attribute.

For all modes, the work attraction is the highest attraction attribute. It indicates that for the 'work' trip purpose, the attraction of the zone has the biggest influence on the probability of the alternatives. For education, the attraction attributes are relatively low. Travel time is more important in this destination choice.

6.1.2. Nested Logit Model

Alternatives are nested to determine the correlation between alternatives. As mentioned in Chapter 2, there are different options to nest the alternatives. The alternatives are correlated with a parameter, which will be estimated, and which indicates to what extent the non-observed attributes are similar. This parameter, μ, will range from 1 to infinite, where 1 indicates no correlation, and a high value indicates a strong correlation. A high correlation will signify a substitution of the correlating alternatives. Instead of multiple alternatives, there is first the choice of the overarching alternative, either the mode or destination. μ cannot be lower than 1. If that is the case, the model will be rejected. Hence μ will have a lower bound of 1.

First, the alternatives are nested per destination. Hence the alternatives zone x-car, zone x-PT and zone x-bike are nested. Each nest is a destination nest, hence all the nests have the same μ. The nesting of alternatives with common destination tests if the mode choice is subordinate to the destination choice. The results of the NL-destination are presented in Table 6.2.

Table 6.2: Estimated NL-destination model

Attributes	**Car**	**PT**	**Bike**
ASC	1.13**	-2.83**	0 (ref)
Travel time	-0.185**	-0.0107**	-0.123**
Age 0-17	-	-	3.7**
Age 18-34	-	-	-
Age 35-65	-	-0.598**	-
Age 66+	0.393**	-	-
Work full-time	0.608**	-	-
Work part-time	0.393*	-	-
Work nan	-0.773**	-	-0.609*
Income high	-	-	-
Income middle	-	-	-
Income low	-	-	-
Urban destination 1	-0.55**	0.802**	-
Urban destination 2	-	-	-
Urban destination 3	-	-	0.296**
Urban destination 4	-	-1.18**	-
Urban destination 5	-0.39**	-1.26**	-
Urban origin 1	-	1.19**	-
Urban origin 2	-	-	-
Urban origin 3	-	-	-
Urban origin 4	-	-	0.358**
Urban origin 5	0.251*	-1.00**	0.752**
Log(attraction) B	0.296**	-	0.659**
Log(attraction) D	0.549**	-	0.345**
Log(attraction) E	-	0.236**	0.328**
Log(attraction) S	0.414**	0.292**	0.523**
Log(attraction) W	0.726**	0.653**	0.933**
Log(attraction) O	0.409**	0.279**	0.581**
μ zone: 1			
$\overline{\rho}^2 = 0.577$			
$*$ = p<0.05, $**$ = p<0.01			

From the results, the μ is the most important one. $\mu = 1$ indicates no correlation among alternatives with the same destination. Because $\mu = 1$ will influence nothing in the utility functions, the rest of the attributes will remain the same. Hence the NL for the TBDCM will be rejected.

The alternatives can also be correlated per mode. When this is the case, it is assumed that the individuals first choose their mode and the destination based on that mode. The mode choice differs per nest; hence a μ is estimated per mode. The nest contains all alternatives with the specific mode to all the destinations. The results are presented in Table 6.3.

Table 6.3: Estimated NL-mode model

Attributes	**Car**	**PT**	**Bike**
ASC	1.04**	-3.01**	0 (ref)
Travel time	-0.178**	-0.0104**	-0.113**
Age 0-17	-	-	3.63**
Age 18-34	-	-	-
Age 35-65	-	-0.601**	-
Age 66+	0.391**	-	-
Work full-time	0.579**	-	-
Work part-time	0.263*	-	-
Work nan	-0.805**	-	-0.612*
Income high	-	-	-
Income middle	-	-	-
Income low	-	-	-
Urban destination 1	-0.529**	0.806**	-
Urban destination 2	-	-	-
Urban destination 3	-	-	0.296**
Urban destination 4	-	-1.18**	-
Urban destination 5	-0.378**	-2.63**	-
Urban origin 1	-	1.14**	-
Urban origin 2	-	-	-
Urban origin 3	-	-	-
Urban origin 4	-	-	0.328**
Urban origin 5	-	-0.97**	0.705**
Log(attraction) B	0.284**	-	0.616**
Log(attraction) D	0.537**	-	0.308**
Log(attraction) E	-	0.244**	0.299**
Log(attraction) S	0.396**	0.283**	0.486**
Log(attraction) W	0.698**	0.637**	0.878**
Log(attraction) O	0.392**	0.272**	0.542**
μ mode	1.04**	1.00***	1.15**
$\overline{\rho}^2 = 0.577$			

** = p<0.05, ** = p<0.01, *** = active bound*

From the results, it can be concluded that the behaviour per mode is quite different. The μ for PT is 1, which indicates no correlation among the alternatives. The choice for PT is not because individuals choose PT first and the destination second. Individuals rather choose PT and the destination simultaneously.

For the car and bike, this is different. The μ is higher than 1, which indicates a correlation. The alternatives for these nests have common unobserved attributes. The choice for these modes is made before the destination choice. However, the μ for the car is 1.04, and the μ for the bike is 1.14. The correlation among in the nests is not strong, on the contrary, it is weak. Furthermore, the $\overline{\rho}^2$ does not show any improvement compared with the MNL model. An equal $\overline{\rho}^2$ indicates that the added nesting does not contribute to a higher predictability of the choices.

Although the $\overline{\rho}^2$ remains the same, the calculated utility of the alternatives is different. It is important to validate both models because they perform the same on the estimation data set, but there is no guarantee it will be the same for a different data set.

Cross Nested Logit

To further elaborate the nesting of the alternatives, the destination nests and mode nests are combined. This combination of nests creates a CNL model. Each alternative is nested in its destination nest and its mode nest. As a result, two utility functions contain each alternative. Each μ parameter is coupled with an α to differentiate between the nests. The differentiation indicates which nest is more dominant.

Table 6.4: Estimated CNL model

Attributes	Car	PT	Bike
ASC	1.13**	-2.83**	0 (ref)
Travel time	-0.185**	-0.0107**	-0.123**
Age 0-17	-	-	3.7**
Age 18-34	-	-	-
Age 35-65	-	-0.598**	-
Age 66+	0.393**	-	-
Work full-time	0.608**	-	-
Work part-time	0.393*	-	-
Work nan	-0.773**	-	-0.609*
Income high	-	-	-
Income middle	-	-	-
Income low	-	-	-
Urban destination 1	-0.55**	0.802**	-
Urban destination 2	-	-	-
Urban destination 3	-	-	0.296**
Urban destination 4	-	-1.18**	-
Urban destination 5	-0.39**	-1.26**	-
Urban origin 1	-	1.19**	-
Urban origin 2	-	-	-
Urban origin 3	-	-	-
Urban origin 4	-	-	0.358**
Urban origin 5	0.251*	-1.00**	0.752**
Log(attraction) B	0.296**	-	0.659**
Log(attraction) D	0.549**	-	0.345**
Log(attraction) E	-	0.236**	0.328**
Log(attraction) S	0.414**	0.292**	0.523**
Log(attraction) W	0.726**	0.653**	0.933**
Log(attraction) O	0.409**	0.279**	0.581**
μ mode	1.00***	1.00***	1.00***
μ zone = 1.00***			
$\bar{\rho}^2 = 0.577$			

** = p<0.05, ** = p<0.01, *** = active bound*

When the nesting between the alternatives is elaborated, it does not lead to a model with higher predictability. On the contrary, it over-complicates the model. The μ parameter for each nest is 1. Thus, it can be concluded that it does not estimate a CNL but rather a MNL model.

6.1.3. Conclusion

Based on the literature, four different types of TBDCMs were estimated: an MNL, an NL destination, an NL mode and a CNL. However, the NL destination and CNL did not yield a correlation among the alternatives. The NL mode did yield a correlation among the mode alternatives. Because the NL destination and CNL are rejected they will not be used in the remainder of this thesis. However, the MNL and NL mode yield significant results; hence, those models are used in the following analysis in this thesis.

6.2. Model Analysis

In the literature, multiple methods to analyse the choice model were identified. This section analyses the estimated MNL and NL models. These models are estimated based on the choice set from OViN 2015 - 2017. These models are validated by applying them to a different choice set. First, the choice set that was used to estimate the models is also used to perform the in-sample validation. The used choice set is split up in multiple slices. Each of these slices are used for the in-sample validation. The results indicate the consistency of the choice model.

The out-of-sample validation requires an independent choice set, a choice set different than the choice set which was used for estimating the choice model. The MNL and the NL are applied to this choice set to determine the reproducibility of the choice models. the TBDFM are also subjected to this validation method to compare the results with the TBDCMs.

Lastly, the sensitivity of the parameters is determined of the MNL, NL and the TBDFM to analyse how the probability of alternatives changes. The results of estimating the MNL and NL are a list of the attributes. A reference is needed to be able to compare the effect of the attributes. The travel time is the only attribute present in all the alternatives; hence it is used as a reference for the other attributes. There are also two different types of attributes: continuous and categorical attributes. Each type of attribute requires a different approach to determine the sensitivity.

6.2.1. In-Sample Validation

For the in-sample validation, the consistency of the TBDCMs is determined. The choice set which was used to estimate the TBDCMs is split up into five, evenly large slices. For every slice, the DCM is re-estimated without the slice. The estimated DCM is used to predict the choices of the slice, which results in the log-likelihood. The variation in the log-likelihood between the slices indicates the consistency of the DCM within the original choice set, which indicates the in-sample validation. The results are presented in Table 6.5.

Table 6.5: In-sample validation.

	MNL		NL Mode	
slice	Log likelihood	Δ	Log likelihood	Δ
1	-1,246.42	-5.51%	-1,246.24	-5.54%
2	-1,145.87	3.00%	-1,146.60	2.90%
3	-1,165.30	1.35%	-1,165.36	1.31%
4	-1,163.63	1.50%	-1,163.74	1.49%
5	-1,185.28	-0.33%	-1,182.27	-0.12%
Average	-1,181.30		-1,180.84	

Based on the results, it can be concluded that the MNL and NL mode perform similarly. The choices of slice 1 are difficult to predict. The log-likelihood of slice 1 is significantly lower than the other slices. This higher log-likelihood is predicted for both the MNL and the NL mode. The log-likelihood for each of the slices is almost identical for the MNL and the NL mode.

Similar results on the in-sample validation for the MNL and NL mode are expected as the results for estimating the MNL and NL mode were also very similar. Based on the in-sample validation, it can be concluded that the MNL and NL mode are indeed very similar.

6.2.2. Out-of-Sample Validation

For the out-of-sample validation, OViN data from 2018 is taken as an independent data set which contains 2,996 observations. This data set was not used for estimating the TBDCM; hence it is "out of sample". Based on Chapter 3, it was determined that for the non-chosen destinations, 20 random zones were sampled. Thus, for each choice 60 non-chosen alternatives are added. These 60 alternatives are the same destination-mode combinations which were randomly sampled for estimating the TBDCM.

For all the alternatives, the probabilities are calculated. These probabilities are ordered from high to low, and the order of the chosen alternative is taken as a result of the model. For each choice, the probabilities of each alternative are calculated, resulting in Figure 6.2.

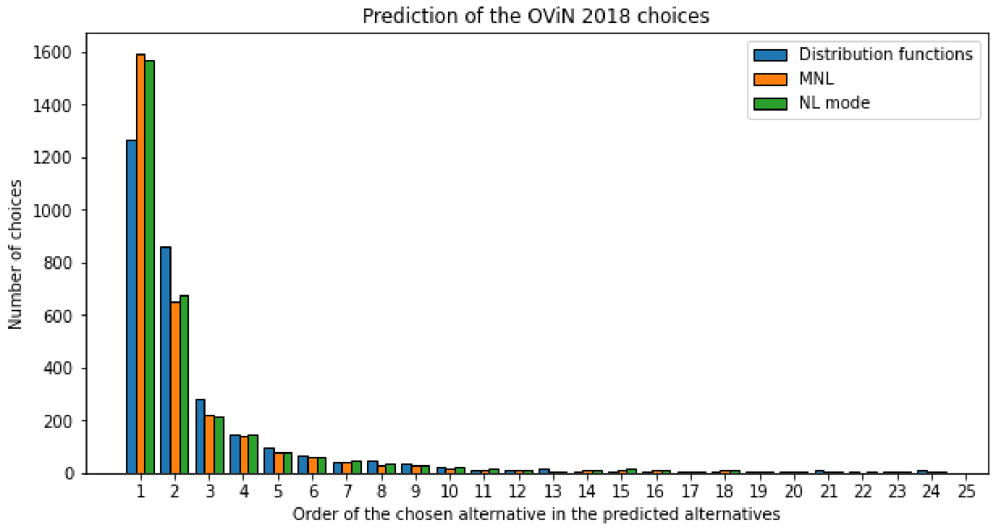

Figure 6.2: Prediction of the OViN 2018 choices

From the Figure 6.2, it can be concluded that the MNL and the NL perform significantly better than the TBDFM. The TBDCMs succeed in assigning the chosen alternative the highest probability in 1,600 of the 3,000 observations. the TBDFM predict the correct alternative only 1,250 times. When the MNL is compared to the NL, the MNL scores marginally better. There are a few choices where the MNL predicts it correctly, and the NL predicts it the second highest.

The equal performance of the MNL and NL was expected because the $\overline{\rho}^2$ of both models is the same. The equal $\overline{\rho}^2$ indicated that the models performed equally for the 2015-2017 observations. Hence an equal performance is expected for the 2018 observations.

The choice can be split into the destination and mode choices to dive deeper into the performance difference between the MNL and NL. The aim is to analyse if the models perform differently within the mode choice or destination choice, or both. These results are presented in Figure 6.3 and Figure 6.4.

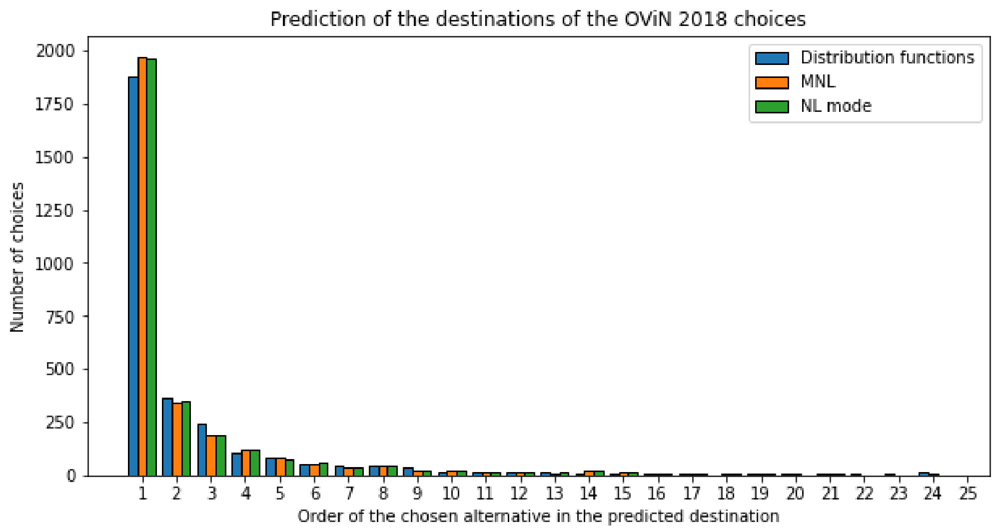

Figure 6.3: Prediction of the OViN 2018 destinations.

Base on Figure 6.3, it can be concluded that the TBDCMs perform slightly better than the TBDFM. Both predict the correct destination for the majority of the observations, but the MNL and NL predict the destinations slightly better. The decay is similarly as all choice models have roughly the same number of observations for the second, third, forth order and so forth.

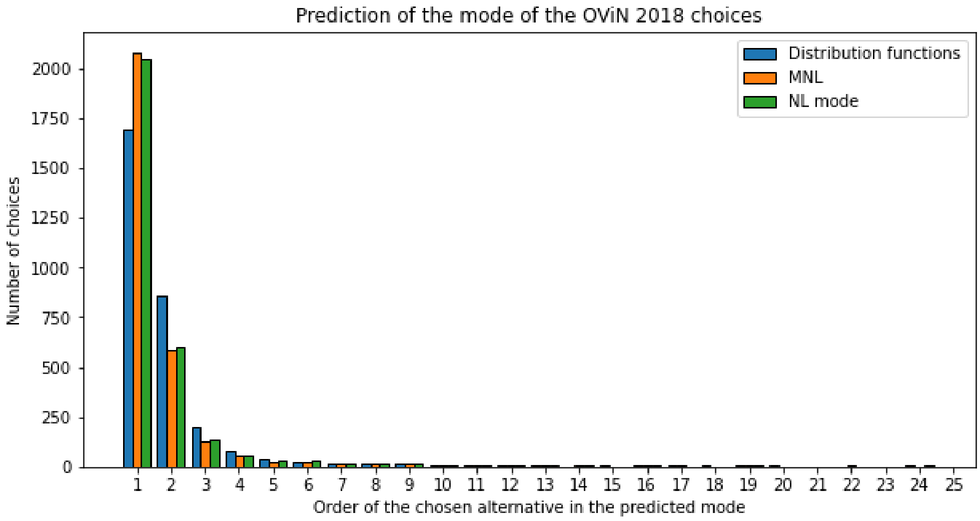

Figure 6.4: Prediction of the OViN 2018 modes.

For predicting the correct mode, the TBDCMs perform significantly better than the TBDFM. They correctly predict the mode around 2,100 times, while the TBDFM predict the correct mode 1,750 times. The MNL also scores better than the NL. This difference is larger than the difference for the destination choice. Thus, the MNL performs better mainly due to the mode choice.

The modal split is derived to investigate the differences between the mode choice. The modal split is taken from the modes of the predicted alternatives and compared with the observations. The results are presented in Figure 6.5.

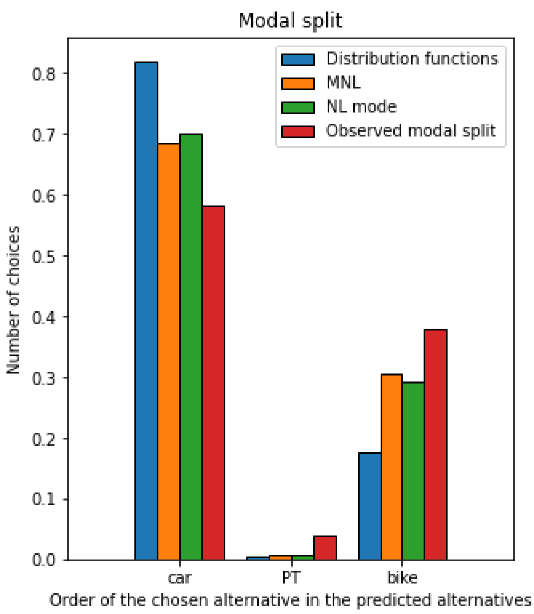

Figure 6.5: Prediction of the OViN 2018 modal split.

The results show that the TBDFM predict a significant amount of car trips, which does not correspond with the modal split of the observed data. The MNL and NL predict fewer car trips, but there is still a strong preference for the car. All models predict almost none PT trips, apparently there is in most cases a non-PT alternative which is more attractive. The observed data shows significantly more PT trips, which suggests that individuals do not always choose the most attractive alternative or the car is not always accessible for them. Between the MNL and NL models, the MNL slightly outperforms the NL. It predicts fewer car trips and more bike trips but is not comparable with the observed modal split.

From the out-of-sample validation, it can be concluded that the MNL and NL perform better than the TBDFM mainly due to predicting the mode choice better. When comparing the MNL and NL, the MNL has a slight advantage.

6.2.3. Sensitivity Analysis

The result of estimating a DCM is a list of attributes with values representing the utility of those attributes. Some of the attributes are related to the trip's distance, and some are related to the characteristics of the individual, origin or destination. Utility influences the probability of the alternatives. It is important to analyse how those attributes influence the alternatives as they influence the utility differently. It demonstrates the attribute' influence on the probability distribution among the alternatives.

The attributes each influence the probability differently. Continuous attributes such as the travel time and attraction differ for almost every OD pair. Hence these different occurrences have to be compared. Categorical attributes cause a shift in the probability of the continuous attributes. The travel time is the characteristic that varies the most per trip. Hence it will be analysed first, and the influence of the other attributes is compared with the travel time attribute as the travel time is the only LoS characteristic. An important note is the occurrence of travel times as presented in Figure 6.1.

Travel Time

One of the most influential attributes is the travel time. Every destination has a different travel time per mode. Hence the influence of the travel time is different for each OD pair, while all the characteristic attributes are fixed regardless of the travel time of the trip.

The result of varying the travel time is calculated to analyse the sensitivity of travel time. A trip with a fixed travel time will be compared to a trip with a varying travel time. The probability that the varying trip will be chosen will change because the travel time will increase or decrease. The extent of the increase or decrease will determine the sensitivity of the travel time. A disadvantage of the TBDFM is that a distribution function is an estimate for every persona. If every persona has a separate distribution function, the effect of an increased travel time cannot be separated from the effects of the other attributes.

Moreover, the TBDFM are discrete, which means it is a list of coordinates. The values in between are interpolated based on a linear interpolation. To analyse the sensitivity for the TBDFM, all the functions will be analysed together to determine an overall sensitivity. This comparison is performed with each mode as each has a separate attribute estimated for the travel time.

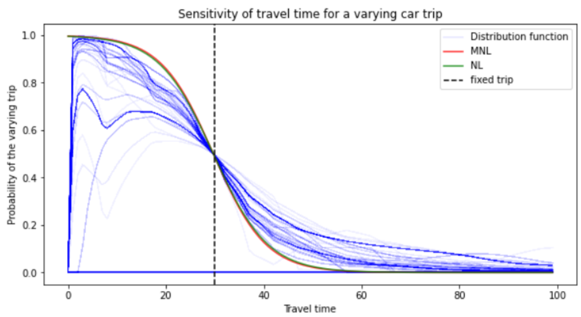

Figure 6.6: Travel time sensitivity for the car

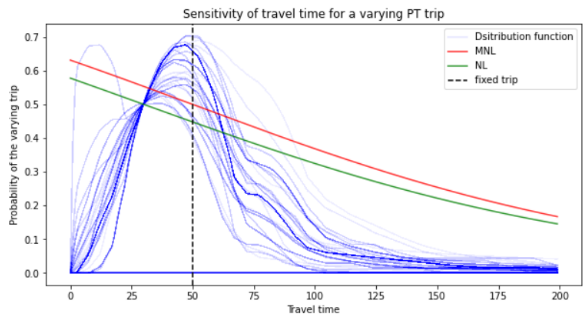

Figure 6.7: Travel time sensitivity for PT

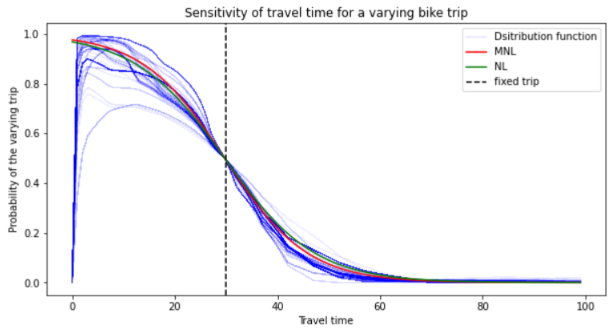

Figure 6.8: Travel time sensitivity for the bike

Based on Figure 6.6, Figure 6.7 and Figure 6.8, it can be concluded that the sensitivity varies significantly per mode. The MNL and NL are more sensitive to the travel time for the car than the TBDFM are. A change in travel time will greatly affect the probability of the trip. A disadvantage of using a DCM is the limitation of one variable to describe the travel time. The probability always increases for a shorter travel time, but the car is rarely used for trips with a very short travel time. the TBDFM are discrete. Hence the probability for very short trips can be reduced to distribute the trips for the car on longer travel time trips.

When comparing the sensitivity for the bike, the MNL, NL and the TBDFM are comparable. The probability decreases approximately at the same rate when the travel time increases.

For PT, there is quite a difference between the TBDCMs and the TBDFM. The main difference is the shape. Because the TBDFM are discrete, the probability of short travel times is reduced significantly. However, as presented in Figure 6.1, no short trips exist for PT. Hence the reduced probability for short PT trips is unnecessary. the TBDFM are nonetheless much more sensitive to the travel time than the MNL and NL. The probability of long-distance trips decreases significantly compared to the probability of the MNL and NL.

To further analyse the influence of the travel time attributes of the TBDCM, a comparison is made between two trips with each a different mode. The travel time for each trip is increased step-wise. The probability is calculated for both alternatives to determine the sensitivity of the travel time. The results are visualised in a heat-map which indicates the probability distribution among the two trips. This comparison is made for each mode combination: car-bike, car-PT and PT-bike. The probability for each alternative is calculated for each combination of travel times.

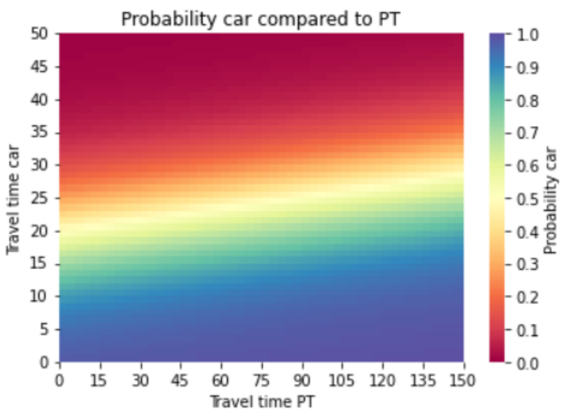

Figure 6.9: Travel time influence of car versus pt.

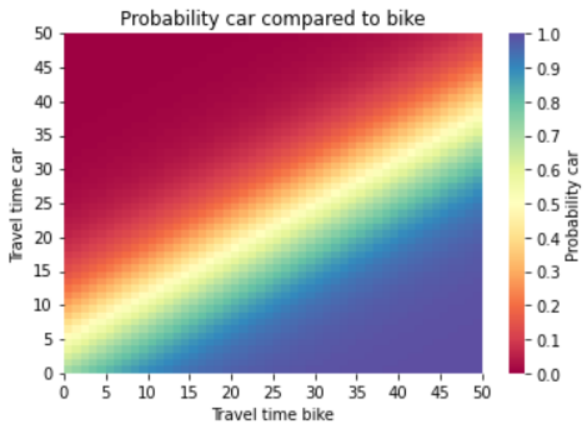

Figure 6.10: Travel time influence of car versus bike.

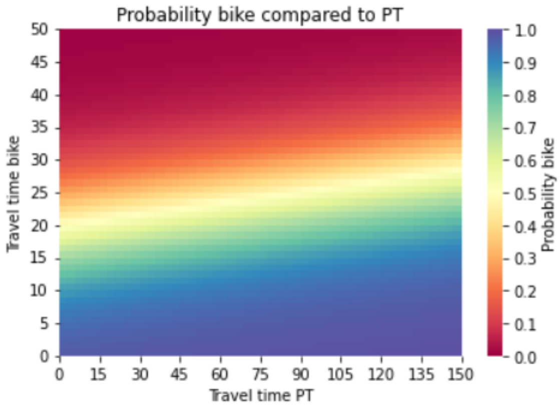

Figure 6.11: Travel time influence of pt versus bike.

As mentioned before, the results of Figure 6.9, Figure 6.10 and Figure 6.11 are the probabilities of the alternative with mode x compared to the alternative with mode y.

Based on Figure 6.9, Figure 6.10 and Figure 6.11, it can be concluded that the travel time of PT has barely any influence on the probability. If the travel time significantly increases for PT, there is only a small increase in travel time needed for the different mode to maintain a 50%-50% distribution among the alternatives. The travel time attribute is so small compared to the ASC that it hardly changes the probability of the alternative. The comparison between the car and the bike seems distorted. The probability of choosing the car is very low for a trip with a travel time of 50 for both modes. But as mentioned before, the car is significantly faster, so this combination of travel times is extremely rare.

To be able to effectively evaluate policies, the elasticities of the travel times are analysed. The travel times of each mode is increased by 1%. The shift in the modal split is measured to determine the elasticity of the travel time attribute.

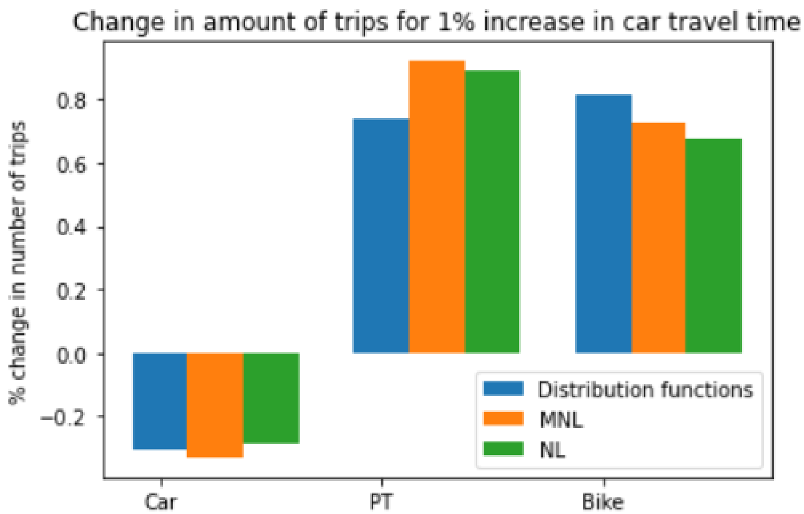

Figure 6.12: Elasticity car.

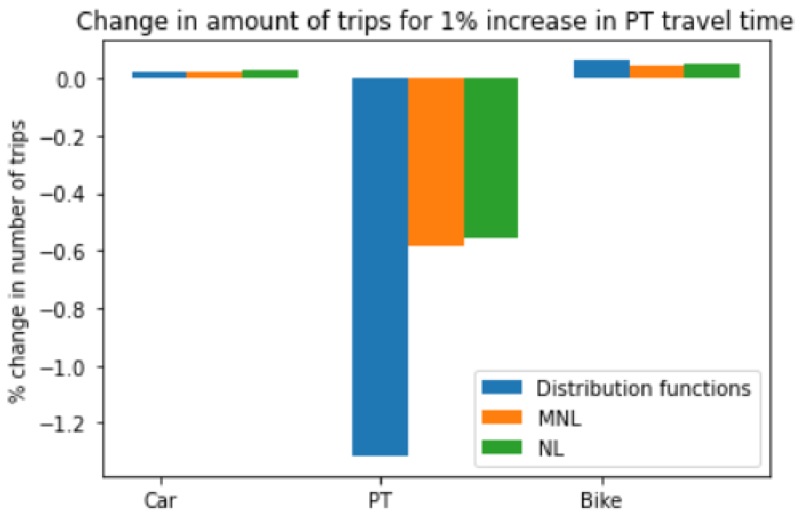

Figure 6.13: Elasticity PT.

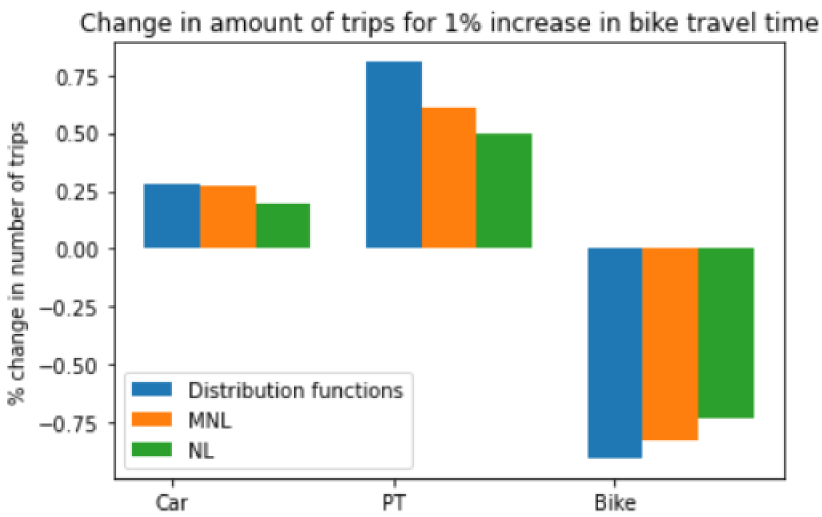

Figure 6.14: Elasticity bike.

Wardman, 2012 performed a meta-analysis on time elasticities of travel demand. It was concluded that the average travel time elasticity for car is -0.30, $\sigma = 0.06$, for bus is -0.63, $\sigma = 0.16$, and for train is -0.69, $\sigma = 0.03$.

Based on Figure 6.12, Figure 6.13 and **??**, it can be concluded that there are differences between the travel time elasticities of the TBDCMs and the TBDFM.

The car elasticity for an increase in car travel time is similar among the choice models. The car elasticities are also comparable with the literature. But there is a difference for the cross elasticities. The TBDCMs predict a larger cross elasticity for PT while the TBDFM predict a higher cross elasticity for the bike.

The PT elasticity is significantly different for the TBDCMs and the TBDFM. The TBDCMs recorded around -0.53 while the TBDFM recorded -1.31. According to the literature, an average of -0.63 - -0.69 was found. It can be concluded that the TBDFM is too sensitive for a change in PT travel time.

For the bike elasticity there is a small difference between the TBDCMs and the TBDFM. All choice models assign the majority of the trips to PT and only a small portion to the car.

Categorical Attributes

The attributes for age, employment and urbanisation are categorical. Based on the characteristics of the individual or zone, a fixed amount of utility is added or subtracted from the corresponding alternative. To analyse the sensitivity for each attribute, the amount of utility is compared to the travel time. How much longer or shorter can the trip be to maintain the same amount of utility?

Table 6.6: sensitivity of categorical attributes.

	MNL			NL		
Attributes	**Car**	**PT**	**Bike**	**Car**	**PT**	**Bike**
ASC	+6.11	-264.49	ref	+5.84	-289.42	ref
Age 0-17	-	-	+30.08	-	-	+32.12
Age 18-34	-	-	-	-	-	-
Age 35-65	-	-55.89	-	-	-57.69	-
Age 65+	+2.12	-	-	+2.20	-	-
Work full-time	+3.29	-	-	+3.25	-	-
Work part-time	+2.12	-	-	+1.48	-	-
Work nan	-4.18	-	-4.95	-4.52	-	-5.42
Urban destination 1	-2.97	+74.95	-	-2.97	+77.50	-
Urban destination 2	-	-	-	-	-	-
Urban destination 3	-	-	+2.41	-	-	+2.62
Urban destination 4	-	-110.28	-	-	-113.46	-
Urban destination 5	-2.11	-245.79	-	-2.12	-252.88	-
Urban origin 1	-	+111.21	-	-	+109.61	-
Urban origin 2	-	-	-	-	-	-
Urban origin 3	-	-	-	-	-	-
Urban origin 4	-	-	+2.91	-	-	+2.90
Urban origin 5	+1.36	-93.46	+6.11	+2.12	-93.27	+6.24

The following conclusions can be drawn based on Table 6.6. Firstly, acPT is very sensitive to a few attributes. Especially the urbanisation of the zones has a big impact on the mode. The change in travel time needed to compensate for the attributes is significant. Secondly, the sensitivity for the car seems low, but as mentioned before, the travel times for the car are lower than those of PT and bike. Third, the car is the most sensitive for the employment of individuals. People who are employed prefer the car, while people who are unemployed avoid the car. Lastly, the bike is dominant for the 0-17 age group. This makes sense, as children are not permitted to drive a car. Hence the bike is the only alternative with a certain flexibility.

Attraction

The attraction of a zone varies greatly. High-urbanisation zones often have more facilities than rural zones. Hence the attraction attribute is estimated as a continuous variable. Furthermore, the attraction attribute is estimated on a logarithmic scale for the TBDCMs. To analyse the sensitivity of the attraction, it is compared with the travel time.

Because both variables are continuous, the comparison can only be made for a specific mode and persona. For each travel time and attraction, the partial component of the probability function is calculated to compare the effect of the TBDCMs and the TBDFM. For the TBDCMs, this is $e^{utility(traveltime,attraction)}$, and for the TBDFM, it is $distribution_function(traveltime) * attraction$. An increase in the partial component will have the same effect on the probability for the TBDFM as the TBDCMs. Hence the results can be interpreted in the same way.

Based on the figures below, the sensitivity of the attraction versus the travel time is quite different per mode and trip purpose. As mentioned before, the MNL and NL parameters are quite similar; hence the sensitivity of the attraction is also similar. The biggest difference is between the TBDCMs and the TBDFM.

Work

Based on Figure A.1, Figure A.2 and Figure A.3, it can be concluded that for work trips, the sensitivity is different per mode for the TBDCMs and the TBDFM. The TBDCMs predict a low sensitivity for the work attraction compared to the travel time. The travel time seems to be more important in choosing the destination for a work trip than the attraction of the destination. the TBDFM predict it and vice versa. the TBDFM predict a significant increase in probability if the attraction increases. Consequently, individuals are prepared to travel much further for a larger zone.

Based on Figure A.4, Figure A.5 and Figure A.6, it can be concluded that for PT, the TBDCMs predict a significant sensitivity for the attraction. An increase in attraction leads to a significant increase in the probability of that alternative. The travel time has a similar sensitivity but not as much as the attraction. the TBDFM predict a small sensitivity of the attraction. The probability of the destination increases when the attraction is larger but decreases as the destination is further away.

Based on Figure A.7, Figure A.8 and Figure A.9, it can be concluded that the TBDCMs predict a medium sensitivity when compared to the sensitivity of the car and PT. Especially when the attraction is small, there is a large sensitivity, but when the attraction increases, the sensitivity decreases. However, a certain sensitivity remains. the TBDFM predict a similar sensitivity between the travel time and attraction

Business

Based on Figure A.10, Figure A.11 and Figure A.12, it can be concluded that for business trips, the car is barely sensitive to a change in attraction for the TBDCMs. A large increase in the attraction translates to a slight increase in travel time. the TBDFM are, however, very sensitive to the attraction. An increase in attraction significantly increases the extra travel time individuals are willing to travel.

The TBDCMs have no estimated value for the business attraction for PT; hence there is no sensitivity. Based on Figure A.13, it can be concluded that the TBDFM are not quite as sensitive for PT as for the car. There is a slight increase in travel time for a low attraction, but at a certain value, the travel time becomes more dominant.

Based on Figure A.14, Figure A.15 and Figure A.16, it can be concluded that for the bike, the sensitivity is almost equivalent for the TBDCMs and the TBDFM. The extra travel time individuals are willing to travel increases at the same rate. The only difference is that for the TBDCMs, the travel time has a bigger impact for zones with the same attraction.

Education

The TBDCMs have no estimated value for the attraction of education for the car. Hence there is no sensitivity. Based on Figure A.17, it can be concluded that the TBDFM predict a sensitive attraction for smaller zones. For larger zones, the TBDFM predict a rigid sensitivity. A change in travel changes the probability more than a change in attraction.

Based on Figure A.18, Figure A.19 and Figure A.20 it can be concluded that for PT the sensitivity of the attraction is very similar for the TBDCMs as for the TBDFM. The travel time is more important than the attraction for all models. Increasing the attraction has less effect on the probability than decreasing the travel time.

Based on Figure A.21, Figure A.22 and Figure A.23, it can be concluded that for the bike the TBDCMs predict a rigid sensitivity for the attraction. Increasing the attraction has a limited effect on choosing that destination. Changing the travel time has a large effect. the TBDFM predict a more sensitive attraction, but the travel time remains dominant.

Shopping

Based on Figure A.24, Figure A.25 and Figure A.26, it can be concluded that for shopping trips the sensitivity of the attraction for the TBDCMs is very low. There is a slight change in travel time for a large increase in attraction. the TBDFM are more sensitive to an increase in shopping attraction, but this is still very minimal.

Based on Figure A.27, Figure A.28 and Figure A.29, it can be concluded that the TBDCMs estimate a high sensitivity for PT when the attraction is low. But as the attraction increases, the extra travel time individuals are willing to travel decreases. The same phenomenon occurs for the TBDFM. However, for the TBDFM, the sensitivity decreases as the attraction grows.

Based on Figure A.30, Figure A.31 and Figure A.32, it can be concluded that the TBDCMs predict a low sensitivity for the bike. A low travel time is more important than a large zone. The extra travel time individuals are willing to travel for a larger zone is small. The same holds for the TBDFM. Only for short distances does the attraction have an impact on the choice.

Bring/get

Based on Figure A.33, Figure A.34 and Figure A.35, it can be concluded that for bring/get trips the sensitivity for the car of the TBDCMs is minimal. An increase in attraction barely increases the probability of the alternative. the TBDFM predict a more sensitive relation, but travel time remains dominant.

For both the TBDCMs and the TBDFM, bring/get trips are not possible with PT. Hence there is no sensitivity.

Based on Figure A.36, Figure A.37 and Figure A.38, it can be concluded that the TBDCMs estimate a rigid sensitivity for the bike. The sensitivity for the TBDFM is quite similar. The extra time individuals are willing to travel is very limited. The travel time has a significantly bigger impact on the probability of the alternative.

Other

Based on Figure A.39, Figure A.40 and Figure A.41, it can be concluded that the TBDCMs estimate a very rigid sensitivity for the attraction of other trips. The increase in probability of the alternatives barely increases; even if the attraction increases tenfold, the extra travel time individuals are willing to travel is no more than 5 minutes. the TBDFM predict a more sensitive attraction. There is a significant increase in the travel time for zones with a bigger attraction. Travel time remains the dominant factor, but attraction has a significant influence.

Based on Figure A.42, Figure A.43 and Figure A.44, it can be concluded that for PT, the TBDCMs estimate a sensitive attraction. There is a significant increase in the extra time individuals are willing to travel for a larger zone. the TBDFM predict a similar effect. The probability increases significantly if the attraction increases.

Based on Figure A.45, Figure A.46 and Figure A.47, it can be concluded that the sensitivity for the bike of the TBDCMs is comparable to the sensitivity of the TBDFM. For a low attraction, the extra time individuals are willing to travel is significant. However, it decreases as the attraction of the zone increases.

6.2.4. Conclusion

Based on the model analysis, it can be concluded that the TBDCMs perform better in predicting the chosen alternative than the TBDFM. When the TBDCMs and the TBDFM predict the correct choices for the choice set of OViN 2018, the TBDCM assign more often the chosen alternative the highest probability among the alternatives. the TBDFM predict less often the correct mode. the TBDFM assign too often a car-alternative the highest probability when compared to the observed modal split.

It can also be concluded that the sensitivity is significantly different for the TBDFM than for the TBDCMs. The TBDCMs are more sensitive for a change in the car travel time while the TBDFM are more sensitive to a change in PT travel time. Also for the attraction values is the sensitivity different. Per mode-attraction combination, is the sensitivity of the TBDCMs more often significantly different than the TBDFM' sensitivity, then not.

6.3. Performance Analysis

In Chapter 2 different KPIs were identified that indicate the performance of a choice model. In this section, these KPIs of the choice models will be analysed. First, the models will be subject to a choice containing all the

Limburg zones as a possible alternative. There are 544 zones in Limburg, which is a significant increase from the 21 used to estimate the TBDCMs.

The TBDCM and the TBDFM are implemented in VML. A complete run with VML is executed to determine the performance on the ground truth observations. VML will determine the link flow on the specific links which have a ground truth observation.

6.3.1. Complete Destination set

The previous section validated the models based on the sampled destinations. In this section, the aim is to analyse how the models distribute the probability among alternatives. Therefore, a larger set of alternatives is used to apply a choice set with more similar alternatives. This method will produce a more accurate trip length distribution and mode choice for an average trip.

To achieve a representative trip length frequency, each zone in VML is used as an origin for the trip. The destination alternatives consist of every VML zone in Limburg. A choice set with all the zones as an alternative ensures the availability of zones which are close by are zones which are further away. The population varies between the zones. To compensate for high-density zones, the weight is adjusted based on the urbanisation degree of the zone. The trips originating from high-density zones will weigh heavier than those originating from rural areas.

There is also a variation in personas and trip purposes. To create a representative set of personas and trip purposes, they are drawn from the 2018 OViN data set, which ensures a representative selection.

The results are the probabilities for each trip. Combined with the corresponding trip length, these probabilities result in the trip length frequency. The trip length frequency will indicate the deterrence of the travel time. The trip length frequency is split up per mode. Furthermore, the modal split is determined to compare it with the observed modal split. The results are presented below.

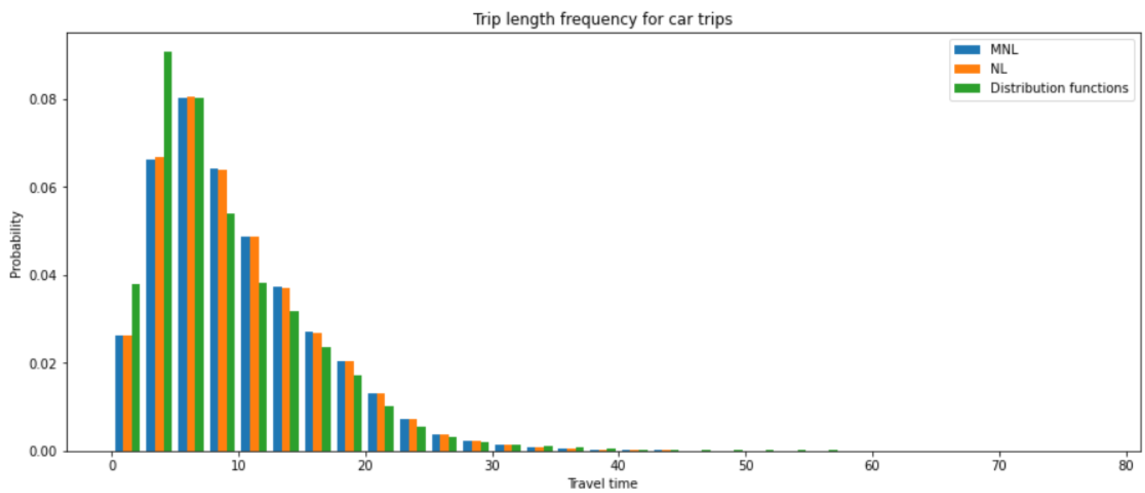

Figure 6.15: Trip length frequency car

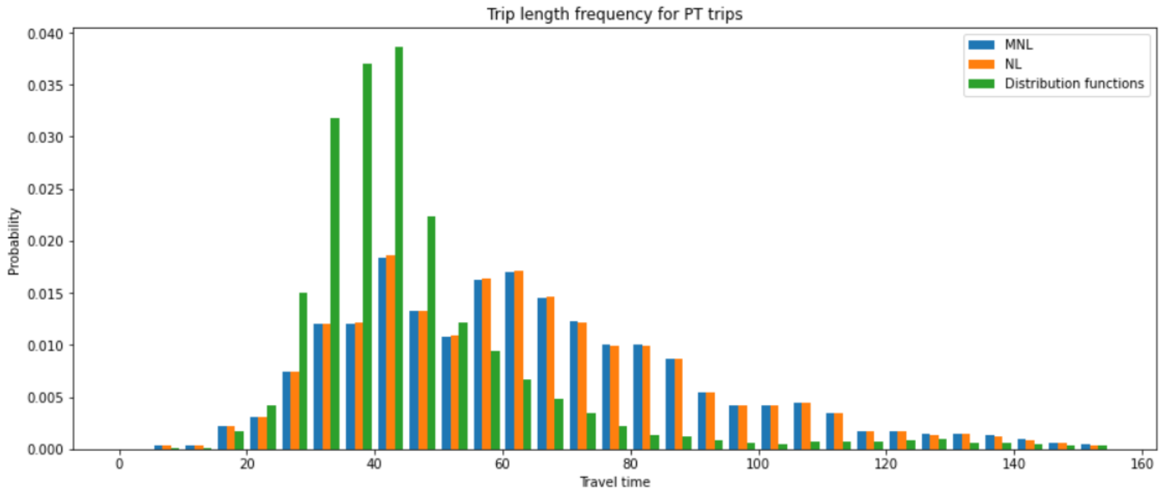

Figure 6.16: Trip length frequency PT

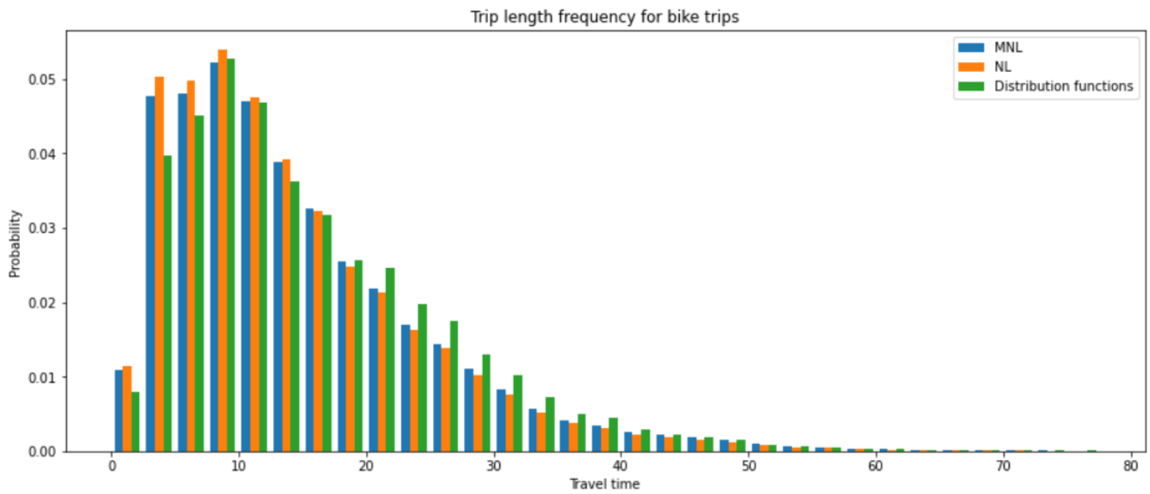

Figure 6.17: Trip length frequency bike

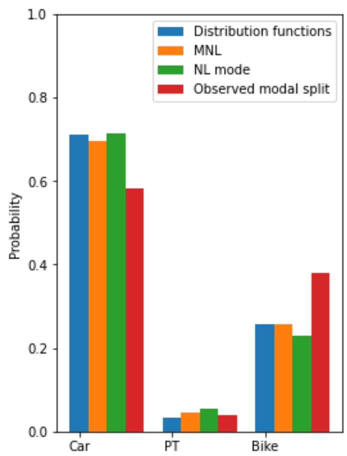

Figure 6.18: Predicted modal split

Based on Figure 6.15, Figure 6.16 and Figure 6.17, multiple conclusions can be drawn. There is a significant difference between the TBDFM and the MNL and NL model. For car trips, the TBDFM tend to predict more short trips than the TBDCMs. For PT, the TBDFM' tendency for short trips is more extreme. the TBDFM predict most of the trips within a small travel time range. The TBDCMs distribute the PT trips among a larger travel time range. For the bike, the trip length frequency is quite similar.

The difference between the MNL and NL is visible. The different parameters influence the trip length frequency. The nest parameter for the bike was the largest, hence the difference between the MNL and NL is the most visible. The NL tend to predict short travel time trips because of the correlation among the alternatives. For the car, the correlation was small; hence the difference between the MNL and NL is also small.

There is also a shift in the modal split. The modal shift for the out-of-sample validation was shown in Figure 6.5. However, when the probabilities of each alternative are taken into account instead of the alternative with the highest probability, the modal shift is as presented in Figure 6.18.The most significant change is the similarity of PT with the observed modal split. PT is not often correctly predicted as an alternative with the highest probability, but in the choice set, the total probability is similar to the observed modal split. For the car, there is still an overestimation. the TBDFM and the TBDCMs all assign significantly more probability to the car when compared to the observed modal split. For the bike, it is the opposite. The models assign significantly less probability to the bike, even less than the probability found in Figure 6.5.

6.3.2. Case Study

For the case study only the MNL is implemented in VML. A complete run of VML is executed, consisting of the trip generation, trip distribution, modal split and assignment. A run is completed with the MNL and a run with the TBDFM as module to estimate the mode & destination choice. Three KPIs were identified to determine the performance of the choice models in a large-scale transport model: modal split, trip length frequency and ground-truth observations.

Modal Split

As mentioned in Chapter 4, VML has a study area, influence area and peripheral area. The traffic within and between the zones in the influence and peripheral are is inaccurate because the zones are quite large. The intrazonal travel time is significantly smaller than the travel time to the neighbouring zone; hence the traffic is roughly estimated. To focus only on the detailed estimated traffic, Only the traffic within the study area is considered to focus only on the traffic that is estimated accurately.

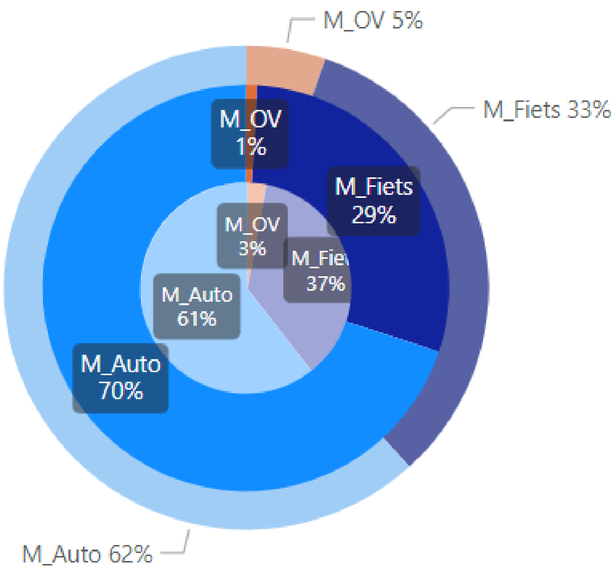

Figure 6.19: Modal split with VML.
Outer ring: OViN
Middle ring: TBDCM
Inner ring: the TBDFM

Based on Figure 6.19, multiple aspects can be concluded. The biggest offset is that the MNL estimates very few PT trips. The MNL models only 1% while the TBDFM estimate 3% but OViN observed 5%. 1% instead of 5% is a quite significant decrease. The MNL also estimates very few bike trips. The MNL estimates 29% while the TBDFM estimate 37% and OViN observes 33%.

These results are not what was expected. Especially the few number of trips for PT is significantly different than then probability for the trips in Figure 6.18. There is however one important aspect that is different in VML than for the complete destination set test. Tours versus trips. The complete destination set calculated the probability for single trips where VML calculates the probability for complete tours. The difference in utility doubles if also the reversed trip is taken into account, assuming trip characteristics between the outbound and inbound trip. This bigger difference in utility is also magnified by the exponent, which calculates the probability of the alternative. As a result, the probability for PT tours drops significantly; hence the modal split for PT is significantly lower than the observed modal split.

Trip Length Distribution

The distance travelled is important to analyse. Not only will it indicate the deterrence of the travel time, it will also influence the observations of the link flows later on. After the assignment the link flows can be analysed. If more distance is travelled by the trips, a single trip will pass more locations which have ground truth observations.

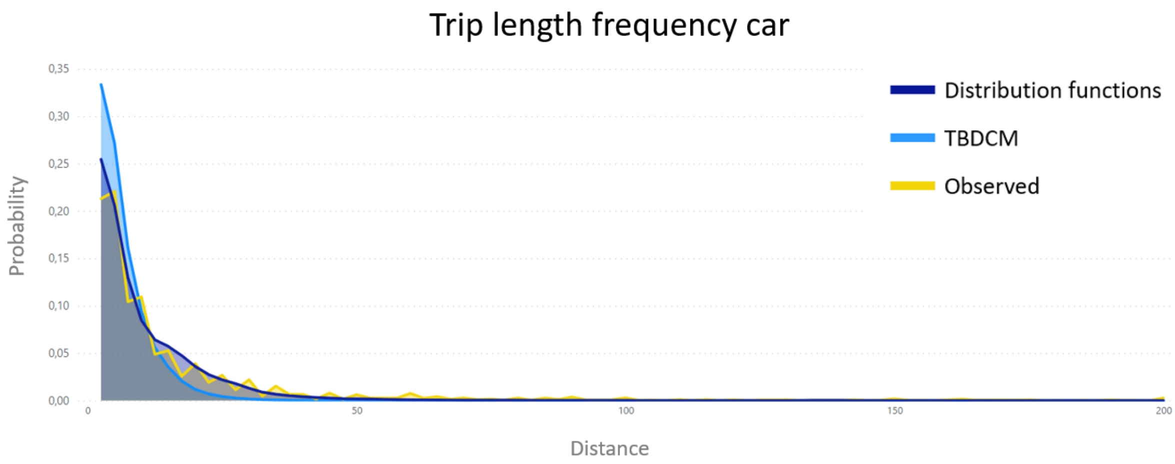

Figure 6.20: Trip length distribution for the car in VML.

Based on Figure 6.20 that there is a subtle difference between the TBDCM and the TBDFM. The TBDCM estimates a higher occurrence for shorter trips while the TBDFM estimate a higher probability for longer trips. the TBDFM are more similar to OViN compared to the TBDCM.

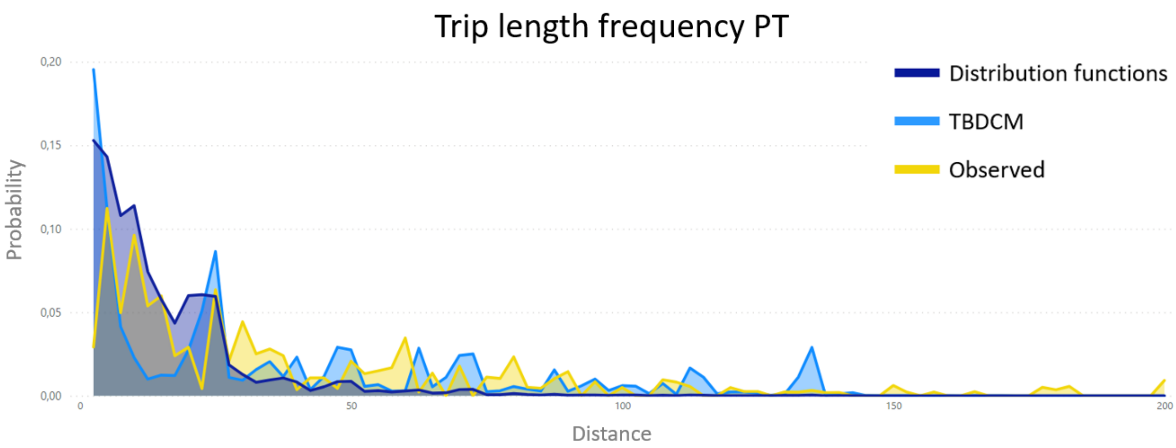

Figure 6.21: Trip length distribution for PT in VML.

Based on Figure 6.21 it can be concluded that there are significant differences between the TBDCM, the TBDFM and OViN. The trip length distribution of PT is difficult to analyse because the distance travelled varies a lot between the alternatives. There a big differences within PT as the train, metro, tram and bus each have different characteristics. Hence the trip length distribution is expected to be various on the distance.

The TBDCM and OViN vary a lot between certain distances. The TBDCM varies the most which is mainly because of the sensitivity of the attributes as presented in Table 6.6. As a result, the distribution is less among the alternatives; hence the trip length distribution has a strong variation. the TBDFM estimate a more constant trip length distribution. the TBDFM also estimate shorter distance trips. Trip with the train often have a high trip distance, which are not reflected by the TBDFM. The TBDCM is more similar to OViN than the TBDFM as the TBDCM contain more long distance trips.

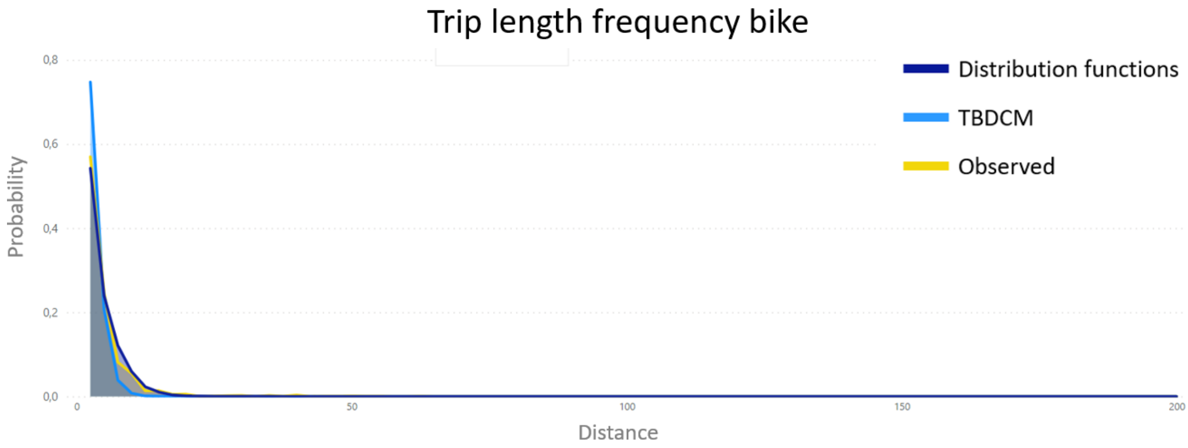

Figure 6.22: Trip length distribution for the bike in VML.

Based on Figure 6.22 it can be concluded that again there is only a subtle difference between the TBDCM and the TBDFM. The TBDCM estimate a higher occurrence for shorter distance trips while the TBDFM estimate a higher occurrence for longer distance trips. the TBDFM are quite similar to the observation of OViN.

Ground Truth Observations

The assignment can be executed if the OD-matrix is modelled. Traffic will be assigned to the network based on certain algorithms. For the car assignment the shortest path is used including congestion functions. Congestion functions are added to links to increase the travel time if a link flow exceeds the capacity of that link. For the PT assignment a multi-routing algorithm is used to determine the travel time based on the different routes one can taken by using PT. For the bike assignment the average of three assignments is used. First, an assignment is executed based on the travel time. Second, an assignment is based on the distance and third an assignment is based on the combination of travel time and distance. The average of those assignments is the assignment of the bike.

As mentioned before, RHDHV has ground truth data on numerous links. RHDHV has a total 11000 observations of different links of different time periods. For each observation the t-value is calculated which determines if the estimated volume of VML is within the boundaries of the ground-truth observation. The results of all these observations will indicate if the TBDCM of the TBDFM approximate reality better. To determine if the modelled flow, X_b, is comparable with the ground truth, X_w, a t-test is done, which is depicted in Equation 6.2.

$$T = ln\left(\frac{(X_b - X_w)^2}{X_w}\right) \tag{6.2}$$

The modelled flow is *good* if $T < 3.5$, *acceptable* if $3.5 < T < 4.5$ and *bad* if $T > 4.5$.

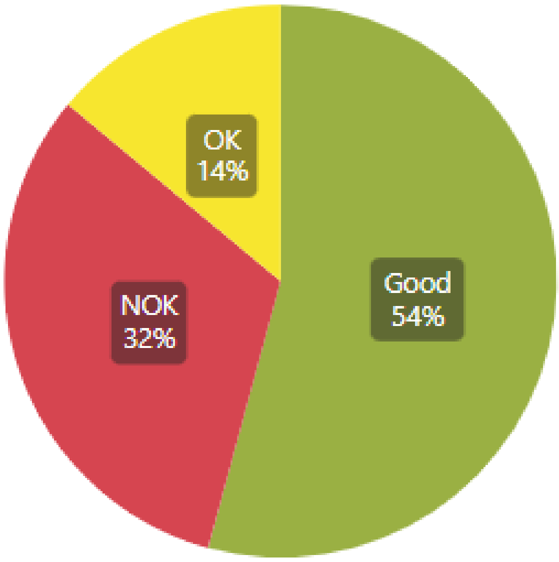

Figure 6.23: Car ground truth observations TBDCM

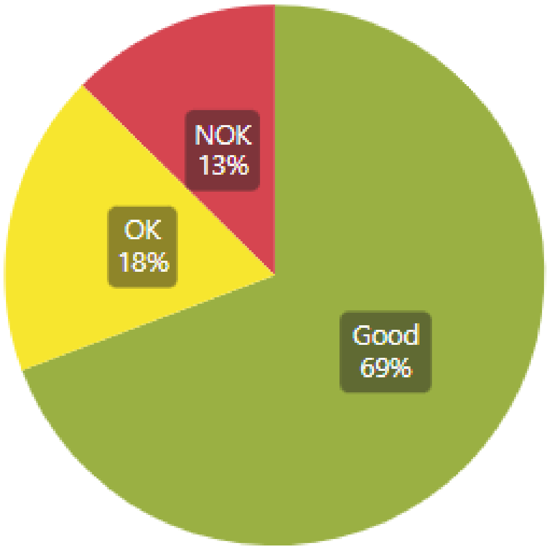

Figure 6.24: Car ground truth observations the TBDFM

Based on Figure 6.23 and Figure 6.24 it can be concluded that the TBDFM perform significantly better for the car. the TBDFM score for morning peak, evening peak and day time around 85% good and around 8% acceptable. Compared to the TBDCM, which scores around 50% - 60% good and 15 % acceptable, the TBDFM perform much better. For the 24h observations the performance is significantly lower than the other time periods.

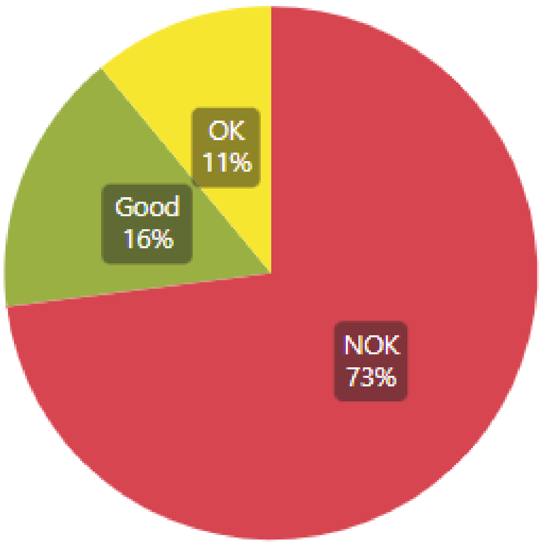

Figure 6.25: PT ground truth observations TBDCM

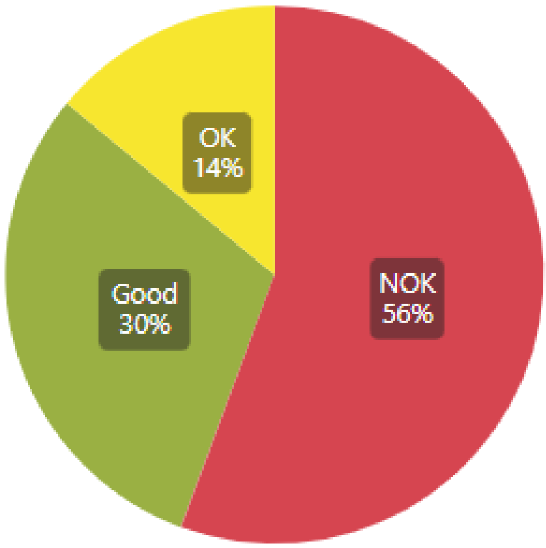

Figure 6.26: PT ground truth observations the TBDFM

Based on Figure 6.25 and Figure 6.26 it can be concluded that both the TBDCM and the TBDFM fail to approximate the ground truth observations. the TBDFM do succeed in correctly predicting more observations but overall it is insufficient.

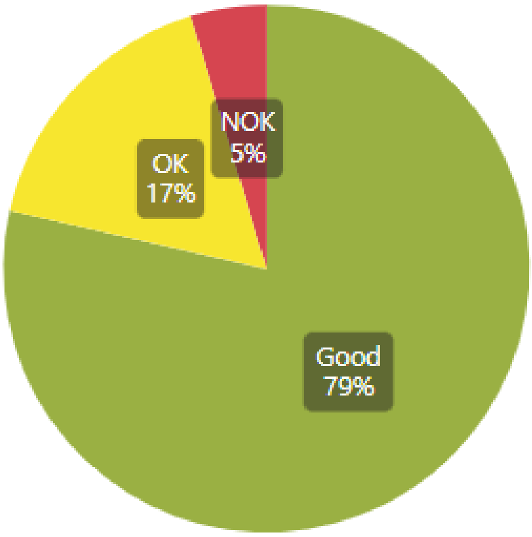

Figure 6.27: Bike ground truth observations TBDCM

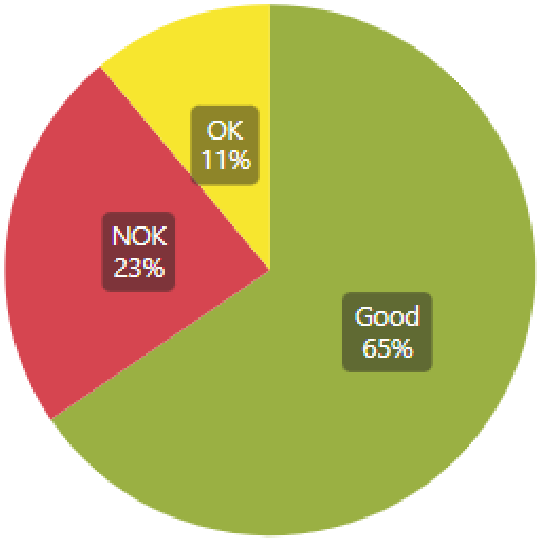

Figure 6.28: Bike ground truth observations the TBDFM

Based on Figure 6.27 and Figure 6.28 it can be concluded that the TBDCM performs significantly better. For the time periods morning peak, evening peak and day time the TBDCM scores roughly 80% good and 17% acceptable. Compared to the 60% good and 14% acceptable of the TBDFM, the results of the TBDCM are an improvement.

6.4. Conclusion

Although different variables are available to estimate the TBDCM, many variables were not estimated significantly. For the MNL, the majority of the possible attributes did not significantly impact the choice. However, the $\overline{\rho}^2$ of the MNL did not suffer from it as it performs well with the estimated attributes. When the different nesting structures were applied to elaborate the MNL, only the NL mode yielded significant results. The NL destination and CNL estimated no correlation among the alternatives. The NL mode estimated a correlation between the car and bike alternatives, but the correlation is minimal. Also, the NL mode did not predict the choices better as the $\overline{\rho}^2$ was not higher than the $\overline{\rho}^2$ of the MNL.

When comparing the performance of the TBDCMs and the TBDFM on the out-of-sample validation, the TBDCMs perform significantly better. The TBDCMs assign the highest probability to the chosen alternative for significantly more observations. The TBDCMs better predict the correct mode. the TBDFM tend to assign the highest probability to the car. The TBDCMs also tend to predict the car as the chosen alternative too often. Assigning the highest probability to the chosen alternative indicates it can predict the choices better. Hence it can be concluded that the TBDCMs perform better on the out-of-sample validation.

However, it is important to notice that the out-of-sample validation used a sampling strategy for sampling the destination alternatives. Hence not every destination in Limburg was presented as an alternative. It simplifies the choice and neglects the probability distribution among the alternatives.

Therefore, a sensitivity analysis is performed on the TBDCMs and the TBDFM. The effect of the parameters was determined to analyse how the utility is affected by a change in the characteristics of the alternative. This sensitivity analysis found that the TBDCMs are generally more rigid. Travel time seems to be the most important factor in determining the utility of the car and bike alternatives. The zonal characteristics have a significantly higher influence for PT when determining the utility of the alternatives.

To analyse how the choice models distribute the probability among similar alternatives, a test is performed with the complete set of destinations. All the zones in VML are considered as an alternative. Based on this test, it can be concluded that the TBDCMs and the TBDFM perform similarly. The trip length frequencies for the car and the bike are comparable. The TBDCMs predict a small number of car trips with a longer travel time, while the TBDFM predict a small number of bike trips with a longer travel time. For PT, the trip length frequency for the TBDCMs is more distributed and over longer travel time trips. the TBDFM estimate shorter travel time trips for PT. The overall modal split is similar for both the TBDCMs and the TBDFM. All choice models estimate more car trips and fewer bike trips than what is observed from OViN.

To analyse how the choice models perform in a large-scale transport model, the MNL and the TBDFM were implemented in VML. A complete run with both choice models was executed. The results are quite different from the complete destination set analysis results. There is a significant change in the trip length frequencies and modal split of the MNL. In VML, the MNL estimates a very low modal split for PT. It also estimates significantly shorter travel times for the bike and the car.

7. Discussion, Conclusion and Recommendations

In this final chapter the results of this thesis are discussed. Limitations of this thesis are identified. Additionally, the sub-questions are answered which leads to the final conclusion. Based on all of the results, the main research question is answered. To finalise this thesis, recommendations for future research are proposed to elaborate on this thesis' topic.

7.1. Discussion

Based on the results presented in Chapter 6, it was concluded that the TBDCM and the TBDFM perform significantly differently. The TBDCM performs better in predicting the correct choice, while the TBDFM distribute the trips better on multiple alternatives. Also, the choice models' performance differs from the ground truth observations. The TBDCM scores higher on the bike observations, while the TBDFM perform better on the car and marginally on PT.

7.1.1. Limitations of this thesis

Multiple sources were required to facilitate this thesis. OViN was used as a data source for trip observations. VML was used as a case study and source for the LoS data. Literature was used as a foundation to construct a TBDCM and to determine how to define the performance of the choice models. All these sources are effectively input for this thesis. Limitations in these inputs trickle down in this thesis and influence the results. The quality of the choice set influences the estimated attributes of the TBDCM. Hence it is important to discuss the limitations.

Limitations of the input data

OViN is a great data source for RP observations. A disadvantage of RP is that the non-chosen alternatives are missing. Finding a source of non-chosen alternatives is difficult. Especially because it requires the LoS data for every OD pair. There are different sources which contain this information but they are not open-source.

Even if a source is found, it is highly unlikely it is compatible with OViN. The travel times of OViN are inaccurate. The inaccuracy is such that the usage of these travel times will significantly decrease the quality of the TBDCM. In this thesis the travel times of VML were used. It limited the area of which the observations could be used to Limburg. This limitations decreased the number of observations and the variation of travel behaviour of the individuals. Thus, the TBDCM is estimated with a lower precision and confidence then it would have if all of the OViN observations could be used.

VML serves as a case study and a data source for the LoS attributes of the chosen and non-chosen alternatives. The TBDCM is limited by VML as it is used as a case study. Only the personal and zonal characteristics present in VML can be used as an attribute in TBDCM, which limits the possible attributes in TBDCM. Characteristics, not present in VML, may describe the mode & destination choice better.

The LoS data of VML is used for the chosen and non-chosen alternatives. The method to model these skim matrices may not be accurate. For PT, an access and egress time is modelled, but for the car, it is not. In reality, parking the car right in front of the origin and destination isn't easy, especially in high urban zones. The absence of this extra travel time will increase the usage of the car for the short distance trips. In reality the car is almost never used for short distance trips.

Limitations of the Tour-Based Discrete Choice Model

The study area of VML is Limburg. As a result, the zones are detailed within Limburg, but outside of Limburg, the zones are coarse. The use of the LoS of VML impacts the area in which the observations can be used. OViN contains the observations for individuals travelling throughout The Netherlands, but due to the usage of VML, only observations within the area of Limburg can be used. This reduces not only the number of observations but also the characteristics of the trip changes.

Firstly, long-distance trips to a different province are absent in the reduced observation set. To compose a choice set between OViN and VML, only the area of Limburg could be used due to the difference in zoning. The absence of long-distance trips may influence the travel time variable. The TBDCM does assigns probability to the alternatives based on the observations. If there are no long-distance trips, the travel time attribute does not take these trips into account. Hence it may not predict any long-distance trips because these trips are filtered out. If there are observations for long-distance trips, the estimation software needs to consider these trips; hence the travel time attribute is estimated as such. Secondly, the usage of only Limburg observation influences the heterogeneity of the individuals. Due to the characteristics and behaviour of individuals who live in Limburg, may behave differently than individuals who live in different areas. If Limburg is less urbanised and has an undeveloped PT network, there would be less observations for PT in comparison with a different province. Hence the usage of only Limburg may influence the estimated attributes.

The TBDCM is designed to estimate a trip. By estimating multiple trips, a tour is constructed. But the ASC of the modes are taken into account for each trip. If a TBDCM is estimated for tours, the ASC for the modes may differ as the behaviour for longer tours may differ for shorter tours. The alternatives would be no longer a trip but a complete tour. Estimating a TBDCM for tours would complicate the data preparations as for each tour-observation the non-chosen alternatives would also be whole tours.

7.2. Conclusion

This thesis aimed to determine if a DCM or the current the TBDFM should be used for modelling the mode & destination choice. A methodology was developed to determine which choice model should be used. A TBDCM was developed, validated and applied to analyse its performance. The performance was compared to the performance of the TBDFM. To conclude, the best performing TBDCM is the MNL as presented in Table 6.1.

7.2.1. Tour-based discrete choice model

A literature study was done on how to develop a TBDCM. The answer to the sub-question *how to construct a discrete choice model for a tour?* was sought out in the literature.

In Chapter 2, an extensive overview of the different DCMs used in the literature was presented. It was concluded that there was no consensus on the approach to estimating the joint mode & destination choice. Correlation among alternatives was not always applied; in theory, it should be. Based on the literature, it was concluded that there are three main structures: MNL, NL and CNL. All the main choice structures found in the literature were estimated to test if this correlation also exists within the choice set.

For all the DCMs that were estimated in the literature, the attributes used were collected and summarised in Table 2.1. Based on this overview, it was concluded that a diverse set of attributes is being used. Also, the occurrence differs significantly per attribute, only a few attributes were present in every model, while most of the attributes were present in some and absent in the other DCMs.

When a DCM is estimated, the non-chosen alternatives are required. For the non-chosen modes, it is straightforward, while the non-chosen destinations are difficult to determine. Taking every possible destination as a non-chosen alternative increases the computational time for estimating the DCM significantly. The literature offered different strategies to tackle this issue. Based on the different strategies, it was concluded that the non-chosen destination was randomly sampled from the complete set of destinations for this research.

In the literature, there was also a difference in how activities were prioritised. Some studies used no prioritisation, while others modelled the tour mode based on the tour's main activity. Based on the disparity of the literature, it was decided that there would be no priority among the activities and that the mode choice was based on the entire tour.

7.2.2. Available data sources

To estimate a TBDCM data on the chosen and non-chosen alternatives are needed. OViN will be used for the chosen alternatives; hence a second data source is needed for the non-chosen-alternatives. As such: *What data sources are available to compose the choice set?*

The data source for the non-chosen alternatives should be compatible with the current data sources: OViN and VML. Hence an overview of the personal characteristics present in OViN and VML was presented in Table 4.1. Also, the trip purpose is a variable in the TBDCM; hence an overview is presented in Table 4.3.

Due to the lack of data on the non-chosen alternatives in OViN, a different source was needed. No feasible data source was found for the non-chosen alternatives. Based on the data present in VML and OViN, it was decided to use the LoS data from VML for the chosen and non-chosen alternatives and the choices themselves from OViN. To facilitate this merge of different databases, only the area of Limburg is considered.

7.2.3. Evaluating the performance

It is key to determine the performance of a choice model if it is developed. Hence the next sub-question: *how is the performance of a choice model defined?*

A DCM is estimated with a maximum likelihood estimation for the choice set. The predictive power of the choice model is based on the total probability of the chosen alternatives. The performance on the choice set is indicated by the $\overline{\rho}^2$. But determining the predictive power based on only the choice set that was used for the estimation is not sufficient.

An independent choice set is required to determine the choice model's performance. Either the decision makers, temporal or spatial resolution, must be different from the choice set used for estimating the choice model.

Based on the literature, it was concluded that it is important for a choice model to be able to predict the chosen alternative. Still, it is also important for a choice model to correctly distribute the probability among all the alternatives. To quantify the distribution among the alternatives, an OD matrix must be modelled by the choice model. The performance of the choice model can be determined based on the modal split, trip length distribution and ground truth observations.

After developing the TBDCM and determining how the performance is defined, the performance of the TBDCM can be compared with the performance of the TBDFM.

Based on the out-of-sample validation, it can be concluded that the TBDCM performs better than the TBDFM in predicting the correct alternative. The chosen alternative is, on average higher in the order of probability within the choice of each individual. If the choice is split into the mode choice and destination choice, it can be concluded that the TBDCM especially performs better in the mode choice. When the modal split is analysed of the predicted alternatives, it can be concluded that the TBDCM and the TBDFM assign very few times a PT alternative the highest probability, compared to the observed trips. Also, the TBDCM and the TBDFM predict too many car trips and too few bike trips. This deviation from the observations is bigger for the TBDFM than for the TBDCM. Overall it can be concluded that the TBDCM performs better in predicting the chosen alternative than the TBDFM.

To evaluate policies in transport models it is important to analyse the elasticities of the travel time. The skim matrices of one mode was increased by 1%. The change in the number of trips indicates the elasticity for that specific mode. It was concluded that the car and bike elasticity for the TBDCM and the TBDFM comparable with the literature. The PT elasticity for the TBDCM was also comparable with the literature but for the TBDFM it was significantly higher than the literature.

To analyse the performance of the choice models on the probability distribution of similar alternatives, the choice models are subjected to choices with a complete destination set as alternatives. Based on the trip length distribution of this analysis, it can be concluded that the TBDCM and the TBDFM have similar trip length distributions for the car and the bike. But for PT, the trip length distribution is significantly different. The TBDCM predicts a more variety of trip lengths, including longer trips, while the TBDFM predict shorter trips and a smaller range of different trip lengths. For the modal split, the TBDCM and the TBDFM perform similarly. For both models, the PT trips is now similar to the observed number of PT trips and still significantly more trips for the car and less for the bike.

A case study is executed to analyse the choice models' performance. The TBDCM and the TBDFM are implemented in VML. VML models the mode & destination choice for different tours and an entire population. Based on the modal split, it can be concluded that the TBDCM predicts very few PT trips and too many car trips. the TBDFM also predict fewer PT trips than observed but a similar representation of the observed modal split overall. An explanation for the very few PT could be that the car is not always accessible to the individual. The TBDCM assumes that the car is always accessible within the household, but in real life, a different member of the household could have already taken the car. Hence the car was never an available alternative for the trip. Thus less individuals take the car in real life, which in turn causes more trips for PT and the bike.

The trip length distribution of the TBDCM and the TBDFM are very similar for the car and the bike. the TBDFM predict longer distances while the TBDCM predicts shorter distances. That the TBDCM predict shorter trips for the bike and the car could relate to the travel time sensitivity. As concluded in Chapter 6, the TBDCM is more sensitive to a changer in travel time than the TBDFM. Comparing the trip length distribution with the observed trip length distribution, the TBDFM approximate reality closer. For PT, the trip length distribution is quite volatile for the TBDCM and the observed data. the TBDFM predict a more stable trip length distribution. But when compared to the observed trip length distribution it can be concluded that the TBDFM predict too many short distance trips.

For the ground truth observations performs, the TBDCM performs mediocre for the car, which is compared to the TBDFM, not an improvement based on Figure 6.23 and Figure 6.24. For the bike, the TBDCM performs well, compare to the TBDFM as presented in Figure 6.27 and Figure 6.28. For PT, both the TBDCM and the TBDFM perform badly. But the TBDCM performs worse compared to the TBDFM depicted in Figure 6.25 and Figure 6.26.

7.2.4. Overall performance

In this section, the main research question is answered. The main research question is: *how does the performance of modelling the OD-matrix with a discrete choice model compare to the OD-matrix modelled with deterrence functions in the utility functions, in a joint destination and mode choice?*

From this thesis, it can be concluded that the TBDFM perform better in a large-scale transport model. Still, the TBDCM performs better in predicting the correct chosen alternatives. It is important to note that the TBDFM were specifically designed for the case study used in this thesis; hence the TBDFM have an advantage on the TBDCM.

Thus, is predicting the correct alternative more important or approximating the ground truth observations? Predicting the correct alternative is important to be able to replicate the choices individuals make. But it is also important to predict the correct second, third, and fourth alternative and so forth. the TBDFM are better in predicting the overall order of the alternatives as they perform better on the ground truth observations.

The TBDCM is estimated only on predicting the chosen alternative and not the correct distribution of the alternatives. Nonetheless, the TBDCM performs well on the ground truth observations.

All in all, taking into account that the TBDFM are specifically designed for the case study, the TBDCM is a generalised choice model. the TBDFM' performance for the ground truth observations is not overpowering the TBDCM's performance. The TBDCM has a higher potential for the performance for modelling the mode & destination choice.

7.3. RECOMMENDATIONS FOR FUTURE RESEARCH

Based on this thesis, various opportunities to elaborate on this research arise. The potential elements of these opportunities are identified and explained.

The most eye-catching result was the high $\overline{\rho}^2$ for the TBDCM. As mentioned before, a possible explanation could be the usage of VML travel times. It would be valuable to know if this is the case. If it is not the case, it should be investigated why the $\overline{\rho}^2$ is high.

VML is only one of the transport models RHDHV is developing. The TBDCM in this thesis was estimated using only observations from Limburg. The travel behaviour of people from Limburg might differ from those from the rest of The Netherlands. This difference in travel behaviour will influence the attributes of the TBDCM. To apply the TBDCM in a different transport model the input data should contain observations from the study area of that transport model. But taking only observations from a certain region will decrease the number of observations as in this thesis. Thus it is valuable to determine if the TBDCM performs better with all the observations from The Netherlands or only the observations from the study area.

Second, for this thesis VML was used for the non-chosen alternatives because RHDHV supplied it. Generally this data would not be available; hence there would be no data base for the non-chosen alternatives. It is valuable to research more on the availability of non-chosen alternatives as it is a troublesome, but essential task in estimating any DCM.

Third, it is interesting to elaborate on the structure of the TBDCM. In this thesis, each observation was seen as an independent trip. But as VML models tours, the TBDCM could be modelling a complete tour instead of sequentially modelling each trip. The most important aspect of this change would be the change of the ASC for each mode. The mode choice may be different for different tour lengths. This elaboration does require more input data and data preparation. More data of the different tour lengths is required to be able to estimate the tours.

Also the non-chosen alternatives are not trips anymore but complete tours. Determining the non-chosen alternatives become more troublesome. 20 zones can still be randomly sampled but the number of tours, hence also the non-chosen alternatives, grow exponentially with every extra destination. It would be valuable to know if this extra effort of data preprocessing is worth the extra detail of the TBDCM.

Forth, it is valuable to analyse the potential of the current TBDCM. Based on the results of the ground truth observations, a method can be developed to tailor the TBDCM. the TBDFM were also tailored to increase the results on the ground truth observations. The calibration of the TBDCM was not incorporated in the scope of this thesis; hence it would be valuable to calibrate the TBDCM and analyse the improvement on the ground truth observations.

A. Attraction Sensitivity

In this appendix the figures for the sensitivity analysis of the attraction are presented.

Work

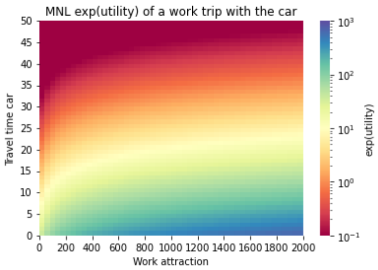

Figure A.1: MNL work attraction sensitivity for the car.

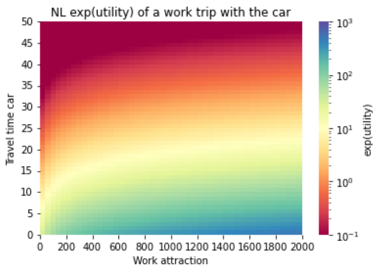

Figure A.2: NL work attraction sensitivity for the car.

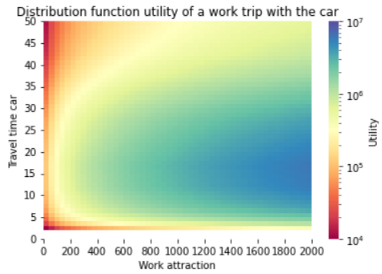

Figure A.3: the TBDFM work attraction sensitivity for the car.

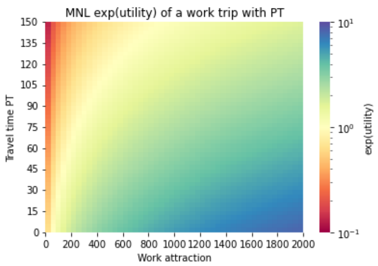

Figure A.4: MNL work attraction sensitivity for PT.

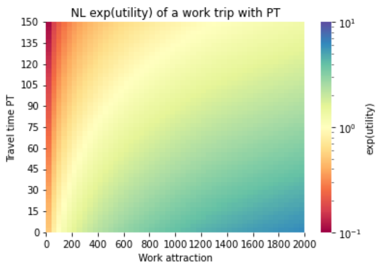

Figure A.5: NL work attraction sensitivity for PT.

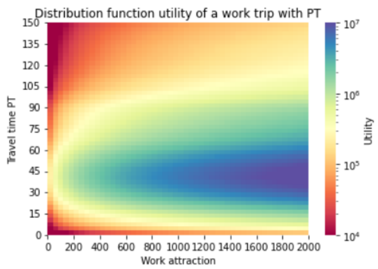

Figure A.6: the TBDFM work attraction sensitivity for PT.

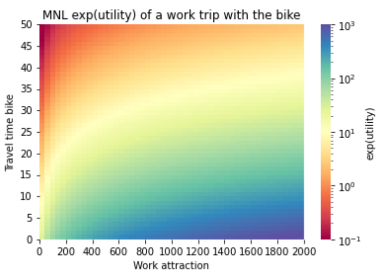

Figure A.7: MNL work attraction sensitivity for the bike.

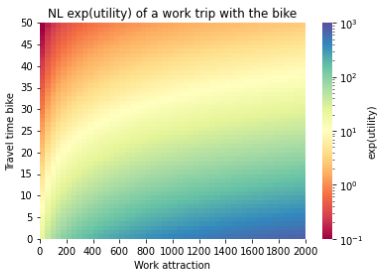

Figure A.8: NL work attraction sensitivity for the bike.

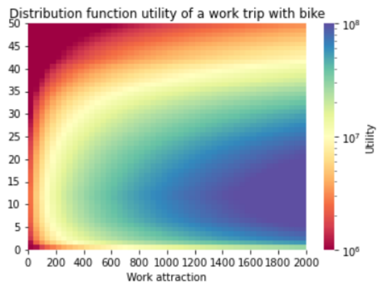

Figure A.9: the TBDFM work attraction sensitivity for the bike.

Business

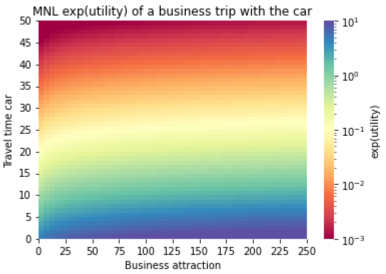

Figure A.10: MNL business attraction sensitivity for the car.

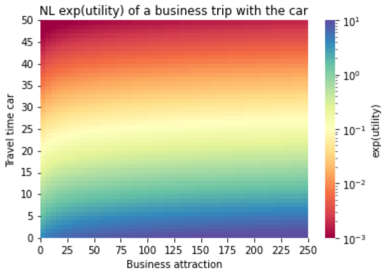

Figure A.11: NL business attraction sensitivity for the car.

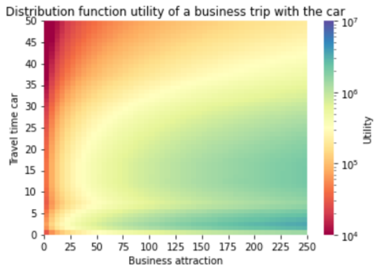

Figure A.12: the TBDFM business attraction sensitivity for the car.

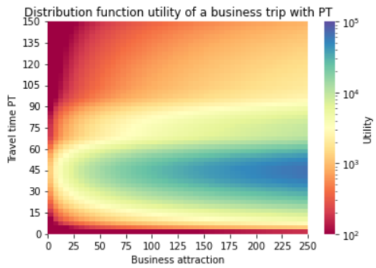

Figure A.13: the TBDFM business attraction sensitivity for PT.

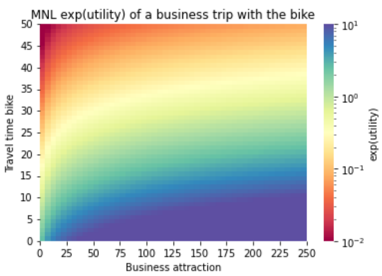

Figure A.14: MNL business attraction sensitivity for the bike.

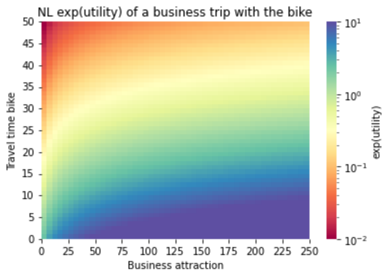

Figure A.15: NL business attraction sensitivity for the bike.

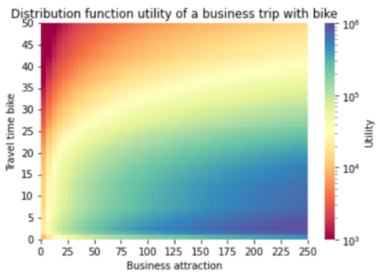

Figure A.16: the TBDFM business attraction sensitivity for the bike.

Education

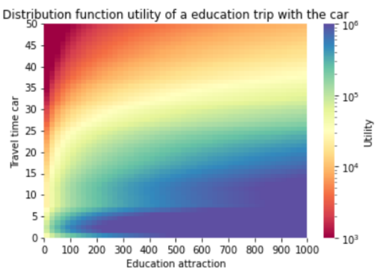

Figure A.17: the TBDFM education attraction sensitivity for the car.

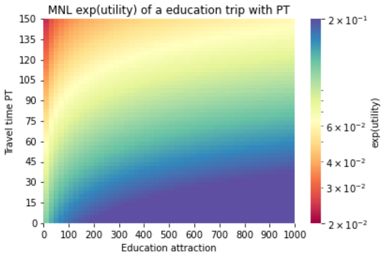

Figure A.18: MNL education attraction sensitivity for PT.

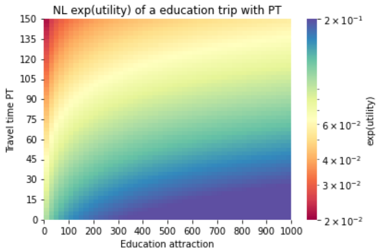

Figure A.19: NL education attraction sensitivity for PT.

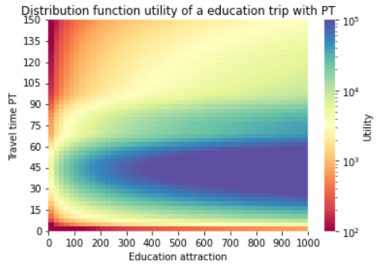

Figure A.20: the TBDFM education attraction sensitivity for PT.

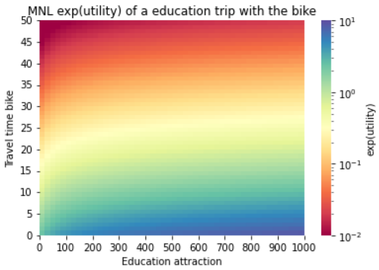

Figure A.21: MNL education attraction sensitivity for bike.

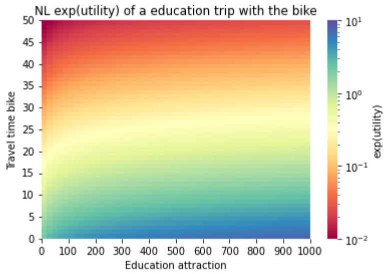

Figure A.22: MNL education attraction sensitivity for bike.

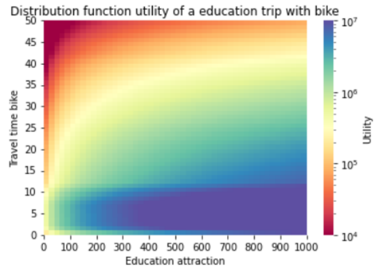

Figure A.23: MNL education attraction sensitivity for bike.

Shopping

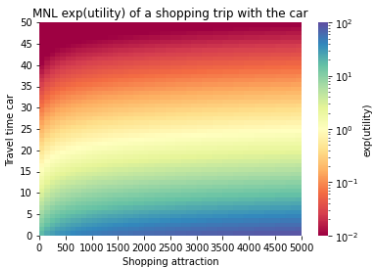

Figure A.24: MNL shopping attraction sensitivity for the car.

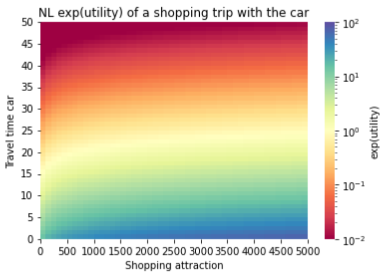

Figure A.25: NL shopping attraction sensitivity for the car.

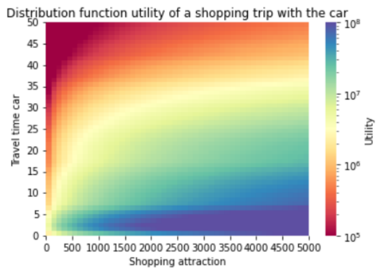

Figure A.26: the TBDFM shopping attraction sensitivity for the car.

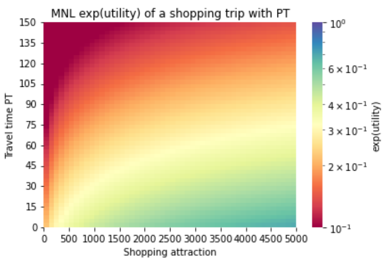

Figure A.27: MNL shopping attraction sensitivity for PT.

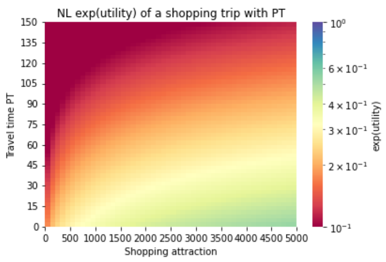

Figure A.28: NL shopping attraction sensitivity for PT.

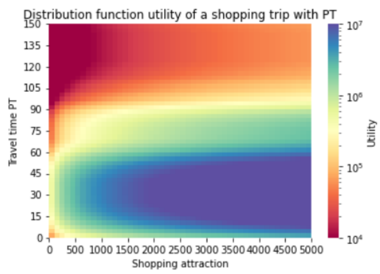

Figure A.29: the TBDFM shopping attraction sensitivity for PT.

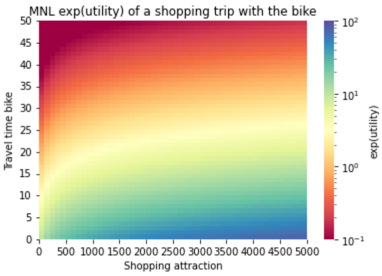

Figure A.30: MNL shopping attraction sensitivity for bike.

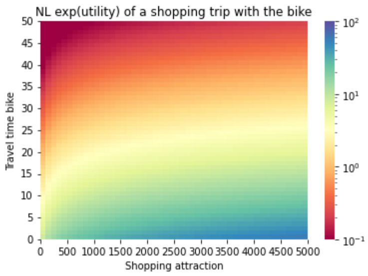

Figure A.31: MNL shopping attraction sensitivity for bike.

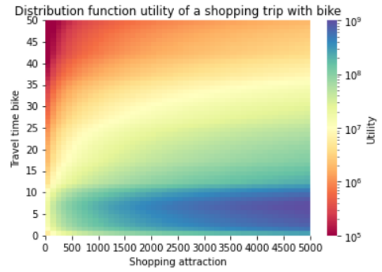

Figure A.32: MNL shopping attraction sensitivity for bike.

Bring/get

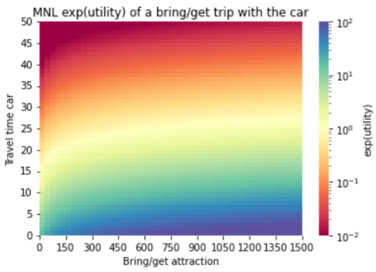

Figure A.33: MNL bring/get attraction sensitivity for the car.

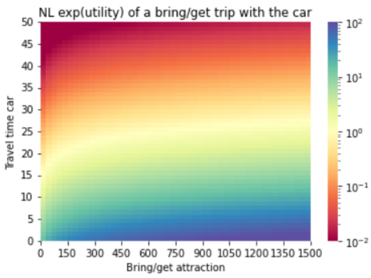

Figure A.34: NL bring/get attraction sensitivity for the car.

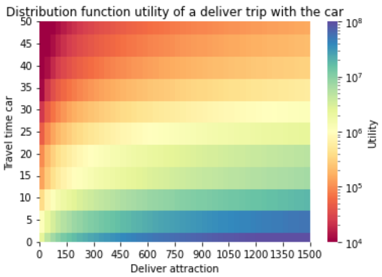

Figure A.35: the TBDFM bring/get attraction sensitivity for the car.

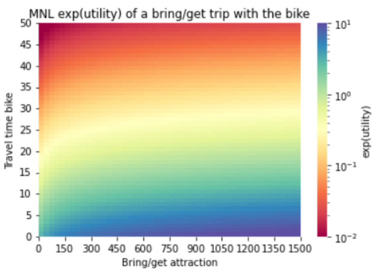

Figure A.36: MNL bring/get attraction sensitivity for the bike.

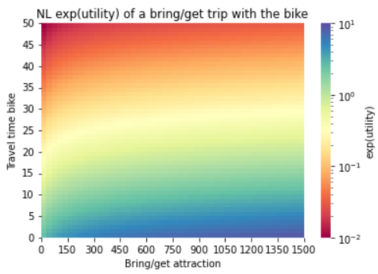

Figure A.37: NL bring/get attraction sensitivity for the bike.

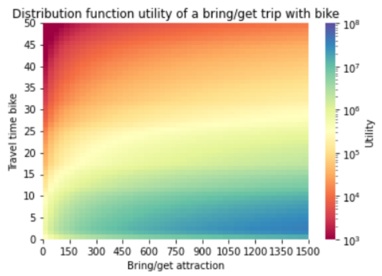

Figure A.38: the TBDFM bring/get attraction sensitivity for the bike.

Other

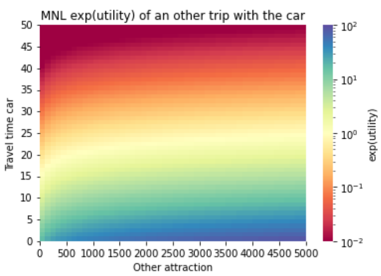

Figure A.39: MNL other attraction sensitivity for the car.

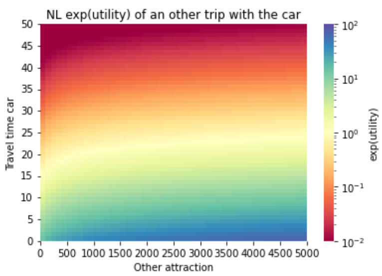

Figure A.40: NL other attraction sensitivity for the car.

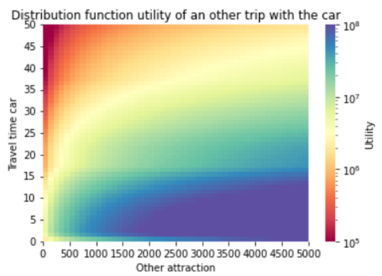

Figure A.41: the TBDFM other attraction sensitivity for the car.

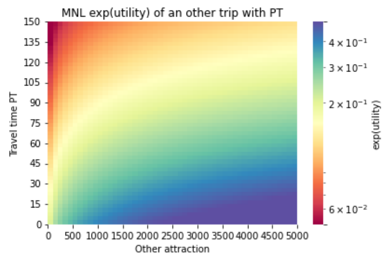

Figure A.42: MNL other attraction sensitivity for PT.

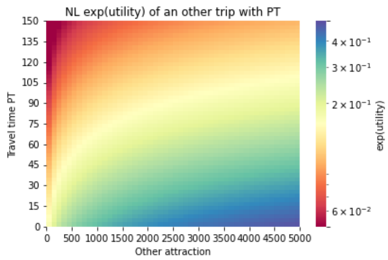

Figure A.43: NL other attraction sensitivity for PT.

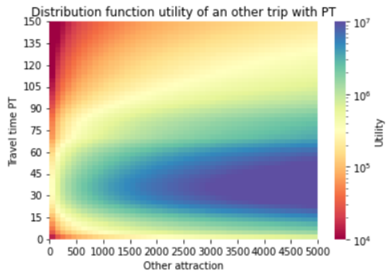

Figure A.44: the TBDFM other attraction sensitivity for PT.

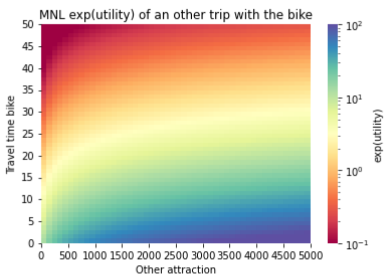

Figure A.45: MNL other attraction sensitivity for the bike.

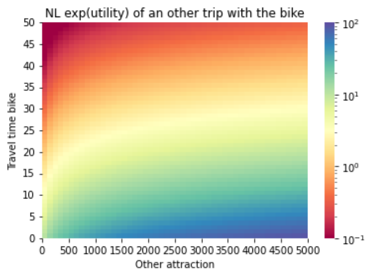

Figure A.46: NL other attraction sensitivity for the bike.

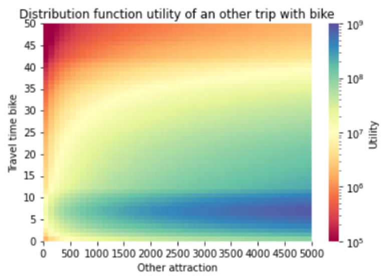

Figure A.47: the TBDFM other attraction sensitivity for the bike.

Bibliography

Arentze, T. A., & Timmermans, H. J. (2004). A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological*, *38*(7), 613–633. https://doi.org/https://doi.org/10.1016/j.trb.2002.10.001

Barff, R., Mackay, D., & Olshavsky, R. W. (1982). A Selective Review of Travel-Mode Choice Models. *Journal of Consumer Research*, *8*(4), 370–380. https://doi.org/10.1086/208877

Ben-Akiva, M. E. (1973). *Structure of passenger travel demand models.* (Doctoral dissertation). Massachusetts Institute of Technology.

Bierlaire, M. Biogeme: A free package for the estimation of discrete choice models. In: *Swiss transport research conference.* (CONF). 2003.

Bowman, J. L., & Ben-Akiva, M. E. (2001). Activity-based disaggregate travel demand model system with activity schedules. *Transportation research part a: policy and practice*, *35*(1), 1–28.

Bradley, M., Bowman, J. L., & Griesenbeck, B. (2010). Sacsim: An applied activity-based model system with fine-level spatial and temporal resolution. *Journal of Choice Modelling*, *3*(1), 5–31. https://doi.org/https://doi.org/10.1016/S1755-5345(13)70027-7

Chorus, C. G., Arentze, T. A., & Timmermans, H. J. (2008). A random regret-minimization model of travel choice. *Transportation Research Part B: Methodological*, *42*(1), 1–18. https://doi.org/https://doi.org/10.1016/j.trb.2007.05.004

Curtis-Ham, S., Bernasco, W., Medvedev, O. N., & Polaschek, D. L. (2021). The importance of importance sampling: Exploring methods of sampling from alternatives in discrete choice models of crime location choice. *Journal of Quantitative Criminology*, 1–29.

de Luca, S., & Cantarella, G. E. Validation and comparison of choice models. In: *Travel demand management and road user pricing.* Routledge, 2016, pp. 57–78.

Ding, C., Xie, B., Wang, Y., & Lin, Y. (2014). Modeling the joint choice decisions on urban shopping destination and travel-to-shop mode: A comparative study of different structures. *Discrete Dynamics in Nature and Society*, *2014*, 1–10.

Eluru, N., Pinjari, A. R., Pendyala, R. M., & Bhat, C. R. (2009). *A unified model system of activity type choice, activity duration, activity timing, mode choice, and destination choice* (tech. rep.). Working Paper, The University of Texas at Austin, Texas.

Hasnine, M. S., & Habib, K. N. A dynamic discrete choice model for tour-based mode choices. In: *98th annual meeting of trans. res. board, washington, dc.* 2019.

Hasnine, M. S., & Nurul Habib, K. (2021). Tour-based mode choice modelling as the core of an activity-based travel demand modelling framework: A review of state-of-the-art. *Transport Reviews*, *41*(1), 5–26.

Hess, S., Daly, A., & Batley, R. (2018). Revisiting consistency with random utility maximisation: Theory and implications for practical work. *Theory and Decision*, *84*(2), 181–204.

Hilty, L., Meyer, R., & Ruddy, T. A general modelling and simulation system for sustainability impact assessment in the field of traffic and logistics. In: 2001, March, 167–185.

Hörl, S., Balać, M., & Axhausen, K. W. Pairing discrete mode choice models and agent-based transport simulation with matsim. In: *2019 trb annual meeting online.* Transportation Research Board. 2019, 19–02409.

Horni, A. (2013). *Destination choice modeling of discretionary activities in transport microsimulations* (Doctoral dissertation). ETH Zurich. Zürich, ETH. https://doi.org/10.3929/ethz-a-010006641

Horowitz, J. L. (1985). Travel and location behavior: State of the art and research opportunities. *Transportation Research Part A: General*, *19*, 441–453.

Jonnalagadda, N., Freedman, J., Davidson, W. A., & Hunt, J. D. (2001). Development of microsimulation activity-based model for san francisco: Destination and mode choice models. *Transportation Research Record*, *1777*(1), 25–35. https://doi.org/10.3141/1777-03

Kane, L., & Behrens, R. (2002). Transport planning models–an historical and critical review. *SATC 2002.*

Kitamura, R., Fujii, S. et al. (1998). Two computational process models of activity-travel behavior. *Theoretical foundations of travel choice modeling*, 251–279.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, *92*(368), 805–824.

McFadden, D. (1977). *Modelling the choice of residential location* (Cowles Foundation Discussion Papers No. 477). Cowles Foundation for Research in Economics, Yale University. https://EconPapers.repec.org/RePEc:cwl:cwldpp:477

McFadden, D. et al. (1973). Conditional logit analysis of qualitative choice behavior.

Mishra, S., Wang, Y., Zhu, X., Moeckel, R., & Mahapatra, S. (2013). *Comparison between gravity and destination choice models for trip distribution in maryland* (tech. rep.).

Molloy, J. (2016). Development of a destination choice model for ontario. *Technical University of Munich*, *74*.

Newman, J. P., & Bernardin, V. L. (2010). Hierarchical ordering of nests in a joint mode and destination choice model. *Transportation*, *37*(4), 677–688.

Parady, G., Ory, D., & Walker, J. (2021). The overreliance on statistical goodness-of-fit and under-reliance on model validation in discrete choice models: A review of validation practices in the transportation academic literature. *Journal of Choice Modelling*, *38*, 100257. https://doi.org/https://doi.org/10.1016/j.jocm.2020.100257

Shiftan, Y. (1998). Practical approach to model trip chaining. *Transportation Research Record*, *1645*(1), 17–23. https://doi.org/10.3141/1645-03

Shiftan, Y., & Ben-Akiva, M. (2011). A practical policy-sensitive, activity-based, travel-demand model. *The Annals of Regional Science*, *47*(3), 517–541.

Shiftan, Y., Ben-Akiva, M., Proussaloglou, K., de Jong, G., Popuri, Y., Kasturirangan, K., & Bekhor, S. Activity-based modeling as a tool for better understanding travel behaviour. In: *10th international conference on travel behaviour research*. 2003, 10–15.

Sælensminde, K. (2006). Causes and consequences of lexicographic choices in stated choice studies. *Ecological Economics*, *59*(3), 331–340. https://doi.org/https://doi.org/10.1016/j.ecolecon.2005.11.001

Vajjarapu, H., Verma, A., & Allirani, H. (2020). Evaluating the climate change mitigation potential of sustainable urban transport measures in india. *Available at SSRN 3715492*.

van Cranenburgh, S., & Chorus, C. G. (2018). Does the decision rule matter for large-scale transport models? [Transportation and Traffic Theory: Behavior and Planning Applications]. *Transportation Research Part A: Policy and Practice*, *114*, 338–353. https://doi.org/https://doi.org/10.1016/j.tra.2018.01.035

Västberg, O. B., Karlström, A., Jonsson, D., & Sundberg, M. (2020). A dynamic discrete choice activity-based travel demand model. *Transportation Science*, *54*(1), 21–41. https://doi.org/10.1287/trsc.2019.0898

Wardman, M. (2012). Review and meta-analysis of uk time elasticities of travel demand. *Transportation*, *39*(3), 465–490.

Yagi, S., & Mohammadian, A. K. (2010). An activity-based microsimulation model of travel demand in the jakarta metropolitan area. *Journal of Choice Modelling*, *3*(1), 32–57.