A Flexible Approach towards Public Transport Modelling in Travel Demand Models

Jeroen van der Gun

Thesis report

July 2013

TU Delft

significance

Thesis author

Jeroen van der Gun BSc
Civil Engineering MSc student
Transport & Planning track
Delft University of Technology

Graduation committee

Prof.dr.ir. Bart van Arem (chair)
Faculty of Civil Engineering and Geosciences
Delft University of Technology

Dr.ir. Rob van Nes
Faculty of Civil Engineering and Geosciences
Delft University of Technology

Eric Kroes MSc
Significance B.V.

Dr. Eric Molin
Faculty of Technology, Policy and Management
Delft University of Technology

Ir. Paul Wiggenraad
Faculty of Civil Engineering and Geosciences
Delft University of Technology

Cover: composition of own photos, public domain photos and Figure 33.

Preface

This thesis report is the final product of my graduation project. The idea for this project originated from the current ANTONIN model that has been developed by Significance, which, as described later in this report, contains a rather complicated mode choice structure. This raised the natural question whether clever alternatives to this structure exist, which is particularly interesting since Significance also develops the GroeiModel, which uses a vastly different, train-centric approach of public transport modelling, yet faces remarkably similar modelling problems as ANTONIN.

First of all, I'd like to thank my graduation committee, chaired by Bart van Arem and further consisting of Rob van Nes, Eric Kroes, Eric Molin and Paul Wiggenraad, for monitoring my work from the initial working plan to this final report.

Specifically from Significance, I'd like to thank Eric Kroes for providing the great opportunity to carry out this project, including generating the idea for this project, supporting the execution and providing the necessary data sets and computer facilities. I'd also like to thank the other employees of Significance for providing a pleasant workplace and help when needed, in particular Marco Kouwenhoven for his help with the estimation of the case study model and Jan Gerrit Tuinenga for answering my questions about the inner workings of ANTONIN.

From the TU Delft, I'd specifically like to thank Rob van Nes for being a great daily supervisor, providing very useful and detailed feedback on intermediate reports and during discussions of intermediate results. Finally, I thank Gijs van Eck for sharing insights about his doctoral research project.

Jeroen van der Gun

Summary

A Flexible Approach towards Public Transport Modelling in Travel Demand Models

Abstract

Public transport systems create large numbers of possible combinations of public and private modes that travellers can choose to use. As a consequence, current travel demand models face problems related to flexibility in the modelling of how travellers value and use modes, consistency in the choice process and computational efficiency. To resolve these problems, on theoretical grounds, a supernetwork approach is found necessary, where a hybrid form of the network GEV model and the path size logit model should be used to model the multi-modal route choice. A case study for mode and route choice in Île-de-France shows that this proposed approach is feasible in practice and confirms that the constructed flexible modelling framework indeed possesses most of its suspected advantages over current travel demand models.

Keywords

public transport, multi-modal transport, travel demand model, supernetwork, network GEV, branch and bound

---

1. Introduction

Without public transport, models for the behaviour of travellers in transport systems would be relatively simple: in general, travellers would only use a single mode for their trip. Public transport is however important due to the existence of congestion, environmental pollution and traffic accidents caused by other means of transport. This does pose a modelling problem though: travellers choose multi-modal routes, in which for example train and bus are combined, and those combinations of public transport modes may for example be further extended with cycling or car usage. In dense transport systems, large numbers of such possible combinations exist.

For modifications to the transport system, policy makers use travel demand models to predict how the transport system will be used, that is, the number of tours made, their destinations and their modes. In case of capacity planning, the routes of the tours also need to be predicted. Travel demand models therefore must model the choice process correctly. Currently existing travel demand models however face multiple problems in the modelling of choices among multi-modal routes. These problems relate to flexibility, consistency and efficiency.

First, there are three problems related to flexibility, that is, flexibility in the modelling of how modes are valued and used by travellers. First of all, in reality, a large number of modes may be available, but only a limited number of distinct modes is typically included in the travel demand model. This aggregation of modes may lead to loss of accuracy in modelling, as differences in valuation between modes are neglected. Secondly, restrictive assumptions are made regarding how and in which orders the different modes can be combined in a trip, i.e. which permutations of modes are permitted, and this gets troubling if the aggregation of modes is reduced. Thirdly, the addition of new modes to the model is not trivial: if new modes are considered distinct from existing modes in either valuation or permitted permutations, the addition of the new modes becomes quite difficult.

Next, travel demand models are usually based on the principle of utility maximisation, using logit models defining stochastic utilities for the alternatives. Nested logit models group alternatives to take positive correlations among their utilities into account. Here, the so-called logsum should be used as the measure of overall utility for each such group of alternatives: this takes into account the freedom of choice of travellers to select their optimal alternative within the group according to their personal perception, which makes the alternatives as a group more attractive. On the other hand, this benefit of freedom of choice should not be overestimated by neglecting further positive correlations between the alternatives within each group. Current travel demand models do not always take this freedom of choice into account in a correct way – its consequences for the attractiveness of groups of alternatives are under- or overestimated – and these models therefore lack consistency in how the choice processes work.

Finally, some models have an efficiency problem, that is, a high computation time. In particular, this problem can occur in models that handle each possible mode combination separately, due to the large number of them.

For this research project, it is assumed that the flexibility, consistency and efficiency of the model are to be maximised. Given these problems, the research objective is to find the theoretically best mode and route choice model structure and to assess its performance in practice. For the latter, a case study is used.

2. Design of the ideal model

In order to design a theoretically ideal model, first of all, attention must be paid to the interface between mode and route choice, the representation of the networks and the route set generation and route choice. Each of these aspects is handled below.

Interface between mode and route choice

Regarding the interface between mode and route choice, it is found that the possible interfaces can be classified according to two dimensions: how networks are combined by the traveller and whether networks contain multiple modes. In case there exist multiple networks, this implies choice processes for the traveller determining the networks or network combinations to be used and the origins and destinations within these networks and network combinations. The route generator subsequently generates one or more possible routes for each of these possible networks or network combinations, which are used for route choice.

The possible interfaces between mode and route choice then are:

1. having only a single network in the model, used by all travellers, without prior choice process;
2. having multiple networks in the model, where travellers choose one of the networks;

3. having multiple networks in the model, where travellers choose a main network as well as access and egress options for the chosen main network;
4. having multiple networks in the model, where travellers choose a combination of networks.

In each case, networks may span multiple modes (B), for example a full public transport network, or not (A), in which case each mode is represented by its own network.

In this framework, a single supernetwork (1B), a PT network with access and egress options (3B) and combinations of networks, where PT is a single network (4B), achieve the flexibility objective, without getting problems in case of a large mode count.

Regarding the consistency objective, there appear to be three sources of positive correlations among utilities of route alternatives:

- there can be route overlap, i.e. routes sharing network links, like the same train connection with bus access and walk access;
- there can be modal overlap, i.e. routes having modes in common, like two distinct bus routes;
- there can be mode similarities, i.e. routes having not identical but similar modes, like a bus route and a train route, which are both public transport routes.

Looking at the flexible structures, it is found that the structures 3B and 4B cannot correctly take correlations across mode combinations into account and therefore lack consistency; only structure 1B does not have this problem and is both flexible and consistent.

Network representations

Next, regarding the representation of networks, various representations exist, each of which is described as a directed graph of nodes and links. They range from a simple level-of-service matrix where links directly connect origins and destinations, a traffic network where the links represent roads, a frequency-based service network where the links represent public transport in-vehicle time, boarding movements and alighting movements to a timetable-based service network that includes specification of PT timetables, where all nodes also represent a point in time.

Furthermore, the service networks can be pre-processed by creating direct stop-to-stop links combining the original boarding, in-vehicle and alighting links into single links. This increases the total number of links in the network, but reduces the number of links per route. For frequency-based service networks, pre-processing allows subsequent merging common lines for each pair of boarding and alighting stops: this merges these lines into a single alternative, which reduces the eventual route set size and increases the realism of the choice model. For example, in a situation with two bus lines from the same bus stop to the same train station, these two bus lines are then no longer modelled as separate alternatives.

If a full timetable is known for all public transport, one can formulate a time-dependent choice model in which the choice set depends on the departure time within the timetable period. If only parts of the PT network are timetable-based and other parts are frequency-based, the timetable information can be used to improve some of the waiting time attributes of routes generated in an integrated frequency-based service network.

Route set generation and route choice

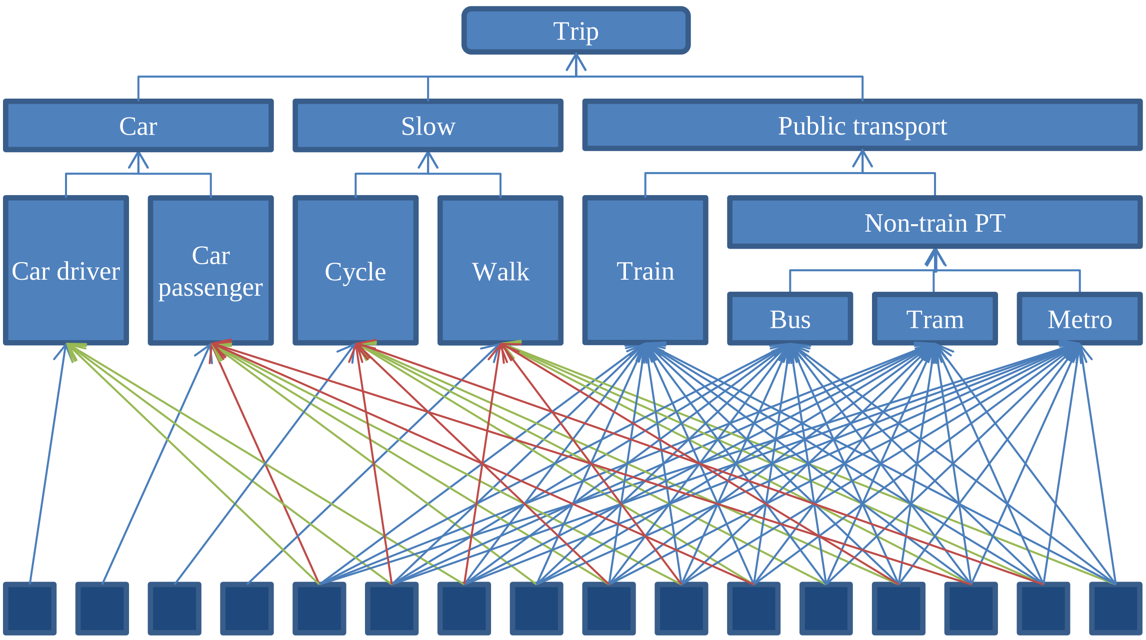
Next, regarding the route set generation, it is advantageous in terms of efficiency and consistency to generate a route choice set in advance instead of finding routes during network loading. Several algorithms exist for this, which differ in computational efficiency, the coverage of observed routes and the variety and comparability in the generated choice sets.

For the public transport part of the network, the branch-and-bound algorithm is preferable, as it can handle a large number of links like in pre-processed service networks. For private modes on the other hand, Monte Carlo labelling and simulation is a more useful algorithm, since many road links need to be combined in a single route. For private modes for which route choice is not deemed important, the Dijkstra algorithm can simply be used to generate just the shortest route. The latter may also be the case for access and egress legs for public transport. These three algorithms can be linked together by splitting the route generation at the boundary of the PT system.

Lastly, regarding the route choice, there is no clearly best choice model for handling route overlap. The path size logit model and the extended path size logit model however have a simple structure, and either the theoretical quality or the practical quality is high.

The ideal model

As stated above, model structure 1B (supernetwork) is necessary to get a flexible and consistent model. For consistency, the generated multi-modal route alternatives must however be cross-nested according to the used modes. This leads to a network GEV path size logit model for mode and route choice, which is a combination of the path size logit model for route overlap, the cross-nested logit model for modal overlap and the nested logit model for mode similarities. The following figure gives an example of such a choice tree:



In this model, mode-specific boarding penalties are used to express average modal preferences in the utilities of the generally multi-modal routes.

Other choices than mode and route choice may also be integrated in the choice process. Choices within the trip such as destination choice can be added by simply extending the network GEV tree with extra decision levels. The trip frequency choice can be based on the logsum of the trip choice tree. Choices such as car ownership can be based on the change in logsum of the trip choice tree.

By concatenating outbound routes and return routes, the model can be extended into a tour-based model. It is also possible to define a pivot-point procedure to enhance the model outcomes with base matrices. Finally, it is possible to add congestion and crowding feedback to the model as well.

The theoretical objectives of the ideal model are satisfied: aggregation of modes is no longer necessary, all permutations of modes are permitted in the choice tree, changes in the number of modes can be handled, consistency in the choice process is achieved and the computation time dependency on the number of permitted mode permutations is removed.

3. Case study for Île-de-France

The proposed modelling methodology is tested in a case study for the Île-de-France region. As a simplification, the case study only focuses on modelling mode and route choice; also, only a single route (part) is considered for non-PT legs.

The case study uses the ANTONIN network and revealed preference data from the Enquête Globale Transport (EGT) to estimate a flexible choice model for morning peak home-work trips in the city Paris and its inner ring Petite Couronne. The modes included in the case study are walk, Transilien, RER, metro, tram, RATP Paris bus, RATP banlieue bus, Optile bus, car driver, bicycle, motor driver and car/motor passenger.

The route set generation algorithm operates in two main steps. First, non-PT segments are constructed using the Dijkstra algorithm. After that, the branch-and-bound algorithm is used to combine the PT and non-PT segments into routes. The branch-and-bound algorithm uses a tolerance constraint specifying maximum deviations from the minimum travel time, several logical constraints and, in the end, also a dominance constraint eliminating routes which are inferior on all aspects compared to one of the other routes.

After mapping the route descriptions from the EGT to the ANTONIN network, the generated and observed routes can be compared in three steps. First of all, 22% of the PT routes from the survey match exactly with a route in the choice set. Secondly, 26% is matched based on the sequence of modes and line numbers, meaning only boarding and alighting stops differ. Finally, 38% of the observed routes are dominated by routes in the generated choice set; these dominated routes may still be attractive in practice due to timetable effects and network errors. This brings the total coverage of PT routes to 86%, which is deemed sufficiently high.

Due to software limitations, no network GEV model could be estimated. Therefore, instead, nested logit models have been used, which have also been combined with path size factors. The availability of vehicles, driving licences and PT discounts has been taken into account during all estimations. The basic model includes as attributes private mode time, PT in-vehicle time, waiting time, PT costs, the number of legs per mode, a dummy for PT usage without discount and a domination size variable indicating the number of routes dominated by each route alternative. As during the estimation it was found that timetable effects appear to play a role, the waiting time has been maximised at 7.5 minutes per PT boarding.

Results and discussion

Several significant differences are found among the boarding penalties of the different PT modes. Overall, the boarding penalties seem logical, with the rail modes more attractive than bus modes. The mutual differences among RATP Paris bus, RATP banlieue bus and Optile bus were significant each, and the difference between Transilien and RER is also slightly significant. As the difference between RATP buses in Paris and suburbs may be suspected to be caused by competition with rail-based urban modes in Paris, handling this effect via nesting would be preferable, but the nested logit model cannot capture this effect completely, because it has limitations compared to a network GEV model. However, even then the Optile bus remains significantly different from the RATP buses, so it is concluded that aggregation of modes may indeed lead to errors, even more when the choice model cannot take all correlations between alternatives into account.

Considering the mode permutations that are used in the ANTONIN travel demand model for this region, only 4% of the PT routes cannot be modelled using the mode combination approach since they use an unavailable mode combination. This could be further reduced with a small change to ANTONIN. Regarding the problem of permutations of modes, the advantage provided by the flexible model is thus rather small.

New modes can be added to the flexible model using stated preference research, which is one of its theoretical advantages. This has been tested in practice by adding T Zen and express bus as two new modes, thereby estimating the boarding penalties and position in the choice tree using a stated preference data set. Although the corresponding questionnaire was not designed for this purpose, the estimation yielded plausible results, with both new modes being more attractive than the existing bus modes. This way of flexibly extending the model hence seems feasible.

The benefit of consistent modelling should be assessed in two aspects. Firstly, the integration of mode and route choice can be assessed by splitting the model into a main mode choice model and a route choice model, where only a single best route represents each main mode in the main mode choice, rather than the logsum of all routes. When both these models are jointly estimated, it follows that the log-likelihood is worse than an integrated multinomial logit model for mode and route choice, even if the scale of utilities is permitted to differ between the route choice level and the main mode choice level. It is therefore concluded that the integration of mode and route choice using logsums improves the fit of the model.

Secondly, the corrections for correlations due to route overlap, modal overlap and mode similarities should be assessed. Regarding route overlap, the path size coefficient is unexpectedly negative, suggesting that ad hoc route choice may play a role. Regarding modal overlap, a significant nest is found for metro/tram as main mode, and there is an indication that PT routes sharing access/egress modes are also positively correlated. Regarding mode similarities, a significant nest is found for all PT routes. As a network GEV model can capture more correlations than the nested logit model, there may exist more positive correlations in reality.

Contrary to the expectations, the computation time of the flexible model is worse than that of the ANTONIN model. There are two possible reasons: by repeatedly generating routes for different mode combinations, the desired spatial variety can be lower without sacrificing modal variety in the overall choice set, and the flexible model may have better coverage of observed routes than ANTONIN, since the coverage of ANTONIN is currently unknown.

4. Conclusions and recommendations

A theoretical framework has been constructed for flexible, consistent and efficient modelling of public transport in travel demand models, and an application for mode and route choice in Île-de-France has shown this to be feasible for practical usage. The case study confirmed the existence of most of the suspected advantages of this flexible model compared to current travel demand models.

Recommendations for further research include assessment of network loading results, investigation of ad hoc route choice behaviour, optimisation of the branch and bound algorithm, integration with other choice processes, estimation of a network GEV model, application of a timetable-based model and investigation of the robustness of the procedure to add new modes to the model. For existing travel demand models, simple enhancements appear possible by taking modal differences into account in the deterministic part of the utility function and by using logsums of the route sets in the mode choice.

Table of contents

1. Introduction

In a world without public transport, the behaviour of users in national and regional transport systems and their demand for travel would be relatively simple to model. Except for some car-poolers, a traveller would generally use a single mode for his whole tour, such as driving a car, being a car passenger, cycling or walking. While some of these modes may share similarities, like driving a car and being a car passenger, these effects are not very difficult to incorporate in a model. The travel demand model would include a choice process for each inhabitant of the region to determine tour frequencies, destinations, modes and routes, possibly supplemented with time-of-day choices and vehicle and licence ownership choices.

Due to congestion, environmental pollution and traffic accidents, it is desirable to have other means of motorised transport than just private vehicles. Luckily, there exists public transport, but the presence of public transport unfortunately complicates the construction of travel demand models. Now, different modes such as train and bus may suddenly be combined within a single multi-modal route and these chains of public transport modes may be complicated further because private modes such as cycling or car usage may be used to access and egress the public transport network. In a dense transport system, the number of available mode chains is large.

In order to optimise the transport system and assess the benefits of improvements, policy makers use travel demand models to predict how a transport system will be used, that is, how many trips people will make to which destinations using which modes. For capacity planning, it is also necessary to predict more specifically which roads and which public transport lines will be used. It is therefore important that travel demand models correctly model the choice process mentioned above.

However, currently existing travel demand models face problems in the modelling of choices among multi-modal routes. Differences in the valuation of modes are being neglected, the combinations and orders in which modes are used are constrained and there is no straightforward way to add new modes to the model; these three problems may be summarised as a lack of flexibility in how modes are perceived and used by travellers. Furthermore, in the choice models, the benefits of each traveller having the freedom to choose his personally optimal multi-modal route, are not estimated correctly, leading to inconsistencies in the choice process and incorrect choice probabilities. These problems may lead to incorrect predictions of the effects of future changes to the transport system, such as the introduction of new public transport lines or modes. Finally, in some cases, the computation time of a model also becomes cumbersome.

1.1. Research question

The purpose of this research project therefore is to investigate how a travel demand model can be constructed that avoids the problems mentioned above, such that usage of the public transport system can be modelled correctly. Due to the nature of these problems, there is a strong focus on the mode and route choice components of such a travel demand model, and the research project is limited to static models. The main research question can hence be formulated as follows:

What is the theoretically best flexible mode and route choice model structure for modelling public transport within a travel demand model and how does it perform in practice?

In order to answer this research question, first, the interface between mode and route choice, the representation of networks and the route set generation and route choice will be considered in detail to construct a theoretically ideal model. Secondly, as a case study, it will be assessed whether, and if so, what modifications are necessary to estimate parameters of this model for the Île-de-France region using revealed preference data. Finally, the results from the case study will be used to check whether the specified problems are indeed solved by the newly developed model.

1.2. Structure of this report

The structure of this report is as follows. First, the problems with current models will be introduced in more detail in Chapter 2. Then, Chapter 3 introduces the methodology to find the theoretically best model to solve these problems, which is subsequently applied in Chapters 4-6. In Chapter 7, the various aspects of the model structure will be integrated into an full, theoretically ideal travel demand model. Chapter 8 subsequently contains a case study which turns the theoretical model into a practical model for the Île-de-France region and analyses its properties. Finally, conclusions and recommendations will be given in Chapter 9.

2. Problem description

As said in the introduction, existing travel demand models pose a number of problems relating to the modelling of public transport, that limit the general applicability of their model structures with respect to multi-modal trips. This chapter describes these problems in more detail.

The problems will be illustrated by the example cases of the GroeiModel (Significance, 2012a) and ANTONIN (Syndicat des transports d'Île-de-France, n.d.; Willigers & Tuinenga, 2007). The GroeiModel is the engine for the Dutch national and regional travel demand models, while ANTONIN is the travel demand model for the French Île-de-France region. Descriptions of these models that can be used to trace the causes of problems mentioned here, are given in Appendix A.

2.1. Aggregation of modes

Both the GroeiModel and ANTONIN aggregate modes by merging their networks. The GroeiModel merges bus, tram and metro, while ANTONIN classifies a large number of systems as either train, metro or bus. The following tables illustrate this:

Modes in reality	Modes in GroeiModel
Train	Train
Bus	
Tram	Bus/tram/metro
Metro	

Modes in reality	Modes in ANTONIN
Transilien	Train
RER	
Metro	Metro
Tram	
RATP Paris bus	
RATP banlieue bus	Bus
Optile bus	

Table 1: aggregation of modes in GroeiModel and ANTONIN

While this simplifies the model structure, intermodal differences within each aggregated mode are lost; this may lead to loss of accuracy in modelling.

In the GroeiModel for example, it is not possible to differentiate the valuation of bus and tram systems, while Dutch stated preference research shows that such a difference does exist (Bunschoten et al., 2012). In this example, the number of travellers in the model would decrease after upgrading a bus line to a tram line – the service frequency is usually lowered in such cases – while the usually more important increase in comfort is neglected. The GroeiModel can also not predict that in cases with competing bus and tram lines, the tram line has an intrinsic advantage.

In Île-de-France, stated preference research indicates that the metro might have more in common with buses than with trams (Significance, 2013); this raises questions about whether the aggregation of modes in ANTONIN is correct.

2.2. Permutations of modes

Because the GroeiModel models access and egress for trains instead of for PT as a whole, and because ANTONIN in some cases excludes car access/egress, in both models several mode permutations (i.e. orders in which modes are used) are unavailable. The following table illustrates this with some examples:

Mode permutation			GroeiModel support	ANTONIN support
Car	Train	Metro	✓	*
Car	Train	Bus	✓	✗
Car	Metro	Bus	✗	✗
Car	Metro	Train	✗	*
Train	Metro	Train	✗	✓

* Supports car drivers, not car passengers.

Table 2: examples of mode permutations with three modes for outbound trips of tours

The table shows that both models contain assumptions on how combinations of modes are used that restrict the possible choices. For ANTONIN, it is visible that the choice for an end leg (e.g. bus) may even influence the start leg choice set (e.g. car becomes unavailable), while the middle leg stays the same (e.g. train).

For the GroeiModel, it seems that the permitted permutations correspond to logical permutations[1]. For ANTONIN, some little used permutations are unavailable (e.g. car passenger – train – metro), while some more rare permutations are also available (e.g. train – metro – train).

However, if in the current models the previously mentioned aggregation of modes is reduced, permitting all permutations of modes gets more difficult due to the large number of them and explicitly specifying all mode combinations is both inefficient and restrictive (Van Eck, 2011).

2.3. Changes in the number of modes

Travel demand models are frequently used to assess the impacts of projects to improve the transport system. Problems arise when a new public transport mode is made available by such a project. There are basically two strategies to add the services to the model:

- Add the new PT services to a network of an similar existing mode. In this case, one cannot take differences between the existing and the new mode into account.

[1] Note that one should take differences between the Netherlands and Île-de-France into account when judging what is logical.

- Add a new mode to the model, thereby estimating the parameters of the new mode using a stated preference experiment. However, the interaction with existing modes then becomes the next problem. There may be similarities with existing mode alternatives that should be taken into account to prevent distortion of the choice model, and the model structure needs to be adapted such that all new combinations of modes become available. This is quite difficult to solve.

For example, for the IJmeerlijn, a possible future rail connection in the Netherlands between Amsterdam and Almere, there has been discussion whether in the GroeiModel it should be modelled as a metro service, thus bus/tram/metro, or as a train service (Werkmaatschappij Amsterdam-Almere, 2012). This binary decision determines the model parameters that determine the attractiveness of the connection as well as the possibilities for access and egress. Instead inserting the IJmeerlijn as a new separate mode to the model structure with its own valuation would be very cumbersome. Using the GroeiModel to predict the number of travellers of the IJmeerlijn thus requires harsh assumptions about the properties of the mode, making it difficult to get an accurate prediction.

2.4. Consistency in the choice processes

Many travel demand models are based on the principle of utility maximisation and this is also true for the GroeiModel and for ANTONIN. More specifically, so-called logit models are often used, in which the alternatives in a choice set have Gumbel-distributed stochastic utilities. The utilities may be positive or negative, since only the mutual differences are important. Negative utility can be referred to as disutility, where utility maximisation equals disutility minimisation.

The utilities are stochastic variables to account for heterogeneity in preferences among travellers (Daly, 2012), meaning that individual decision-makers may attach different utilities to a particular alternative, as well as to account for unobserved factors influencing the decision-making process (Train, 2002), meaning that the researcher does not know all factors taken into account by the decision-makers.

In nested logit models, similar alternatives are grouped in so-called nests to take the positive correlations among their utilities into account – similar alternatives have positively correlated utilities, since a personal preference for one of the alternatives suggests this person may also have a preference for similar alternatives. This is necessary to correctly calculate the probabilities that certain alternatives are chosen.

Specifically in case of travel demand models with public transport, the eventual alternatives that are to be chosen, are (multi-modal) routes, which thus possess a particular utility as a function of the route characteristics. Other chosen trip characteristics, such as destination, the time-of-day and the modes to be used, are all reflected in the choice for a particular route. This means that these other trip characteristics can and are used to group the eventual route alternatives. The travellers, who are the decision-makers, are then assumed to make rational choices from the available routes, based on how they perceive the utilities of the alternatives.

In general, in nested choice models, there are three ways in which a group of options with utilities each can be aggregated to an overall utility of the particular nest of options. These are:

$$average \le maximum \le logsum$$

The average represents a situation in which there is no freedom of choice: people just have to accept a random one of the options and the associated utility. If a new option is introduced, the average utility may go both up and down. In a situation with freedom of choice, the utility would never go down, since people would not choose an inferior option. Instead, people would choose the best option, so the maximum is then a better measure of overall utility.

However, once heterogeneity in preferences is taken into account, the best option is not the same for everyone. In a logit model, this is represented by the Gumbel error term in the utility of each option. This results in a logsum for overall utility that is higher than the maximum utility of each of the options[2], particularly when the utilities of the options have similar utilities. This is realistic, since people have more options to choose from and the heterogeneity in preferences makes people benefit more from this freedom (Daly, 2012). Alternatively, from the perspective that the error terms represent unobserved factors, this higher overall utility is also appropriate (Train, 2002), since a larger group of options increases the chance that its best option contains an unobserved benefit. Hence the logsum is the ideal method of calculating the utility of a group of alternatives.

On the other hand, the benefit of having freedom of choice, such that each decision-maker can select his own optimal alternative, taking observed and unobserved factors into account, should not be overestimated by neglecting further positive correlations within the group itself between the utilities of multi-modal route alternatives (Van Nes & Bovy, 2008); the logsum of alternatives sharing physical network links or sharing some of the used modes is not as large as fully independent alternatives with similar travel times and costs, i.e. overlap should be taken into account. For example, for a group of public transport routes, the benefit of having a choice between two buses is not as large as the benefit of having a choice between a bus and a metro, and this difference should be taken into account.

In the GroeiModel, for train users, the assignment is incorporated in the nested logit choice model. This means that the logsum of the train services works its way up into the mode, destination and other choices. Since the travellers have freedom of choice, this method is correct. However, for the bus/tram/metro network, the shortest route algorithm unrealistically implies that there is no heterogeneity in preferences, contrary to the other choice processes, possibly leading to underestimation of the attractiveness of PT. Also, route overlap is not accounted for, neither for bus/tram/metro route parts nor for train route parts, possibly leading to overestimation of the attractiveness of PT.

For ANTONIN, the route choice logsum from the assignment is not propagated to the rest of the choice structure; instead, the weighted average utility of the route set is propagated. This means that there is heterogeneity in preferences at both levels in the model, but that this heterogeneity at the lower level is not consistently taken into account in decisions at the higher level. A new alternative at the lower level may even deteriorate the utility at the higher level. Practically, this means that the PT market share may deteriorate after opening a new PT

[2] ‘Maximum’ here refers to the maximum of the expected utilities $\max_{i\in I} \mathrm{E}\left[U_i\right] = \max_{i\in I} V_i$ and is not to be confused with the expected maximum utility, that is the logsum $\mathrm{E}\left[\max_{i\in I} U_i\right] = \ln \sum_{i\in I} e^{V_i}$.

line, and that the number of available PT alternatives has in general no positive effect on the total attractiveness of the PT system. Hence the model is inconsistent at this point. Similarly, the choice model for PT pass ownership is not consistent with the gain in logsum in the travel choices. Again, physical route overlap is not taken into account.

2.5. Computation time

Finally, one has to keep an eye on computation time. In ANTONIN, the required computation time for calculating the level-of-service and performing the assignment is quite high, since this is done for each mode combination separately. For the GroeiModel, this is currently not problematic, but it may become problematic if the other mentioned problems are solved in an inefficient way.

2.6. Conclusion

The above problems of current models seem to be divided over three categories. Models incorporating public transport lack *flexibility* in the modelling of how travellers value and use the available modes (Sections 2.1-2.3), lack *consistency* in how the choice processes work (Section 2.4), and/or lack *efficiency* with respect to computation time (Section 2.5). While it is probably possible to identify more problems of current models, this research project focuses on solving the previously mentioned problems. By doing so, the area of application of travel demand models is extended.

It should be noted that the flexibility problems relate to the ambition with respect to intended applications of the model. If the modeller is only interested in some particular mode, less correct treatment of the other modes may be acceptable. The GroeiModel for example focuses on the correct modelling of car and train, and not on bus, tram and metro. Also, if the available (e.g. network) data for some of the modes is of low quality, or if there is not enough stated or revealed preference data to estimate the parameters of the designed model, a theoretically more accurate model may not lead to more useful results.

For this research project, it is assumed that one indeed wants to maximise the flexibility of the model, additional to maximising the consistency and efficiency. Therefore, the next chapter proposes a methodology to design a model that avoids the problems mentioned in this chapter. The problems can thus be reformulated as requirements for the model to be designed:

- aggregation of modes must not be necessary;
- all possible permutations of modes must be available;
- changes in the number of modes must be possible;
- the choice process must be consistent;
- the computation time must be reasonable.

3. Model design methodology

As stated in Chapter 1, a mode and route choice model structure for a travel demand model should be designed, thereby avoiding the flexibility, consistency and efficiency problems described in Chapter 2. This short chapter states how the theoretically best model structure will be sought for in subsequent chapters.

The model design problem can be subdivided into three smaller problems. These problems are related since the chosen solution for one of these problems influences the other problems. The problems are:

- *The interface between mode and route choice*. This determines the main structure of the model, i.e. what choices are made at network level and above network level, including what the choice models above network level should look like and whether they include full, partial or no mode choice; it should be designed subject to the posed flexibility requirements and taking correlations between alternatives into account.

- *The representation of the networks.* This determines the network structures that should be able to represent both private and public transport modes at the maximum precision that the varying data availability allows, thereby specifying how timetables of public transport should be handled if they are known, that should facilitate all relevant network-level mode choices, that should facilitate efficient path-finding and that should define a measure of overlap between routes.

- *The route set generation and route choice.* This determines how reasonable sets of reasonable paths are found in the generally multi-modal network structures, subject to the constraints posed by non-network-level choices, and how the choice between these routes is made using the measure of overlap corresponding to the network, taking timetables of public transport services into account.

In Chapters 4-6, these three problems are respectively addressed, enumerating and comparing solutions for each of them, based on existing models, scientific literature, and, incidentally, brainstorming. This discussion results in an ideal model framework that is subsequently presented in Chapter 7.

4. Interface between mode and route choice

In this chapter, the possibilities for connecting mode choice and route choice in a travel demand model will be investigated. As stated in Section 2.4, the selection of a multi-modal route is the final product of the choice process of a particular traveller; it is the final step in the choice tree. This is a choice at network level: the route alternatives are generated in networks. This chapter looks at the various options that exist for inserting a mode choice step before turning to the networks, which is common in many existing travel demand models, in order to find the choice model structure that suffers the least from the problems mentioned in Chapter 2.

Section 4.1 starts with an overview of the various potential model structures, based on literature and existing travel demand models. Each of these structures is further analysed in Sections 4.2-4.5, including examples of existing models that utilise the particular structure. Finally, Section 4.6 will assess and compare the structures based on whether the problems set out in Chapter 2 occur.

4.1. Overview of model structures

According to Fiorenzo-Catalano (2007), there exist three main approaches to construct a travel demand model with multiple transport modes[3]: the classical uni-modal approach, in which a main mode is chosen; the extended classical approach, in which a combination of modes is chosen; and the supernetwork approach, in which the mode and route choices are simultaneous. Additionally, an advanced version of the classical uni-modal approach is obtained if access/egress mode choices are nested within the main mode choice. The following figure summarises these four approaches:

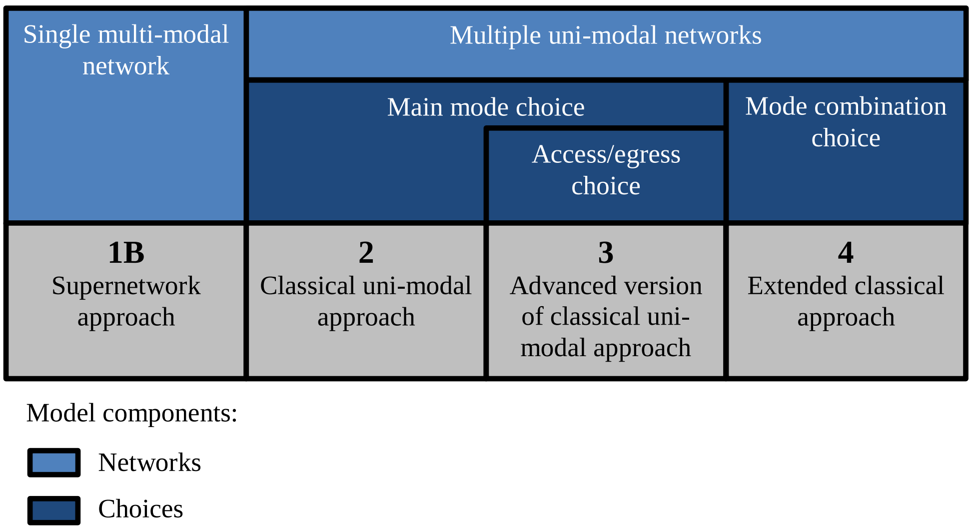


Figure 1: theoretical model structures as may be derived from Fiorenzo-Catalano (2007)

[3] Fiorenzo-Catalano (2007) identifies five approaches, but the other two differ only in the way congestion on the network(s) is handled.

However, this framework has a shortcoming when looking specifically at public transport modelling. Some models contain public transport as a single mode, while others have multiple modes for different PT systems. Since this has a large impact on the way public transport is modelled, a further extension of this framework is necessary.

The required extension can be realised by replacing ‘mode’ with ‘network’ in the descriptions of the classical uni-modal and extended classical approaches, where a single network may span multiple modes coded as distinct links, and *defining public transport to be a single, multi-modal network*. Next, a distinction is made regarding the number of modes that may occur in a network.

A mode then is anything smaller than or equal to a network with one or more distinguishing characteristics with respect to other modes. This may range from vehicle modes, i.e. systems characterised by production technique, to service modes, i.e. systems characterised by level of service (Combes & Van Nes, 2012). Because of the variety in definitions of ‘mode’ in existing models, no further restrictions will be placed on this definition here.

The above reasoning extends the number of model structure categories from four to seven (plus one structure without multiple modes at all), as shown in the following figure:

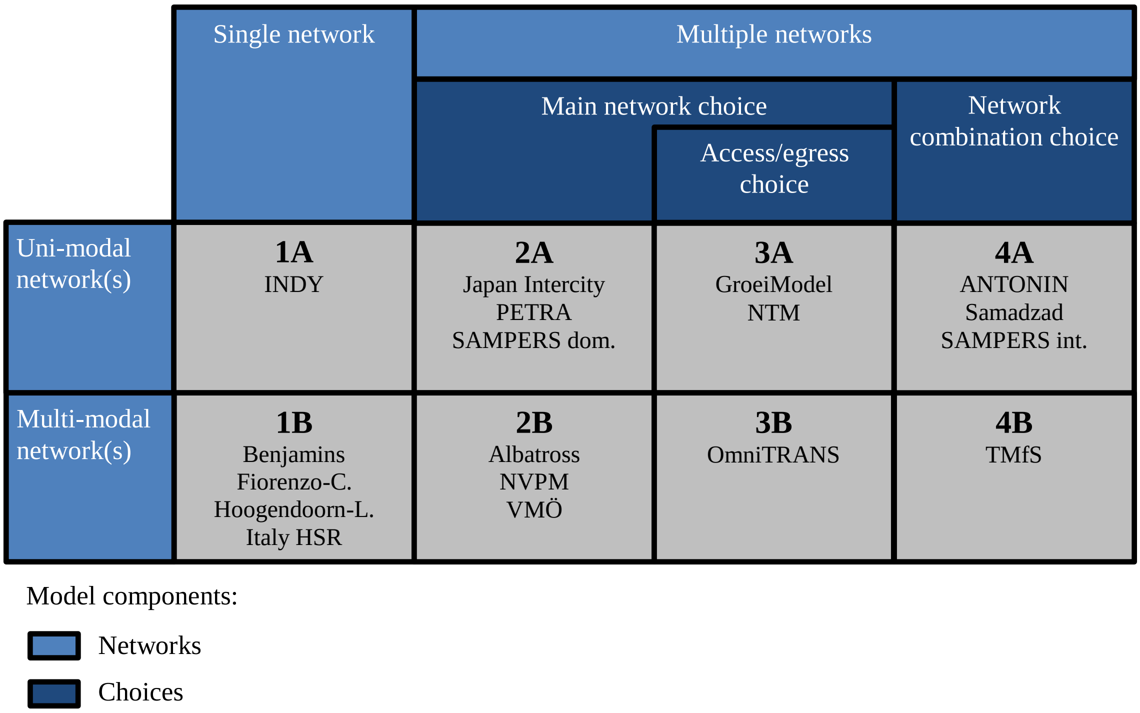


Figure 2: theoretical interfaces between mode and route choice, including examples

The choice processes in this figure determine the chosen networks or network combinations and the chosen origins and destinations within these networks or network combinations – access/egress choice here includes both mode choice and transfer station/stop choice. Consequently, all choices in the figure must be iterated over to perform route set generation, i.e. one or more routes are generated for each iterated choice separately. Additional choice processes are allowed to be added to the model, for example to group the generated routes according to certain criteria.

All choice trees in this chapter thus indicate some part of the choice tree that can be drawn before the routes have been generated. Afterwards, the route set generator iterates over all final decisions in this partial tree, together *defining the interface between mode and route choice*, to generate zero, one or more routes for each of these possible decisions to complete the choice tree – if no routes are found, the corresponding decision is simply eliminated from the tree. The generated routes are inserted into the choice tree below the decision they were constructed for[4], and the utilities of the routes do not depend on characteristics of routes generated for other decisions.

To verify the completeness of this framework, several existing models have been consulted. A list of these is provided in Appendix B. All of these models could indeed be classified; they are listed in the figure as examples.

In the following sections, the possible model structures are described in more detail. However, before proceeding, it is advisable to understand how logit models work, which are commonly applied in each of the structures. A detailed overview of logit models can be found in Appendix C; for the purposes of this chapter, the following ones are important as they occur in examples of the various model structures:

- The *multinomial logit* model assumes independent utilities of the alternatives available to choose from.
- The *nested logit* model groups similar alternatives together in nests, such that their utilities become positively correlated.
- The *cross-nested logit* model extends the two-level nested logit model by allowing alternatives to belong to multiple nests. This is useful if the alternatives can share different characteristics with each other.
- The *network GEV* model extends the cross-nested logit model with multiple levels of (cross-)nesting, thereby facilitating even more complex choice trees.

4.2. Classical uni-modal approach (1A/2A/3A)

In the classical uni-modal approach, a single main mode is determined for each trip. The structures 1A, 2A and 3A fall into this category. Each of these is described below.

Single uni-modal network (1A)

Structure 1A is the simplest possible structure within this framework. It contains all models having a single network for a single mode, without any mode choice. This includes assignment-only models (e.g. INDY (Bliemer et al., 2004)), but also uni-modal elastic demand models.

Due to its nature, this structure is obviously unsuitable for multi-modal demand modelling.

[4] Insertion into the choice tree does not necessarily mean that route choice is included in the nested logit model of the prior choices. ANTONIN and the GroeiModel for example use a completely separate program component for route choice (see also Section 2.4). In case of such separation, the route choice is said to be 'network modelling' while the prior choices are said to be 'demand modelling'.

Multiple uni-modal networks (2A)

Structure 2A is an extension that includes a main mode choice for travellers to decide which network to use. It does not include a choice for access and egress modes.

However, generalised costs for access and egress may still be taken into account in the utility functions. In this case, the model must assume predefined access/egress modes (e.g. walking, or bus if the main mode is train) and predefined stations or stops (e.g. closest to the origin/destination) for all users.

The choice of the main mode itself is usually a multinomial logit model, but it may also be a nested logit model to account for correlations between similar modes. For example, the following figure shows the nested mode choice structure of the local and regional trip model of SAMPERS (Algers et al., 2000):

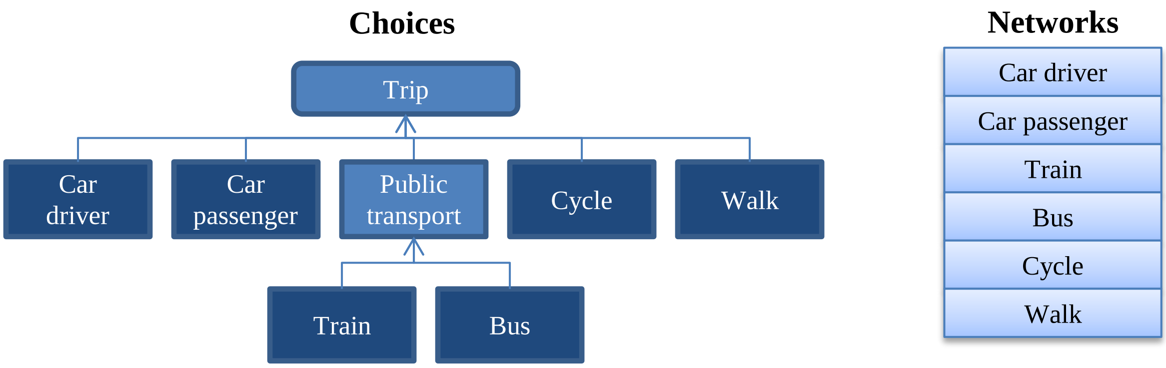


Figure 3: nested mode choice structure of the local and regional component of SAMPERS

Although this structure allows travellers to choose a mode, it does not allow the usage of multiple modes in a single trip.

Multiple uni-modal networks with access/egress (3A)

In structure 3A, access and/or egress mode and transfer stop choice is appended to the choice tree for some main modes, where a stop may also refer to a station. The result is a nested logit model.

Note that this involves the generation of choice sets for mode and stop choice. This can be based on distance to the origin/destination, impact on travel time and/or stop characteristics and will be further discussed in Section 6.2.

After the modes and stops are chosen, separate uni-modal assignment procedures are used for the access leg (from origin to access stop using access mode), the main leg (from access stop to egress stop using main mode) and the egress leg (from egress stop to destination using egress mode).

An example is the GroeiModel structure (Significance, 2012a), which models access and egress if train is the main mode. For both access/egress mode and stop choice, it combines the choice sets at the home-end and the activity-end. There is a logical constraint that excludes car driver from being used as activity-end mode. The choice model is shown in the following figure:

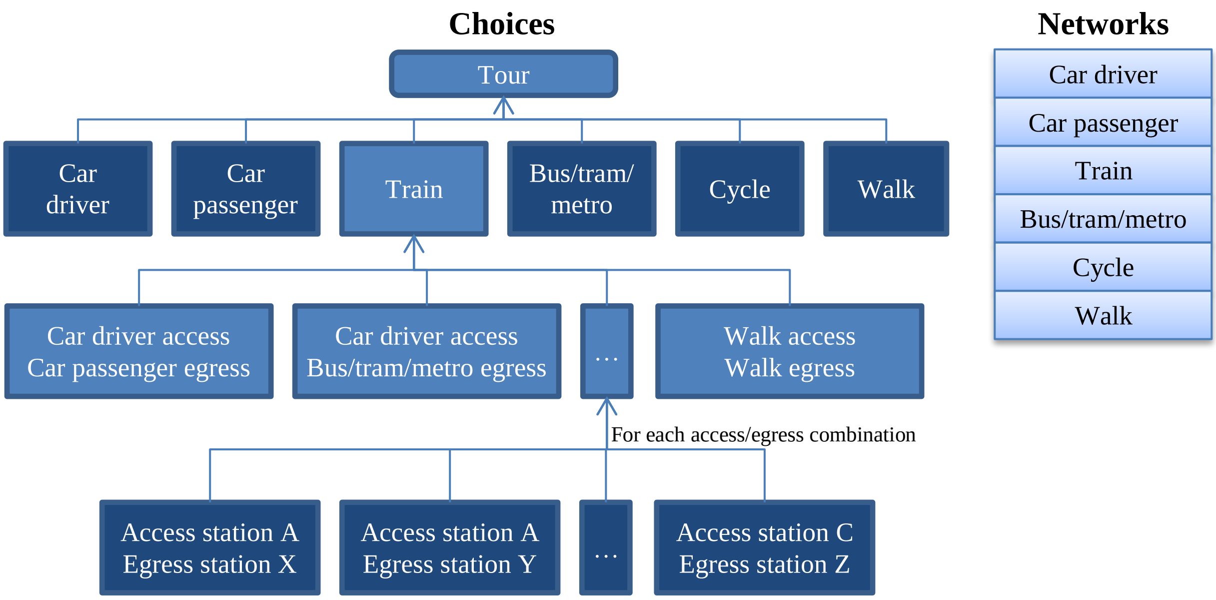


Figure 4: GroeiModel mode choice structure, where 'access' refers to the home-end and 'egress' refers to the activity-end (other choices have been removed for clarity)

This structure allows multiple modes to be used in a single trip, although this is necessarily restricted to three legs.

The structure also assumes a fixed hierarchy in modes: in the GroeiModel example, train is placed higher in the hierarchy than bus/tram/metro. This implies that the train leg is the most important part of the trip; if not, it will be problematic that the overlap between access/egress legs of one main mode with another main mode is not accounted for. This problem is illustrated in the following figure:

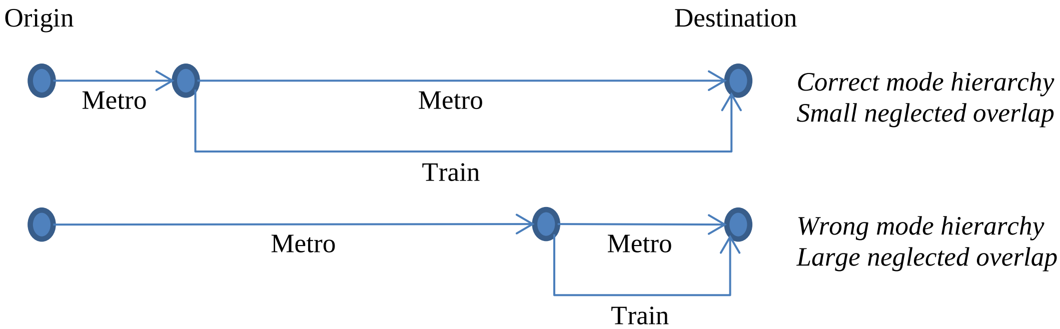


Figure 5: illustration of neglected overlap between alternatives, if train is higher in the mode hierarchy than metro

In situations in which a large number of modes is available (e.g. split the joint bus/tram/metro mode into three separate modes), the choice tree becomes quite complex and both the restricted number of legs and the hierarchy of modes may become problematic.

Note that the GroeiModel structure could be extended to account for modal correlations between alternatives with identical access legs and identical egress legs by inserting additional nesting levels using the cross-nested logit model, as illustrated in the following figure:

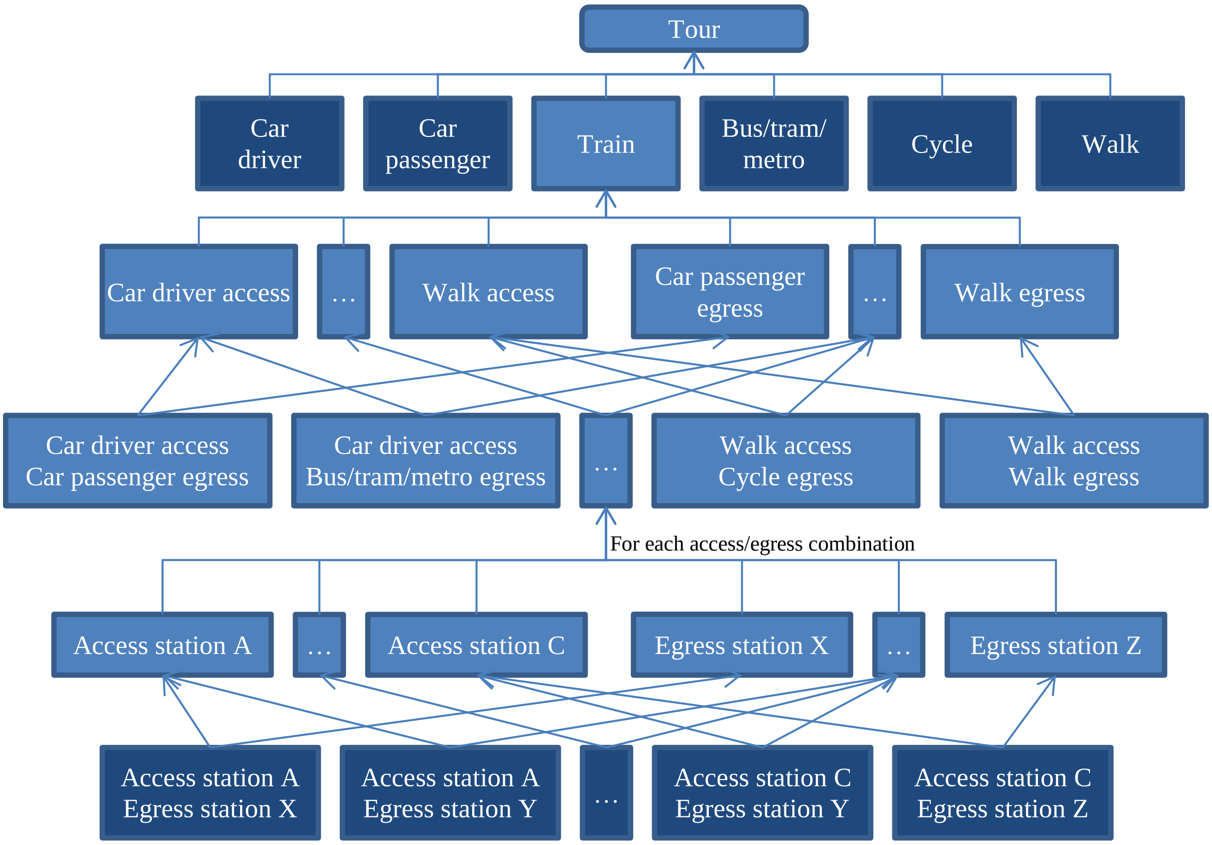


Figure 6: fictional cross-nested logit model for main mode choice with access/egress choice

The access/egress mode choice levels of this structure match the cross-nested logit model of Hoogendoorn-Lanser (2005), which is classified as structure 1B (see Section 4.4) because the choice tree is only constructed to classify the already known multi-modal routes of the route set. This model outperformed all other nesting structure, implying that this structure might indeed be better than the current GroeiModel structure.

4.3. Extended classical approach (4A)

Instead of adding access and egress modes besides a main mode, combinations of modes may also be inserted into the choice process as new artificial modes. This means the original networks of several modes are joined into a new network combination by inserting transfer links. Contrary to structure 3A, no single main mode is chosen by travellers, but only a combination of modes. The subsequent assignment uses the corresponding network combination. This is structure 4A, the extended classical approach.

In general, the order in which modes are used in the trip is not specified. However, by disabling certain transfer links and connectors in the combined network, the mode combination choice may contain constraints on which mode is used first or last.

A simple example of structure 4A is the international model of SAMPERS, which contains a combination of car driver and ferry as an additional alternative to the single modes:

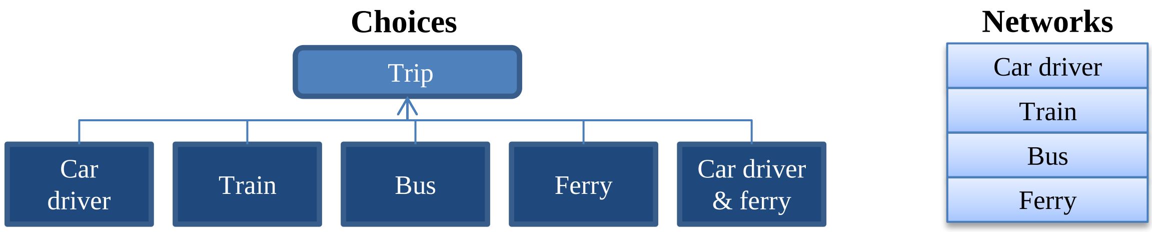


Figure 7: mode combination choice structure of the international component of SAMPERS

Alternatively, this could be designed as a cross-nested logit model[5] in which the combination of car driver and ferry belongs to both the car driver nest and the ferry nest. This can be used to capture the correlations between the individual modes and the mode combination:

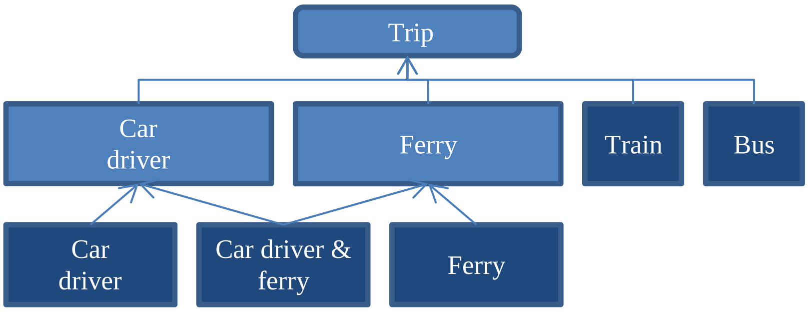


Figure 8: fictional cross-nested logit model for mode combination choice

Although the idea seems simple, the complexity of this structure increases rapidly with the number of modes that need to be combined. An example of this is ANTONIN (Syndicat des transports d'Île-de-France, n.d.; Willigers & Tuinenga, 2007; Citilabs, 2008), which contains many combinations of modes. In case of public transport, ANTONIN also contains restrictions on the first and last modes in the chain (i.e. access/egress modes). The model structure is displayed below:

[5] For a description, see Appendix C.4.

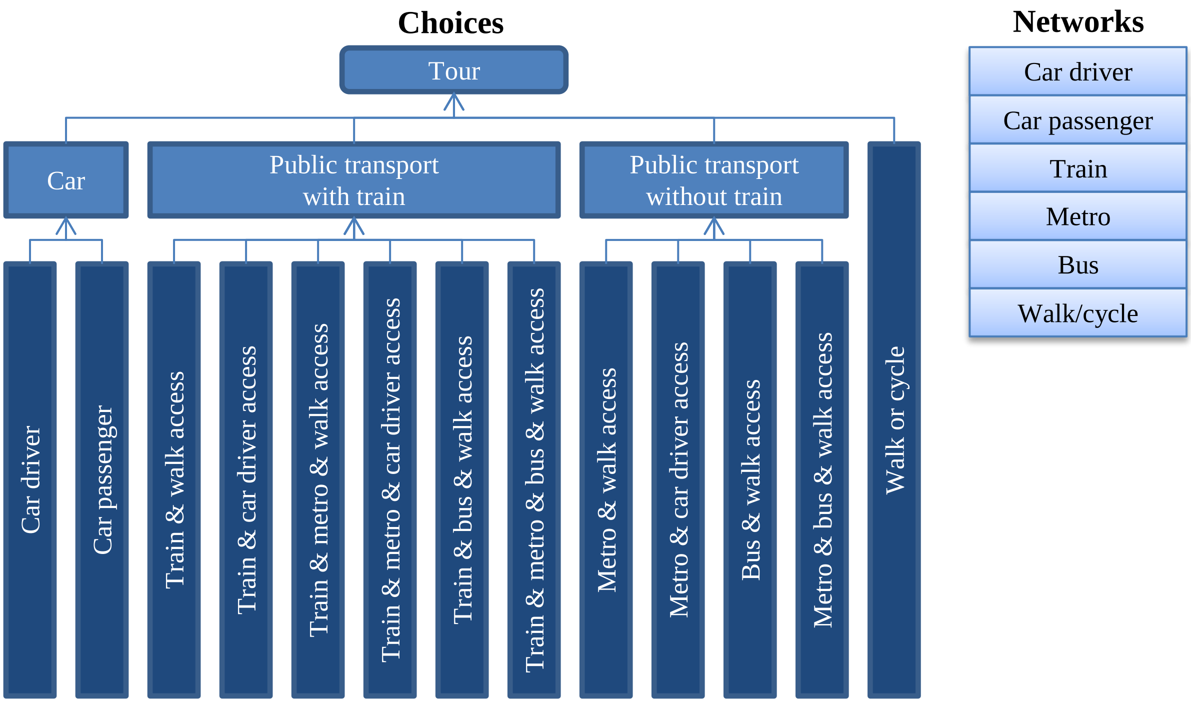


Figure 9: mode combination choice structure of ANTONIN (public transport always has walk egress)[6]

Although the extended classical approach does not have a mode hierarchy problem from structure 3A, the choice tree gets very complicated with a large number of modes and mode combinations may have to be eliminated to reduce computation time, as in ANTONIN.

It may also be difficult to capture correlations between mode combinations correctly. Nesting structures as in ANTONIN may to some extent capture correlations in preferences for mode combinations, but not route overlap, since all nests must be independent. Figure 5 applies to structure 4A as well, since metro on its own and the combination of train and metro are modelled as independent mode combinations.

4.4. Supernetwork approach (1B)

Structure 1B represents the supernetwork approach. In the pure supernetwork approach (e.g. Benjamins (2001), Fiorenzo-Catalano (2007)), the mode choice is entirely delegated to the route choice; this means that there is no mode choice process prior to the assignment, as there is a single network that all travellers are effectively forced to choose:

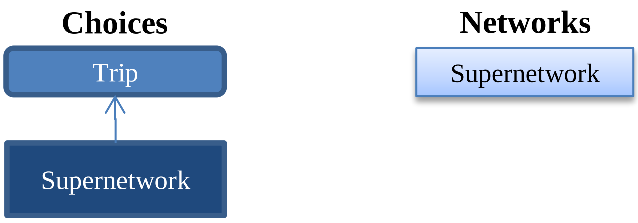


Figure 10: choice structure of supernetwork models

[6] Some mode combinations are unavailable for some purposes. All nests have the same nest coefficient (the reason for this is unclear; theoretically the correlation between mode combinations can differ per nest); for most purposes, the nesting has been removed by setting the coefficients to one.

The route finding process thus does not have more stringent route constraints than the eventual constraints that apply to all routes; this means that all modes can be used in any order, as long as the route set generation process creates a route set with sufficient modal variety. Similar to the extended classical approach, the route finding takes place in a joint network; however, this joint network (the supernetwork) includes all modes.

Because there is no mode choice separate from route choice, mode-specific constants need to be attached directly to routes. They can be placed on transfer links in the network (e.g. Benjamins). Other than that, travel times and distances on normal links can easily be valued differently depending on the mode they occur in.

Alternatively, the found routes may be classified according to the used modes after the routes have been generated (e.g. Italy HSR (Cascetta & Coppola, 2012), cross-nested logit model of Hoogendoorn-Lanser (2005)). For example, a cross-nested logit model may be constructed for access and egress modes (e.g. Hoogendoorn-Lanser). This allows for modelling of overlap between mode combinations, while the route set generation is not repeatedly constrained to get different mode combinations, which is the crucial difference with the classical approaches.

4.5. Hybrid approaches (2B/3B/4B)

The supernetwork approach may also be combined with one of the classical approaches; in that case, modes are grouped together (e.g. a public transport group). The classical approach is used to choose groups instead of modes. The mode choice within this group is delegated to route choice in the multi-modal supernetwork of the group.

Note that a group may also contain just one member mode. A typical approach is to group all public transport modes together, but leave the private modes separate. Within the public transport network, different modes can be valued differently according to supernetwork principles.

Multiple multi-modal networks (2B)

Structure 2B is an improvement over structure 2A in that within a group of modes, multiple modes can be combined in a single route. Hence, not one of the modes, but one of the networks is chosen by the travellers. This typically means there is a single public transport network, such that multiple public transport systems can be combined within this network. However, public and private modes still cannot be combined, since they are part of different groups.

Note that although it is possible to attach different values to usage of different PT modes in the supernetwork, some models simply don’t consider differences between PT modes (e.g. Albatross).

Multiple multi-modal network with access/egress (3B)

Structure 3B is an improvement of structure 3A that is obtained by substituting a mode for the public transport network as a whole. This means that for the public transport network as a whole, private access and egress modes may be chosen (e.g. OmniTRANS (Veitch & Cook, 2010)). Hence the number of legs is no longer a problem, since the number of public transport modes is unlimited and, if short transfer walking links between nearby stops are included in the PT supernetwork, one private mode on each end of the PT system is sufficient.

The mode hierarchy problem is also largely solved, because PT will generally be the main network, if it is chosen; the only exception perhaps being park & ride situations where a long distance may be travelled by car. The problem of route overlap between different alternatives is reduced (e.g. route overlap in Figure 5 can be accounted for since both routes are in the same nest), but not completely eliminated (e.g. route overlap between car and public transport with car access is not accounted for).

Network combination approach (4B)

Structure 4B is an extension of structure 4A: a network combination is chosen by travellers, where one of the networks is the PT system. This is very similar to structure 3B, where main, access and egress networks are chosen.

Because of the reduced number of networks that need to be combined, the structure is much simpler than structure 4A. An example is TMfS (Johansson, 2009; Robinson & Pollard, 2009), which contains a park & ride alternative:

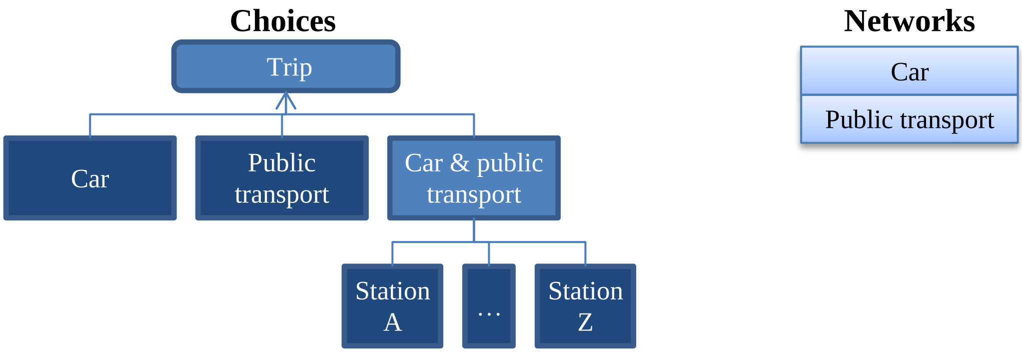


Figure 11: network combination choice structure of TMfS including a park & ride option

Note that although a transfer station choice is included here, a transfer station choice before the assignment is not necessary; this remark also applies to structure 4A. Like structure 4A, a cross-nested logit model might also have been used:

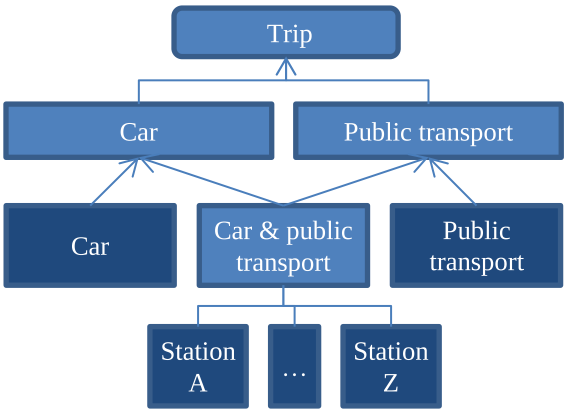


Figure 12: fictional cross-nested logit model for network combination choice

Although the number of required nests is smaller, the problem of structure 4A that the route overlap between nests is not modelled, persists.

4.6. Comparison of structures

Now that an overview of possible model structures has been given, these will be checked against the flexibility and consistency requirements stated in Chapter 2.

Comparison regarding flexibility

Combined usage of modes should be possible and this should also be possible once a new mode is added, i.e. the model structure should be flexible in supporting multi-modal routes.

The supernetwork and hybrid approaches appear to be ideal for this situation: the modeller only needs to add the new network to the public transport network and properly set the parameters for the valuation of the new PT mode. On the contrary, the classical and extended classical approaches require careful consideration on which modes can be used in which chains and specifically the extended classical approach requires large sacrifices in computation time.

Structures 1B, 3B and 4B are thus the most flexible with regard to public transport modelling. This is indicated in the following figure:

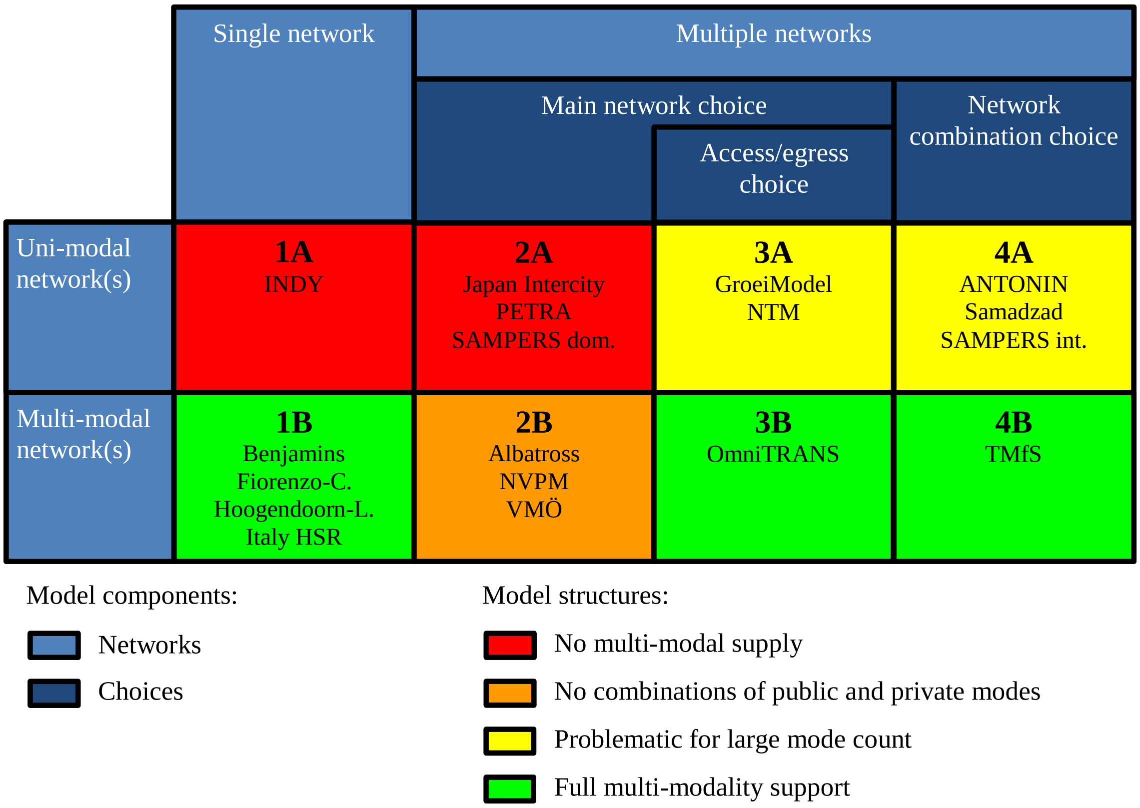


Figure 13: theoretical interfaces between mode and route choice assessed on flexibility

Remember that the choices in this figure are iterated over in order to generate routes. The route generators of the flexible models thus iterate over the following choices of the choice model:

- nothing, since there is a single supernetwork (1B); or:
- networks with access and egress options, where PT is a single network (3B); or:
- combinations of networks, where PT is a single network (4B).

Comparison regarding consistency

Regarding consistency, so far mentioned two types of overlap have been mentioned that may exist between route alternatives: route overlap (i.e. routes share network links) and modal overlap (i.e. routes have a mode in common). An example of route overlap is a single train connection with either bus access or walk access, which share the network link of the train connection. An example of modal overlap is two distinct bus routes, which have the bus mode in common.

Modal overlap may exist without route overlap, but not the other way around, assuming the influence of infrastructure being shared by multiple modes may be neglected. The following figure illustrates both types of overlap:

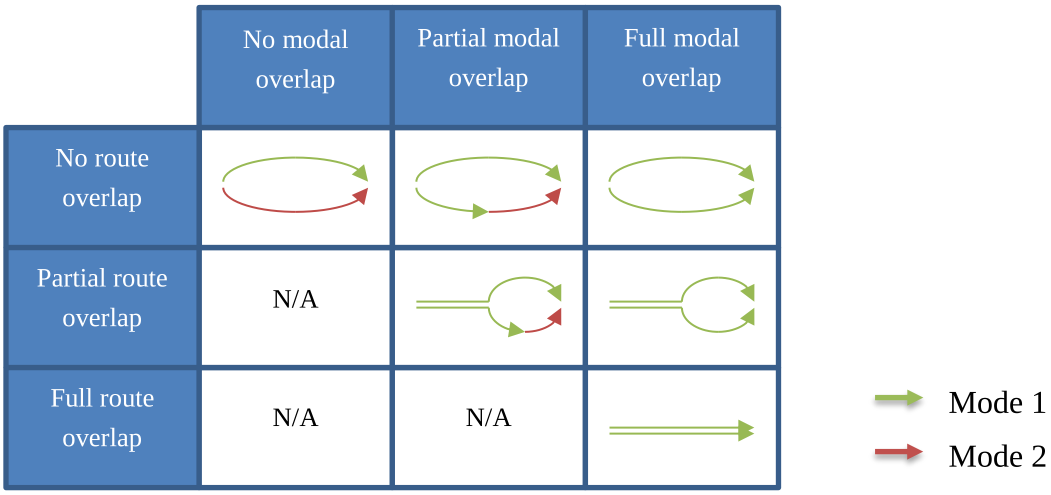


Figure 14: modal overlap and route overlap for a pair of multi-modal routes

Ideally, both are taken into account in the model. If modelling only one of them and subsequently estimating the model parameters, it will effectively serve as a proxy for both. However, if only one of these overlap types can be included in the model, the results of Hoogendoorn-Lanser (2005) suggest that including modal overlap is more important than route overlap.

To complicate things further, there exists a *third type of correlation* between route alternatives, namely similarities between modes themselves, which is why some models shown in this chapter use nested logit models for mode choice. A classic example of this is the red-bus-blue-bus problem (Sobel, 1980). A route with a red bus and a route with a blue bus do not have modal overlap, because the red bus and the blue bus would be considered different modes, but there still is correlation because the modes are similar. Another example is a bus route and a train route, which are similar because they both are public transport routes.

Although modelling all three types of correlation is trivial for uni-modal models, multi-modal trips make this more difficult by introducing partial modal overlap. While handling of route overlap in route choice is discussed in Chapter 5 and Section 6.6, there should be a defined way to incorporate this into the larger choice model structures like in this chapter, i.e. integrating this with handling of modal overlap and mode similarities. Models in current practice only incorporate part of these three types of correlations.

For structures 3B and 4B, modal overlap is best taken into account by using a cross-nested logit model; in practice, normally nested logit models are common. However, they cannot take modal and route overlap across mode combinations into account, but only within them: for example, a train with bicycle access and a train with walk access, could only be correlated because they both use public transport, but not more specifically because it is a train in both cases or because the trains have route overlap. Although flexible, these structures are thus lacking full consistency regarding correlated alternatives.

While the pure supernetwork approach of structure 1B fully takes route overlap into account, it does not consider modal overlap nor mode similarities by default. However, as said, there is a variant of structure 1B that groups generated routes according to mode usage, in order to handle modal overlap; the cross-nested logit model of Hoogendoorn-Lanser is an example. If combined with a method to handle route overlap, both modal overlap and route overlap are taken into account, and nesting of modes can complete this list with mode similarities. This is discussed further in Chapter 7.

5. Network representations

Depending on the modes in a network, the structure of the model and the available data, different network representations may be chosen. This chapter reviews the possible network representations that can be used. Whenever the choice of a specific network representation has consequences for the multi-modal route choice model, this chapter also discusses these consequences.

Each network representation in this chapter can be described as a directed graph consisting of nodes N and links M; while some illustrations in this chapter contain bi-directional links, these may be read as two uni-directional links. Travellers enter and leave the network in origins and destinations $O = D \subseteq N$. Each link $m \in M$ contains various attributes $\mathbf{X_m}$; the attributes $\mathbf{X_i}$ of each route $i \in C_{od}$ from origin $o \in O$ to destination $d \in D$ using links R_i equals the sum of the attributes of the used links:

$$\mathbf{X_i} = \sum_{m \in R_i} \mathbf{X_m} \quad \forall o \in O, d \in D, i \in C_{od}$$

These route attributes may then be used in a choice model. Note that this does not imply that linear valuation of these attributes is required, as a non-linear transformation can still be applied to the final $\mathbf{X_i}$.

The chapter starts with simple network representations with low data requirements and gradually moves to advanced network representations with high data requirements. Section 5.6 lists conclusions.

5.1. Level-of-service matrix

The most simple network representation in a travel demand model is a level-of-service matrix. This matrix contains the disutility of travelling from each origin to each destination. It is often equivalent to the characteristics of a single optimal route, which minimises some linear combination of route attributes $\boldsymbol{\beta}^{\mathrm{T}}\mathbf{X_i}$ (e.g. Albatross, bus/tram/metro network in GroeiModel, road network in Hoogendoorn-Lanser model), but the logsum of a route set may also be used to account for heterogeneity among travellers.

No physical locations can be identified in this network representation, other than the origins and destinations: the network is a black box. The representation is visualised below:

Figure 15: level-of-service matrix network representation

If it is used, a level-of-service matrix is typically derived from another, more complex network representation. These other network representations are discussed below.

5.2. Traffic network

The traffic network representation explicitly contains the infrastructure that vehicles can use. Additional connector links are added to connect the zone centroids to the infrastructure. An example is shown below:

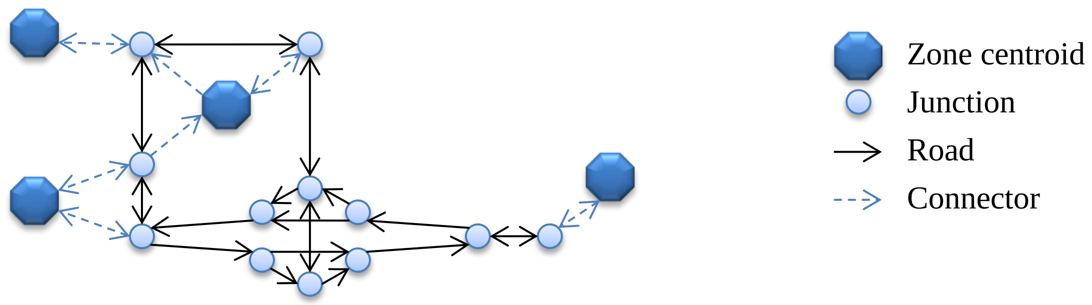


Figure 16: traffic network representation

The traffic network representation is suitable for private modes. However, for public transport, a traveller cannot use the traffic network freely; instead, the traveller can freely use a service network[7] (Bovy et al., 2006). Possible service network representations are discussed below.

Measure of route overlap

As seen in Appendix A, for the creation of a route choice model, one needs to define which z_m is used in defining the amount of overlap between two routes. In case of private modes, common expressions of overlap are road length and free-flow travel time, where free-flow travel time gives the best modelling results (Ramming, 2002). Using actual travel time instead of free-flow travel time would be problematic because the amount of overlap then depends on the conditions of the road network (Hoogendoorn-Lanser, 2005).

5.3. Frequency-based service network

In a frequency-based service network[8], public transport lines are represented as links between stations or stops (hereafter collectively referred to as stops) that are represented as nodes (Bovy et al., 2006). The only used information about the timetable is the frequency of all lines, which equals the inverse of the mean headway (Friedrich et al., 2001).

The stops are connected to the lines via boarding and alighting links (Bovy et al., 2006); the boarding link can contain the expected waiting time, which is usually assumed to equal half of the headway time of the line that is to be boarded (Friedrich et al., 2001). The lines themselves are represented by in-vehicle links, more specifically a combination of running

[7] An exception to this would be demand responsive transport systems (e.g. taxi, Superbus), which like other public transport systems do have a waiting time for boarding, but after that use the traffic network instead of a predefined service network, possibly according to predefined schedule constraints, with possible detours to share the vehicle with other travellers (Van Nes & Bovy, 2008). Such systems might generally be modelled using level-of-service matrices with expected waiting times and expected travel times.
[8] Frequency-based PT networks are sometimes called line-based networks in literature (Friedrich et al., 2001).

links between stops and dwelling links at stops. The modelling of PT lines is thus similar to the modelling of highways in road networks: the on- and off-ramps are simply replaced by boarding and alighting links.

The following figure shows an example network; note that it offers both an express service and a local service from the first stop to the third stop:

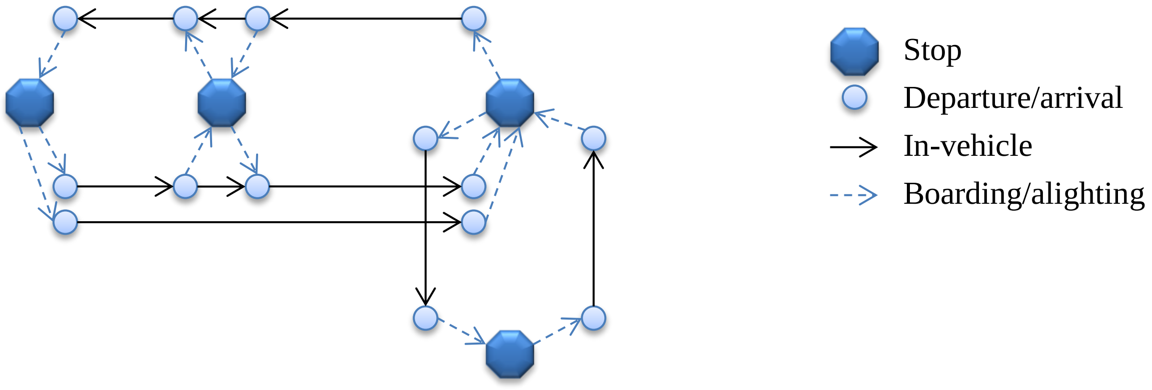


Figure 17: frequency-based public transport network representation

ANTONIN currently uses the frequency-based representation, because detailed timetable information is not known for the future.

Pre-processing

To speed up the route generation, the public transport network may be pre-processed to reduce the number of links that are necessary to construct a route (Friedrich et al., 2001). This involves replacing all links with single links that connect the pairs of boarding and alighting stops directly. These new links are called route segments.

The pre-processed version of the previous example network is shown below:

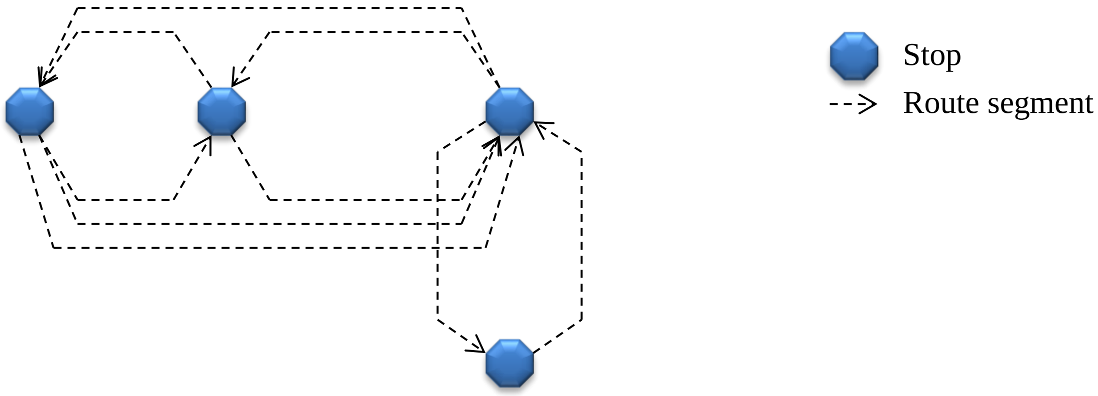


Figure 18: frequency-based public transport network after pre-processing

Whether the pre-processed version is more efficient than the original version, depends on the route generation algorithm (see Chapter 6). The pre-processed version has more links in total but uses fewer links per route.

Merging common route segments

The pre-processed network of Figure 18 may be modified further by merging multiple route segments between identical transfer stops to a single route segment (Bovy et al., 2006). This simplifies the network and thus route set generation, prevents the necessity to generate a possibly unrealistic large number of similar routes to cover all travel alternatives and seems a more realistic way of modelling travel behaviour due to the similarities between the alternate route segments. More precisely, the benefit of having multiple route segments available is modelled as decrease in expected waiting time rather than increase in choice freedom. For example, in a situation with two bus lines from the same bus stop to the same train station, these two bus lines are then no longer modelled as separate alternatives.

The question then is how the properties of the aggregate route segment should be specified. In literature, this problem is known as the 'common lines' problem (Kurauchi et al., 2003), since each route segment between two fixed stops represents a different line. A typical way of dealing with this is to replace each set of route segments by a single 'strategy[9] segment' (Cominetti & Correa, 2001). At a larger scale, this implies that the route choice is replaced with a strategy choice where a deterministic line choice within each transfer-free strategy segment is separated from the stochastic route choice. This leads to the following version of the example network:

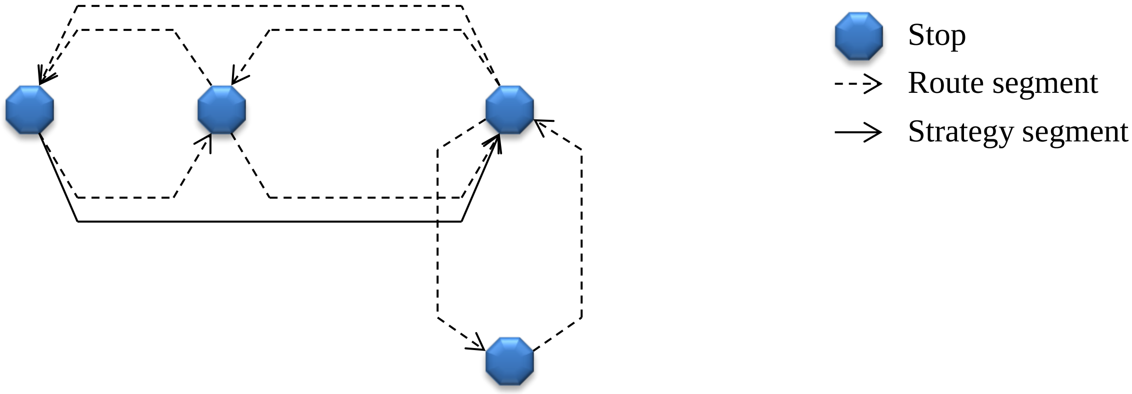


Figure 19: frequency-based public transport network after pre-processing and merging common route segments

Some of the lines may be so unattractive (e.g. due to detours) that travellers may prefer to wait longer for runs of other lines. The strategy segment is then determined by the optimisation problem answering the question which route segments of set C should be included in strategy segment $S \subseteq C$.

Assuming line usage proportional to their frequencies, the probabilities p_i of lines $i \in C$ and the waiting time t_w (Cominetti & Correa, 2001; Lam et al., 2002):

[9] In literature, the term 'hyperpath' is also used instead of 'strategy' (Cominetti & Correa, 2001).

$$p_i = \begin{cases} \dfrac{f_i}{\sum_{j\in S} f_j} & \text{if } i \in S \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in C$$

$$t_w = \frac{\phi}{\sum_{i\in S} f_i}$$

If travellers arrive randomly and board the first departing vehicle, line usage proportional to frequencies may represent both random departures and fixed-interval departures while travellers arrive randomly at the stop. The corresponding values of ϕ are 1 and $\frac{1}{2}$ respectively (see Appendices D.3 and D.6 respectively or Lam et al. (2002)), while other situations might be approximated by values of $\phi \in (0,1]$ (Cominetti & Correa, 2001). Note that, contrary to the random departures interpretation, the fixed-interval departures interpretation leads to inconsistencies in case of partially overlapping lines with different frequencies as in the following figure:

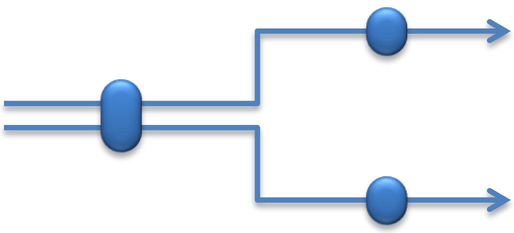


Figure 20: if both PT lines have different frequencies, fixed-interval departure of PT vehicles is only possible at two out of these three stops

Line usage proportional to frequencies can however still be a reasonable approximation.

The following equation defines the utility $\bar{V}$ of the strategy segment, including linearly valued waiting time, based on the route segment utilities V_i excluding waiting time:

$$\bar{V} = \beta_{t_w} t_w + \sum_{i\in C} p_i V_i = \beta_{t_w} \frac{\phi}{\sum_{i\in S} f_i} + \sum_{i\in S} V_i \frac{f_i}{\sum_{j\in S} f_j} = \frac{\beta_{t_w}\phi + \sum_{i\in S} V_i f_i}{\sum_{i\in S} f_i}$$

This can be used to define the following algorithm to find the optimal $S \subseteq C$ (Cominetti & Correa, 2001):

- Set $\varnothing \rightarrow S$, $\infty \rightarrow \bar{V}$.
- While $\min_{i\in C\setminus S} V_i < \bar{V}$:
    - Set $S \bigcup \operatorname{argmin}_{i\in C\setminus S} V_i \rightarrow S$.

- Set $\frac{\beta_{t_w}\phi + \sum_{i \in S} V_i f_i}{\sum_{i \in S} f_i} \to \bar{V}$ [10].

Now that S is known, the strategy segment attribute values $\bar{\mathbf{X}}$ can be determined from the route segment attribute values $\mathbf{X_i}$:

$$\bar{X}_a = \begin{cases} \sum_{i \in S} f_i & \text{if } a = \text{frequency} \\ \sum_{i \in S} p_i X_{ia} & \text{otherwise} \end{cases} \qquad \forall \text{segment attributes } a$$

The replacement of route segments with strategy segments may be problematic in two cases:

- If the to be merged route segments differ in utility components that are not valued linearly over the route (e.g. pricing units in a non-linear fare structure), the computed strategy segment is not necessarily optimal.
- If the choice tree distinguishes public transport modes and not all merged route segments belong to the same mode, the resulting strategy segment should be allocated to multiple modes. Furthermore, such a merge neglects potential individual preferences for certain modes, since within each strategy segment, travellers are assumed to pick the optimal route segment deterministically.

Because these problems challenge the utility maximisation principle, it could be better to not merge route segments in these two cases. In all other cases, the stated advantages at the beginning should make implementation worthwhile.

Relation between frequency and utility

Public transport frequency is typically embedded in utility functions in the form of waiting time or another function of the frequency (Hoogendoorn-Lanser, 2005). For the valuation, a distinction should be made between initial waiting time and transfer waiting time. The valuation may also depend on the mode one is waiting for (Hoogendoorn-Lanser, 2005).

Several possible utility component specifications – some based on waiting time, some based on the logarithm of the frequency – are discussed below. In all cases, the utility is defined at line-level; for routes consisting of multiple lines, the utilities at line-level $\bar{V}$ should be summed. However, because this depends linearly on the original utilities V of the runs, further manipulation of utilities at route-level is allowed (e.g. non-linear costs).

Waiting time or Weber function

For small initial waiting times, travellers may be assumed to arrive randomly at the first stop, implying an expected waiting time $t_w = 1/2f$ that can be placed in the utility function.

[10] Although $\beta_{t_w}\phi = 1$ in the proof of the algorithm by Cominetti & Correa (2001), this proof can be easily extended to this more general case, since $\bar{V}$ remains a weighted average of the previous $\bar{V}$ and the newly added V_i.

However, if the expected initial waiting time is large when travellers start their trip at a random moment in time, travellers may adjust their departure time to reduce the initial waiting time. This may be modelled by a Weber function that increases linearly for small waiting times and less than linearly for large waiting times (Weber, 1966). This could for example be one of the following functions that cap the waiting time at T:

$$\hat{t}_w = T\left(1 - e^{-\frac{t_w}{T}}\right) = T\left(1 - e^{-\frac{1}{2fT}}\right)$$

$$\hat{t}_w = \min\left(t_w, T\right) = \min\left(\frac{1}{2f}, T\right)$$

For the transfer waiting time, it makes sense to assume uniform arrivals of travellers (i.e. a waiting time of $1/2f$ where f is the frequency of the next leg), since travellers cannot freely choose when they arrive at the transfer point, unless the modeller wishes to correct for timetable synchronisation that is not included in the model.

Log-frequency

The alternative log-frequency utility component is based on run choice. If multiple runs for the same line are offered with regular frequency $f = n/T$ (i.e. there are n runs in the timetable with period T) with independent stochastic utilities with identical expectations $V = \mathrm{E}\left[U\right]$, the combination of these runs (the line) has an aggregate expected utility $\bar{V} = \mathrm{E}\left[\bar{U}\right]$ according to the logsum[11]:

$$\begin{aligned}\bar{V} &= \mathrm{E}\left[\max\left\{V + \varepsilon_1, \ldots, V + \varepsilon_n\right\}\right] \\ &= \ln ne^V = \ln\left(e^{\ln n} e^V\right) = \ln e^{V + \ln n} = V + \ln n\end{aligned}$$

This formulation is used in modelling framework OmniTRANS (Veitch & Cook, 2010). A similar formulation, but not based on the logit model, is used by modelling framework Cube (Citilabs, 2008; Kroes & Tuinenga, 2012). In this formulation, it does not matter whether a line is entered into the model twice with frequency one or once with frequency two. It implies that when merging route segments as described earlier, the strategy segment gets a utility equal to the logsum of the route segment utilities if these are equal.

There also is a mode of operation in which $V = 0$, that can be interpreted as travellers always boarding the first line, regardless of travel time or costs (Kurauchi et al., 2003). In this case, the model cannot really be seen as a logit model since there is no choice process. This interpretation holds assuming either fixed-interval or random departures of PT vehicles. This leads to a situation similar to the merging of common route segments described above. It is currently used by ANTONIN (Kroes & Tuinenga, 2012).

---

[11] Note that the choice probabilities do not depend on the size of the period. Substitution of $n = fT$ yields $\bar{V} - V = \ln fT = \ln f + \ln T$; because $\ln T$ is identical for all alternatives, it can be removed from the utility function.

Weighted log-frequency

The log-frequency theory is problematic because runs of the same line overlap, which is not accounted for. In other words, it is not possible to specify a weight for the valuation of frequency. This may be solved by switching to a nested logit model (see Appendix C.3) with nesting coefficient θ that can act as a weight:

$$\bar{V} = V + \theta \ln n$$

Note that in theory the overlap may have a positive influence ($\theta > 1$), because n vehicles on a single line offer a more regular service than n independent lines – for reducing the expected waiting time, fixing the headways is as effective as doubling the frequency (see Appendix D.6) – although in practice these lines would typically not be independent.

By substituting $f = n/T$ and $t_w = 1/2f$ (i.e. $n = fT = T/2t_w$) and subtracting the constant $\ln(T/2)$ from all utilities, one can see that this is not a Weber function since it is not linear for small values of the expected waiting time t_w:

$$\bar{V} = V - \theta \ln t_w$$

No practical applications of this weighted log-frequency have been found.

Comparison of utility components for frequency

The log-frequency approach with $V \neq 0$ does not offer a solid interpretation since overlap is not accounted for; its weighted version lack known practical applications. A Weber function for the initial leg and waiting times for the subsequent legs have a solid theoretical interpretation and have also been used in practice, and are thus the most promising way to incorporate frequency.

Measure of route overlap

For the expression of overlap between routes, the traffic network overlap measures distance and time can theoretically also be used in public transport networks, although distance cannot capture overlap in waiting periods: since walking at transfers is negligible compared to in-vehicle distances, this would mean overlap is only possible in in-vehicle legs between stops. For public transport, the number of legs is another possibility, where a leg is defined as a transfer-free in-vehicle period (a route segment) (Hoogendoorn-Lanser, 2005).

Using the number of legs results in the best model for multi-modal trips, with time as the second-best option. Using distance as the measure of overlap has a negligible effect compared to not considering overlap at all (Hoogendoorn-Lanser, 2005). Note that the number of common legs between two routes cannot be derived directly from common links in the non-pre-processed network.

5.4. Timetable-based service network

In a timetable-based service network the nodes do not represent points in space, but points in space at specific times (Hoogendoorn-Lanser, 2005; Bovy et al., 2006). This network structure allows precise calculations of waiting times; this leads to much more acceptable results in rural areas where frequencies are low and transfers are coordinated (Friedrich et al., 2001).

Since nodes in the graph now represent both space and time, the graph is called diachronic. Stops now consist of multiple nodes representing different time instants, connected by out-of-vehicle waiting links.

The network between the first three stops of the previous frequency-based service network example is shown below as a timetable-based network, with the time dimension displayed vertically[12]:

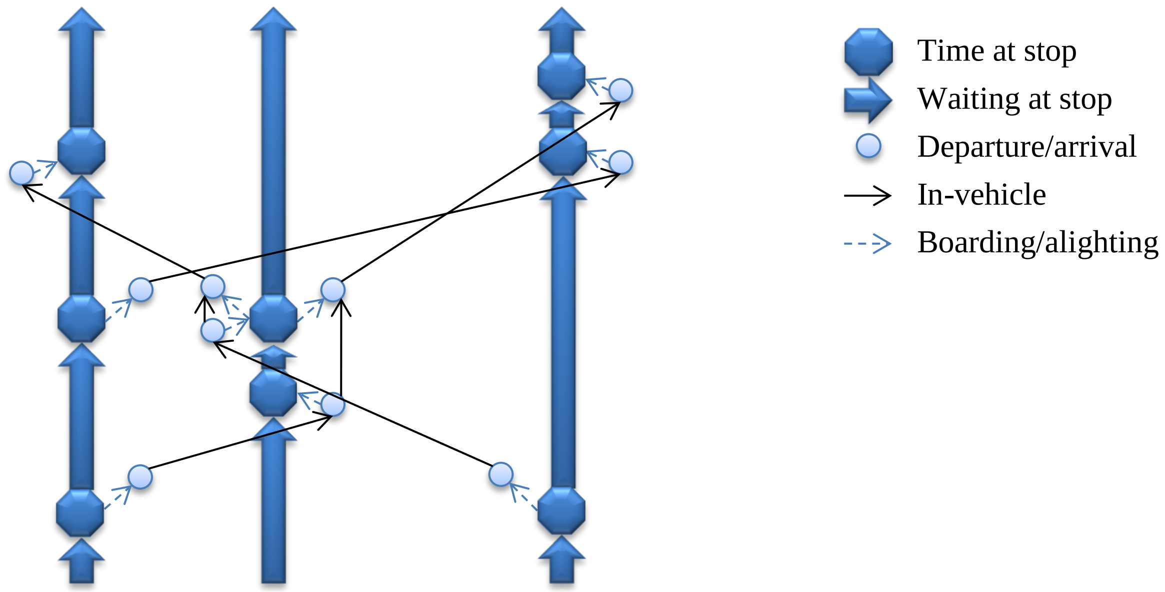


Figure 21: timetable-based public transport network representation

Although this network representation can be used for dynamic assignments, it can also be used for static assignments during a particular period of the day. In static assignments, there is not time choice, so it does not matter at what time the traveller visits the first and last stop; only the time in between matters. If the timetable is periodic, stop nodes at the end of the timetable could be connected to the corresponding nodes at the beginning of the timetable.

The GroeiModel currently uses the timetable-based representation for the train network[13].

Pre-processing

Like the frequency-based service network, the timetable-based service network may be pre-processed by replacing the links with connection segments, as has been done in the following figure:

[12] In a dynamic model, time-dependent origins and/or destinations may be included in the network separately from the stops, such that travellers may enter and/or leave the network at specific times (Hoogendoorn-Lanser, 2005). However, this project considers static models only.

[13] Note that the route sets are generated by the separate programme TRANS, so the GroeiModel uses the timetable-based network indirectly (Significance, 2012a).

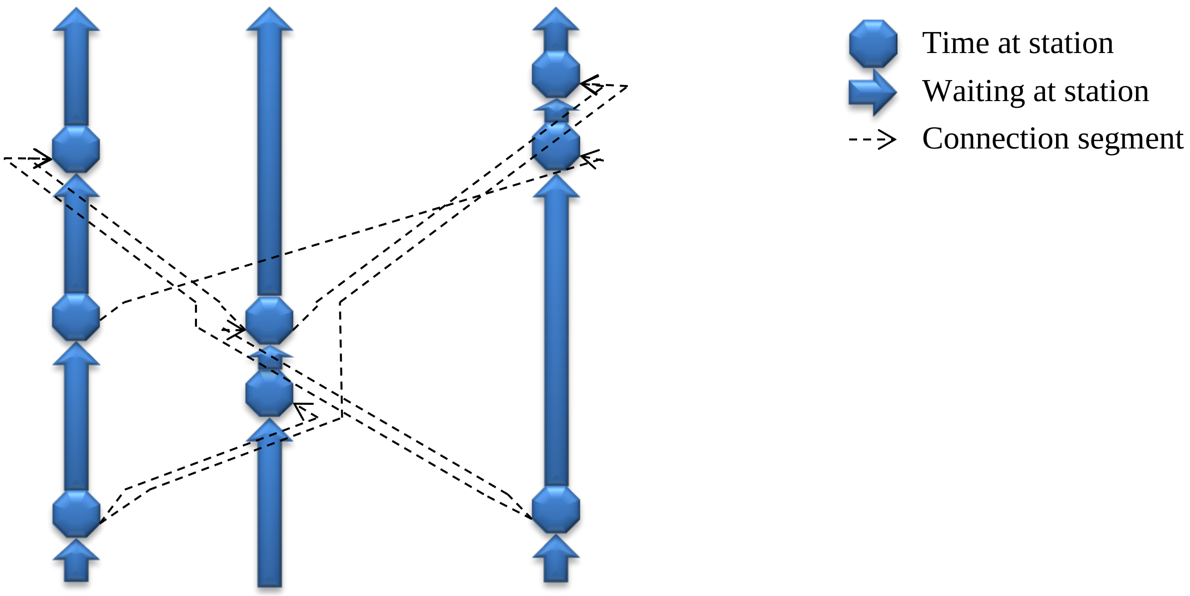


Figure 22: timetable-based public transport network after pre-processing

Consequences for choice model

If multiple lines are scheduled in the same timetable, a time-dependent choice model such as the one proposed in Appendix D.2 is needed to account for synchronisation effects between the competing services at the trip origin, even if the model is static (i.e. without departure time choice). Appendix D.2 shows that such a time-dependent choice model has a closed-form solution if waiting times are included linearly in the utility functions: in this case, the utilities become piecewise linear functions of the time relative to the periodic timetable. The choice probabilities of PT routes are then functions of time which are constant between the departure times of the routes; by integrating these over the timetable period, the total choice probabilities are found. Note that this can be trivially extended to include private access and egress for public transport and private-only alternatives (see Appendix D.8).

An alternative to a time-dependent model is to modify the utilities to correct for the synchronisation effects based on the headways with the previous and next running services (Bel, 2013; Hague Consulting Group, 1996). Unlike the time-dependent model proposed here, this enables travellers to adapt their behaviour to the timetable (Bel, 2013). However, the utility correction does not take the utilities of the other services into account, but only their departure times, which seems problematic in the current case where a variety of modes may be available.

Relation with frequency-based approach

To avoid the common line problem described earlier, each run in the time-dependent choice model can eliminate later similar runs from the choice set. The choice model proposed in Appendix D allows such domination constraints.

Appendix D also shows that the frequency-based approach is a special case of the timetable-based approach where all lines have independent timetables, assuming decision-making based on expected utilities including waiting times.

Considerations similar to the frequency-based approach hold for the handling of waiting time in utility functions. However, the expected transfer waiting time is replaced with the exact transfer waiting time, if there are services that span multiple lines and services are being chosen instead of lines. Also the expected initial waiting time can be calculated in more detail. The model in Appendix D requires linear valuation of waiting time.

The measure of route overlap is not different from the frequency-based approach.

5.5. Supernetworks

A supernetwork is a network "augmented with virtual (dummy) links to represent several choice dimensions" other than just route choice (Sheffi, 1985). Here, this definition is narrowed to the integration of modal choices in the network. Note that this does not require the mode choice to fully take place at network level, like in the pure supernetwork approach as described in Section 4.4.

To construct the supernetwork, the networks of individual modes are connected to each other via transfer links[14]:

Figure 23: construction of a supernetwork

In the supernetwork, either some (e.g. public transport modes only) or all modes are combined. If all modes are combined, the origins and destinations are typically (but not always) placed in the pedestrian network layer and there are no direct transfers between two non-pedestrian network layers (Van Eck, 2011; Fiorenzo-Catalano, 2007). In the figure, the middle layer could represent the pedestrian network.

Measure of route overlap

If public transport networks are combined with the pedestrian network, walking between stops (or platforms within the same station, which are then modelled as different stops) may also be counted as a leg for the purposes of expressing route overlap (Hoogendoorn-Lanser, 2005).

[14] Note that more difficult supernetworks also exist, such as the state-augmented multi-modal network (Lo et al., 2003). However, such a network is very large and all mode combinations have to be specified explicitly (Fiorenzo-Catalano, 2007). The problems this complicated network solves can also be solved in other ways: route set generation prior to route choice (see Chapter 6) solves the problem of non-linear utilities, while splitting route generation (see Section 6.2) and eliminating routes (see Section 6.3) can prevent unrealistic transfers.

In case private and public modes are mixed in the supernetwork, the expression of route overlap becomes more problematic, since for private modes the (free-flow) travel time is optimal while for public modes the number of legs is optimal. No literature dealing with this dilemma has been found.

Merging different service network types

If some modal networks are governed by a timetable and others are not, or if different networks are governed by separate, unsynchronized networks, or if a level-of-service matrix is the only available accurate description of certain modal networks, it is not a trivial task to combine them into a single supernetwork. In case of mixed timetable-based and frequency-based networks, some nodes just represent space while others also represent a certain time. The following figure illustrates this challenge:

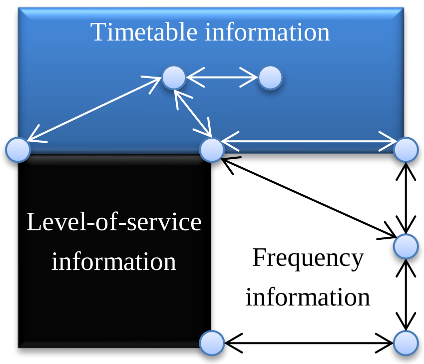


Figure 24: the problem of merging different public transport network types

There seem to be three ways of dealing with this problem: a bottom-up approach, a top-down approach and a horizontal approach. Each of these is discussed below.

Bottom-up approach

If not all services are in the same timetable, the choice problem as stated in Appendix D.8 becomes nested and all utility components need to be linear. This nested choice problem first looks for services within timetables and then, using these results, for routes that traverse multiple independent timetables using aggregate timetable logsums representing the combined utility of all services within the timetable. The approach can be called bottom-up since one first looks at detailed timetable information and then at less detailed level-of-service and frequency information.

The resulting logsum-based supernetwork may contain a very large number of services, since a new virtual service needs to be inserted for each origin-destination pair of lower-level timetabled networks. This may make the route search quite slow, although this procedure has to be performed only once[15]. Also, because during the process, the original service utilities are transformed into a logsum, route-level transformation of utilities (e.g. non-linear costs) is not possible. Additionally, this nesting has a consistency problem as the upper level problem

[15] For example, for the branch and bound algorithm (see Section 6.1), a possible simplification is to set a constraint that at each number of service legs (i.e. at each tree depth level), only one route to each node is accepted to keep the tree width fixed at the number of nodes. However, this only offers route set variety in the number of service transfers, removing mode and transfer stop variety from the choice set (De Jong et al., 2008; Baak, 2013).

works with expected values of the logsum while the lower level problem works with actual utilities.

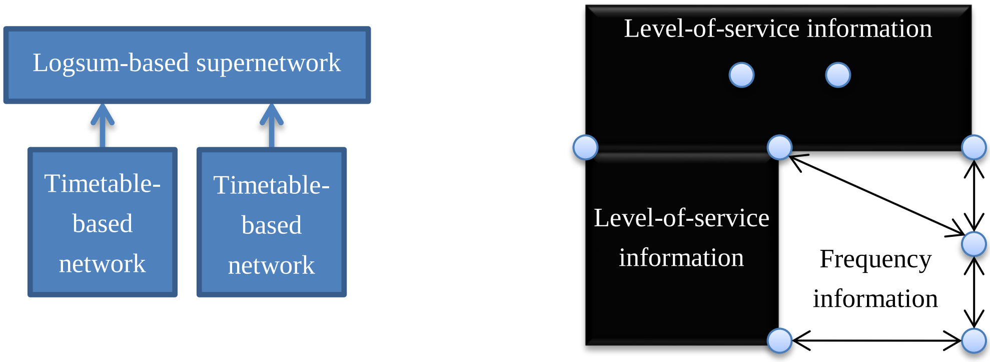


Figure 25: on the left the bottom-up approach (arrows indicate direction of computation); on the right an example logsum-based supernetwork

Top-down approach

Alternatively, for the purposes of route set generation, the timetable-based network parts may be replaced with a frequency-based ones. Then, after all routes have been generated, the utility of the routes is re-calculated by looking up the line numbers and transfer stops in the original timetable-based networks. This method thus works top-down, by first looking at frequency information and secondly at more detailed timetable information.

Unlike the previous bottom-up method, the choice problem occurs at a single level and waiting time and other utility components do not need to be linear since they are defined at route level (e.g. non-linear costs are possible). The simplification of the network during route set generation does not likely lead to problems in practice.

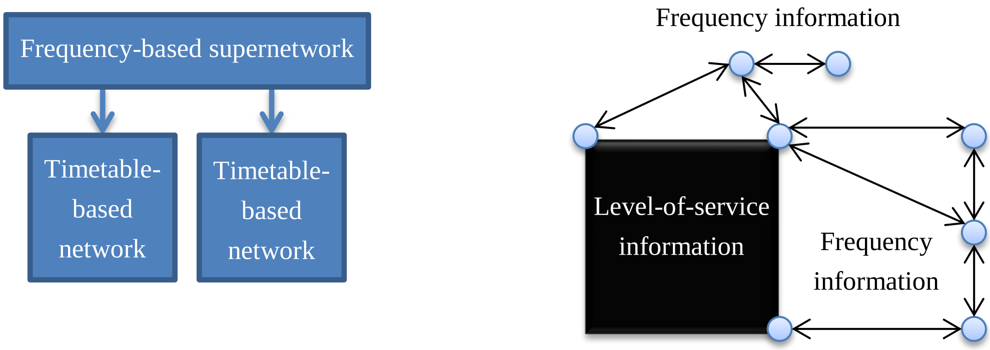


Figure 26: on the left the top-down approach (arrows indicate direction of computation); on the right an example frequency-based supernetwork

Horizontal approach

Finally, different route parts can be generated in the different modal networks which are combined afterwards (Hoogendoorn-Lanser, 2005). This avoids explicit construction of a supernetwork and is equivalent to splitting up the problem as described in Section 6.2.

Again, waiting times and other utility components are defined at route level and can be non-linear. The problem how to combine route parts to construct routes, which is hard to solve without sacrificing the multi-modal flexibility gain of the supernetwork (see Section 4.6), particularly if the possible transfer locations are not located close to the origin or destination, makes this approach more difficult than the top-down approach.

Figure 27: the horizontal approach (arrows indicate possible directions of computation)

5.6. Conclusions

Networks are generally expressed as directed graphs, but the shape of this graph differs considerably depending on the network type. Particularly public transport networks require some special attention.

For frequency-based public transport networks, pre-processing followed by merging of common lines increases the realism of the model. Pre-processing also simplifies the recognition of overlapping legs, which is the best method to express route overlap. The route set generation algorithm (see Chapter 6) should therefore preferably be optimised for networks with a large number of links but relatively few links per route. For timetable-based public transport networks, a time-dependent choice model as in Appendix D can be used.

Transfer waiting time should be included linearly in the route utility function, unless timetable synchronisation effects play a role and the network is frequency-based. For the initial waiting time, a Weber function provides more realism, although this is not possible if the timetable-based route choice model from Appendix D is to be used.

Finally, supernetworks with timetable-based components are best constructed by using frequencies from the timetable for the route generation and recalculating the utilities of found routes using the more detailed full timetable information.

6. Route choice sets

Specifically in route choice, the number of alternatives available to the decision maker may be very large, if not unlimited. The obviously also holds for multimodal route choice in a supernetwork. This chapter investigates how to select relevant route alternatives and model the route choice.

A traditional all-or-nothing assignment procedure can be used to select the best route, but unlike a choice model assumes that all travellers pick the same route, which is particularly unrealistic in multi-modal networks. Monte Carlo simulation using the multinomial probit model (see Appendix C.1) can be used to take personal preferences into account, but due to the nature of the Monte Carlo simulation this slows the network loading down considerably.

It can therefore be more efficient to generate a route choice set once in advance that is applied repeatedly in the iterations of the model (Bliemer et al., 2004; Fiorenzo-Catalano et al., 2004). This also allows for non-additive and route-based utility components in the assignment (Fiorenzo-Catalano, 2007), like they are currently used in the choice models of the GroeiModel (Significance, 2012a) and ANTONIN (Willigers & Tuinenga, 2007). This improves the consistency between the assignment and the higher-level choice models (Daly, 2012); if a logit model using the generated choice set is used for the assignment, this also allows to calculate a logsum for route choice that can be used in higher-level choices. A route choice set also offers more flexibility for the structure of the choice model (Fiorenzo-Catalano, 2007).

A disadvantage is that there is no guarantee that all eventually used routes are included in the route set. It is therefore important that the choice set generation process is of sufficient quality.

Different types of choice sets exist with different amounts of alternatives. For forecasting purposes, the set of known alternatives (Fiorenzo-Catalano et al., 2004) or the set of considered alternatives (Bekhor et al., 2006), should be generated; in the current state-of-the-art, these sets may be assumed to be identical (Fiorenzo-Catalano, 2007). To be more precise, the union of these sets for individuals should be generated, since each individual knows and considers different alternatives (Fiorenzo-Catalano et al., 2004). All types of choice sets are listed in the table below (Hoogendoorn-Lanser, 2005; Fiorenzo-Catalano, 2007; Van Nes & Bovy, 2008):

Choice set	Alternatives
Universal set	Existing alternatives
Objective master set	Logical alternatives
Objective choice set	Feasible alternatives
Subjective choice set	Known alternatives
Consideration set	Considered alternatives
	Chosen alternatives

Table 3: choice sets from the perspective of the researcher (from large to small)

Since explicitly enumerating all possible routes is infeasible in any network of reasonable size, an algorithm is needed to find the set of relevant route alternatives C_{od} from origins $o \in O$ to destinations $d \in D$ (i.e. origin-destination pairs $od \in OD$). Usually, the construction of the choice set takes place in two steps: firstly, routes are being generated, and secondly, routes are being eliminated to get the desired choice set as defined above[16]. The methods for the first step known in literature are discussed in Sections 6.1 and 6.2, followed by a discussion of the second step (Section 6.3) and a comparison of all methods (Section 6.4). Finally, Sections 6.5 and 6.6 discuss models to choose a route from the generated route set.

6.1. Basic set generation algorithms

To get started, this section will provide an overview of path finding algorithms. The algorithms included are Dijkstra, A*, K-shortest path, biased random walk, branch and bound, Monte Carlo simulation and labelling.

Dijkstra

The mother of all route finding algorithms is of course the Dijkstra algorithm: using link impedances z_m for links $m \in M$ it finds a single, shortest path (Dijkstra, 1959). The algorithm works as follows:

- $\forall o \in O$:
  - Set $\{o\} \to X$, $R_{oo} = \varnothing$, $z_{oo} = 0$.
  - While $D \setminus X \neq \varnothing$:
    - Select a link $m \in M$ connecting a node $x \in X$ to a node $d \notin X$ minimizing $z_{ox} + z_m$. The candidates for this selection are maintained in a priority queue.
    - Set $X \bigcup \{d\} \to X$, $R_{ox} \bigcup \{m\} \to R_{od}$, $z_{ox} + z_m \to z_{od}$.

The Dijkstra algorithm can be used on its own (to generate a route set consisting of a single route $C_{od} = \{R_{od}\}$, i.e. an all-or-nothing assignment) or as part as a more complicated choice set generation process, as is described below.

A*

The A* algorithm for routing is an enhancement of the Dijkstra algorithm to speed up the discovery of the shortest path in case of a single destination (Hart et al., 1968; Wikimedia, 2013a). It defines a heuristic function $h(o,d) \leq z_{od}$ as an estimate for z_{od}, typically based on crow-fly distance, and uses this to change the order in which nodes are explored to find the shortest path faster. The algorithm works as follows:

---

[16] Between these two steps, there may be a another step that checks whether new shortest routes can be found in user-equilibrium conditions that were not found using the initial link impedances. This check for missing routes is however not considered here since detailed treatment of congestion and crowding effects is outside the scope of this project.

- $\forall od \in OD$:
  - Set $\{o\} \to X$, $R_{ood} = \varnothing$, $z_{oo} = 0$, $y = o$.
  - While $y \neq d$:
    - Select a link $m \in M$ connecting a node $x \in X$ to a node $y \notin X$ minimizing $z_{ox} + z_m + h(y,d)$. The candidates for this selection are maintained in a priority queue.
    - Set $X \cup \{y\} \to X$, $R_{oxd} \cup \{m\} \to R_{oyd}$, $z_{ox} + z_m \to z_{od}$.
  - Set $R_{odd} \to R_{od}$.

If the algorithm is based on travel time, $v_{\max}$ is the maximum speed in the network and (x_n, y_n) are the coordinates of node n, the following heuristic function, representing the minimum crow-fly travel time, can be used:

$$h(o,d) = \frac{\sqrt{(y_d - y_o)^2 + (x_d - x_o)^2}}{v_{\max}}$$

K-shortest path

The K-shortest path method finds the first K shortest paths for each origin-destination pair (Bekhor et al., 2006). Two popular heuristic algorithms repeatedly apply the Dijkstra or A* algorithm with modifications so that in further iterations new routes are found. The link penalty algorithm adds penalties to the links in found paths, while the link elimination algorithm removes links in found paths from the network. Other algorithms exist as well (Van der Zijpp & Fiorenzo-Catalano, 2005).

Whether circuitous routes can be generated, depends on the algorithm.

Biased random walk

The biased random walk method starts at the origin and repeatedly adds a next link randomly based on the shortest path through that link to the destination, until the destination is reached (Frejinger et al., 2009). This procedure is repeated to generate multiple routes. The algorithm is as follows:

- Set $\varnothing \to C_{od}$ $\forall od \in OD$.
- Repeat multiple times $\forall od \in OD$:
  - Set $\varnothing \to R_{od}$, $o \to x$.
  - While $x \neq d$:

        - Randomly pick a link m connected to node x based on the length of the shortest path via link m to destination d (short lengths are favoured).
        - Add m to R_{od}.
        - Set x to the other end of link m.
    - Set $C_{od} \cup \{R_{od}\} \to C_{od}$.

Within the algorithm, shortest paths via a certain link m can be determined by the A* algorithm starting at the end of link m.

Note that this method may generate circuitous routes.

Branch and bound

The branch and bound approach (Friedrich et al., 2001) starts at an origin node and systematically explores all links as long as the resulting routes are considered reasonable. Here, a segment is considered reasonable for a particular branch if constraints regarding transfer time, relevance, tolerance and loops are satisfied. The algorithm works as follows, where X_i represents the set of elements in the tree at level i, and N is the maximum number of segments in the route:

- Set $\varnothing \to C_{od} \quad \forall od \in OD$.
- $\forall o \in O$:
    - Set $0 \to i$, create a new tree with a start node $\{o\} \to X_0$.
    - While $i < N$:
        - $\forall x \in X_i$, for all reasonable segments from node x to any node d:
            - Insert a sub-branch for this segment to node d into the tree under x.
            - Add this branch as a route to C_{od}.
        - Set $i + 1 \to i$.

To reduce computation time, this method preferably first uses a pre-processing step in which combinations of consecutive links are aggregated into segments; for example, a segment can consist of all links between a particular boarding and a particular alighting stop of a public transport line (see Section 5.3). The creation of segments reduces the depth of the search tree, so that N can be considerably lower.

This branch and bound method has been applied to multimodal supernetworks in literature (Hoogendoorn-Lanser, 2005) and is also applied in Cube Voyager (Citilabs, 2008), which is used by the current version of ANTONIN.

The method cannot efficiently handle large, dense networks (e.g. road networks) since the search tree depth becomes too large in this case[17] (Hoogendoorn-Lanser, 2005). For public transport, one can work around this problem by pre-processing the network as mentioned. However, there is no trivial way to pre-process road networks. The method can be applied to supernetworks, but the private mode subnetworks then need to be replaced by level-of-service matrices, such that the route choice inside these subnetworks is removed. A similar strategy could perhaps be followed for road networks by distinguishing motorway and non-motorway subnetworks.

Monte Carlo simulation
Similar to a multinomial probit assignment, a Monte Carlo approach can be used for route set generation (Bliemer et al., 2004). In a fixed number of iterations N, shortest routes are found using increasingly stochastic link impedances[18]. The algorithm is as follows:

- Set $\varnothing \to C_{od} \quad \forall od \in OD$, $0 \to \lambda$ and $0 \to i$.
- While $i < N$:
  - Set $z_m^0 + \varepsilon_m \to z_m \quad \forall m \in M$ with ε_m sampled from a positive random distribution satisfying $\mathrm{Var}\left(\mathrm{E}_m\right) = \lambda z_m^0$.
  - Use Dijkstra algorithm to find shortest paths $R_{od} \quad \forall od \in OD$ using link impedances z_m.
  - Set $C_{od} \bigcup \left\{R_{od}\right\} \to C_{od} \quad \forall od \in OD$.
  - Increase λ.
  - Set $i+1 \to i$.

Labelling
The labelling method distinguishes multiple groups of travellers U with different preferences and finds the best route for each group (Fiorenzo-Catalano et al., 2004; Bekhor et al., 2006). The union of these routes is the choice set. This deterministic algorithm is as follows:

- Set $\varnothing \to C_{od} \quad \forall od \in OD$.
- $\forall u \in U$:
  - Use Dijkstra algorithm to find shortest paths $R_{od} \quad \forall od \in OD$ using link impedances z_m^u.

[17] Although Prato & Bekhor (2006) successfully applied the branch and bound method to a non-pre-processed road network, this network contained only 1427 links. This is small compared to ANTONIN with 33 thousand road links (Significance, 2012b) and the Landelijk ModelSysteem dataset of the GroeiModel with 75 thousand road links (De Jong, 2011).
[18] Because this is a route set generation procedure and not an assignment procedure, the stochastic component of the link impedances do not need to be realistic. The distribution can have bigger tails to find non-trivial routes faster (Van Nes, 2011).

- Set $C_{od} \bigcup \{R_{od}\} \rightarrow C_{od} \quad \forall od \in OD$.

The labelling method may also be combined with another route set generation method. The Dijkstra algorithm inside the labelling method is then replaced with this other method. The other method works as usual with link impedances z_m^u instead of z_m.

The groups of travellers used by the labelling method may also be randomly generated (Monte Carlo labelling), such that the link impedance is a randomly weighted combination of travel time, costs, etc. Literature shows that in case these labels are randomly generated, the combination of Monte Carlo simulation with Monte Carlo labelling yields better route sets than each method individually (Fiorenzo-Catalano et al., 2004). Monte Carlo simulation may also be combined with a predefined set of (non-random) labels (Benjamins, 2001).

Since ANTONIN distinguishes a mode combination choice and a route choice, the route choice may be interpreted as a labelling method where the prescribed mode permutation functions as the label[19] (Syndicat des transports d'Île-de-France, n.d.). ANTONIN thus uses an inefficient combination of the labelling method and the branch and bound method[20].

6.2. Splitting the problem

In some models, the route generation problem in large multi-modal networks is split into multiple smaller problems in smaller subnetworks by enumerating possible transfer locations. This can be done for both behavioural and computation time considerations (Hoogendoorn-Lanser, 2005).

Some model structures presented in Chapter 4 dictate that the problem should or should not be split. Structures 1A, 2A and 2B have a uni-modal route generation problem that cannot be split on modal boundaries. In structures 3A and 3B on the other hand, the route generation must be split into an access part, a main part and an egress part, since the access and egress stops are prescribed by the prior choice model. For the other structures, splitting the route generation problem is optional. The following figure illustrates this, including all example models from Chapter 4:

[19] Note that a separate choice set is generated for each label; these choice sets are not merged as usual.

[20] The branch and bound method in general does not use a single impedance value for each link, but looks at individual disutility components. It therefore doesn't make sense to combine the branch and bound method with the labelling method; a properly configured branch and bound method on its own already generates alternatives that are attractive to different groups of travellers.

	Implied by model structure	Decided independently from model structure
Integral route generation	**1A & 2** INDY Japan Intercity PETRA SAMPERS dom. Albatross NVPM VMÖ	**1B & 4** Benjamins Fiorenzo-C.
Split route generation by enumerating PT access and egress locations	**3B** OmniTRANS	**1B & 4** Italy HSR ANTONIN TMfS
Split route generation by enumerating transfer locations for part of the PT system	**3A** GroeiModel NTM	**1B & 4** Hoogendoorn-L. SAMPERS int. Samadzad

Route splitting approaches:

Limited mode chain flexibility

Full mode chain flexibility

Figure 28: relation between model structures and the ability to split route generation, including model structure numbers and example models

For example, ANTONIN must generate routes that use a set of public transport modes in an arbitrary order with a private access mode (car driver or walk) and a private egress mode (walk). To do this, the model, having structure 4A, splits the problem of enumerating possible transfer stations/stops between private and public modes and separately finding sub-routes for each of these three route parts (Citilabs, 2008).

Regarding the desired flexibility in how modes are combined in multi-modal a chain, a distinction is necessary between splitting at PT access and egress locations and splitting at transfer locations for part of the PT system, e.g. train access and egress locations. The latter case reduces flexibility since it assumes that each route can be split into e.g. a non-train part, a train part and another non-train part. This distinction has been indicated in the figure.

Transfer location enumeration

Of course, splitting the route generation requires enumeration of possible transfer locations where the route parts may be concatenated to routes. For example, the GroeiModel examines the train stations $s \in S$ within a radius r_z from the origin or destination zone $z \in Z$ based on the distance to the closest station $\min_{s \in S} d_{zs}$, a factor $C > 1$, a maximum search distance $D_{\max}$ and a minimum search distance based on the intrazonal distance x_z, according to the following formula (Significance, 2012a):

$$r_z = \max \begin{pmatrix} \min \begin{pmatrix} C \min\limits_{s \in S} d_{zs}, \\ D_{\max} \end{pmatrix}, \\ \min\limits_{s \in S} d_{zs}, \\ x_z \\ \end{pmatrix} \qquad \forall z \in Z$$

Local train and express train stations are considered separately with different parameters and hence have different maximum distances. Furthermore, stations that are geographically positioned behind other stations are excluded (Significance, 2012a).

6.3. Elimination of routes

If the algorithm may generate circuitous routes, such routes should be eliminated instead of added to the choice set (Fiorenzo-Catalano, 2007).

Before adding to the choice set, a found route may be rejected if it has a large overlap with existing routes in the set (Bliemer et al., 2004; Fiorenzo-Catalano, 2007). In car traffic this can prevent routes using an off-ramp directly followed by an on-ramp. Whether such a criterion is necessary in a public transport network would depend on the network structure. However, this criterion can also be used to reduce a large choice set to acceptable proportions.

Elimination of routes can also be applied after the whole choice set has been generated. For example, this can be used to remove largely overlapping routes if it is not guaranteed that the shorter of two largely overlapping routes is generated first.

In case of the branch and bound algorithm, the constraints for adding connections to intermediate nodes in the route may simply be repeated for the route as a whole, but with more restrictive settings (Friedrich et al., 2001; Hoogendoorn-Lanser, 2005).

The branch and bound algorithm includes a constraint that eliminates dominated alternatives[21], i.e. alternatives that are inferior on all aspects compared to one of the other alternatives. According to literature, a large detour is a valid reason to eliminate a route alternative (Fiorenzo-Catalano, 2007); obviously such routes are usually dominated. Elimination or penalization of dominated alternatives has also been shown in literature to improve destination choice modelling (Cascetta & Papola, 2009). Note that 'weird' routes cannot be generated by methods based on the Dijkstra algorithm using realistic weights in the utility function (Van Nes, 2011).

6.4. Comparison of set generation algorithms

Criteria

An important indicator of the quality of a choice set is computation time. Algorithms can be classified as one-to-one (generates routes from one origin to one destination) or one-to-many (generates routes from one origin to all destinations). One-to-one algorithms have a

[21] Dominated alternatives may also be called irrelevant alternatives in literature (Friedrich et al., 2001).

computation time of $\mathrm{O}(|O||D|)$ while one-to-many algorithms have a considerably smaller computation time of $\mathrm{O}(|OD|)$. Algorithms may also be classified as deterministic (one procedure generates all routes) or stochastic (a route generation procedure is repeated with different random numbers each time); the stochastic procedures are typically slower due to the repetition. Note that for one-to-many stochastic algorithms, the stochastic error terms are generated on a many-to-many basis, but the subsequently applied algorithm works on a one-to-many basis.

The coverage of observed routes is another important indicator (Bekhor et al., 2006). It is the percentage of observed routes that are present in the choice set, or, alternatively, the percentage of observed routes that have a sufficiently high overlap with a route in the choice set. If the coverage is too low, the model cannot properly replicate the route choices people make in reality. It is better to erroneously include an unused route than to erroneously exclude a used route (Fiorenzo-Catalano, 2007). Note that there will always remain some routes missing: in literature, really good coverage is about 90% and usually requires a combination of route generation algorithms (Frejinger et al., 2009; Bekhor et al., 2006). The obtained coverage of course depends on the invested computation time (i.e. number of iterations, number of labels or flexibility of constraints) (Bekhor et al., 2006).

Variety, both in terms of space and preferences, is also an important indicator (Fiorenzo-Catalano, 2007; Van Nes, 2011). Travellers must have something to choose from based on their personal preferences and the positioning of the origin and destination within their respective zones. In a multimodal context, the variety in preferences implies variety in modes in the choice set.

Additionally, the routes also need to be comparable (Fiorenzo-Catalano, 2007; Van Nes, 2011). This implies that routes have no large detours, excessive travel time or costs, etc. (e.g. there are no dominated alternatives).

Finally, the choice set size should be limited (Fiorenzo-Catalano, 2007). However, as mentioned above, additional elimination procedures can be applied to the generated choice set to solve problems of large choice sets.

Assessment

Based on literature (Bekhor et al., 2006; Fiorenzo-Catalano, 2007; Fiorenzo-Catalano et al., 2004; Bekhor & Prato, 2009), the following table compares the mentioned route set generation algorithms using the criteria defined above:

Method	Classification	Coverage	Variety	Comparability
Dijkstra	One-to-many deterministic	Low	None	N/A
A*	One-to-one deterministic	Low	None	N/A
K-shortest path	One-to-one deterministic	Medium[22]	Low	High
Biased random walk[23]	One-to-one stochastic	Medium	Medium	Low
Branch and bound	One-to-many[24] deterministic	High	High	High
Monte Carlo simulation	One-to-many stochastic	Medium	Medium	High
Labelling method	One-to-many deterministic	Medium[25]	High	Low
Monte Carlo labelling	One-to-many stochastic	High	High	Low
Monte Carlo labelling & simulation	One-to-many stochastic	High	High	High[26]

Table 4: comparison of route choice set generation methods

Based on this overview, the branch and bound method and the Monte Carlo labelling and simulation method seem the most promising options. The branch and bound method is not very suitable for private mode networks, such that Monte Carlo labelling may be recommended there. For pre-processed public transport networks as discussed in Section 5.3, the branch and bound method may be expected to be most efficient due to the number of links. Finally, for private modes for which route choice is not deemed important, the Dijkstra algorithm may simply be used, i.e. considering only the shortest paths to all destinations. These different algorithms can be successfully combined if the route generation is split at the boundary of the PT system, as discussed in Section 6.2. Note that the Dijkstra algorithm may also be used for access and egress legs for public transport, since here too this sub-route choice may not be deemed important.

6.5. Route choice models

Once a route choice set has been generated, a choice model is needed to divide the travellers over the alternatives. Possible choice models are discussed in detail in Appendix C; this

[22] Better coverage than Monte Carlo simulation may be achieved, but the computation time required for this is really huge (Bekhor et al., 2006).
[23] The biased random walk is seldom mentioned in literature; the scores therefore have been guessed based on the description of the algorithm.
[24] In general, the branch and bound algorithm is one-to-many (Hoogendoorn-Lanser, 2005). However, some implementations include directionality constraints to bound branches; in this case, it turns into a one-to-one algorithm (Prato & Bekhor, 2006).
[25] Very high coverage is possible but requires definition of many different labels, which is difficult (Bekhor et al., 2006).
[26] Since this method is very efficient in generating routes, route elimination procedures may be applied to improve the quality of the resulting set.

section summarises which of these models can be applied to route choice. Route choice models need to be able to handle overlap between routes; possible measures of overlap have been discussed in Chapter 5.

As already stated in Chapter 5, the availability or unavailability of timetable information affects the utility function, and may even make utility time-dependent as in Appendix D. This section however focuses on how to model the route choice, once the utilities of the available routes have been provided.

For this purpose, the following models can be considered:

- The *multinomial probit* model is based on stochastic link impedances. Due to the required Monte Carlo simulation to repeatedly search for new paths, its computational efficiency is low. Another disadvantage is that linear valuation of route attributes is required.

- The *multinomial logit* model assumes independent error terms for the utilities of route alternatives. It is simple and efficient. However, the consequent independence of the utilities means route overlap is not handled. This is a theoretical weakness and leads to worse performance compared to more advanced models.

- The *cross-nested logit* model extends the multinomial logit model with nests for all links in the network to account for route overlap. While it improves the fit to actual choice behaviour, it has the theoretical weakness that it models route choice as the choice for a link and a route passing that link.

- The *paired combinatorial logit* model extends the multinomial logit model with nests for all pairs of routes to account for route overlap. The theoretically required pairwise evaluation of available routes is still not ideal for large numbers of alternatives, but again the model fit improves compared to multinomial logit.

- The *joint network GEV* model represents route choice as a sequence of link choices. Unlike the other models, it is unable to work with an explicit set of routes, instead, it must use a set of permitted pairs of consecutive links. A disadvantages of this approach is the rather complex mathematical structure that requires linear valuation of route attributes and independent the link impedances.

- The *C-logit* and *path size logit* models correct the utility functions to take overlap with other routes into account, and are therefore relatively simple to implement. The path size logit is the superior of the two due to its theoretical foundation, but both have limited benefits for the model fit compared to multinomial logit. The *extended path size logit* model improves the model fit by assigning a higher penalty for overlap to long routes, but undermines its theoretical foundation in doing so.

- The *expanded path size logit* model is another modification of the path size logit model; this one takes into account the probability that the route is generated. Its fundamental assumption that travellers choose a route directly from the universal route set instead of a subset of considered routes, can however be regarded as a theoretical problem. The principle of each possible route having a specific probability of being generated also puts severe constraints on the choice set generation algorithm, making this model difficult to use.

- The *logit-kernel model* uses a full covariance matrix for route overlap to improve the model fit considerably and also has a good theoretical foundation, However, this model requires Monte Carlo simulation – it is thus not very efficient – and there are issues with the parameter estimation process.

6.6. Comparison of choice models

The table below compares the models for route choice modelling listed above, based on their theoretical qualities (i.e. theoretical foundation for the formulas in relation to choice behaviour), practical qualities (i.e. goodness of fit to actual choice behaviour), computational efficiency (speed) and simplicity, particularly when applied in a multi-modal context:

Model	Theoretical quality	Practical quality	Speed[27]	Simplicity
Multinomial probit	Medium	Medium	Low	Low
Multinomial logit	Low	Low	High	High
Cross-nested logit	Low	Medium	Medium	Low
Paired combinatorial logit	Medium	Medium	Medium	Low
Joint network GEV	Medium	Medium	Medium	Low
C-logit	Low	Low	High	Medium
Path size logit	High	Low	High	High
Extended path size logit	Low	High	High	Medium
Expanded path size logit	Low	Unknown	High	Low
Logit-kernel	High	Medium	Low	Low

Table 5: comparison of route choice models

Based on this comparison, there is no clearly best method. However, the path size logit model and extended path size logit model seem to be most promising, because they have simple structures – which is important, since route overlap is only one of the three sources of correlation from Section 4.6 – and at least either high theoretical or practical quality (but unfortunately not both). If the path size coefficient is not significantly different from zero, the model falls back to the multinomial logit model.

The next chapter will further extend the (extended) path size logit model, which itself handles route overlap, to handle modal overlap and mode similarities as well.

[27] Refers to the computational efficiency. "Low" means Monte Carlo simulation is required. "Medium" means a large logit tree.

7. Ideal model

As stated in Section 4.6, the ideal model should, besides being flexible, take both modal overlap, mode similarities and route overlap into account. This automatically leads to model structure 1B (supernetwork) from Section 4.4 with classification (i.e. cross-nesting) of the generated multi-modal routes according to the used modes.

Such a classification method will be proposed in Section 7.1. Subsequently, network modelling and route generation are discussed in Section 7.2 for different amounts of available data. Section 7.3 then explains how other choice processes can be added to the mode and route choice component to complete the choice model. Next, Sections 7.4-7.6 discuss how other features that improve the quality of existing travel demand models can be added to the model proposed here; this included tour-based modelling, pivot-point procedures and congestion feedback. Finally, Section 7.7 evaluates the proposed model using the problem statement of Chapter 2.

The proposed model will be illustrated by an example with various public and private modes. It is assumed that all trips either start and end with a private mode and use public modes in between, or use only a single private mode.

7.1. Mode and route choice

The mode and route choice component is a network GEV path size logit model. This is a hybrid form of the network GEV model for handling modal overlap and mode similarities and the (extended) path size logit model for handling route overlap (see Section 4.6 for definitions). The network GEV model itself is a generalisation of the cross-nested logit model to allow multiple levels of nesting. A mathematical definition of the network GEV path size logit model is given in Appendix E.

Possible classifications

Although a model of structure 1B is designed here, the other structures of Chapter 4 can serve as an approach to find possible classification methods. Here, two main ideas can be distinguished for modelling multi-modal trips: trips with public transport as main network and private access and egress modes (structure 3B), or simply trips as combinations of modes (structures 4A and 4B)[28]. This leads to two potential classifications.

An advantage of the access-/egress-based classification, which is described by Appendix F, is that access modes on the one hand and egress modes on the other hand may have different similarities and that different weights may be attached to access and egress parts of a route (Hoogendoorn-Lanser, 2005). However, it has important disadvantages:

- Since this classification defines access and egress relative to the public transport system as a whole, multi-modal route alternatives for the same origin-destination pair

[28] Structure 3A, with a main public mode with possibly also public access and egress modes, is not flexible enough due to fundamental constraints on how modes are used in a chain. While constraints also apply to structure 4A, these are not as fundamental since all possible mode combinations can theoretically always be enumerated, although their number may be very large.

may differ substantially in access and egress length: for example, a traveller may walk a hundred meters to a bus stop or two kilometres to a train station. This way, there may not be strong correlations between alternatives sharing an access or egress mode (in this example walk access).

- Similarly, also within a single route, the access and egress parts may not be of comparable length; this makes an interpretation of the access/egress nest memberships (see Appendix F) difficult.

- Furthermore, the classification does not contain separate nests for individual PT modes; consequently, modal overlap is not handled at the level of PT modes. This means for example that when a traveller can choose between a train route and several bus routes, the probability that the train route is chosen is underestimated.

- Also, both route and modal overlap between access or egress parts of PT-routes and non-PT-routes are neglected due to the structure of the choice tree (see Appendix F). For example, if an individual relatively prefers car driving, then within the public transport alternatives, park & ride is also relatively more attractive, but this positive correlation is neglected by the model. Another example is that an individual can have a preference for cycling that makes both public transport alternatives with cycle access and alternatives with cycle egress more attractive.

Therefore, the classification based on structures 4A and 4B is adopted here.

Adopted classification

The mode-based classification described and adopted here is based on structures 4A and 4B from Chapter 4. Compared to these structures, transfer station choice has not been used such that utilities are defined at route level (e.g. allowing non-linear costs) and overlap may be handled with path size factors. Also, the nests explicitly specifying all mode combinations have been removed such that route overlap can be taken into account across these combinations. The modes are nested to capture similarities among them.

The following figure shows the resulting choice tree, which thus is a combination of an (extended) path size logit model for route overlap, a cross-nested logit model for modal overlap and a nested logit model for mode similarities:

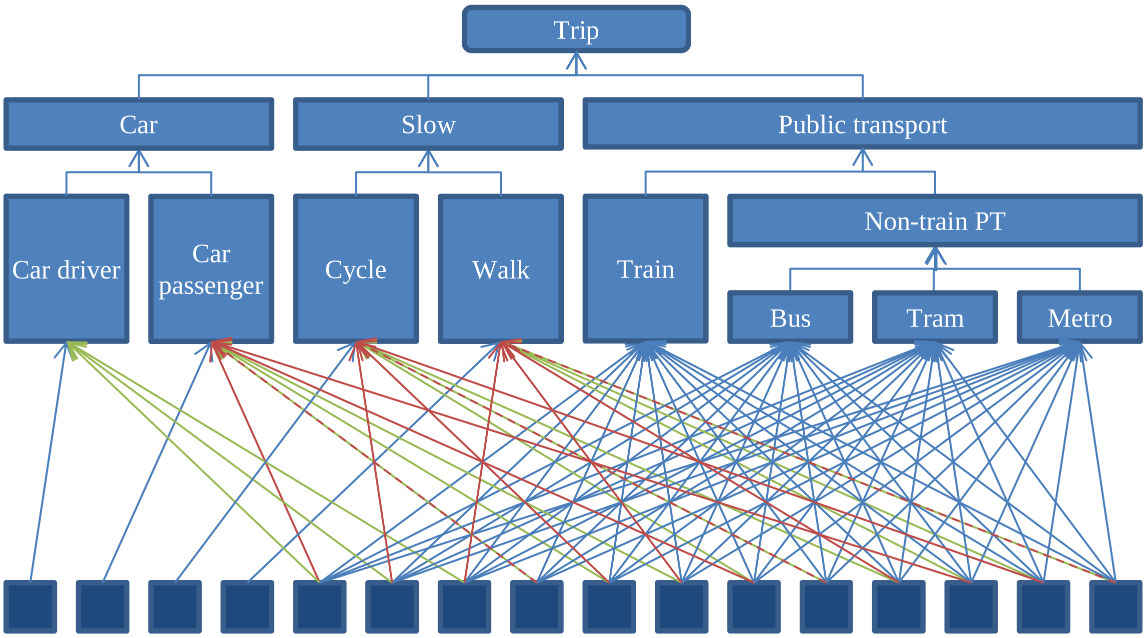


Figure 29: choice tree in case of mode-based classification; each dark box represents an arbitrary number of routes

The tree represents a classical mode choice with nested route choice, but with cross-nesting to account for multi-modal routes. In the example displayed in the figure, a nesting structure for the modes has been assumed as a representation of similarities among modes.

Behaviourally, this model can be interpreted as choosing a mode and then choosing a route that uses that mode. No single main mode is assumed for multi-modal routes, resulting in more flexibility and better modelling of overlap. The mode choice thus cannot be interpreted as a main mode choice, although given a chosen mode, routes which use this mode a lot are more likely to be chosen. In fact, the modelling of modal overlap is directly coupled with the modelling of route overlap, i.e. all route alternatives having one or more modes in common, also have nests in common at the lowest nesting level of the tree, and thus influence each other's path size factors in these nests.

The measure of route overlap could be travel time for private mode nests and number of legs for public mode nests, which is optimal for each of these nests separately. Since the nest memberships represent the measure of modal overlap, it seems logical to use related measures for that. Because the nest memberships have to be allocated top-down through the choice tree (see Appendix E.2), one can first use travel time to divide the route over PT and non-PT nests. Then, for the non-PT nests, travel time can be used to divide the route further over more specific private modes, while for the PT nest, the number of legs can be used to divide the route further over the PT modes.

Required parameters

The following parameters need to be specified for this classification:

- nest coefficients for all nests – the nesting structure itself also has to be specified.
- mode-specific constants for all private modes.

The public transport network may contain mode-specific boarding penalties (see Section 7.2), such that the number of legs using a PT mode is correctly taken into account.

Note that while the modes are allowed to be nested like in Figure 29, cross-nesting modes is not possible due to constraints of the network GEV model. More precisely, the resulting choice tree would not be 'crash safe' nor 'crash free', as is necessary for the normalisation of the model. See Appendix C.6 for more information.

7.2. Network data usage

Usage of timetable information

If all public transport services are operated according to a known timetable, then the network GEV path size logit model can be made time-dependent similarly to the path size logit model in Appendix D.2[29]. This way, synchronisation effects between the offered services at departure stations can be taken into account, allowing precise modelling of the distribution of travellers over PT lines. This does require linear valuation of initial waiting time in the utility function.

If no timetable is known for some or all PT modes[30], the PT route generation should take place in a frequency-based network, possibly incorporating just level-of-service matrices for some modes. Afterwards, the utilities of the found routes are corrected using available timetable data, according to the top-down approach from Section 5.5. In this case, initial waiting time is ideally valued with a Weber function.

Modelling of PT modes

Within the PT network, mode-specific boarding penalties are used to express average modal preferences. In-vehicle time and waiting time can also be valued separately depending on the mode.

At this network level, there is no limitation on the number of modes that may be distinguished. Even in the hypothetical red-bus-blue-bus situation, different boarding penalties, in-vehicle time factors and waiting time factors may be set for both bus types if the parameter estimation process shows significant differences.

If different vehicles are used on the same PT line, the valuation of the line can be the weighted average of the vehicle valuations where the frequencies are used as weights, just like in the common line problem described in Section 5.3. Similarly, for routes with strongly similar modes, route segments of different but very similar modes can be merged to strategy segments. Since the modes are averaged in this case, this is appropriate as long as the diversity benefit is negligible (Daly, 2012), implying that it is unnecessary to create separate nests in the choice tree. Otherwise, the modes should be specified as separate lines and their route segments should not be merged to strategy segments.

[29] For the alternative mentioned in Section 5.4, a time-independent model with adjusted utilities, it is not known how to combine this with partial nest memberships due to cross nesting and path size factors.
[30] Or if the linear valuation of waiting time is considered problematic (e.g. due to many lines with low frequencies) or if the time-dependent model is not efficient enough regarding computation time.

At the classification level, one should note that route overlap can only be taken into account with path size factors if overlapping routes are a member of the same lowest-level nest of the tree. This matches the definition of Section 4.6 that route overlap cannot occur without modal overlap. Modal overlap, and indirectly mode similarities, negatively adjust the choice probabilities of routes through the nesting structure, and an additional negative adjustment is caused by the path size factors if there is route overlap in addition to modal overlap (see Appendix E). Consequently, separate nests for all modes should not be created in the choice tree if route overlap considerations are expected to dominate diversity benefit effects.

In some cases, PT users may walk a short distance between two stops. Nearby stops should therefore be connected with short walk links in the PT network. If the PT network is timetable-based, these walk links should be repeated in the diachronic graph for every arrival at the alighting stop.

Route generation

Routes for private-mode only trips are generated with the Monte Carlo labelling and simulation approach, or, if route variety is considered unimportant for some modes (e.g. cycling or walking), simply with the Dijkstra algorithm, i.e. all-or-nothing assignment. Route generation for trips with both private and public modes, is split at the transfers between both. In the public transport network, stop-to-stop route generation is performed by the efficient branch and bound algorithm in the pre-processed version of the network with route segments merged to strategy segments.

Private access and egress legs are be generated separately with the Dijkstra algorithm by iterating over all access/egress stops within a reasonable radius from the origin/destination zone, where the radius depends on the stop density for the particular PT mode, while it also depends on the private mode (i.e. car driver access offers a larger range than walk access). The Dijkstra algorithm is chosen here because generating multiple access/egress sub-routes per stop seems unnecessary and also creates route overlap that cannot be handled by the public transport overlap measure, which is the number of common legs. The latter is also a reason against using Monte Carlo labelling in a supernetwork of all modes. By temporarily combining access sub-routes of different modes to the same stop, the branch and bound algorithm can save computation time since it only needs to search routes from a specific origin via a specific access stop once, such that routes for all access modes are found simultaneously. After concatenation of all legs, each route as a whole is re-evaluated to eliminate illogical alternatives.

7.3. Other choices

Since mode and route choice are not the only choices to be made in a travel demand model, this section shows how the previously proposed mode and route choice model should be extended into a complete travel demand model, taking the desired consistency of the model into account.

Other choices within trips

Other choices within the trips, such as destination choice and possibly time-of-day choice are included in the same network GEV model as the mode and route choice. The network GEV model is flexible enough to insert these other choices as a new level between any existing levels of the choice tree.

Note that this may lead to a large choice tree with cross-nesting. Since there is little experience with network GEV models, parameter estimation may become a cumbersome process. This might be avoidable by keeping mode and route choice at the bottom of the tree and estimating this part separately.

Trip frequency choice

Only for trip frequency choice, a separate logit model may be used, since choosing a frequency larger than one implies that multiple trips need to be chosen. A stop/repeat model, which is displayed below, is a typical solution for trip frequency choice used by both the GroeiModel and ANTONIN (Significance, 2012a; Bovy et al., 2006; Significance, 2012b):

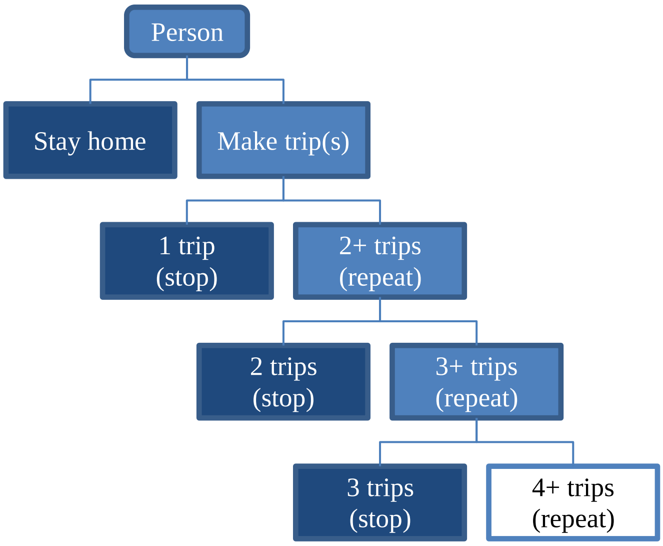


Figure 30: the 0/1 model followed by stop/repeat models for trip frequency choice (here visually cut off after three trips)

Note that the stop/repeat model is not a nested logit model, but a repeated multinomial logit model without internal logsum propagation. However, the utilities of making more trips in this structure can and should depend on the logsum of the nested logit model of the choices within trips in order to get a consistent model that incorporates elastic demand effects.

Ownership choices

Ownership choices may exist in the model at an even higher level than trip frequency choice, such as car ownership choice, driving licence ownership choice or PT pass ownership choice: the choices relating to the person or household instead of to specific trips. These choices influence the logsums of the main nested logit tree of the model, because they change the availability or costs of trip alternatives. From a consistency perspective, such ownership models should therefore ideally be based on the change in logsum of this tree.

Note that in some cases, ownership choices could also be implemented as a lower-level choice. This is for example necessary for ownership choices of PT passes for a specific destination (e.g. for home-work tours). In that case, pass ownership could be nested below route choice, where the utility of the route equals the logsum of the various possible payment methods for that route.

7.4. Trips versus tours

A distinction should be made between trip- and tour-based models. A trip is a journey from an origin to a destination, while a tour is a sequence of trips that ends at the same location it originated from (Significance, 2012a). A tour-based model simultaneously models pairs of outbound and return trips, possibly with the insertion of additional trips to visit secondary destinations, while a trip-based model considers each trip separately.

Tour-based modelling has theoretical advantages compared to trip-based modelling: it allows to explicitly attach outbound and return trips to the same travel purpose, to guarantee preservation of travellers and private vehicles during a modelled day and to return all travellers and private vehicles to their original positions in the network by the end of a day (Significance, 2012a). The ideal travel demand model is therefore tour-based. Both the GroeiModel and ANTONIN are examples of tour-based models (Significance, 2012a; Syndicat des transports d'Île-de-France, n.d.).

Differences for tour-based models

The models described above are trip-based, but the following simple modifications turn it into a tour-based model, realising the corresponding theoretical advantages:

- “access” and “egress” should be replaced with “home-end” and “activity-end” (i.e. the return trip starts with the egress mode of the outward trip and ends with the access mode of the outward trip);
- routes should be replaced with pairs of outbound routes and return routes; these combined routes should use the same networks for both trips, but in reversed order, and the transfers between different networks should be at the same locations.

The latter modification makes sure that all private vehicles are returned to their original positions, but at the same time allows different usage of the public transport system in both trips. More restrictive formulations that require the return trip to use exactly the same route in reverse are also possible (Hoogendoorn-Lanser, 2005); in this case, the routes do not need to be replaced.

Secondary destinations

Tour-based models may also add secondary and higher-order destinations to tours. There exist two types of non-primary destinations, both of which are included in the GroeiModel for the mode car driver (Significance, 2012a):

- *Nested tours*. An additional nested tour visiting the additional destination is inserted between the outbound trip and the return trip. In this case, the tour choice model can be repeated with the primary tour destination as origin, with different parameters and with the availability of private modes derived from the chosen primary tour. The logsum of such secondary destinations could be incorporated in the utilities of route alternatives in the primary choice model to get a single consistent choice model in which availability of attractive secondary destinations influences the primary tour choice.

- *Detours*. The outbound or return trip is, possibly repeatedly, split into two separate trips to visit the additional destination, resulting in a detour[31]. This requires a different choice model that compares the utility of the additional destination with the disutility of the required detour.

The detours are more difficult to integrate in the primary tour choice model:

- a separate detour model means that the additional destination is chosen after a direct route to the primary destination has been chosen independent of any potential additional destinations;
- a joint route and additional destination choice level at the bottom level of the primary tour choice model, where routes with additional destinations are added to the choice set[32], leads to complications in defining route overlap;
- adding combinations of destinations to the alternatives at destination choice level requires attention to be paid to destination combination overlap.

Although it is relatively easy to take nested tours into account, allowing detours in any of the proposed model structures requires further theoretical research.

7.5. Base matrix usage

In a classic model with separate network choice and route choice, before the assignment, a pivot-point process may be used to improve the quality of the network-specific origin-destination matrices that serve as input to the assignment, with base matrices. The GroeiModel does this for car driver and train (Significance, 2012a), while ANTONIN does this for car driver, car passenger, public transport and walk (Syndicat des transports d'Île-de-France, n.d.).

The integration of mode and route choice and the fact that routes are multi-modal[33] require some changes to this procedure. Therefore, Appendix G derives a modified pivot-point procedure that, after the full choice model has run, uses base matrices for parts of the supernetwork (e.g. car and train base matrices as in the GroeiModel) to enhance the quality of the multi-modal route usage numbers.

7.6. Congestion and crowding feedback

Inclusion of congestion (car) and crowding (PT) effects in a travel demand model implies the presence of a feedback loop to the choice modelling: one wants to see how travellers respond to the congestion or crowding. After a number of iterations, this should lead to an equilibrium

[31] An alternative option is to simply insert another trip on top of the existing tour, like ANTONIN does (Willigers, 2007). Although it allows simple re-use of the primary tour choice model structure, this method is physically illogical and is therefore not considered here.
[32] For routes with public transport, a requirement could be set to use only walking between the PT system and the additional destination. This seems logical and avoids complications with nest memberships.
[33] In violation of the conservation of travellers at train stations, the GroeiModel pivot-point procedure does not take into account that train trips may have access/egress bus/tram/metro legs (Significance, 2012a).

state. For the models proposed here, having explicit route choice, inclusion of congestion and crowding amounts to changing the stochastic uncongested model into a stochastic user-equilibrium (SUE) model (Bovy et al., 2006).

The utilities of the choice model may be modified iteratively: after the choice model has run, an updated level-of-service subject to congestion and crowding is calculated[34] and using these new utilities (i.e. travel times) of route alternatives, the choice model is run again. The link costs of the iterations so far can be weighted according to the method of successive averages[35] (Van Nes & Bovy, 2008). Note that hard capacity constraints like in Lam et al. (2002) are not trivial to add to the model proposed here: it is not trivial to attach a maximum number of users to a link, which is represented by a group of route alternatives spread over separate, advanced logit models.

For congestion, simple BPR functions may be used for the determination of utilities (Bovy et al., 2006; Van Nes & Bovy, 2008) or more sophisticated quasi-dynamic traffic models that take into account spillback effects, such as the traffic flow theory-based STAQ (Brederode et al., 2010) or the heuristic link travel time calculation component of QBLOK (Significance, 2012a). For crowding, utility components as proposed by Bel (2013) may be inserted.

Note that congestion with spillback effects does not yield hard capacity constraints on the number of travellers that choose particular routes: instead, they imply that the modelled time period is extended. In STAQ, this extension of the time period is modelled explicitly as an additional period with no network inflow (Brederode et al., 2010). This means that including spillback effects does not require non-trivial modifications to the choice model.

Due to the integration of route choice with the other choice processes, the iteration scheme is different from a classic model in which these choices are separated. This is shown in the following figure:

[34] Note that if strategy segments are used (see Section 5.3), their attributes should be recalculated since crowding may affect optimal line choice.

[35] Link-level averaging is typically much faster than averaging at route level: the Landelijk ModelSysteem dataset of the GroeiModel for example has 1.9 million OD pairs, while its road network only contains 75 thousand links (De Jong, 2011). Cost averaging may be faster than flow averaging in case of non-separable link costs (e.g. spillback situations) (Van Nes & Bovy, 2008).

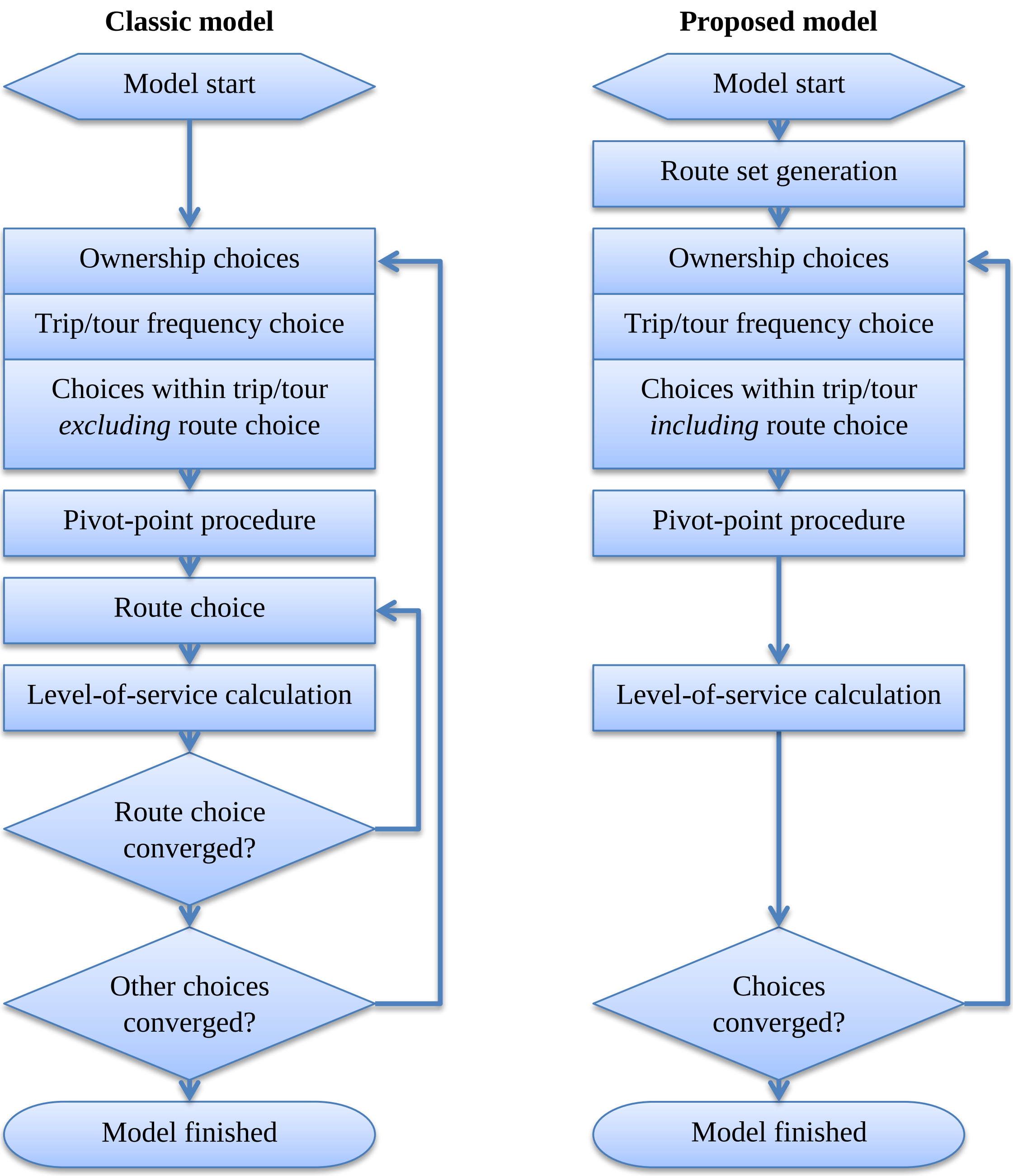


Figure 31: comparison of iteration schemes between a classic model and the proposed model

A classic model has two loops: a demand model loop and a network model loop, where the network model typically uses more iterations and less user classes (Daly, 2012). The model that is proposed here can only have a single feedback loop.

7.7. Satisfaction of objectives

One can easily verify that the model proposed in this chapter solves the problems posed in Chapter 2 and thus satisfies its theoretical objectives:

- ✓ *Aggregation of modes* is no longer necessary, because the proposed model works with an arbitrary number of PT modes.

- ✓ All *permutations of modes* are permitted in the choice tree: a route simply consists of an arbitrary number of modes. Even combinations of private modes (e.g. carpooling) are theoretically possible in the choice tree, although the proposed route set generator can only create routes with either a single private mode only or with a private mode followed by an arbitrary number of PT modes (including short walk transfers) followed by another private mode. This allows all potential logical uses of the public transport system.

- ✓ *Changes in the number of modes* can be handled. New PT modes can be added with only slightly more computation time compared to adding the same new services to a similar existing mode – only the nesting structure in the choice tree needs to be extended. New private modes could be added as well, although this does increase computation time because the route set generation procedure has to be extended. Maximum likelihood estimation based on stated preference research can be used to find the parameters of a new mode as well as the nesting coefficient for the new nest in the choice tree. If the position of the new mode in the choice tree is disputed or unclear, multiple possible positions can also be statistically tested and evaluated this way. If the mode appears to be highly similar with an existing mode, it should not be inserted into the choice tree, such that differences with the existing mode are only made at network level; common route segments can be merged to strategy segments in this case.

- ✓ *Consistency in the choice processes* is achieved by placing explicit route choice in the same nested logit tree as the other choice processes in the same trip or tour and by propagating logsums to the frequency and ownership choices. Route overlap, modal overlap and mode similarities are all taken into account.

- ✓ The *computation time* dependency on the number of permitted permutations of all modes is removed, leaving only a limited dependency on the number of private modes. The number of public modes does not affect the route set generation, but only the number of nests in the choice model. This results in better scalability with the number of modes.

8. Île-de-France case study

In this chapter, the proposed modelling methodology will be tested with a case study for the Île-de-France region, which is the region around Paris in France that is also modelled by ANTONIN. This case study attempts to estimate the parameters of an integrated mode and route choice model for home-work trips in this region, in order to evaluate how the developed theoretical framework performs in practice, particularly with respect to the requirements from Chapter 2.

The case study contains simplifications compared to the ideal proposal in Chapter 7. The case study only considers trips, not tours, and assumes a fixed OD matrix, no secondary destinations, fixed vehicle ownerships, fixed driving licence ownerships and fixed PT pass ownership. Furthermore, only a single route (part) is considered for non-PT legs. Finally, there is no pivot point-process and no congestion and crowding feedback to the choice model. Hence, the case study creates a mode and route choice model according to the principles of Sections 7.1 and 7.2 and the theory in earlier chapters.

The road network, including zonal connectors and transfer links to the rail network, as well as the frequency-based PT network representation, have been imported from the 2007 version of the ANTONIN2 model. Both networks are plotted in the next figure:

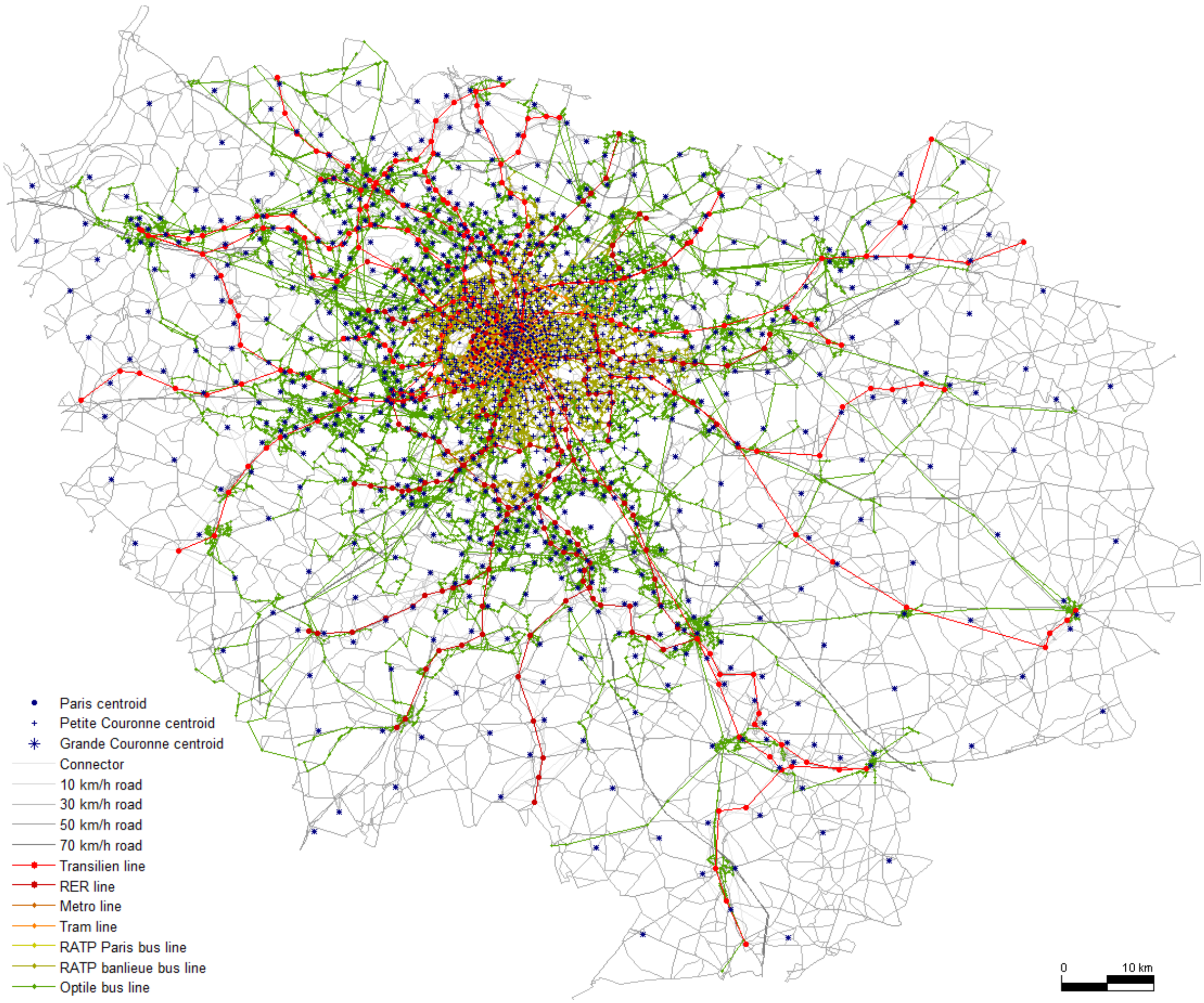


Figure 32: Île-de-France road and PT networks as in the 2007 version of ANTONIN2 for the morning peak

The following table provides some key numbers about this network:

Number of zones	1.342
Number of rail stations	936
Number of bus stops	10.978
Number of other road nodes	56.407
Total number of nodes	**69.663**
Number of zone connectors	21.336
Number of station connectors[36]	10.546
Number of road links	261.518
Number of PT transfer links	15.054
Total number of links	**308.454**
Number of rail lines	198
Number of bus lines	2.494
Total number of PT lines	**2.692**

Table 6: main network characteristics in numbers; the numbers of connectors and links are even since all connectors are bidirectional and all links are bidirectional for at least pedestrians

[36] There are separate station connectors for cars/motors and pedestrians/cyclists.

To illustrate the scalability with respect to the number of modes, contrary to ANTONIN, PT modes were not aggregated in this case study (see Section 2.1) and additional modes for motor and bicycle have been added compared to ANTONIN.

For the purpose of choice model estimation, revealed preference data is preferred over stated preference data, since the number of available multi-modal route alternatives in reality is usually way too large for a stated preference survey, and reducing the size of the choice set would make it harder to distinguish correlations between the utilities of alternatives. Furthermore, revealed preference surveys provide a better guarantee for inclusion of all aspects relevant to the decision-making.

The case study model is therefore estimated based on revealed preference data from the Enquête Globale Transport (EGT) 2010, which is the household travel survey for the Île-de-France region held in 2009-2011 (l'Observatoire de la mobilité en Île-de-France, 2012). The mode combination and destination choice component of ANTONIN2 has been calibrated using an earlier version of this survey (the EGT 2001). The survey data set contains a lot of information. For this case study, the description of all chosen routes as sequences of transfer-free legs is the primary information that has been used. Driving licence ownership of the traveller and vehicle ownership of his household have also been used. The estimation has been targeted specifically at trips with a of purpose of home-work and with the mean of the departure and arrival time in the morning peak (7:30–9:30 AM); all other observed trips have been filtered out.

The Île-de-France region can be divided into three parts: the city Paris (department 75), its inner ring Petite Couronne (departments 92/93/94) and its outer ring Grande Couronne (departments 77/78/91/95) (Wikimedia, 2013c). Since the PT lines are fully described only in Paris and Petite Couronne (Tuinenga et al., 2006), trips with an origin or destination in the Grande Couronne region have been ignored in the estimation of this case study model. The following figure zooms in on the Paris and Petite Couronne parts of the networks that together contain 727 of the 1342 zones:

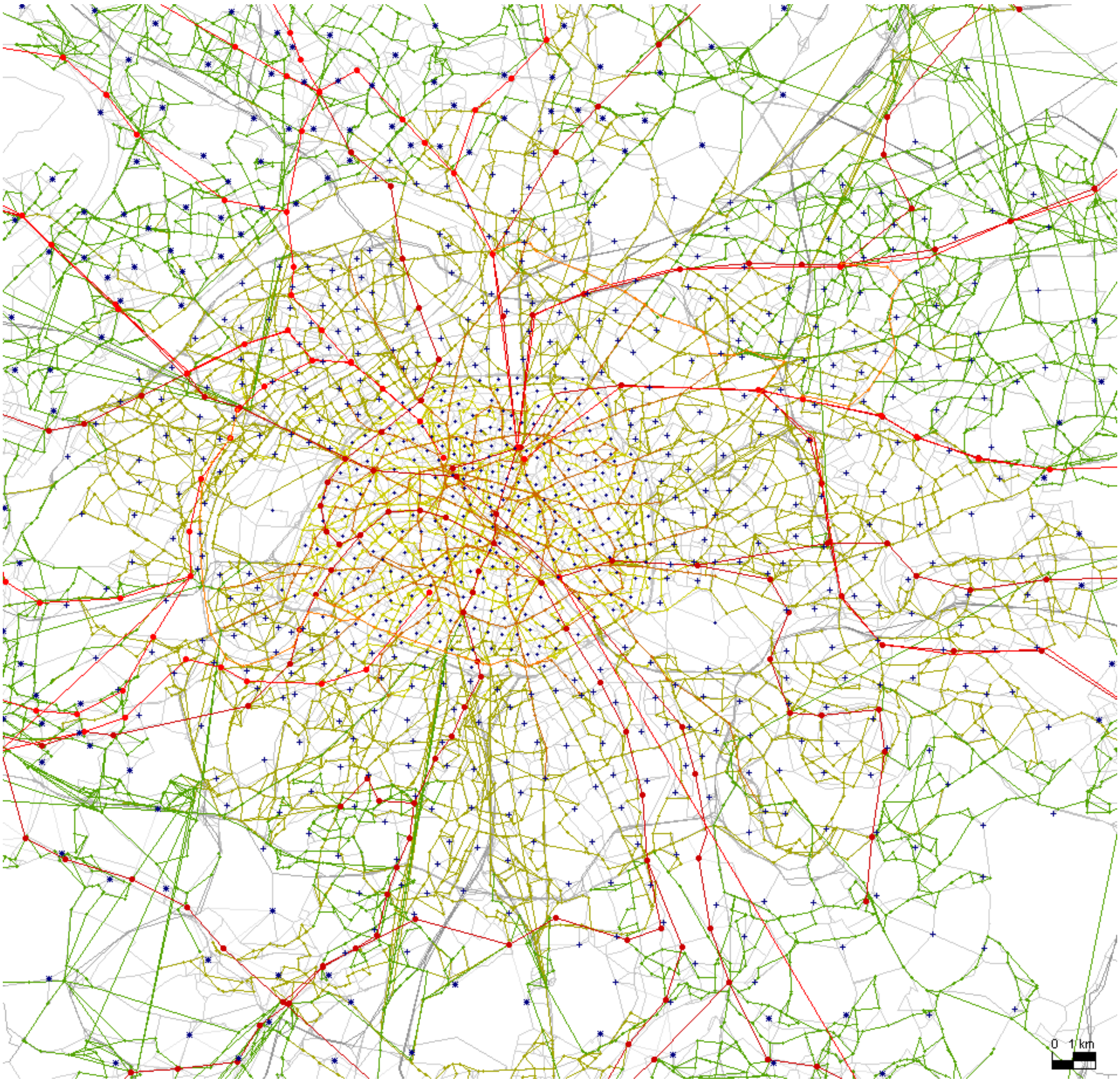


Figure 33: networks zoomed in to Paris and Petite Couronne

This chapter is organised as follows. Section 8.1 describes the route set generation algorithm that has been implemented. Next, Section 8.2 describes how the survey data has been interpreted and compares the observed routes to the choice sets that were generated using the algorithm of the preceding section. Section 8.3 then estimates the choice model based on the generated and observed routes. Finally, Section 8.4 reflects on the requirements stated in Chapter 2.

8.1. Route set generator implementation

The used route set generation algorithm consists of two main steps. First, the non-PT-segments are constructed using the Dijkstra algorithm. After that, the branch-and-bound algorithm is used to combine the PT and non-PT segments into routes, which means there is a split at the PT system boundary as described in Section 6.2. Both components have been implemented as one-to-many algorithms to achieve good scalability. The route set generator, which was created specifically for this case study, has been programmed in C++11 and was compiled using Microsoft Visual Studio Express 2012.

Calculation of intrazonal travel times

Before generating routes, for each zone and private mode, an intrazonal travel time is calculated. This is defined as half the travel time of the shortest route from the zone to the most nearby other zone. This shortest route is computed with the Dijkstra algorithm. The results of this computation are used later on in route generation.

Creation of PT segments

First, the PT line data is converted into route segments by iterating over all possible combinations of boarding and alighting nodes, conform Section 5.3. During this process, for each PT node, a list of line numbers of PT lines stopping at that node is recorded. These lists will be used in the creation of non-PT segments below.

Then, for each PT mode separately, all route segments with identical start and end nodes are merged into a single strategy segment conform Section 5.3. For simplicity, this process of finding the optimal subset of route segments that should be included in the strategy segment, assumes identical valuation of waiting time and in-vehicle time.

Creation of non-PT segments

The route set generation algorithm then starts in the road network by generating direct, non-PT routes, which are called direct segments, and access and egress segments that will be used for PT routes later on. This step simply uses the Dijkstra algorithm based on travel time. Note that walk transfer segments are not generated here, but have been imported directly from ANTONIN, which internally uses an algorithm based on crow-fly distance between PT nodes supplemented with a database of observed rail-to-rail transfers with measured walking times (Syndicat des transports d'Île-de-France, n.d.).

To get access and direct segments, a shortest path tree is created around the origin zone. Next, for the egress segments, a shortest path tree is created around the destination zone, using the road network links in reversed directions. The Dijkstra algorithm thus runs twice. The following rules determine when segments, representing concatenations of road links found so far, are saved:

- For each zonal centroid discovered by the Dijkstra algorithm, a direct segment is saved if it was found in the first run, which was the run intended to find access and direct segments. Since the second run is only intended to find egress segments, direct segments found in that step are not saved, since these were already found in the previous run.
- For each new PT node discovered by the Dijkstra algorithm, it is checked whether this node gives access to a new PT line. The segment is only saved if this condition is satisfied, such that for each line only the access stop with the least access time from the zonal centroid is stored. Furthermore, the line numbers of those new lines are stored on the segment to prevent the branch-and-bound algorithm from accessing a line with an access segment that was intended to access other lines. The same rule applies in the second, reversed run for egress segments.
- For other nodes, no segments are saved since these would be useless.

The following figure illustrates this with an example for access segments. First, an access segment is found to node A that gives access to lines 1 and 2. Next, an access segment to line B will be found, but it is rejected because it does not give access to a new line. Finally, an access segment to node C is found, but this access segment may only be used for access to line 3, not for access to line 1.

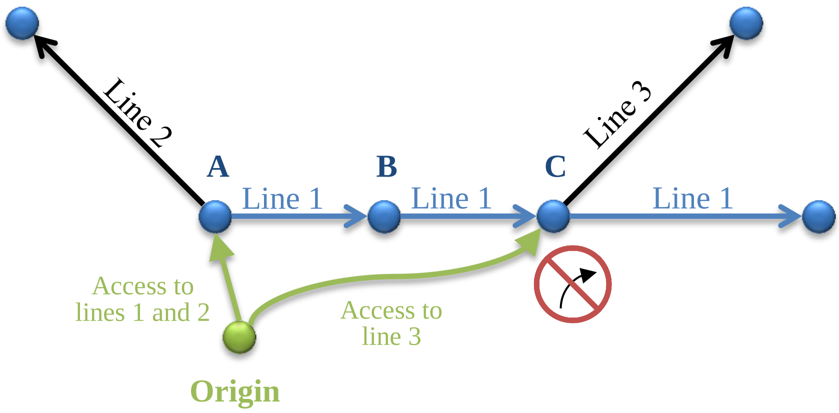


Figure 34: generated access segments in an example situation; for clarity, this figure displays the original in-vehicle links instead of the strategy segments for the PT network

The procedure is repeated three times: for walking, for cycling and for motorised modes. The walking and cycling speeds are fixed at 5 km/h and 15 km/h respectively, while the link-dependent speeds from ANTONIN2 have been used for motorised modes, varying from 10 km/h to 70 km/h during peak hours (Syndicat des transports d'Île-de-France, n.d.). The slow modes on the one hand and the motorised modes on the other hand have separate transfer links to the rail network. For both walking and cycling, one-way street links may be used in the opposite direction at walking speed. In the step for motorised modes, three copies of each found segment are saved: one for car driving, one for motor driving and one for car/motor passenger. Based on vehicle availability considerations, the modes car driver and motor driver are not available for egress.

Some additional constraints are applied to avoid routes from being created in the choice set that, while seeming theoretically possible, appear to be rare in practice, that is, in the EGT data for home-work trips, leading to problems during the estimation of the choice model (see Section 8.3 and Appendix H.1). The following segments are therefore not created:

- motorised access/egress segments to/from bus stops;
- car/motor passenger egress segments.

Note that the rareness of these segments in the survey data is the only reason why they are excluded from the model; due to the flexibility of the model structure, they could simply be included had there been more observations in the survey.

Furthermore, no more access and egress segments are generated beyond a radius of 30 minutes travel time. For direct segments, the maximum travel time is 60 minutes for walking and cycling and infinite for motorised modes. To account for differences in zone size, these limits are incremented with the intrazonal travel time of the considered zone and mode. The number of access legs and the number of egress legs are also limited to a maximum of 25 for slow modes and 3 for motorised modes.

All found segments are grouped according to the start node and end node they are attached to. Each segment in such a group thus represents a different private mode. The collective travel time T_G of such a group G of segments is defined as the minimum of the travel times of the individual segments. The number of vehicle legs L_G is defined 0 if the group contains a walking segment and 1 otherwise.

The purpose of this grouping is to reduce the width of the branch-and-bound tree in the next step: this is achieved by creating only a single branch for each segment group. The properties of the group therefore represent the optimal values of the properties of the individual segments in the group. Once complete routes have been found, the groups are again split up such that routes are created for all access/egress modes that each group consists of (see below). Note that the benefit is created by grouping access segments; the grouping of direct and egress segments is only done because this simplifies the implementation.

Route finding

The route finding process now comes down to concatenating the segments found in the previous step. The branch and bound algorithm is used for this. This algorithm requires constraints to be defined to determine whether a node should be added to the search tree or not. Because this branch and bound algorithm is used in a one-to-many setting, these constraints can only be based on the positioning of the considered node with respect to the origin, instead of with respect to a predefined destination.

The most important constraint is the tolerance constraint that is applied at each node in the network. In fact, there are two tolerance constraints: a constraint on travel time T (including any waiting time) and a constraint on the number of vehicle legs L. Neither value may exceed a certain threshold above the shortest path from the origin to the considered network node; the branch is discarded if one or both of these conditions are not met. However, both conditions are related: branches with less travel time are allowed to use more legs and vice versa. The exact formulation is given below, where B indicates whether the branch is further developed (i.e. the branch is discarded if $\neg B$):

$$B = \begin{pmatrix} \left(L \le \min\left(L_{ref} + 1, 4 \right) \wedge T \le \min\left(T_{ref} + 0.25\Delta T, T_{ref} + 20 \text{ minutes} \right) \right) \\ \vee \left(L \le \min\left(L_{ref}, 4 \right) \wedge T \le \min\left(1.2 T_{ref} + 0.25\Delta T, T_{ref} + 20 \text{ minutes} \right) \right) \\ \vee \left(L \le \min\left(L_{ref} - 1, 4 \right) \wedge T \le \min\left(1.4 T_{ref} + 0.25\Delta T, T_{ref} + 20 \text{ minutes} \right) \right) \end{pmatrix}$$

In this formula, T_{ref} and L_{ref} are the travel time and number of vehicle legs of the fastest PT path to the considered node. To compute these reference values, the Dijkstra algorithm is run prior to the branch and bound algorithm. Furthermore, ΔT represents the intrazonal walk time of the origin zone and also includes the intrazonal walk time of destination zones in case the considered node is a destination centroid. Note that no tolerance constraint is applied to routes consisting of a single segment (i.e. non-PT routes).

Besides the tolerance constraints, there are a few logical constraints. These prevent incorrect usage of access and egress segments (see above), enforce that transfer segments are preceded and superseded by PT segments and re-boarding the same PT line[37] directly after alighting. They also prevent going back to a previously visited node (i.e. cyclic routes) or any node that could have been visited by alighting earlier during any previously used PT strategy segment.

For example, in the following figure, the current branch cannot be extended to node A because this node was previously visited, and neither to node B because this node could have been visited by alighting earlier during the segment from A to C. The branch can also not be extended to node E because the branch just alighted from line 1. This means that only line 3 to node D is an acceptable extension to the current branch.

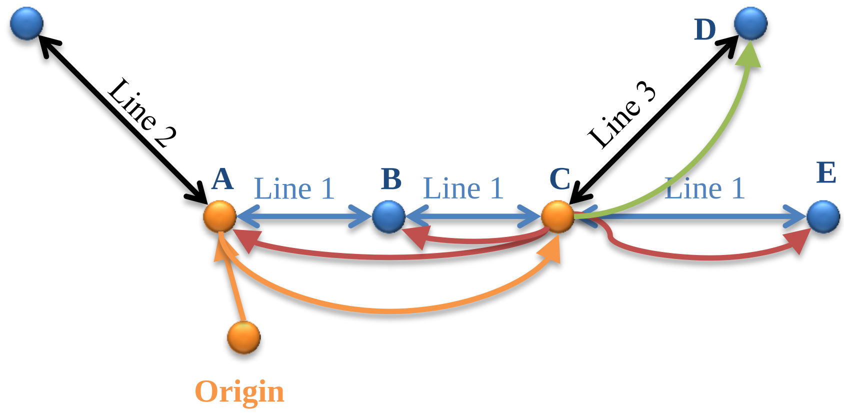


Figure 35: example of the logical constraints in a branch-and-bound situation where the segments of the current branch are indicated in orange

Route evaluation

Once a branch has been completed to one of the destinations, it is converted into routes. In this process, the groups of access and egress segments created above are unpackaged by iterating over all possible access and egress combinations. Then, for each individual route, the tolerance constraint is checked once again, now using the actual T and L for the access and egress segments instead of T_G and L_G for the segments as a group. Note that the same parameters for the tolerance constraint are used as during the development of the branches; this guarantees that at least one of the routes represented by the branch satisfies this constraint.

Finally, the routes satisfying the tolerance constraint are added to the route set, provided that two other constraints are met: the rareness constraint and the dominance constraint. The rareness constraint eliminates PT routes with neither walk access nor walk egress, which appear to be very rare in practice leading to problems during the estimation of the choice model (see Section 8.3 and Appendix H.1); again, this is not a limitation of the flexible model itself. The dominance constraint is used to check whether the new route is a useful addition to the choice set. This constraint also works the other way around: if the new route is useful (i.e.

[37] Note that the programme interprets both directions of a bidirectional PT line as separate lines. If strategy segments represent multiple lines, the programme checks whether the intersection of the sets of line numbers of both segments is empty.

not dominated), then any previously found routes that are dominated by the new route are removed from the choice set.

The dominance constraint is defined as follows:

$$D_{ba} = \begin{pmatrix} t_{ta} \geq t_{tb} \\ \wedge t_{wa} \geq t_{wb} \\ \wedge c_a \geq c_b \\ \wedge \bigwedge_{m \in M \setminus \{\text{walk}\}} n_{ma} \geq n_{mb} \end{pmatrix}$$

Here, D_{ba} indicates whether route b dominates route a, t_{ax} is the total travel time of route x, t_{wx} the waiting time, c_x the costs without any discounts (see Appendix K.1) and n_{mx} the number of legs of mode $m \in M$. In order to not be dominated and eliminated, each route must, compared to each other route, have a better score for at least one of these attributes.

Route set generation results

Without considering the actual routes that travellers choose, the generated route sets can be analysed based on their size and the used computation time. These analyses are made in this section; a comparison with observed routes will be made in Section 8.2 after the survey data has been interpreted.

Route set size

The following figure shows boxplots of the sizes of the generated route sets per type of OD relation:

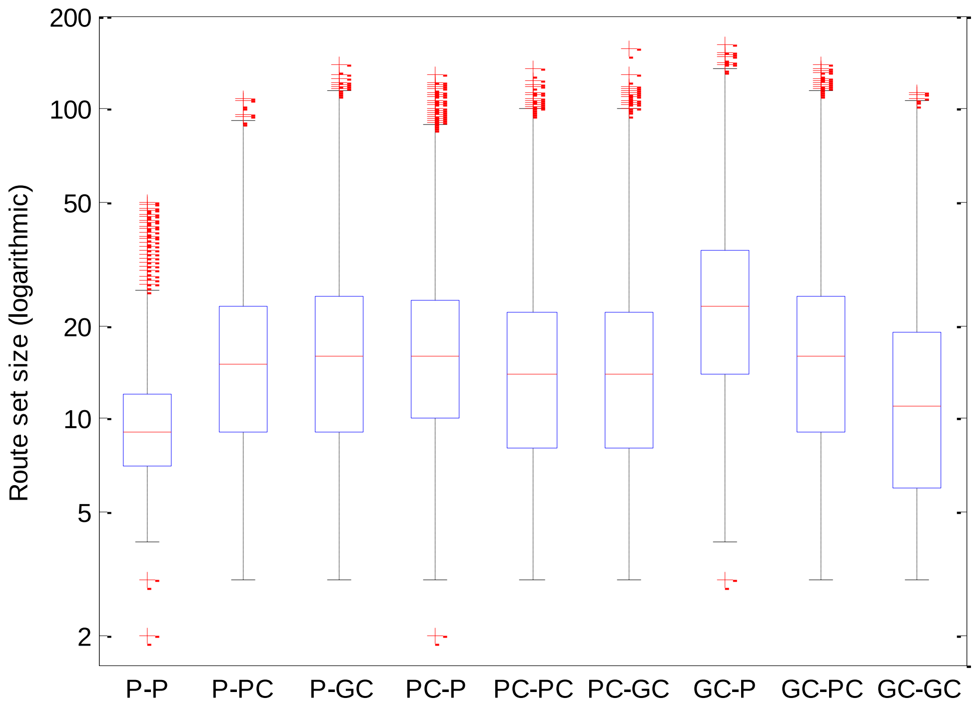


Figure 36: boxplots of route set size per OD pair assuming all private vehicles (car, motor, bicycle) are available; P stands for Paris, PC for Petite Couronne and GC for Grande Couronne

There appear to be considerable variations in the route set size between zone pairs. Also, relatively many route alternatives seem to be generated for trips from Grande Couronne to Paris: this may indicate that routes create a bubble of spare travel time under the tolerance constraint in the rural areas and that this bubble pops when the routes reach the dense urban network of Paris. The permitted tolerance should therefore perhaps not depend linearly on the reference travel time of the Dijkstra algorithm. The route sets are smaller for trips with both origin and destination inside Grande Couronne; a possible explanation is that not all destinations are reachable within four PT legs, as ANTONIN2 uses a maximum of five PT legs.

Computation time

The following figure shows boxplots of the computation time of the branch and bound algorithm (the computation time of the segment creation is negligible):

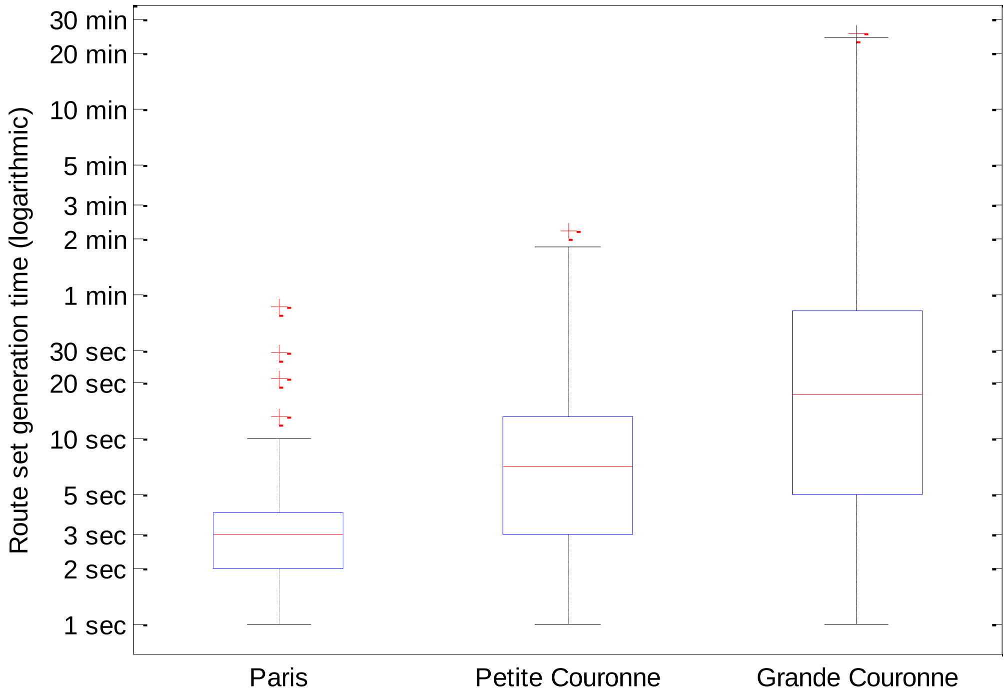


Figure 37: boxplots of branch and bound computation time used per origin zone to generate route sets to all other zones; all time registrations have been rounded down to whole seconds; time registrations with less than 1 second computation time are not displayed in the logarithmic figure

Like the choice set size, the computation time appears to vary considerably from zone to zone, although there is a clear pattern that zones further away from the centre of the network on average need more computation time. This may be related to the large choice sets that are created for trips from Grande Couronne to Paris. The constraints may therefore require further tweaking to get lower computation times in these cases. The following table shows the total computation time:

Area	Computation time	Cumulative computation time
Paris	0:17 h	0:17 h
Petite Couronne	1:27 h	1:43 h
Grande Couronne	12:08 h	13:52 h

Table 7: total computation time

For the current choice model estimation for Paris and Petite Couronne, the computation times are acceptable, especially since not all zones occur as origins in the survey data set (see Section 8.2). Furthermore, simultaneous parallel execution of the algorithm for multiple origin zones is possible, such that a large reduction in computation time is easily achieved on multi-core PCs like the computation PCs available at Significance.

However, compared to the route generation time in ANTONIN2 of about one hour, this algorithm performs worse even if the relatively high computation times for Grande Couronne are neglected. This might be caused by the density of the urban PT network of Paris (see Figure 33). Generating routes for specific mode combinations allows ANTONIN2 to use a more stringent tolerance constraint for each such combination while maintaining modal diversity: its tolerance factor is 1.05 while this case study uses a tolerance factor of up to 1.4 depending on the number of legs. On the other hand, it should be noted that the actual coverage of chosen routes, which is important for the choice model estimation of this case study (see Section 8.2), has not been studied for ANTONIN2, which has been estimated based on observed mode combinations instead of observed routes.

8.2. Survey data conversion

In order to calibrate the model, it is necessary to know which of the routes in the route sets are used and in what proportions. This requires more detail in the data than the calibration of ANTONIN2, for which only the usage proportions of mode combinations are used (Willigers & Tuinenga, 2007). The household travel survey data from the EGT 2010 contains enough information to do this[38], provided that only a single route (part) is created for non-PT legs and that the PT network is pre-processed and common route segments are merged to strategy segments (see Section 5.3). The route set generation algorithm described above has been designed to satisfy these requirements.

In this survey, all trips are split into transfer-free legs. For each leg, the mode is known, as well as the line number for RER, metro and tram legs. Since, additionally, for each leg, the boarding and alighting locations are given as carroyages, the real-world routes can be 'assigned' to the ANTONIN network by 'reverse-engineering' the approximate boarding and alighting locations to specific public transport stops. The carroyages form a grid of squares on the map with a size of 100 m.

The following algorithm to do this has been created for this case study in MATLAB R2012b. First, all legs of the trip are loaded into memory. The trip is formatted as a sequence of an odd number of legs, where all even legs are PT legs and all odd legs are private mode legs. If necessary, zero-length dummy walking legs are inserted between consecutive PT legs and walking access and egress legs are added. The first and last nodes of the trip are set to the origin and destination centroids respectively, which have been found by matching the carroyages with zone shapes. PT legs with identical start and end carroyages are cut out of the

[38] Note that ANTONIN2 has been calibrated on the EGT 2001 using a 2001 network. Major advantages of the EGT 2010 are that boarding/alighting locations have been recorded at 100 m precision instead of 300 m precision and that line numbers have been recorded for RER, metro and tram legs, allowing more precise reconstruction of chosen routes. Furthermore, the 2001 network is less detailed. On the other hand, disadvantage of using the new EGT is that the most recent network data available is for 2007.

route, since there is no way to tell what the boarding and alighting stops of this short PT leg were.

Next, the boarding and alighting stops of PT legs have to be determined. For legs of modes with lower stop densities, this is done first[39], since these stops can be determined with more certainty. For all legs, the algorithm iterates over all known PT lines from ANTONIN2 searching for the PT line that minimises the crow fly distance from the centre of the first carroyage to the boarding stop and from the alighting stop to the centre of the second carroyage[40]. If a neighbouring PT leg has been matched previously, the corresponding centre of the carroyage is replaced with the position of the ANTONIN2 node (i.e. minimising transfer walking distance) and a constraint is created that an ANTONIN2 transfer link between both legs must exist, limiting the set of possible boarding or alighting nodes for the current leg. If no line could satisfy the transfer link existence constraint, and the mode was one of the bus modes, then other bus modes are tried instead. If there still are zero-length walking legs after the whole process completed, these are removed from the route.

Errors and warnings

The algorithm described above may not always convert each route perfectly. More specifically, the following errors can occur in this algorithm that leave the route unusable:

- multiple private modes other than walking are used consecutively without a PT mode in between, e.g. carpooling[41];
- the route includes modes that are unsupported by the model, i.e. modes other than walk, Transilien, RER, metro, tram, RATP Paris bus, RATP banlieue bus, Optile bus, car driver, bicycle, motor driver and car/motor passenger – a short overview of the usage of these supported modes is provided in Appendix H;
- the route contains transfers involving walking between different PT stops for which no transfer links exist in the ANTONIN2 network;

Additionally, the following non-fatal warnings may occur, indicating that accuracy of the route conversion may be questioned:

- a short PT leg had to be eliminated due to identical boarding and alighting carroyages;
- a bus mode had to be modified to find a possible node pair;
- more than one node pair could be possible for a certain leg, that is, had a crow fly distance $d \leq 2D\sqrt{2}$ where D is the carroyage size (100 m);

The frequencies of these errors and warnings and their consequences on the total number of routes available for choice model estimation are listed in the tables below:

[39] The used order is: Transilien, RER, metro, tram, Optile bus, RATP banlieue bus, RATP Paris bus.
[40] The iteration over all lines is not strictly necessary since only a node pair needs to be determined, not a line (due to the usage of strategy segments). However, it is useful because it makes sure that a transfer-free leg between both nodes of the specified mode actually exists.
[41] If a walking leg and a non-walking leg are combined, e.g. somebody describes walking to his parked car, the walking leg is removed from the route and the algorithm continues.

Error or warning	Count	Percentage of all routes	Percentage of PT routes[42]
Consecutive private modes error	1	0.0%	N/A
Unsupported mode error	70	2.4%	N/A
Transfer link error	25	0.9%	1.5%
All errors combined	**96**	**3.3%**	**1.5%**
Short PT leg elimination warning	2	0.1%	0.1%
Bus mode modification warning	44	1.5%	2.6%
Multiple possible node pairs warning	6	0.2%	0.4%
All errors and warnings combined	**148**	**5.1%**	**4.6%**

Table 8: survey route conversion errors and warnings; if multiple warnings occur in an observed route, the route is only counted once for the warning listed first in this table.

	All routes	PT routes
Original observations	2874	1664
Observations with conversion errors	96	25
Remaining observations	**2778**	**1639**

Table 9: number of observed routes remaining available for choice model estimation

Overall, only 1.5% of the PT routes is unusable for model estimation. 95.4% of the PT routes could be interpreted without doubts. These results are deemed acceptable.

Comparison of generated and observed routes

Once the observed routes from the survey data have been converted, they can be compared to the routes in the choice set generated according to the algorithm of Section 8.1. In order to do so, the algorithm in Section 8.1 has been adapted such that additional access and egress segments are generated. These additional segments are not used to generate routes; instead, they are simply stored such that the attributes, such as travel times, waiting times and costs, of any observed route can be calculated, even if the observed route contains segments that would never be considered by the route generation algorithm. For this purpose, the only constraint on the segments is that the travel time is less than 90 minutes for motorised modes or 60 minutes for slow modes.

However, some observed segments are still not generated in this case. These are all access, egress and direct segments, since a missing PT or transfer segment would lead to a warning or error in the route conversion described above. The following table shows the number of cases in which the attributes of the observed route could not be obtained and the type of the missing segment that caused this:

[42] Routes with an error other than the transfer link error are discarded before it is determined whether the route was a PT route. The transfer link error is therefore the only error that can occur in a PT route.

Type of missing non-PT segment	Count	Percentage of all routes	Percentage of PT routes
Egress segment	5	0.2%	0.3%
Access segment	13	0.5%	0.8%
PT routes total	**18**	**0.6%**	**1.1%**
Direct segment	1	0.0%	N/A
All routes total	**19**	**0.7%**	**N/A**

Table 10: missing segments leading to failure in obtaining the attributes of observed routes; if both access and egress segments were missing for the same route, the route is only counted once for the egress segment.

	All routes	PT routes
Original observations	2874	1664
Observations with conversion errors	96	25
Observations with missing non-PT segments	19	18
Remaining observations	**2759**	**1621**

Table 11: number of observed routes remaining available for choice model estimation

The comparison then happens in three steps. These are described below.

1. First, it is checked whether any of the generated routes matches all segments of the observed route exactly.
2. Secondly, it is checked whether the modes and line numbers of all legs of the observed route match with any of the generated routes[43]. The primary reason for this is that the exact boarding and alighting stop may differ due to spatial disaggregation of origins and destinations within zones, which is not accounted for in the route set generation as could be seen in Figure 34 in Section 8.1. Transfer legs are ignored in this comparison. If there are multiple matching generated routes, the route with the least total travel time is selected.
3. Finally, it is checked whether any of the generated routes dominates the observed route according to the formula in Section 8.1. This makes sense because the route set generator removes these from the route set if they are generated. In practice, dominated routes may still be attractive to use due to timetable effects, for example efficiently scheduled transfers, which the frequency-based network model does not take into account. Additionally, travellers may seem to pick dominated routes due to network improvements, since the network is for the year 2007 while the survey was held in 2009-2011. Again, if there are multiple dominating generated routes, the route with the least total travel time is selected.

The results of this comparison give the following PT route coverage percentages for the route set generation programme:

[43] When strategy segments are being compared, it is checked whether the intersection of the sets of line numbers of both segments is non-empty.

Matching criterion	PT coverage count	PT coverage percentage	Overall coverage count	Overall coverage percentage
Exact	361	22.3%	1499	54.3%
Modes and line numbers	423	26.1%	423	15.3%
Dominance	610	37.6%	610	22.1%
Total matched	**1394**	**86.0%**	**2532**	**91.8%**
No match	227	14.0%	227	8.2%
Total	**1621**	**100.0%**	**2759**	**100.0%**

Table 12: route coverage; each observed route is only counted for the first matching criterion it satisfies; observations with unsupported mode combinations have been excluded from the overall coverage

The coverage of 86.0% for PT routes is deemed sufficiently high. In case the observed route and matched generated route are not exactly equal, the choice model can be estimated as if the matched generated route was actually chosen, so that the generated choice sets are not modified based on observations. Observations which could not be matched are therefore excluded from the choice model estimation, such that the number of observations remaining available is now as follows:

	All routes	PT routes
Original observations	2874	1664
Observations with conversion errors	96	25
Observations with missing non-PT segments	19	18
Observations not in generated choice sets	227	227
Remaining observations	**2532**	**1394**

Table 13: number of observed routes remaining available for choice model estimation

8.3. Choice model estimation

In this section, the parameters of various integrated mode and route choice models will be estimated. These choice models use the route choice set generated by the algorithm in Section 8.1 and the estimation procedure is based on the survey data as converted in Section 8.2. All estimations are maximum likelihood estimations, i.e. estimations that maximise the probability of replicating the choices that were observed in the survey results, which have been carried out in the model estimation software package ALOGIT 4.3.

For simplicity, the estimation does not contain weights for the observations. In this case, weights could be used for two purposes: to make the estimation procedure more representative by weighting the survey respondents using socio-economic variables, and to avoid bias against routes with PT and routes with many legs, since these observations were more likely to be eliminated in Section 8.2. The addition of weights might be interesting for further research.

For all models, vehicle availability is taken into account by removing route alternatives from the personal choice set of an individual based on the licence ownerships of the person and vehicle ownerships of the household the person belongs to[44]:

[44] As reported in the EGT. The ownership decisions themselves have not been modelled in this case study.

- routes containing a car driver leg are unavailable if the person has no car driving licence or his household doesn't own a car;
- motor driver routes are unavailable if the household of the person doesn't own a motor[45];
- routes containing one or more bicycle legs are unavailable if the household of the person doesn't own a bicycle.

The routes surviving these constraints form the personal subset of the generated choice set. 8 people chose the direct car driver alternative while this mode was unavailable according to the above definition; these observations have been eliminated. Finally, one observation had a personal subset consisting of only the chosen walk alternative; this observation was therefore also useless.

The above results in the following final number of observations that is available for model estimation:

	All routes	PT routes
Original observations	2874	1664
Observations with conversion errors	96	25
Observations with missing non-PT segments	19	18
Observations not in generated choice sets	227	227
Observations not in personal subset	8	0
Observations with singleton personal subset	1	0
Observations for estimation process	**2523**	**1394**

Table 14: number of observed routes that are used for choice model estimation

This section will start with the estimation of a simple multinomial logit model, which will gradually be extended with more detailed route attributes and corrections for correlations using nesting and path size factors.

Note that in the following text, significance generally refers to a p-value being smaller than 0.05. The precise test statistics are included in the accompanying tables.

Multinomial logit

Boarding penalties

To get started with the estimation process, a simple model is proposed with only travel time (including any PT waiting times) in the utility function, together with boarding penalties for train, metro/tram, bus and one for each of the private modes except walking. Here, the division of PT into train, metro/tram and bus matches the ANTONIN modes (see Table 1). By design, walking has no boarding penalty, since there is no vehicle to board and since it is a very plausible mode for very short distances in the mode chain, in which case a penalty

[45] It is possible to check for motor licence ownership as well, but not all motorised two-wheelers require a driving licence. The data set contains 50 travellers who use a motorised two-wheeler but don't have a motor licence.

unrelated to the distance does not make sense. The model as described is shown as MNL-1 in the table below:

	MNL-1	MNL-2	MNL-3	MNL-4
Log-likelihood	-3304.9	-3277.5	-3257.0	-3256.8
ρ^2	0.348	0.355	0.359	0.359
Observations	2523	2523	2523	2523
Free coefficients	8	12	15	16
Total travel time	-6.60 h^{-1} (-26.3)	-6.71 h^{-1} (-26.6)	-6.64 h^{-1} (-26.1)	-6.64 h^{-1} (-26.1)
Transilien legs	-0.37 (-5.2)	-0.20 (-1.4)	-0.16 (-1.1)	-0.16 (-1.1)
RER legs		-0.38 (-4.8)	-0.32 (-4.1)	-0.33 (-4.1)
Metro legs	-0.37 (-8.1)	-0.43 (-8.9)	-0.39 (-8.0)	-0.39 (-8.0)
Tram legs		-0.43 (-2.7)	-0.41 (-2.5)	-0.41 (-2.5)
RATP Paris bus legs	-1.61 (-26.2)	-2.19 (-20.7)	-2.16 (-20.3)	-2.16 (-20.3)
RATP banlieue bus legs		-1.21 (-16.4)	-1.19 (-15.9)	-1.19 (-15.9)
Optile bus legs		-1.68 (-6.2)	-1.67 (-6.1)	-1.68 (-6.2)
Access car driver legs	-2.06 (-17.9)	-2.06 (-18.0)	-2.64 (-8.8)	-2.65 (-8.8)
Direct car driver legs			-1.96 (2.3↲)	-1.96 (2.3↲)
Direct motor driver legs	-1.73 (-10.5)	-1.75 (-10.7)	-1.67 (-10.1)	-1.67 (-10.1)
Access bicycle legs	-3.54 (-25.3)	-3.54 (-25.3)	-5.62 (-7.9)	-5.97 (-5.9)
Egress bicycle legs				-5.10 (0.6↲)
Direct bicycle legs			-3.27 (3.3↲)	3.27 (2.7↲)
Access passenger legs	-5.45 (-28.9)	-5.44 (-29.0)	-3.76 (-9.4)	-3.76 (-9.4)
Direct passenger legs			-5.48 (-4.0↲)	-5.48 (-4.0↲)
Model improvement p-value (χ^2, df)		0.000 (67.8, 4)	0.000 (41.0, 3)	0.546 (0.4, 1)
New model adopted		Yes	Yes	No

Table 15: model estimation for various levels of boarding penalty disaggregation

In this table and in consecutive tables in this section, each column indicates a model structure, listing goodness-of-fit characteristics and the estimated values of the coefficients, where the values between parentheses indicate the t-ratios of the coefficients indicating whether the coefficients are significantly different from zero[46].

As a next step, MNL-2 disaggregates the boarding penalties by separating Transilien and RER, metro and tram and the three bus modes from each other. This yields a significantly better fit compared to MNL-1, which is indicated by the χ^2-test on the log-likelihood improvement[47]. MNL-2 is therefore adopted as the new best multinomial logit model, as indicated at the bottom of the table.

Looking at the t-ratios of the coefficients, the improvement is clearly attributable to the bus modes, where there is a large difference between RATP Paris bus and RATP banlieue bus. A possible explanation could be that the bus in Paris has more competition from the rail-based

[46] Except for nest coefficients and for t-ratios marked with an arrow (↲), as will be explained later on.
[47] The χ^2 test statistic equals the improvement in log-likelihood multiplied with 2; the degrees of freedom (df) variable is equal to the number of parameters that were added to the model. These two variables together determine the p-value.

urban modes than in suburbs, eliminating the bus alternatives from individual consideration sets for some travellers (see introduction of Chapter 6). The fact that buses in Paris are primarily used by tourists and elderly people supports the high boarding penalty for commuters. This difference between buses supports the idea that aggregation of modes may lead to problems (see Section 2.1). Section 8.4 will investigate this in more detail.

Next, MNL-3 makes a distinction for private modes between when they are used directly from origin to destination or as a means of access to or egress from public transport. In the table, some t-ratios are now marked with an arrow (↲), which means that the t-value indicates the difference from a previous coefficient instead of the difference from zero. This shows that car driving directly is significantly more attractive than driving to public transport and that direct cycling is significantly more attractive than cycling to or from PT, but that being a car/motor passenger is significantly more attractive if it is only as part of a larger PT route.

Distinguishing between bicycle access and bicycle egress as in MNL-4 has the counter-intuitive result that cycling on the activity-end is more attractive, but the model is neither significantly better than MNL-3 and hence not adopted.

Travel time components

Next, it is checked whether all travel time components are valued equally as in the models so far. This is not the case, as MNL-5 shows in the following table by splitting the travel time into walk/bicycle time, car/motor time, PT in-vehicle time and PT waiting time:

<table>
<tr><th></th><th>MNL-5</th><th>MNL-6</th><th>MNL-7</th><th>MNL-8</th></tr>
<tr><td>Log-likelihood</td><td>-3188.3</td><td>-3187.6</td><td>-3183.3</td><td>-3183.3</td></tr>
<tr><td>ρ^2</td><td>0.372</td><td>0.372</td><td>0.373</td><td>0.373</td></tr>
<tr><td>Observations</td><td>2523</td><td>2523</td><td>2523</td><td>2523</td></tr>
<tr><td>Free coefficients</td><td>18</td><td>19</td><td>19</td><td>18</td></tr>
<tr><td>Walk/bicycle time</td><td>-7.60 h^{-1} (-27.1)</td><td>-7.61 h^{-1} (-27.1)</td><td>-7.61 h^{-1} (-27.0)</td><td>-7.60 h^{-1} (-27.2)</td></tr>
<tr><td>Car/motor time</td><td>-6.80 h^{-1} (2.1↲)</td><td>-6.78 h^{-1} (2.2↲)</td><td>-6.75 h^{-1} (2.2↲)</td><td>-6.75 h^{-1} (2.2↲)</td></tr>
<tr><td>PT in-vehicle time</td><td>-4.42 h^{-1} (8.0↲)</td><td>-4.39 h^{-1} (8.0↲)</td><td>-4.37 h^{-1} (8.1↲)</td><td>-4.36 h^{-1} (8.1↲)</td></tr>
<tr><td>Short waiting time</td><td rowspan="3">-1.44 h^{-1} (7.6↲)</td><td>-5.33 h^{-1} (0.7↲)</td><td rowspan="2">-6.91 h^{-1} (0.4↲)</td><td rowspan="2">-7.00 h^{-1} (0.8↲)</td></tr>
<tr><td>Medium waiting time</td><td rowspan="2">-1.09 h^{-1} (1.2↲|)</td></tr>
<tr><td>Long waiting time</td><td>0.57 h^{-1} (3.1↲|)</td><td>0</td></tr>
<tr><td>Transilien legs</td><td>-1.01 (-5.8)</td><td>-0.70 (-2.3)</td><td>-0.42 (1.6)</td><td>-0.41 (-2.4)</td></tr>
<tr><td>RER legs</td><td>-1.18 (-9.1)</td><td>-0.87 (-3.1)</td><td>-0.69 (-3.4)</td><td>-0.68 (-5.3)</td></tr>
<tr><td>Metro legs</td><td>-0.80 (-12.8)</td><td>-0.73 (-8.9)</td><td>-0.71 (-10.3)</td><td>-0.71 (-11.5)</td></tr>
<tr><td>Tram legs</td><td>-0.91 (-5.4)</td><td>-0.76 (-3.6)</td><td>-0.72 (-4.0)</td><td>-0.72 (-4.2)</td></tr>
<tr><td>RATP Paris bus legs</td><td>-2.88 (-23.0)</td><td>-2.68 (-13.0)</td><td>-2.60 (-16.9)</td><td>-2.59 (-21.2)</td></tr>
<tr><td>RATP banlieue bus legs</td><td>-2.06 (-18.7)</td><td>-1.80 (-7.5)</td><td>-1.66 (-9.8)</td><td>-1.65 (-15.6)</td></tr>
<tr><td>Optile bus legs</td><td>-3.11 (-9.9)</td><td>-2.83 (-7.3)</td><td>-2.69 (-7.8)</td><td>-2.68 (-8.5)</td></tr>
<tr><td>Access car driver legs</td><td>-2.92 (-9.6)</td><td>-2.92 (-9.6)</td><td>-2.92 (-9.6)</td><td>-2.92 (-9.6)</td></tr>
<tr><td>Direct car driver legs</td><td>-2.10 (2.6↲)</td><td>-2.11 (2.6↲)</td><td>-2.12 (2.6↲)</td><td>-2.12 (2.6↲)</td></tr>
<tr><td>Direct motor driver legs</td><td>-1.78 (-9.6)</td><td>-1.78 (-9.6)</td><td>-1.79 (-9.6)</td><td>-1.79 (-9.6)</td></tr>
<tr><td>Acc./egr. bicycle legs</td><td>-5.81 (-8.2)</td><td>-5.81 (-8.2)</td><td>-5.82 (-8.2)</td><td>-5.82 (-8.2)</td></tr>
<tr><td>Direct bicycle legs</td><td>-3.20 (3.6↲)</td><td>-3.20 (3.6↲)</td><td>-3.20 (3.6↲)</td><td>-3.20 (3.6↲)</td></tr>
<tr><td>Access passenger legs</td><td>-4.00 (-10.0)</td><td>-4.00 (-10.0)</td><td>-4.00 (-10.0)</td><td>-4.00 (-10.0)</td></tr>
<tr><td>Direct passenger legs</td><td>-5.60 (-3.7↲)</td><td>-5.61 (-3.7↲)</td><td>-5.61 (-3.7↲)</td><td>-5.61 (-3.7↲)</td></tr>
<tr><td>p-value (χ^2, df)</td><td>0.000 (137.4, 3)</td><td>0.000 (138.5, 4)</td><td>N/A (8.4, 0)</td><td>N/A (8.4, -1)</td></tr>
<tr><td>Model adopted</td><td>No</td><td>Yes</td><td>No</td><td>Yes</td></tr>
</table>

Table 16: model estimation with various levels of travel time disaggregation

MNL-5 shows a very low valuation of waiting time, that is, quite lower than in-vehicle time. Due to this counter-intuitive result, MNL-6 is created to split the waiting time into two bins: “short waiting time” that contains the first 5 minutes of waiting time for each leg and “medium and long waiting time” that contains the rest. Now the first 5 minutes are valued higher than in-vehicle time, yet the rest is valued even lower. A likely explanation is that the waiting time is overestimated: initial waiting time may be overestimated due to departure time modification and transfer waiting time may be overestimated by neglecting timetable synchronisation (see Section 5.3).

The boundary value of 5 minutes corresponds to a service frequency of 6 runs per hour. One may question whether this is the optimal boundary; therefore, a boundary value of 7.5 minutes has been tried in MNL-7, which corresponds to a frequency of 4 per hour. This improves the log-likelihood and makes the boundary significant (t-value is 3.1 instead of 1.2), but the “long waiting time” coefficient now has the wrong sign. Eliminating this coefficient in MNL-8 does not visibly worsen the log-likelihood, such that MNL-8 is preferred over MNL-6.

It should be noted that the resulting model does not take service frequency differences into account below 4 per hour. This of course limits possible applications of the model. Note however that if the waiting time boundary were to be kept at 5 minutes, it would also have been statistically better to set the coefficient for waiting time above 5 minutes to zero, resulting in an even less applicable model. It is possible that service frequency effects below 4 per hour have been partially absorbed by the mode-specific boarding penalties.

	MNL-9
Log-likelihood	-3183.3
ρ²	0.373
Observations	2523
Free coefficients	19
Walk/bicycle time	-7.61 h^{-1} (-27.0)
Car/motor time	-6.75 h^{-1} (2.2↲)
PT in-vehicle time	-4.37 h^{-1} (8.1↲)
Non-long waiting time	-6.94 h^{-1} (0.4↲)
Max. long waiting time	-0.08 h^{-1} (-0.0)
Transilien legs	-0.42 (-1.7)
RER legs	-0.69 (-3.5)
Metro legs	-0.71 (-10.4)
Tram legs	-0.72 (-4.0)
RATP Paris bus legs	-2.60 (-17.3)
RATP banlieue bus legs	-1.65 (-10.2)
Optile bus legs	-2.69 (-7.9)
Access car driver legs	-2.92 (-9.6)
Direct car driver legs	-2.12 (2.6↲)
Direct motor driver legs	-1.79 (-9.6)
Acc./egr. bicycle legs	-5.82 (-8.2)
Direct bicycle legs	-3.20 (3.6↲)
Access passenger legs	-4.00 (-10.0)
Direct passenger legs	-5.61 (-3.7↲)
p-value (χ^2, df)	0.972 (0.0, 1)
Model adopted	No

Table 17: model estimation with various levels of travel time disaggregation (continued)

A similar maximum waiting time is present in the ANTONIN2 assignment model, but there it depends on the mode: it limits the waiting time to 7.5 minutes for rail modes and to 15 minutes for bus modes (Syndicat des transports d'Île-de-France, n.d.). In order to restrict the differences between PT modes to the boarding penalties, such a distinction is not made here.

Although the "short/medium waiting time" coefficient is not significantly different from the walk/bicycle time coefficient – the relative t-value equals 0.8 – it has been decided to not merge them into a single coefficient since both components of travel time may be considered fundamentally different: the second represents travelling in a mode while the first does not. The cut-off of waiting time at 7.5 minutes also indicates a fundamental difference. Additionally, constraining the waiting time coefficient could have unintended consequences for the boarding penalties as they may try to correct for changes in the waiting time coefficient.

MNL-9 gives a final attempt to include the effect of frequencies below 4 per hour into the model: it includes an attribute representing the waiting time above 7.5 minutes only for the leg with the largest waiting time, which is then assumed to reflect how often travellers can depart. The results are shown in Table 17. Since the new coefficient is not significant at all, MNL-9 is not adopted.

Monetary costs

Now that boarding penalties and travel time attributes have been handled, monetary costs are added as an attribute to the utility functions. These include fuel costs for car and motor – data on parking costs is not available – and PT costs, taking into account discounts for PT pass owners. Appendix K provides a detailed explanation of what formulas were used to compute these travel costs.

MNL-10 in the following table shows that this is the most significant improvement of the model so far:

	MNL-10	MNL-11	MNL-12	MNL-13
Log-likelihood	-2939.3	-2924.0	-2925.1	-2925.1
ρ²	0.421	0.424	0.424	0.424
Observations	2523	2523	2523	2523
Free coefficients	19	20	19	18
Walk/bicycle time	-8.20 h^{-1} (-26.5)	-8.08 h^{-1} (-25.7)	-8.02 h^{-1} (-25.7)	-8.02 h^{-1} (-25.7)
Car/motor time	-5.63 h^{-1} (6.1↲)	-7.45 h^{-1} (1.1↲)	-7.98 h^{-1} (0.1↲)	
PT in-vehicle time	-4.87 h^{-1} (7.9↲)	-4.51 h^{-1} (8.4↲)	-4.40 h^{-1} (8.6↲)	-4.42 h^{-1} (11.0↲)
Non-long waiting time	-7.35 h^{-1} (1.0↲)	-7.18 h^{-1} (1.0↲)	-7.08 h^{-1} (1.3↲)	-7.09 h^{-1} (1.3↲)
PT costs	-1.26 €$^{-1}$ (-19.0)	-1.42 €$^{-1}$ (-18.4)	-1.43 €$^{-1}$ (-18.4)	-1.43 €$^{-1}$ (-18.6)
Car/motor costs		-0.30 €$^{-1}$ (5.3↲)	0	0
Transilien legs	-0.05 (-0.3)	0 (max)	0 (max)	0 (max)
RER legs	-0.46 (-3.4)	-0.38 (-2.8)	-0.37 (-3.1)	-0.37 (-3.2)
Metro legs	-0.51 (-7.8)	-0.50 (-7.6)	-0.50 (-7.8)	-0.50 (-7.8)
Tram legs	-0.39 (-2.2)	-0.35 (-2.0)	-0.35 (-2.0)	-0.35 (-2.0)
RATP Paris bus legs	-2.39 (-19.1)	-2.40 (-19.0)	-2.41 (-19.6)	-2.41 (-19.7)
RATP banlieue bus legs	-1.39 (-12.4)	-1.36 (-12.1)	-1.36 (-13.2)	-1.36 (-13.7)
Optile bus legs	-2.69 (-7.7)	-2.66 (-7.5)	-2.65 (-7.8)	-2.65 (-7.8)
Access car driver legs	-3.00 (-9.7)	-2.88 (-9.2)	-2.84 (-9.1)	-2.83 (-9.1)
Direct car driver legs	-2.39 (1.9↲)	-2.35 (1.7↲)	-2.32 (1.6↲)	-2.32 (1.6↲)
Direct motor driver legs	-2.25 (-11.5)	-2.21 (-11.2)	-2.18 (-11.1)	-2.17 (-12.2)
Acc./egr. bicycle legs	-5.91 (-8.2)	-5.86 (-8.2)	-5.83 (-8.1)	-5.84 (-8.1)
Direct bicycle legs	-3.57 (3.2↲)	-3.54 (3.2↲)	-3.52 (3.2↲)	-3.53 (3.2↲)
Access passenger legs	-4.41 (-10.7)	-4.03 (-9.7)	-3.91 (-9.7)	-3.91 (-9.7)
Direct passenger legs	-7.04 (-5.8↲)	-5.89 (-4.1↲)	-5.59 (-3.9↲)	-5.59 (-3.9↲)
p-value (χ^2, df)	0.000 (488.1, 1)	0.000 (30.6, 1)	N/A (28.4, 0)	0.920 (-0.0, -1)
Model adopted	Yes	No	Yes	Yes

Table 18: model estimation with travel costs included

When splitting the costs into car/motor costs and PT costs as in MNL-11, the valuation of car/motor costs becomes very low, resulting in a unrealistically high value-of-time for car/motor (25 €/h), while the cost coefficient is not significantly different from zero. This may be caused by the high correlation between the time and cost coefficients of car/motor (-0.657[48]). MNL-11 is therefore not adopted as the new best model.

The boarding penalty for Transilien legs also has become (insignificantly) positive. Since it is theoretically questionable that a positive utility is associated with unnecessarily boarding and

[48] To be precise, this is the correlation between the difference between the coefficients of PT in-vehicle time and walk/bicycle/wait time, and the difference between the coefficients of car/motor costs and PT costs. This is a consequence of how the estimation has been defined in ALOGIT.

alighting a mode without travel and waiting time, it has been constraint to zero as shown in the table. However, to correctly assess the benefit of adding other new attributes to the model, it is still counted as a free coefficient and it will be re-included if it gets positive again in further models[49], since for a comparison of modes it is better to include all possible boarding penalties in the estimation process.

The car/motor costs have been removed from the model in MNL-12, which is better than MNL-10, while MNL-11 only was insignificantly better according to the χ^2-test. The value-of-time for PT now equals about 3 €/h. This is rather low for a travel purpose of home-work, but the inclusion of the cost attribute is a significant improvement for the model. The value-of-time will however improve later on (see Table 19 below).

In MNL-12, the difference between time inside a car or motor is valued and walking or cycling time has become insignificant due to the introduction of PT costs into the model. Therefore, finally, MNL-13 merges the coefficient for car/motor time with the coefficient for walk/bicycle time, resulting in a general coefficient for travel time inside private modes, so that the differences among private modes are embedded exclusively in the boarding penalties, just like the differences among PT modes.

Other attributes

Before switching to path size logit and nested logit models, there is still some tweaking that can be done to the multinomial logit model. First of all, it has to be noted that the matching by dominance played a large role in the preparation of the survey data (see Table 12). One could therefore wonder whether the number of routes that are dominated by a route in the choice set, i.e. the number of routes n that were directly or indirectly eliminated because of that route, influence the probability that the route is chosen. Therefore, the following size variable X_{dom} is introduced as a new attribute:

$$X_{dom} = \ln(1+n)$$

This attribute can be interpreted as a nested choice among the route and its dominated routes, where the utilities of the dominated routes are considered equal to the utility of the dominating route and $\beta_{dom} \in (0,1]$ would be the nest coefficient. MNL-14 in the following table includes this new attribute, leading to a significantly better model:

[49] Each model is initially estimated with the Transilien boarding penalty unconstrained. If it has the wrong sign, the estimation is repeated with this boarding penalty constrained to zero and only the second results are listed in this report.

	MNL-14	MNL-15	MNL-16	MNL-17
Log-likelihood	-2908.3	-2905.9	-2857.6	-2827.7
ρ²	0.427	0.428	0.437	0.443
Observations	2523	2523	2523	2523
Free coefficients	19	20	20	21
Private mode time	-7.27 h^{-1} (-21.8)	-7.12 h^{-1} (-20.9)	-7.59 h^{-1} (-22.2)	-7.11 h^{-1} (-20.2)
PT in-vehicle time	-3.81 h^{-1} (10.5↲)	-3.83 h^{-1} (9.8↲)	-4.11 h^{-1} (10.4↲)	-4.57 h^{-1} (7.1↲)
Non-long waiting time	-5.43 h^{-1} (2.4↲)	-4.89 h^{-1} (2.9↲)	-5.89 h^{-1} (2.0↲)	-5.65 h^{-1} (1.7↲)
PT costs	-1.44 €$^{-1}$ (-18.8)	-1.46 €$^{-1}$ (-18.6)	-0.52 €$^{-1}$ (-4.9)	-0.28 €$^{-1}$ (-2.6)
PT usage w/o discount	0	0	-2.24 (-9.5)	-3.00 (-11.9)
PT transfers	0	-0.27 (-2.2)	0	-1.27 (-7.6)
Transilien legs	0 (max)	0 (max)	-0.02 (-0.1)	0.89 (4.0)
RER legs	-0.31 (-2.6)	-0.16 (-1.2)	-0.42 (-3.1)	0.52 (2.8)
Metro legs	-0.42 (-6.3)	-0.21 (-1.7)	-0.44 (-6.5)	0.62 (4.0)
Tram legs	-0.21 (-1.2)	-0.04 (-0.2)	-0.17 (-1.0)	0.83 (3.6)
RATP Paris bus legs	-2.40 (-19.6)	-2.21 (-14.8)	-2.39 (-19.0)	-1.37 (-7.4)
RATP banlieue bus legs	-1.35 (-13.6)	-1.19 (-9.6)	-1.30 (-11.7)	-0.28 (-1.6)
Optile bus legs	-2.60 (-7.7)	-2.46 (-7.1)	-2.52 (-7.1)	-1.36 (-3.6)
Access car driver legs	-2.76 (-8.8)	-2.75 (-8.8)	-2.82 (-9.0)	-2.84 (-9.0)
Direct car driver legs	-1.98 (2.4↲)	-1.87 (2.7↲)	-2.26 (1.8↲)	-1.88 (2.9↲)
Direct motor driver legs	-1.59 (-7.9)	-1.49 (-7.2)	-1.89 (-9.1)	-1.54 (-7.3)
Acc./egr. bicycle legs	-5.77 (-8.1)	-5.75 (-8.0)	-5.83 (-8.1)	-5.81 (-8.1)
Direct bicycle legs	-3.57 (3.0↲)	-3.46 (3.1↲)	-3.83 (2.7↲)	-3.43 (3.3↲)
Access passenger legs	-3.85 (-9.5)	-3.85 (-9.5)	-3.90 (-9.7)	-3.92 (-9.7)
Direct passenger legs	-5.22 (-3.1↲)	-5.11 (-2.9↲)	-5.48 (-3.6↲)	-5.08 (-2.6↲)
Domination size	0.16 (5.8)	0.16 (5.6)	0.16 (5.6)	0.14 (5.0)
p-value (χ^2, df)	0.000 (33.7, 1)	0.030 (4.7, 1)	0.000 (101.5, 1)	0.000 (59.7, 1)
Model adopted	Yes	No	Yes	No

Table 19: model estimation with several other attributes

Next, MNL-15 introduces an additional attribute representing the number of PT-to-PT transfers in the route. This is an additional penalty on top of the normal boarding penalty for each but the first PT leg.

Although MNL-15 performs significantly better than MNL-14, there are reasons to be cautious adopting this new model structure: there may exist a methodological bias towards PT routes with few legs. Illegal transfer errors, which are more likely to occur on routes with many transfers, and zero-length PT leg warnings in the survey route conversion may have resulted in an estimation bias against PT routes with many legs. Furthermore, and probably more importantly, the domination-based matching may cause chosen routes to have fewer legs in the model than in reality, which may also contribute to over-estimation of the transfer penalty.

Instead of a transfer penalty for each but the first PT leg, one could also use an attribute that represents just the first leg: this would be a boarding penalty for the PT system as a whole. Because the transfer penalty was indeed a penalty, this replacement would result into a constant representing a bonus for PT usage. Since this bonus may have been caused by the large number of people owning a PT pass (see Appendix I.2) – implying they may intend to use at least one PT leg for their home-work trips regardless of the PT level-of-service, since the first PT boarding penalty has virtually already been paid when buying the PT pass – such a bonus for the first leg may not actually exist: it just represents "sunk disutility". Therefore, in addition to the reasons mentioned previously, MNL-15 is not adopted.

However, when this PT usage constant is only estimated for travellers without any PT discounts, as in MNL-16, it is significantly negative: using PT while not having discounts is apparently perceived comparable to an additional boarding of an Optile bus. The value-of-time (8 €/h) increased considerably due to a strong correlation between the new constant and the cost coefficient (-0.774).

Although the model fit improves a lot in MNL-16, it is somewhat questionable whether this really is a causal effect (i.e. having discounts makes PT more attractive, apart from the reduced cost effect, for example due to ease of access) or merely correlates with socio-economic characteristics (e.g. people without discounts are usually more time-sensitive and therefore avoid PT). In the latter case, including the new constant means that the benefit of buying a PT pass would be over-estimated. Socio-economic disaggregation of the model, i.e. creating multiple user classes, may therefore yield better results. However, since such disaggregation is outside the scope of this research project, the new model has been adopted due to the improvement in model fit and value-of-time.

One might wonder whether an additional penalty for travellers without PT discounts should be included once, like in the current model, or should be multiplied with the number of legs, as an increment to the boarding penalties. Here, the former has been chosen, since the latter would imply an even greater correlation between this additional boarding penalty and the travel costs: the cost coefficient seems to be mainly determined by travellers without PT discounts who do not choose a PT alternative.

MNL-17 again adds the number of transfers as an attribute, which now is even more significant than in MNL-15. The explanation for this paradox is that this relatively benefits the first PT boarding, and thus makes all PT alternatives more attractive. Unlike in MNL-16, the constant for travellers without PT discounts compensates for this by getting more negative, leading to the net result that the first PT boarding is more attractive only for people with PT discounts. Considering that the travel purpose is home-work, this is likely not a causal effect, but the other way around: PT discounts (i.e. buying PT passes) are more attractive for travellers who intend to board a PT mode at least once.

Furthermore, MNL-17 produces multiple significant boarding penalties with an incorrect sign. These would need to be eliminated to give a good interpretation to the model (other than the PT discounts effect described above), thereby thus limiting possibilities to investigate modal preferences[50]. For these reasons, MNL-17 has not been adopted despite its significance.

[50] In theory, one could create a full matrix of transfer penalties for each pair of modes, but this of course creates a lot of new parameters and does not solve the problem that PT pass holders distort the estimation.

Path size logit

Section 7.1 states that route overlap should be corrected for by using path size factors. Therefore, the model PSL-1 introduces a (non-extended) leg-based path size factor for routes using PT, in which a PT route consists of one access leg, one or more PT legs and one egress leg, so transfer walk segments are not counted as legs. Since the availability of routes depends on vehicle ownership of the household of the traveller, the path size factor also depends on vehicle ownership. The table below shows the estimation results:

	PSL-1	PSL-2
Log-likelihood	-2853.5	-2853.3
ρ^2	0.438	0.438
Observations	2523	2523
Free coefficients	21	21
Private mode time	-7.61 h^{-1} (-22.3)	-7.64 h^{-1} (-22.3)
PT in-vehicle time	-4.25 h^{-1} (9.9↲)	-4.15 h^{-1} (10.3↲)
Non-long waiting time	-5.80 h^{-1} (2.1↲)	-5.81 h^{-1} (2.1↲)
PT costs	-0.57 €$^{-1}$ (-5.3)	-0.54 €$^{-1}$ (-5.1)
PT usage w/o discount	-2.18 (-9.2)	-2.23 (-9.4)
Transilien legs	-0.09 (-0.5)	-0.07 (-0.4)
RER legs	-0.53 (-3.7)	-0.50 (-3.6)
Metro legs	-0.49 (-7.0)	-0.47 (-6.9)
Tram legs	-0.21 (-1.2)	-0.21 (-1.2)
RATP Paris bus legs	-2.44 (-19.2)	-2.42 (-19.2)
RATP banlieue bus legs	-1.33 (-11.9)	-1.33 (-11.9)
Optile bus legs	-2.58 (-7.2)	-2.60 (-7.3)
Access car driver legs	-2.86 (-9.1)	-2.86 (-9.2)
Direct car driver legs	-2.25 (1.9↲)	-2.25 (1.9↲)
Direct motor driver legs	-1.90 (-9.2)	-1.89 (-9.2)
Acc./egr. bicycle legs	-5.87 (-8.2)	-5.88 (-8.2)
Direct bicycle legs	-3.80 (2.8↲)	-3.81 (2.8↲)
Access passenger legs	-3.92 (-9.7)	-3.93 (-9.7)
Direct passenger legs	-5.48 (-3.6↲)	-5.48 (-3.5↲)
Domination size	0.15 (5.1)	0.16 (5.4)
Path size	-0.34 (-2.8)	0
PT part path size	0	-0.31 (-2.9)
p-value (χ^2, df) compared to MNL-16	0.005 (8.0, 1)	0.004 (8.5, 1)

Table 20: model estimation including a path size factor

In PSL-2, the path size factor has been defined differently: only the PT part of a route is considered for the calculation of the path size factors. Also, if multiple routes share the PT part of the route exactly, that is, only differ in access and/or egress modes, they are counted as a single route when deciding how many routes contain a certain leg. This redefinition of the path size factor eliminates the variety of access and egress modes from the route overlap measure, such that it focuses strictly on the PT part of the route.

The table shows that in both models, the path size coefficient is estimated significantly negatively, while according to the theory in Appendix C.8 it should be positive. One could think of the following explanations for this:

- It is known that path size logit models tend to have a bias towards long routes (see Appendix C.8), such that a negative coefficient could create a relative bonus for the first PT leg, like in MNL-17. However, the constant for PT usage without discount does not get more negative, so this cannot have been used to benefit PT routes specifically for PT pass holders.

- The variety of access and egress modes may also be a benefit, as it provides additional freedom of choice. This could explain why the path size factor would be negative for PSL-1, but not for PSL-2, so this isn't a proper explanation either.

- A possible proper explanation for the negative coefficient is that overlap is simply valued positively, which may be caused by ad hoc route choice behaviour (i.e. travellers still modify their routes after departure based on the actual situation they encounter at PT stops). In the positively valued route overlap context, choosing routes with overlap is a strategy that reduces the risk of missed, delayed or cancelled PT services, as they can be substituted with other services, where these "other services" may appear to be dominated routes according to the model. The large number of travellers choosing seemingly dominated routes (see Table 12) and the significance of the domination size coefficient (see Table 19) would then be explained as a combination of timetable synchronisation effects and ad hoc route choice effects.

However, even in this last case, path size factors are not designed to take such effects into account: it is based on common legs rather than on ad hoc alternatives at intermediate PT stops or stations. This lack of a theoretical foundation for negative path size coefficients in this context also leads to strange consequences in some practical situations, as in the following example:

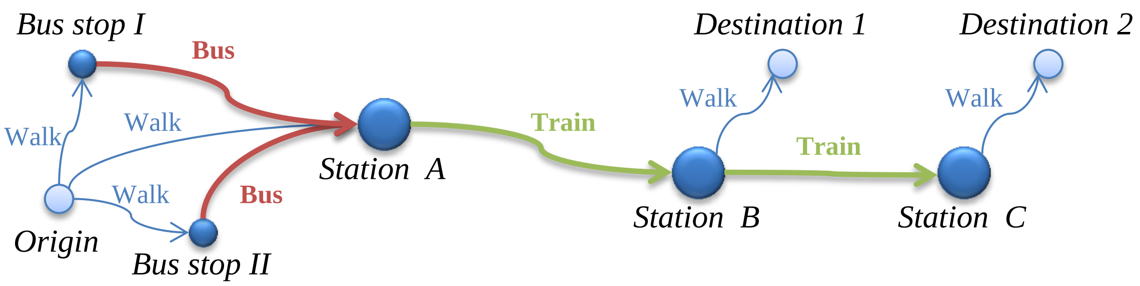


Figure 38: example situation with inconsistencies between the negative path size coefficient and ad hoc route choice behaviour

In this simple example, the overlap between the three routes would be greater for destination 2 than for destination 1, so that the ability to change routes is counter-intuitively valued higher. Furthermore, when the routes are reversed, such that one does not have to walk before boarding a bus, this makes no difference if path size factors are used, while one behaviourally expects ad hoc route choice behaviour to play a larger role in this case.

The path size models are therefore not recommended for use with these negative coefficients, while noting that the influence of ad hoc route choice behaviour could be an interesting subject for further research, both theoretical and practical.

Apart from the fact that the path size coefficient is negative, a reason that the path size coefficient is not positive could be that domination-based matching plays an important role, which is again shown by the large share of survey routes that were matched this way and the domination size coefficient being significantly positive. Because of this, an alternative actually represents more than one physical route, while the path size factor can only measure the amount of overlap in one of these. This makes it hard to consider physical route overlap in such situations, since the knowledge about the physical routes is limited in the choice model.

Nested logit

According to Section 7.1, a network GEV model would be ideal to capture all effects of modal overlap and mode similarities. However, as explained in Appendix J.2, the estimation of such a model turned out to be problematic with current choice model estimation software. Therefore, this network GEV model will be simplified to a nested logit model in this case study. This model structure can take some, but not all, effects of modal overlap and mode similarities into account.

Before proceeding with estimation results, some comments about nest coefficients must be made. Nest coefficients are defined in Appendix C.3. In case a nest with nest coefficient θ contains a single alternative, one can easily see that the utility of the alternative V is multiplied with a factor θ. If other alternatives are not in this nest, their alternatives are not multiplied with θ. The coefficients $\boldsymbol{\beta}$ in the utility functions are therefore not directly comparable (ALOGIT Software & Analysis, 2007), which is a problem is the same coefficient must be used in all of the utility functions, as in the current situation; the travel time coefficient is a simple example of this.

This could be solved by switching to the cross-nested logit and network GEV definition of nest coefficients as in Appendices C.4 and C.6, in which the utilities are divided by θ before a logsum is calculated[51], but this is not supported by ALOGIT (ALOGIT Software & Analysis, 2007). Instead, additional dummy nests have been inserted between the alternatives and their immediate parent nests to multiply their utilities with all nest coefficients with which they would not be multiplied otherwise. This makes sure the same coefficients $\boldsymbol{\beta}$ can be used across nests while allowing nest coefficients $\boldsymbol{\theta}$ to differ per nest.

However, since this means that all coefficients $\boldsymbol{\beta}$ are eventually multiplied with all nest coefficients $\boldsymbol{\theta}$ in the application of the model, the scale of these coefficients $\boldsymbol{\beta}$ is no longer comparable with the previous non-nested models. To correct for this, all coefficient tables below transform the estimated coefficients according to the following formulas, where the original β and θ match the definitions in Appendix C.3:

[51] Note that $e^{V/\theta} = e^{V^{1/\theta}} = \sqrt[\theta]{e^V}$.

$$\beta_{displayed} = \beta \cdot \prod \theta$$
$$\theta_{displayed} = \theta \in (0,1]$$

All t-tests for nest coefficients check whether the θ value is significantly different from one, unlike the t-tests for other coefficients which compare the value to zero. The reason for this is that a value of one indicates the nesting is irrelevant. For positive correlations between the route alternatives, the nesting coefficients should lie between zero and one. See Appendix C.3 for details.

Main modes

As in Section 7.1, the nesting structure will be based on modes. However, unlike the network GEV model, in a nested logit model, each alternative can have only one direct parent nest. This means that each multi-modal route must be allocated to a single mode instead of to multiple modes. To this end, for each route a main mode is defined according to the following rules:

- if the route contains Transilien or RER legs, the main mode is “train”;
- else, if the route contains metro or tram legs, the main mode is “metro/tram”;
- else, if the route contains RATP Paris bus, RATP banlieue bus or Optile bus legs, the main mode is “bus”;
- else the route must be a uni-modal non-PT route and the main mode is the corresponding private mode.

These rules lead to the following nested logit choice tree as a simplification of the more advanced network GEV choice tree:

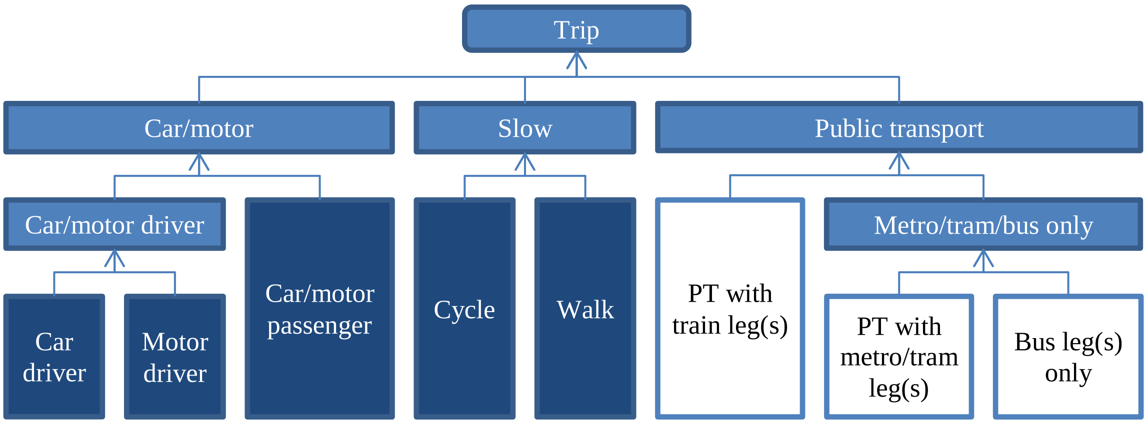


Figure 39: tested nesting structure for nested logit estimation based on main modes

In this choice tree, the private modes represent only a single alternative if they are available, while in the PT part of the tree, multiple generated routes may be nested under each of the three nests representing main modes for PT.

Model NL-1 in the following table has been estimated with all corresponding nesting coefficients unconstrained:

	NL-1	NL-2	NL-3	NL-4
Log-likelihood	-2828.8	-2835.6	-2835.8	-2835.8
ρ^2	0.443	0.442	0.442	0.442
Observations	2523	2523	2523	2523
Free coefficients	28	25	23	22
Private mode time	-7.00 h^{-1} (-3.1)	-6.89 h^{-1} (-4.5)	-6.90 h^{-1} (-6.0)	-6.90 h^{-1} (-12.1)
PT in-vehicle time	-3.70 h^{-1} (3.1↲)	-3.57 h^{-1} (4.4↲)	-3.59 h^{-1} (5.6↲)	-3.59 h^{-1} (9.3↲)
Non-long waiting time	-6.28 h^{-1} (1.0↲)	-6.22 h^{-1} (0.9↲)	-6.20 h^{-1} (1.0↲)	-6.19 h^{-1} (1.0↲)
PT costs	-0.39 €$^{-1}$ (-2.6)	-0.38 €$^{-1}$ (-3.1)	-0.39 €$^{-1}$ (-3.5)	-0.39 €$^{-1}$ (-4.1)
PT usage w/o discount	-2.30 (-3.0)	-2.29 (-4.1)	-2.29 (-5.2)	-2.28 (-7.9)
Transilien legs	-0.03 (-0.2)	-0.02 (-0.2)	-0.02 (-0.2)	-0.03 (-0.2)
RER legs	-0.33 (-2.1)	-0.31 (-2.4)	-0.32 (-2.5)	-0.32 (-2.8)
Metro legs	-0.34 (-2.9)	-0.34 (-3.8)	-0.34 (-4.6)	-0.34 (-6.3)
Tram legs	-0.15 (-1.0)	-0.15 (-1.1)	-0.15 (-1.1)	-0.15 (-1.1)
RATP Paris bus legs	-1.80 (-3.2)	-1.76 (-4.5)	-1.74 (-6.3)	-1.74 (-13.5)
RATP banlieue bus legs	-1.09 (-3.1)	-1.07 (-4.5)	-1.07 (-5.9)	-1.07 (-10.0)
Optile bus legs	-2.06 (-2.9)	-2.03 (-3.9)	-2.02 (-4.6)	-2.03 (-6.6)
Access car driver legs	-2.37 (-3.1)	-2.32 (-4.3)	-2.32 (-5.3)	-2.32 (-8.5)
Direct car driver legs	-2.22 (0.6↲)	-2.16 (0.6↲)	-2.17 (0.6↲)	-2.17 (0.6↲)
Direct motor driver legs	-1.97 (-3.1)	-1.89 (-4.1)	-1.89 (-5.0)	-1.89 (-7.5)
Acc./egr. bicycle legs	-4.53 (-3.0)	-4.44 (-4.1)	-4.43 (-4.9)	-4.43 (-7.5)
Direct bicycle legs	-4.01 (0.8↲)	-3.67 (1.3↲)	-3.67 (1.3↲)	-3.67 (1.3↲)
Access passenger legs	-3.16 (-3.1)	-3.10 (-4.3)	-3.09 (-5.3)	-3.09 (-8.7)
Direct passenger legs	-6.14 (-2.7↲)	-5.37 (-3.4↲)	-5.37 (-4.1↲)	-5.37 (-5.0↲)
Domination size	0.12 (2.8)	0.11 (3.3)	0.11 (3.8)	0.11 (4.7)
Car/motor nest	1.32 (2.3)	1	1	1
Car/motor driver nest	1.01 (0.1)			
PT nest	0.71 (-3.4)	0.70 (-3.7)	0.74 (-4.7)	0.74 (-4.8)
Train nest	1.06 (0.5)	1.06 (0.5)	1	1
Metro/tram/bus nest	1.11 (0.7)	1.11 (0.7)		
Metro/tram nest	0.74 (-3.0)	0.74 (-3.0)	0.77 (-3.9)	0.77 (-4.1)
Bus nest	0.97 (-0.1)	0.98 (-0.1)	1.04 (0.2)	1
Slow nest	1.30 (2.0)	1	1	1
p-value (χ^2, df) compared to MNL-16	0.000 (57.5, 8)	0.000 (44.0, 5)	0.000 (43.5, 3)	0.000 (43.4, 2)

Table 21: model estimation using nesting based on main modes

For the private modes, none of the nesting coefficients is below one, while the car/motor nest coefficient and the slow nest coefficient are significantly above one, meaning that multiple available alternatives may reinforce each other's attractiveness. This might for example indicate that car owners are more likely to travel as car passenger and that cycling is more attractive for destinations that can be reached by walking less than an hour (see Section 8.1).

If it would be desired to take such effects into account, more advanced utility functions should be used. For now, these nest coefficients are constrained to one to get a consistent model (see

Appendix C.3). This has been done in NL-2, where the coefficients of the train nest and the metro/tram/bus nest are still (insignificantly) greater than one. NL-3 sets these coefficients to one as well, moving the bus nest coefficient above one. In NL-4 this coefficient has also been constrained to one, leaving a model with two nesting coefficients significantly below one, providing a significantly better fit than MNL-16, the best multinomial logit model.

Although significant nests have thus been found in NL-4, there are not as many as one could expect based on a theoretical analysis. However, the interpretation of those nests matches with the theory regarding utility correlations in Section 4.6: the metro/tram nest can be seen as the effect of modal overlap, while the PT nest can be seen as the effect of mode similarities. This shows that when adding an extra alternative, an additional public transport route does not add as much benefit as an additional private mode, and an additional metro/tram route does not add as much benefit as another additional PT route, given that the expected value of the utility is equal. The following figure sketches the resulting choice tree:

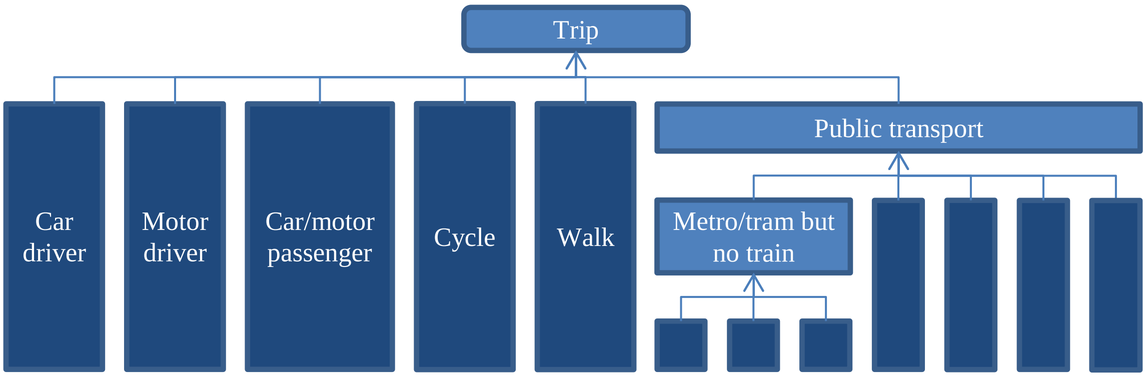


Figure 40: choice tree corresponding to NL-4

In order to compare this result, the following table shows some more nested logit model estimations. The first one, NL-5, includes only a public transport nest without sub-nests for the three main modes:

	NL-5	NL-6	NL-7
Log-likelihood	-2842.5	-2840.8	-2841.0
ρ²	0.440	0.441	0.441
Observations	2523	2523	2523
Free coefficients	21	22	22
Private mode time	-6.86 h^{-1} (-15.2)	-6.87 h^{-1} (-12.6)	-6.88 h^{-1} (-15.1)
PT in-vehicle time	-3.55 h^{-1} (10.3↲)	-3.56 h^{-1} (9.5↲)	-3.56 h^{-1} (10.2↲)
Non-long waiting time	-6.21 h^{-1} (1.1↲)	-6.22 h^{-1} (0.9↲)	-6.20 h^{-1} (1.2↲)
PT costs	-0.37 €$^{-1}$ (-4.1)	-0.38 €$^{-1}$ (-4.1)	-0.37 €$^{-1}$ (-4.1)
PT usage w/o discount	-2.30 (-8.5)	-2.27 (-8.0)	-2.32 (-8.6)
Transilien legs	0 (max)	-0.02 (-0.1)	0 (max)
RER legs	-0.25 (-2.8)	-0.29 (-2.6)	-0.24 (-2.6)
Metro legs	-0.37 (-7.0)	-0.36 (-6.6)	-0.35 (-6.7)
Tram legs	-0.18 (-1.3)	-0.17 (-1.2)	-0.16 (-1.2)
RATP Paris bus legs	-1.73 (-18.9)	-1.68 (-16.0)	-1.79 (-17.3)
RATP banlieue bus legs	-1.06 (-12.8)	-1.05 (-10.4)	-1.08 (-12.6)
Optile bus legs	-1.99 (-7.5)	-2.02 (-6.7)	-1.98 (-7.4)
Access car driver legs	-2.29 (-9.7)	-2.31 (-8.7)	-2.29 (-9.8)
Direct car driver legs	-2.16 (0.5↲)	-2.17 (0.6↲)	-2.15 (0.6↲)
Direct motor driver legs	-1.88 (-8.1)	-1.89 (-7.5)	-1.88 (-8.0)
Acc./egr. bicycle legs	-4.30 (-8.6)	-4.39 (-7.6)	-4.28 (-8.7)
Direct bicycle legs	-3.66 (1.2↲)	-3.67 (1.3↲)	-3.66 (1.2↲)
Access passenger legs	-3.02 (-10.4)	-3.07 (-8.9)	-3.01 (-10.4)
Direct passenger legs	-5.36 (-5.5↲)	-5.37 (-5.2↲)	-5.36 (-5.5↲)
Domination size	0.11 (4.7)	0.11 (4.7)	0.11 (4.7)
Public transport nest	0.68 (-6.7)	0.72 (-5.1)	0.77 (-2.9)
Metro/tram/bus nest	N/A	0.87 (-1.9)	N/A
Train/metro/tram nest		N/A	0.87 (-1.9)
Predecessor	MNL-16	NL-5	NL-5
p-value (χ², df) compared to predecessor	0.000 (30.1, 1)	0.066 (3.4, 1)	0.081 (3.0, 1)

Table 22: model estimation using nesting based on main modes (continued)

NL-5 is also a significant improvement over MNL-16, while NL-4 provides another significant improvement on top of that[52]. Instead of creating a metro/tram sub-nest as in NL-4, it is also possible to create a sub-nest for relatively short-distance modes (metro/tram/bus), as in NL-6, or a sub-nest for rail modes (train/metro/tram), as in NL-7. In both cases, the χ^2-test does not indicate a significant improvement compared to NL-5, but either model does show significance in a one-tailed t-test on the nest coefficient itself[53]. NL-6 slightly outperforms NL-7, but NL-4 clearly outperforms both.

[52] For the improvement of NL-4 compared to NL-5, the p-value is 0.000 (χ^2 equals 13.3, df equals 1).
[53] For nest coefficients, a one-tailed t-test may be used for testing whether nest coefficients are smaller than one, since a nest coefficient should by definition not be larger than one. The critical value then is 1.64 instead of 1.96

PT route overlap

Since the path size logit models failed to represent a correction for route overlap, one may wonder whether route overlap could instead be captured by a nested logit model. To this end, routes with identical PT parts, thus differing only in access and egress modes, have been grouped together in nests. In the table below, NL-8 shows the effect of these nests, while NL-9 combines these nests with a PT nest:

	NL-8	NL-9
Log-likelihood	-2854.9	-2838.0
ρ^2	0.438	0.441
Observations	2523	2523
Free coefficients	21	22
Private mode time	-7.60 h^{-1} (-4.6)	-6.84 h^{-1} (-5.3)
PT in-vehicle time	-4.13 h^{-1} (4.3↲)	-3.56 h^{-1} (4.9↲)
Non-long waiting time	-5.88 h^{-1} (1.8↲)	-6.24 h^{-1} (1.0↲)
PT costs	-0.52 €$^{-1}$ (-3.5)	-0.37 €$^{-1}$ (-3.3)
PT usage w/o discount	-2.24 (-4.2)	-2.30 (-4.7)
Transilien legs	-0.01 (-0.1)	0 (max)
RER legs	-0.42 (-2.6)	-0.24 (-2.4)
Metro legs	-0.44 (-3.8)	-0.36 (-4.3)
Tram legs	-0.17 (-0.9)	-0.17 (-1.3)
RATP Paris bus legs	-2.39 (-4.6)	-1.69 (-5.2)
RATP banlieue bus legs	-1.30 (-4.4)	-1.04 (-5.0)
Optile bus legs	-2.56 (-4.1)	-2.01 (-4.6)
Access car driver legs	-3.56 (-6.2)	-2.91 (-6.8)
Direct car driver legs	-2.26 (2.9↲)	-2.16 (2.2↲)
Direct motor driver legs	-1.89 (-4.2)	-1.88 (-4.6)
Acc./egr. bicycle legs	-8.39 (-6.8)	-6.42 (-7.0)
Direct bicycle legs	-3.84 (3.6↲)	-3.66 (2.8↲)
Access passenger legs	-5.32 (-7.4)	-4.21 (-7.8)
Direct passenger legs	-5.48 (-0.2↲)	-5.36 (-1.3↲)
Domination size	0.17 (3.7)	0.11 (3.6)
PT nest	1	0.66 (-7.2)
Same PT part nests	1.67 (1.9)	1.88 (2.5)
Predecessor	MNL-16	NL-5
p-value (χ^2, df) compared to predecessor	0.020 (5.4, 1)	0.003 (9.0, 1)

Table 23: model estimation using nesting based on routes

In both cases, a significant improvement of the model is found compared to MNL-16 and NL-5 respectively. However, the nest coefficient for the new nests is larger than one rather than smaller.

---

while maintaining just a 5% chance that the nested model is incorrectly preferred (i.e. a type I error). The χ^2-test does not take the allowed values of new coefficients into account.

This may indicate that routes with a diversity of access and egress modes are relatively more attractive. To verify this hypothesis, an attempt has been made to reverse the nesting structure, that is, put the choice for access/egress modes above the choice for PT routes, instead of below. This results in the following choice tree:

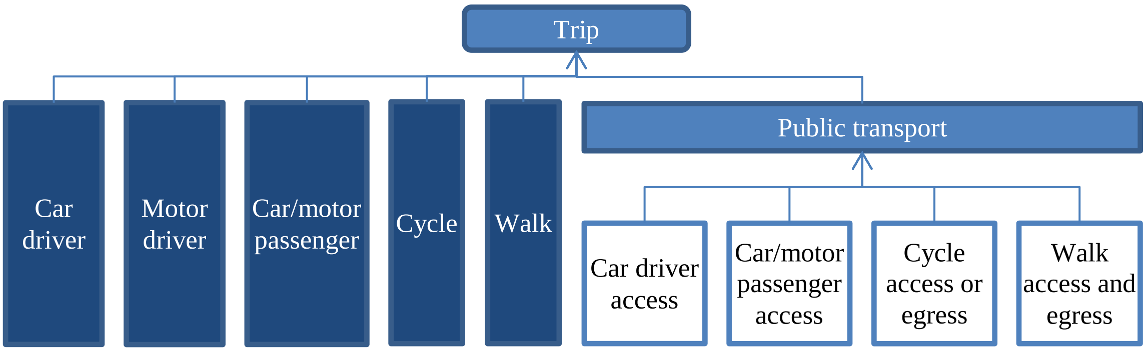


Figure 41: nesting structure based on access/egress modes; note that all alternatives can be placed in this tree unambiguously since routes with both non-walk access and non-walk egress have not been generated by the route set generator

However, estimation of this model results in convergence failure, even in the most simple case when only a single nest coefficient is estimated for all access/egress mode nests without a nest coefficient for the PT nest. This may have been caused by the fact that non-walk access or egress is quite rare in the data set (see Appendix H.1).

Nested path size logit

Finally, the nested logit models and path size logit models can also be combined into nested path size logit models. In the table below, NPSL-1 is a combination of NL-5 and PSL-1, with path size factors for the routes in the PT nests, while NPSL-2 is a combination of NL-9 and PSL-2, with, within the PT nest, nests for routes having the same PT part and path size factors for the route overlap in these PT parts.

	NPSL-1		NPSL-2	
Log-likelihood	-2837.6		-2832.3	
ρ²	0.441		0.442	
Observations	2523		2523	
Free coefficients	22		23	
Private mode time	-6.84 h^{-1}	(-15.4)	-6.84 h^{-1}	(-5.2)
PT in-vehicle time	-3.60 h^{-1}	(9.9↲)	-3.52 h^{-1}	(4.8↲)
Non-long waiting time	-6.15 h^{-1}	(1.0↲)	-6.22 h^{-1}	(0.9↲)
PT costs	-0.41 €$^{-1}$	(-4.5)	-0.38 €$^{-1}$	(-3.4)
PT usage w/o discount	-2.25	(-8.3)	-2.30	(-4.7)
Transilien legs	-0.03	(-0.2)	-0.01	(-0.0)
RER legs	-0.32	(-3.0)	-0.28	(-2.4)
Metro legs	-0.40	(-7.3)	-0.38	(-4.3)
Tram legs	-0.21	(-1.5)	-0.20	(-1.4)
RATP Paris bus legs	-1.75	(-18.5)	-1.68	(-5.2)
RATP banlieue bus legs	-1.07	(-11.8)	-1.05	(-4.9)
Optile bus legs	-2.03	(-7.4)	-2.06	(-4.5)
Access car driver legs	-2.30	(-9.8)	-2.90	(-6.8)
Direct car driver legs	-2.13	(0.7↲)	-2.13	(2.3↲)
Direct motor driver legs	-1.87	(-8.1)	-1.87	(-4.6)
Acc./egr. bicycle legs	-4.32	(-8.6)	-6.37	(-7.1)
Direct bicycle legs	-3.61	(1.3↲)	-3.62	(2.9↲)
Access passenger legs	-3.02	(-10.5)	-4.18	(-7.8)
Direct passenger legs	-5.33	(-5.5↲)	-5.33	(-1.3↲)
Domination size	0.10	(4.1)	0.10	(3.4)
Path size	-0.28	(-3.1)	0	
PT part path size	0		-0.26	(-2.9)
PT nest	0.68	(-6.9)	0.65	(-7.6)
Same PT part nests	1		1.86	(2.5)
Predecessor	NL-5		NL-9	
p-value (χ^2, df) compared to predecessor	0.002 (9.7, 1)		0.001 (11.4, 1)	

Table 24: model estimation using nesting combined with path size factors

The results are not very different form the nested logit models and path size logit models estimated earlier. Yielding again significant improvements in model fit, this suggests that the path size factor, the PT nest, and the same PT part nests each represent different aspects of the choice process. For the PT nest, this matches the expectation that PT routes are positively correlated alternatives; for the same PT part nests, this matches the expectation that routes sharing private access/egress modes are positively correlated, such that the nesting structure should be inverted. Finally, the path size coefficient may indicate ad hoc route choice behaviour as suggested previously.

8.4. Analysis of model requirements

This section will provide an analysis of the requirements of the model that were stated in Chapter 2, based on the model estimation results of the case study, to check whether the flexible model structure indeed possesses the expected advantages over traditional travel demand models.

Aggregation of modes

The following figure visualises the boarding penalties estimated in NL-4, which provided the best fit with all parameters in their expected ranges:

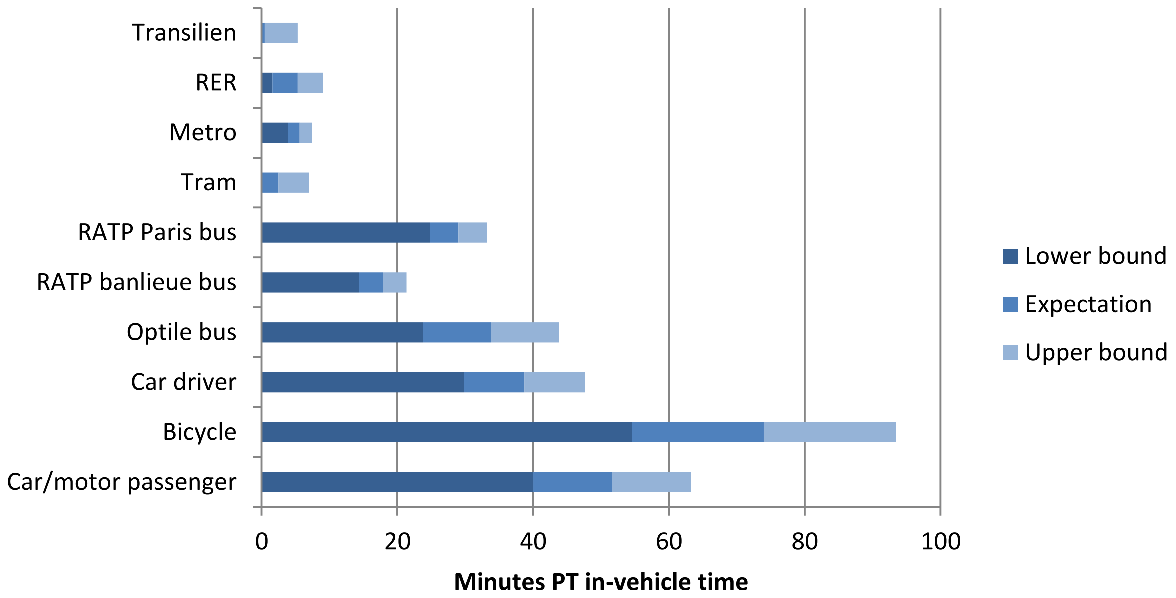


Figure 42: boarding penalties for legs in PT routes as 95% confidence intervals, scaled to minutes PT in-vehicle time

These results seem reasonable: Hoogendoorn-Lanser (2005) found transfer penalties of comparable size and also large differences between metro and non-metro transfers; the existence of a tram bonus matches with Bunschoten et al. (2012). Appendix L shows that there are differences with the stated preference data of Significance (2013), but these could be consequences of the stated preference setup. It is interesting that there is no clear relation between the attractiveness of modes, as indicated in this figure, and the position of modes in network hierarchy, as reported by Combes & Van Nes (2012): regarding attractiveness, metro and tram are like trains, but regarding network hierarchy, metro and tram are like buses.

Based on the figure, one could roughly aggregate all rail modes, but the aggregation of bus modes is problematic. The following table compares the differences between all boarding penalties using t-tests[54]:

[54] If X and Y are two boarding penalties to be compared, these t-tests check whether $X - Y \neq 0$ using $\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y - 2\rho_{X,Y}\sigma_X\sigma_Y$ where the standard deviations σ_X and σ_Y and the correlation $\rho_{X,Y}$ are taken from ALOGIT output.

	Transilien	RER	Metro	Tram	RATP Paris bus	RATP banlieue bus	Optile bus
Transilien		2.1	2.1	0.7	9.9	7.1	6.5
RER	-2.1		0.2	-1.0	11.3	7.1	6.0
Metro	-2.1	-0.2		-1.4	12.7	7.7	5.7
Tram	-0.7	1.0	1.4		9.1	6.0	5.9
RATP Paris bus	-9.9	-11.3	-12.7	-9.1		-6.4	1.0
RATP banlieue bus	-7.1	-7.1	-7.7	-6.0	6.4		3.5
Optile bus	-6.5	-6.0	-5.7	-5.9	-1.0	-3.5	

Table 26: pairwise comparison of boarding penalties of PT modes using t-tests; note that this table by definition is symmetric except for the signs; the colours indicate how modes are aggregated in ANTONIN (blue and red) and how they could be aggregated based on the estimation results (blue and green)

As visible in the table and noted earlier in Section 8.3, the significant differences among the bus modes are most problematic for the mode aggregation in ANTONIN. There also is a slightly significant difference between Transilien and RER, which could however be solved by grouping RER with metro and tram instead of with Transilien.

RATP buses

It only seems partially possible to attribute modal differences to intrinsic mode characteristics: the estimated difference between RATP Paris bus and RATP banlieue bus might be better explained considering competition with other PT modes: for example, according to this hypothesis, the union of a metro and a bus service is still more attractive than just a metro service or just a bus service, but the bus service itself gets less usage because of the metro.

In principle, a nested logit model should be used to capture such mode competition effects: bus and metro alternatives should be nested together so that this sub-choice will be more deterministic, leading to more metro usage since bus has a higher boarding penalty. Indeed, if the RATP buses are aggregated to a single mode, as in NL-10 below, the

	NL-10
Log-likelihood	-2861.3
ρ^2	0.437
Observations	2523
Free coefficients	21
Private mode time	-6.64 h^{-1} (-13.2)
PT in-vehicle time	-3.36 h^{-1} (10.4↲)
Non-long waiting time	-5.63 h^{-1} (1.6↲)
PT costs	-0.37 €$^{-1}$ (-4.1)
PT usage w/o discount	-2.26 (-8.3)
Transilien legs	-0.16 (-1.1)
RER legs	-0.36 (-3.3)
Metro legs	-0.30 (-6.0)
Tram legs	-0.23 (-1.8)
RATP bus legs	-1.28 (-15.3)
Optile bus legs	-2.10 (-7.4)
Access car driver legs	-2.24 (-9.0)
Direct car driver legs	-2.14 (0.4↲)
Direct motor driver legs	-1.88 (-7.8)
Acc./egr. bicycle legs	-4.16 (-7.7)
Direct bicycle legs	-3.62 (1.0↲)
Access passenger legs	-2.96 (-9.2)
Direct passenger legs	-5.35 (-5.7↲)
Domination size	0.09 (4.2)
PT nest	0.67 (-6.4)
Metro/tram/bus nest	0.81 (-3.2)
p-value (χ^2, df) compared to NL-6	0.000 (-40.9, -1)

Table 25: model estimation with a single boarding penalty for RATP buses

metro/tram/bus nest gets more significant than in NL-6, as shown in Table 25.

NL-10 removes one coefficient compared to NL-6, while NL-6 was significantly better according to the χ^2-test. A possible reason is that the nested logit model is not as perfect as in a network GEV model: for example, it does not correlate the alternatives if they also contain a train leg.

On the other hand, some modal image effects may still be present: this is illustrated by the fact that the Optile bus was and is estimated significantly worse than the RATP banlieue bus. It can thus be concluded that the aggregation of modes may indeed lead to errors, even more when the choice model cannot take all correlations between alternatives into account. The flexible model proposed here avoids these errors.

Permutations of modes

Assuming that ANTONIN aggregates the car and motor modes and aggregates the bicycle and walk modes, there remain two situations in which a permutation of modes is available in the flexible model proposed here, but not in ANTONIN: trips containing both a car driver access leg and one or more bus legs, and PT trips in which the access or egress mode is car/motor passenger (Syndicat des transports d'Île-de-France, n.d.). If, on the other hand, car and motor and bicycle and walk are considered separate modes – which seems more reasonable considering the estimated boarding penalties – usage of the modes motor driver and bicycle poses additional problems for ANTONIN.

The following table show the frequency of each of these situations for the whole Île-de-France region (see Appendix H.2 for more details on this data set):

Problem for ANTONIN	Count	Percentage of PT routes	Percentage of all routes
Car/motor driver combined with bus	28	1.0%	0.5%
Car/motor passenger access/egress	90	3.3%	1.6%
Subtotal	**118**	**4.3%**	**2.0%**
Motor driver/bicycle access/egress	18	0.7%	0.3%
Direct motor driver/bicycle trips	279	N/A	4.8%
Total	**415**	**4.9%**	**7.2%**

Table 27: situations that can be handled by the flexible model, but not by ANTONIN

Comparing the subtotal and the total in this table, the limitations of ANTONIN regarding mode usage lie more in the availability of modes (5.2%) than in the availability of combinations of available modes (2.0%). Regarding the permitted permutations themselves, car passenger access seems to be the largest part of the problem; this could be solved simply within the ANTONIN framework without significant additional computation time by copying the routes for car driver access and replacing car driver access with car passenger access in the copy. The proposed flexible model structure thus only provides a rather small advantage regarding this issue.

Changes in the number of modes

The introduction of new modes in the flexible model proposed here, such as the Grand Paris Express, a future rapid automatic metro network connecting the suburbs (Société du Grand Paris, 2012), or the T Zen, a recently introduced comfortable bus on dedicated bus lanes (Syndicat des transports d'Île-de-France, 2011), is one of its advantages. In the current ANTONIN model, it would be difficult to decide whether to classify the Grand Paris Express as a metro or as a train – its project organisation uses these terms interchangeably – and whether to classify the T Zen as a metro or as a bus – its quality is expected to be rated between tram and bus.

In case of the flexible model proposed by this research project, one could set up a stated preference research in which a number of PT routes can be chosen, one of them including the new mode, the others combinations of existing PT modes. The results can be used to estimate the boarding penalty for this new mode. To determine whether the Grand Paris Express belongs to the metro/tram nest or the train nest of the choice tree, or whether the T Zen belongs to the metro/tram nest or the bus nest, one can simply repeat the estimation for both cases and select the case with the best fit to the data.

Alternatively, one could use expert judgement to select a boarding penalty for the new mode. Unlike the current ANTONIN model, this allows picking a boarding penalty from the continuous scale from metro to Transilien, permitting multiple scenarios to be tested. Only the choice for a nesting structure remains a discrete decision.

Of course, with a stated preference survey, one could also check which of the existing aggregated modes fits better for the current ANTONIN model, but this does not allow estimating the attractiveness of the new mode on a continuous scale. The number of possible models is thus limited to two, risking a worse fit to the data.

Note that while the current case study only used PT mode differences in boarding penalties, this approach can be easily extended to include differences in the valuation of in-vehicle time and waiting time.

Example addition of T Zen and express bus

As an example, the stated preference data described in Appendix L are applied to estimate boarding penalties for the new T Zen and express bus[55] modes, such that the NL-4 model can be extended with these modes without changing any existing parameters of the flexible model[56]. As such an extension would typically be used for long-term mode usage forecasts, the old designs for the existing modes are excluded from the estimation.

[55] It is not clear whether express bus lines currently exist in the ANTONIN2 model coded as normal buses, but the express bus does not exist as a separate mode in the EGT, although some express bus lines already existed at that time. The estimation here is carried out as if it is a completely new mode in addition to the existing bus modes.

[56] It is thus assumed that the model has already been calibrated (like NL-4), but that these new modes need to be added afterwards. If the new modes are already a desired feature of the model before its calibration, a joint estimation involving both revealed and stated preference data could be used instead (Daly & Rohr, 1998). Also, note that the stated preference questionnaire did not explicitly distinguish RATP Paris bus, RATP banlieue bus and Optile bus. Since the origin and destination were mentioned explicitly, the bus modes have been coupled to Paris, Petite Couronne and Grande Couronne respectively and the average of the bus modes over all regions

This means that a nested logit model needs to be fitted on the stated preference data with all existing coefficients fixed at their values from NL-4, such that the estimation positions the new coefficients relative to the existing ones. Additionally, a new scale factor is placed on all utility functions to account for possible differences in the variance of the Gumbel error term (Daly & Rohr, 1998), which is estimated by introducing additional dummy nests in ALOGIT. Also, two bias dummy variables have been added to the utility function in order to correct for the way in which alternatives were presented to the respondent. The scale factor and bias coefficients should be present during estimation, but eventually only the new boarding penalties are the numbers of interest: these numbers should be directly transferable to NL-4.

In this example, it is also investigated whether the T Zen belongs to the metro/tram main mode, and thus its nest. The express bus has not been nested. This results in the following estimated extensions to NL-4, where NL-4a has T Zen outside the metro/tram nest and NL-4b inside:

	NL-4a		NL-4b	
Log-likelihood	-5921.0		-5853.0	
ρ^2	0.144		0.153	
Observations	8098		8098	
Free coefficients	5		5	
Express bus legs	-0.72	(-12.8)	-0.73	(-12.7)
T Zen legs	-0.29	(-7.9)	-0.22	(-7.0)
1st alternative bias	0		0	
2nd alternative bias	-0.25	(-7.0)	-0.24	(-7.0)
3rd alternative bias	-0.26	(-7.3)	-0.25	(-7.3)
SP/RP scale factor	0.72	(-14.8)	0.70	(-17.1)
2 Δ log-likelihood compared to NL-4a			135.9	

Table 28: stated preference extensions to NL-4 for express bus and T Zen, with T Zen outside (NL-4a) or inside (NL-4b) the metro/tram nest; the results have been scaled to the revealed preference context and all coefficients not listed here are by definition identical to NL-4; when applying the model in stated preference context instead, the utilities should first be multiplied with the SP/RP scale factor which is significantly different from one

Although the stated preference questionnaire was not designed for this purpose, and, as explained in Appendix L, the data set may not be useful on itself for the estimation of mode-specific constants, these results seem plausible. The T Zen belongs to the metro/tram main mode, is slightly less attractive than the tram (-0.15 in NL-4), but a lot more attractive than the bus modes – even more attractive RER (-0.32) and metro (-0.34). The express bus also outperforms the existing bus modes, but still performs closer to the RATP banlieue bus (-1.07) than to the metro.

These results again support the case for not using one single bus mode. Checking the robustness of these estimates could be a subject of further research.

travelled through has been assumed for the perceived boarding penalty of non-express bus alternatives in the questionnaire.

Consistency in the choice process

Section 2.4 argues that choice modelling should be consistent, and splits this consistency in two aspects: integration of the choice processes and correction for positive correlations among utilities of alternatives. Both aspects will be discussed below.

Integration of mode and route choice

As stated in Section 2.4, the joint destination and mode combination choice of ANTONIN does not take the diversity of the available routes into account: it uses an average utility of the route set instead of the logsum, hence it has disintegrated mode and route choice. Based on the model estimations in this case study, one cannot directly investigate whether this theoretically inconsistent way of modelling is problematic in practice, since in each case, the model that is estimated is a single, consistent model containing both route choice and (implicit) mode choice.

Therefore, two separate multinomial logit models have been estimated instead: one for a mode choice and one for route choice, with the constraint that the parameters in both models are equal. The main mode choice resembles the main mode nesting structure tested in Section 8.3:

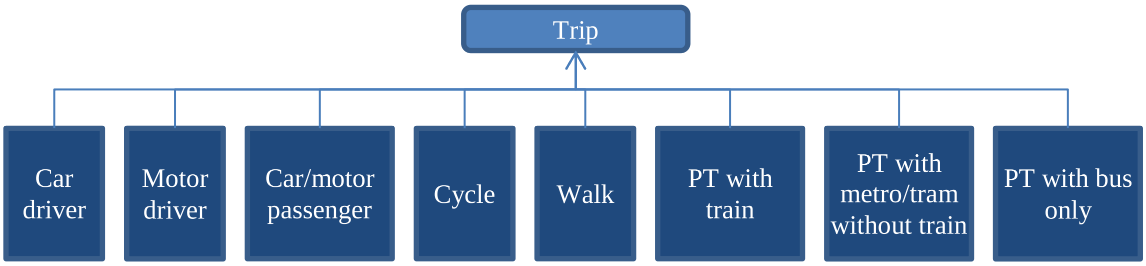


Figure 43: multinomial logit model for main mode choice

For those observations in which a PT route was chosen, the log-likelihood of choosing the correct route within the main mode is added to the log-likelihood of choosing the correct main mode: the total log-likelihood thus contains the complete choice process and can be compared to the models estimated in Section 8.3. This is achieved in ALOGIT by splitting each observation into two observations: one main mode choice observation and one route choice observation conditional to the main mode choice.

The only remaining problem now is how to specify the utility of the PT main modes. A weighted average cannot be used because unlike ANTONIN, the route choice parameters are also to be estimated. Instead, a single best route is selected to represent each main mode at the main mode choice level. Although the utility of this single best route depends on personal characteristics, such as PT discounts and vehicle ownership, the route itself should not depend on personal characteristics. Therefore, the best route per main mode is selected based on previous estimation results for a model similar to MNL-9[57], with an additional preference for walk access and egress.

[57] The only difference with MNL-9 was that the non-long waiting time was constrained to be valued equally as walk/bicycle time. There is no reason for this small difference other than the non-monotonic progress of this research project.

Due to the small number of main modes – as opposed to the number of mode combinations – it is assumed that each main mode will usually have at least one route generated by the route set generator of Section 8.1. In other cases, the main mode can be thought of as having a utility of -∞. So unlike ANTONIN, the route set generator does not specifically search for separate route sets for each mode combination, or in this case, each main mode.

The following table shows the estimation results for this model Inconsistent-1:

	Inconsistent-1	Inconsistent-2
Log-likelihood	-2915.1	-2899.9
ρ^2	0.439	0.442
Observations	3559	3559
Free coefficients	19	20
Private mode time	-8.26 h^{-1} (-26.8)	-7.73 h^{-1} (-16.3)
PT in-vehicle time	-3.86 h^{-1} (14.3↲)	-3.85 h^{-1} (12.6↲)
Non-long waiting time	-7.23 h^{-1} (1.5↲)	-7.27 h^{-1} (0.8↲)
PT costs	-0.36 €$^{-1}$ (-3.5)	-0.30 €$^{-1}$ (-3.2)
PT usage w/o discount	-2.47 (-10.7)	-2.46 (-9.2)
Transilien legs	0 (max)	0 (max)
RER legs	-0.31 (-2.8)	-0.23 (-2.5)
Metro legs	-0.42 (-6.8)	-0.40 (-7.4)
Tram legs	-0.31 (-1.8)	-0.31 (-2.2)
RATP Paris bus legs	-2.47 (-20.3)	-2.00 (-18.3)
RATP banlieue bus legs	-1.45 (-14.7)	-1.25 (-13.6)
Optile bus legs	-2.77 (-8.6)	-2.22 (-8.3)
Access car driver legs	-2.90 (-9.3)	-2.45 (-10.1)
Direct car driver legs	-2.69 (0.7↲)	-2.57 (-0.5↲)
Direct motor driver legs	-2.56 (-14.2)	-2.45 (-11.2)
Acc./egr. bicycle legs	-5.87 (-8.2)	-4.44 (-8.8)
Direct bicycle legs	-3.88 (2.7↲)	-3.79 (1.2↲)
Access passenger legs	-3.92 (-9.7)	-3.14 (-10.5)
Direct passenger legs	-5.95 (-4.7↲)	-5.81 (-6.4↲)
Main mode “nests”	1	0.69 (-6.9)
2 Δ log-likelihood compared to MNL-16	-115.1	-84.6

Table 29: model estimation with separate multinomial logit models for main mode choice and route choice

Inconsistent-1 fits considerably worse to the data than MNL-16, the best multinomial logit model. Both neglecting route diversity and not correctly selecting the best route per main mode may have contributed to this. Note that compared to MNL-16, the domination size attribute has been removed, since it represents route diversity while the purpose of this estimation was to eliminate route diversity benefits from the choice process.

Because the size of the error terms in the utilities may differ between the main mode choice model and the route choice model, model Inconsistent-2 adds a scale factor that is applied to the utility functions in the main mode choice model. Despite this being similar to nesting in a consistent model, Inconsistent-2 is still outperformed by MNL-16. It is therefore concluded that consistent modelling, specifically integration of mode and route choice, improves the fit of the model. The earlier negatively estimated path size coefficients also support this idea that the availability of multiple routes benefits the overall utility.

Integration of other choices

As mentioned in the discussion of MNL-15 and MNL-17, the PT pass ownership choice is related to the mode and route choice conditional to that PT pass ownership choice, since it means that boarding the first PT leg is basically a sunk cost. This effect distorts the model estimations in this case study and therefore supports the theoretical recommendation to integrate other choices, specifically the PT pass ownership choice, with the mode and route choice as well: this may be expected to improve the predictive quality of the model estimation results.

Corrections for correlations

The above reasoning demonstrated that the choice process should be represented by an integrated model. The other aspect of consistent choice modelling is that corrections are applied for positive correlations among the utilities of alternatives. Section 4.6 distinguishes three types of correlations that should ideally be taken into account: route overlap, modal overlap and mode similarities. These are discussed below.

- Although various path size logit models and nested path size logit models have been estimated, none of these formulations found a *route overlap* correction of the correct sign. This may indicate that ad hoc route choice plays a role, such that overlapping routes with overlap are considered more robust and therefore more attractive than when both routes would be judged individually. Further research should tell whether this is indeed the case.

- *Modal overlap* plays a significant role, as can be seen in NL-4 where PT routes with metro/tram as main mode are significantly positively correlated. NL-9 suggests that PT routes sharing access/egress modes might also be positively correlated, although this could not be investigated in detail.

- *Mode similarities* also play a significant role: NL-4 shows a positive correlation among PT routes on top of the positive correlation among metro/tram routes. Alternatively, if this metro/tram nest is removed, positive correlations can be found for relatively short-distance modes or for rail modes. However, NL-1 showed that no mode similarity effects could be found for the private modes.

It should be noted that since the nested logit models can only approximate the ideal network GEV model, it may have captured only part of the positive correlations that exist in reality.

Computation time

As indicated in Section 8.1, the computation time of the flexible model is worse than that of ANTONIN. There are two possible reasons for this: by repeatedly generating routes for different mode combinations, the desired spatial variety can be lower without sacrificing modal variety in the overall choice set, and the flexible model may have a better coverage of observed routes than ANTONIN, since the coverage of ANTONIN is currently unknown.

Further tweaking could perhaps improve the stability of computation time such that it depends less on the position of the origin zone within the network, such that the overall computation time improves. Analysis of the route sets generated by ANTONIN could provide information about its route coverage, so that a better comparison of both models becomes possible.

While noting that ANTONIN may be faster, it should also be stressed that the flexible model still has a clearly feasible computation time, equal to about 14 hours for the full ANTONIN network divided by the number of cores in the PC.

9. Conclusions and recommendations

The purpose of this research project was to find out the theoretically best way to model public transport in terms of the mode and route choice component of travel demand models, taking into account the flexibility, consistency and efficiency of the model structure, and to assess its practical performance. To this end, a theoretical framework has been constructed for flexible travel demand models that handles these issues successfully, and a practical application of this theory has been investigated by implementing such a flexible model for the Île-de-France region as a case study.

9.1. Conclusions from the theoretical framework

First of all, the interaction between mode and route choice component structures of existing travel demand models can be classified according to two dimensions: how networks are combined and whether networks are multi-modal. Several of these structures are flexible regarding how PT modes are used in a chain. However, the three sources of correlation between alternatives, namely route overlap, modal overlap and mode similarities, can only be taken into account if all routes with PT components are generated inside a single multi-modal network with no assumed prior choices about how modes are to be combined within the network.

Several different network types exist and particular attention needs to be paid to the structure of public transport networks, which can be frequency-based or timetable-based. In the frequency-based case, pre-processing the network and merging common lines increases the realism of the model. In the pure timetable-based case, one can formulate a time-dependent choice model, while in hybrid situations the timetable information can be used to improve the waiting time attributes of generated routes.

Once the network structures have been selected and it has been decided to not set prior choices as constraints for routes, the route set generation algorithm itself is needed. Regarding the public transport part of the network, the branch and bound algorithm is fit for this purpose, particularly since frequency-based networks should be pre-processed, leading to a large number of links, but fewer links per route. On the other hand, for private mode networks, the branch and bound tree depth may become excessive, such that the Monte Carlo labelling and simulation method is more useful, while for private modes for which route choice is not deemed important, the Dijkstra algorithm can simply be used. These three algorithms can be linked together by splitting the route generation at the boundary of the PT system.

The final mode and route choice model should ideally be a network GEV path size logit model. This newly proposed model combines existing aspects of existing logit models to take all three sources of correlation into account: the path size logit model for route overlap, the cross-nested logit model for modal overlap and the nested logit model for mode similarities. The model has been shown to be extendable to a fully integrated tour-based travel demand model including pivot-point procedures and congestion and crowding feedback. This model does not require aggregation of modes, permits all logical permutations of modes, allows for changes in the number of modes, is fully consistent and has a computation time independent of the number of PT modes.

9.2. Case study conclusions

In an application for mode and route choice for the Île-de-France region, the theoretical framework has been shown to be a feasible model for a dense, frequency-based PT network, estimated on revealed preference data. However, an important simplification from the network GEV path size logit model to a nested path size logit model was necessary due to limitations of existing model estimation software. Using exact matching, matching using modes and line numbers, and matching based on the domination constraint, 86% of the observed PT routes is covered by the generated route sets.

Most of the suspected problems of existing travel demand models and the corresponding suspected benefits of the implemented flexible model are demonstrated by the case study. First of all, the case study confirms the suspected problem of aggregation of modes, in which case significant differences between modes would be overlooked. Addition of new PT modes to the model by means of a stated preference survey yielded realistic estimation results, so that this method appears feasible. Integration of mode and route choice using logsums has been shown to improve the fit of the model, as well as taking correlations between alternatives into account. Regarding this last aspect, significant effects were found for modal overlap, mode similarities and route overlap, but for route overlap, the estimated coefficient had an incorrect sign: a possible reason might be that travellers appreciate route overlap so that ad hoc route switching becomes possible and the overall experienced travel time therefore is more robust. The large proportion of travellers choosing dominated routes could be explained by timetable effects and network inaccuracies, but may also hint at ad hoc route choice behaviour.

On the other hand, the suspected advantage of permitting all permutations of modes appears to be rather small, when compared to a model based on mode combinations. Furthermore, the computation time of the route set generation process appears to be worse than a classical model based on mode combinations, possibly because a classical model allows generating the same modal variety with less spatial variety, but a fair comparison is difficult to make since the route coverage for this classical model is unknown. However, reducing the aggregation of modes as suggested above may realise these last two suspected advantages of the flexible model as well.

9.3. Recommendations for further research

As the case study was limited to generation of the choice sets and estimation of the choice model, assessment of the results of subsequent network loading using such a choice model is an important recommendation for further research. While this research project focused on the flexibility, consistency and efficiency of the model, they are eventually just a means to estimate usage of the transport system. For example, it should thus be verified that the model indeed yields correct estimations for the usage of specific PT lines, such that it can be confidently applied for capacity planning.

As listed in the conclusions, there were some indications that ad hoc route choice behaviour plays a role in decision-making. It is therefore recommended that this phenomenon is investigated further by developing and testing utility terms that, unlike the path size factor, explicitly represent the ability to change a route half-way[58]. Related to this, it would be desirable to investigate how the definition of the dominance constraint can be modified such that fewer of the observed routes are considered dominated by the model, but without increasing the choice set sizes by large numbers of routes.

Since the branch and bound algorithm showed considerable variation in both choice set size and computation time among origin zones, it is recommended to check whether the tolerance constraints can be modified to reduce the excesses in computation time. One could for example think of replacing the linear part of the tolerance constraint with a concave function.

As the observations from the survey were not weighted during model estimations, the estimated models may be biased against PT alternatives and against routes with many legs. Socio-economic representativeness of the survey respondents has also not been investigated. It can therefore be recommended to repeat the model estimations with weights to correct for these effects and see what impact this has on the estimated coefficients.

As indicated in the theoretical specification of the ideal model structure, the mode and route choice process should be integrated with the other choice processes, for example destination choice and vehicle ownership choice. Since this has not been done in the case study, it remains a recommendation for further research. Specifically for home-work travel, the integration of mode and route choice on the one hand and PT pass ownership choice on the other hand seems important, as neglecting the relation between these choice processes may distort the estimation of the mode and route choice model as was seen in the case study. Also, testing of a tour-based version of the case study model may be recommended, such that the utilities of both the outbound route and the return route affect mode choice.

Due to software limitations, the case study simplified a network GEV model to a more common nested logit model. However, the network GEV model can theoretically capture more correlations among alternatives, so it can be recommended to try to fit such a model as well and see if leads to more significant nest coefficients. It might for example allow removal of the somewhat strange distinction between RATP buses in Paris and in suburbs. Such further research would also be fundamentally interesting as the current number of applications of the network GEV model is very limited. This recommendation does however require that new choice model estimation software is developed or existing software is improved. The improvement of model estimation software also is a recommendation on itself regarding handling of large numbers of alternatives.

It can also be further investigated how the model behaves in cases with more non-walk access to and egress from the PT system. Some of the nesting structures in the case study could not be tested, presumably because such choices are rare in Paris and Petite Couronne.

Since the case study contained frequency-based PT networks only, a practical application with timetable-based PT networks can be recommended to test this part of the theoretical framework as well. In particular, the behaviour of the proposed time-dependent choice model

[58] A joint network GEV model might be better for combining route overlap and ad hoc choice behaviour, but it would undermine the inclusion of modal overlap and mode similarities, which have been shown to play significant roles, and has other issues in multi-modal context.

could be interesting to investigate. Splitting the population into multiple user classes may also improve the model.

Regarding the addition of new modes to the flexible model, it can be recommended to investigate the robustness of the estimation of the new boarding penalties with respect to the design of the stated preference experiment, and to investigate if and how the revealed preference and stated preference data can be merged for a joint choice model estimation.

Finally, for existing travel demand models which were not designed for maximum flexibility and consistency, two relatively simple enhancements can be recommended based on this research that do not require overturning the existing model structure. Firstly, one should take modal differences between into account in the deterministic part of the utility function, even when the model structure does not allow stochastic modal preferences to be included. Secondly, when a route set is generated, a logsum of this route set should be used in the mode choice rather than the minimum or average generalised cost.

References

Abbé, E., Bierlaire, M. & Toledo, T. (2007). Normalization and correlation of cross-nested logit models. *Transportation Research Part B*, 41, pp.795-808.

Algers, S., Beser, M. & Sveder, G. (2000). SAMPERS - Erfarenheter av komplexa modeller i ett komplext utvecklingsprojekt. *Trafikdage på Aalborg Universitet*. Available at: http://www.trafikdage.dk/td/papers/papers00/Dag1/paper/1102.pdf [accessed 17 December 2012].

ALOGIT Software & Analysis (2007). *ALOGIT 4.2 User Manual*. January draft.

Arentze, T. & Timmermans, H. (2004). A learning-based transportation oriented simulation system. *Transportation Research Part B*, 38, pp.613-33.

Baak, J. (2013). Het Deense Logistieke Model. Presentation at Significance, 10 January.

Bekhor, S., Ben-Akiva, M. & Ramming, M. (2006). Evaluation of choice set generation algorithms for route choice models. *Annals of Operations Research*, 144, pp.235-47. Available at: http://web.mit.edu/~rama/Public/papers/paper7.pdf [accessed 21 January 2013].

Bekhor, S. & Prato, C. (2009). Methodological transferability in route choice modeling. *Transportation Research Part B*, 43, pp.422-37.

Bel, N. (2013). *Crowding in train passenger assignment: A study on the implementation of the influence of crowding on train passenger choice behaviour in assignment models*. Significance, TU Delft. Available at: http://repository.tudelft.nl/view/ir/uuid:0d8de327-d29c-48af-b593-1ab215333e94/ [accessed 29 May 2013].

Benjamins, M. (2001). *Het toedelen van verplaatsingen op multimodale netwerken met behulp van hypernetwerken*. TU Delft. Available at: http://repository.tudelft.nl/view/ir/uuid:924894c2-5c40-4791-9015-ae7628eeafc0/ [accessed 22 November 2012].

Bierlaire, M. (2006). A theoretical analysis of the cross-nested logit model. *Annals of Operations Research*, 144, pp.287-300.

Bierlaire, M. (n.d.). Bison Biogeme model specification file. *Biogeme users guide*. Available at: http://biogeme.epfl.ch/doc/syntaxmod.html [accessed 21 May 2013].

Bliemer, M., Versteegt, H. & Bovy, P. (2004). INDY: A new analytical multiclass dynamic traffic assignment model. *Transportation Research Board Annual Meeting*, 83.

Bovy, P., Bliemer, M. & Van Nes, R. (2006). *Course CT4801: Transportation Modeling*. TU Delft.

Brederode, L., Bliemer, M. & Wismans, L. (2010). STAQ: Static Traffic Assignment with Queuing. *European Transport Conference*. Available at: http://www.etcproceedings.org/paper/static-traffic-assignment-with-queuing-staq-static-assignment-revisited [accessed 28 January 2013].

Bundesamt für Raumentwicklung ARE (2006). *Erstellung des nationalen Personenverkehrsmodells für den öffentlichen und privaten Verkehr - Modellbeschreibung*.

Bunschoten, T., Molin, E. & Van Nes, R. (2012). Tram of Bus: Bestaat de Trambonus? *Colloquium Vervoersplanologisch Speurwerk*, 22-23 November. Available at: http://www.goudappel.nl/media/files/uploads/Paper__Tram_of_Bus__Bestaat_de_Trambonus_.pdf [accessed 11 January 2013].

Carlier, K., Fiorenzo-Catalano, S., Lindveld, C. & Bovy, P. (2003). A supernetwork approach towards multimodal travel modeling. *Transportation Research Board Annual Meeting*.

Cascetta, E. & Coppola, P. (2012). An elastic demand schedule-based multimodal assignment model for the simulation of high speed rail (HSR) systems. *EURO Journal on Transportation and Logistics*, 1(1-2), pp.3-27.

Cascetta, E. & Papola, A. (2009). Dominance among alternatives in random utility models. *Transportation Research Part A*, 43, pp.170-79.

Citilabs (2008). *Cube Voyager Reference Guide, Version 5.0*.

Combes, F. & Van Nes, R. (2012). A simple representation of a complex urban transport system based on the analysis of transport demand: the case of Region Ile-de-France. *Procedia - Social and Behavioral Sciences*, 48, pp.3030-39.

Cominetti, R. & Correa, J. (2001). Common-Lines and Passenger Assignment in Congested Transit Networks. *Transportation Science*, 35(3), pp.250-67. Available at: http://www.dii.uchile.cl/~jcorrea/papers/Journals/CC2001.pdf [accessed 5 February 2013].

Courriers de l'Ile de France (n.d.). *Carte Optile Scolaire*. Available at: http://www.cif-bus.com/exec/titre.asp?NumRub=1086&Numfiche=1293&NumMode=1 [accessed 12 April 2013].

Daly, A. (2012). Integrated models of choice and assignment. Presentation at RAND Europe, 8 October.

Daly, A. & Bierlaire, M. (2006). A general and operational representation of Generalised Extreme Value models. *Transportation Research Part B*, 40, pp.285-305.

Daly, A., Fox, J. & Tuinenga, J. (2005). Pivot-Point Procedures in Practical Travel Demand Forecasting. *ERSA Conference*. Available at: http://ideas.repec.org/p/wiw/wiwrsa/ersa05p784.html [accessed 23 January 2013].

Daly, A. & Rohr, C. (1998). Forecasting Demand for New Travel Alternatives. In Garling, T., Laitila, T. & Westin, K. *Theoretical Foundations of Travel Choice Modeling*. Elsevier. pp.451-71.

De Jong, G. (2011). The Dutch National Model System (LMS): methodology. Presentation at TU Delft, 3 October.

De Jong, G., Ben-Akiva, M. & Baak, J. (2008). *Method Report - Logistics Model in the Swedish National Freight Model System (Version 2)*. Significance.

Dijkstra, E. (1959). A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, 1, pp.269-71. Available at: http://oai.cwi.nl/oai/asset/9256/9256A.pdf [accessed 20 June 2013].

Fiorenzo-Catalano, M. (2007). *Choice Set Generation in Multi-Modal Transportation Networks*. TU Delft, TRAIL Research School. Available at: http://repository.tudelft.nl/view/ir/uuid:ef3b9c22-b979-4f46-9b02-110c82d67535/ [accessed 10 December 2012].

Fiorenzo-Catalano, S., Van Nes, R. & Bovy, P. (2004). Choice set generation for multi-modal travel analysis. *Transportation Research Board Annual Meeting*, 83.

Frejinger, E., Bierlaire, M. & Ben-Akiva, M. (2009). Sampling of alternatives for route choice modeling. *Transportation Research Part B*, 43, pp.984-94.

Friedrich, M., Hofsäß, I. & Wekeck, S. (2001). Timetable-Based Transit Assignment Using Branch and Bound Techniques. *Transportation Research Record*, 1752, pp.100-07.

Goudappel Coffeng (2013). *De waaier van Brogt*. Available at: http://issuu.com/gcoffeng/docs/waaier-van-brogt [accessed 17 January 2013].

Hague Consulting Group (1996). *Promise Model System: Technical Documentation*. 10 October draft.

Hart, P., Nilsson, N. & Raphael, B. (1968). A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2), pp.100-07. Available at: http://www.cs.auckland.ac.nz/compsci767s2c/projectReportExamples.d/astarNilsson.pdf [accessed 20 June 2013].

Herriges, J. & Kling, C. (1996). Testing the consistency of nested logit models with utility maximization. *Economics Letters*, 50, pp.33-39.

Hoogendoorn-Lanser, S. (2005). *Modelling Travel Behaviour in Multi-modal Networks*. TU Delft, TRAIL Research School. Available at: http://repository.tudelft.nl/view/ir/uuid:3e013e1d-5bd6-40d6-b1d5-87fc53f3b3d9/ [accessed 21 January 2013].

Hoogendoorn-Lanser, S., Van Nes, R. & Bovy, P. (2005). Path Size and overlap in multi-modal transport networks, a new interpretation. *International Symposium of Transportation and Traffic Theory*.

Hoogendoorn-Lanser, S., Van Nes, R., Hoogendoorn, S. & Bovy, P. (2006). Home-activity approach to multi-modal travel choice modelling. *Transportation Research Board Annual Meeting*, 85, 22 March.

Johansson, B. (2009). *Transport Model for Scotland, TMfS07 National Model Audit*. SIAS. Available at: http://www.transportscotland.gov.uk/files/documents/analysis/LATIS/TMfS07_NationalModelDevelopmentAudit_29102009.pdf [accessed 21 December 2012].

Jovicic, G. (2001). *Activity-based travel demand modelling - a literature study*. Danmarks TransportForskning. Available at: http://wwwdtu2.sitecore.dtu.dk/upload/institutter/dtu%20transport/pdf_dtf/notater/not0801.pdf [accessed 17 December 2012].

Koppelman, F. & Wen, C. (1998). Alternative Nested Logit Models: Structure, Properties and Estimation. *Transportation Research Part B*, 32(5), pp.289-98.

Koppelman, F. & Wen, C. (2000). The paired combinatorial logit model: properties, estimation and application. *Transportation Research Part B*, 34, pp.75-89.

Kroes, E. & Tuinenga, J. (2012). *Model structure of ANTONIN*. Significance.

Kurauchi, F., Bell, M. & Schmöcker, J. (2003). Capacity Constrained Transit Assignment with Common Lines. *Journal of Mathematical Modelling and Algorithms*, 2, pp.309-27. Available at: http://link.springer.com/article/10.1023%2FB%3AJMMA.0000020426.22501.c1 [accessed 5 February 2013].

l'Observatoire de la mobilité en Île-de-France (2012). *Enquête globale transport, La mobilité en Île-de-France, №1*. Available at: http://www.iau-idf.fr/fileadmin/Etudes/etude_954/Enquete_globale_transport_HD.pdf [accessed 13 June 2013].

Lam, W., Zhou, J. & Sheng, Z. (2002). A capacity restraint transit assignment with elastic line frequency. *Transportation Research Part B*, 36, pp.919-38.

Lo, H., Yip, C. & Wan, K. (2003). Modeling transfer and non-linear fare structure in multi-modal network. *Transportation Research Part B*, 37, pp.149-70.

Newman, J. (2008). Normalization of network generalized extreme value models. *Transportation Research Part B*, 42, pp.958-69.

Papola, A. & Marzano, M. (2013). A Network Generalized Extreme Value Model for Route Choice Allowing Implicit Route Enumeration. *Computer-Aided Civil and Infrastructure Engineering*, pp.1-21.

Prato, C. & Bekhor, S. (2006). Applying Branch-and-Bound Technique to Route Choice Set Generation. *Transportation Research Record*, 1985, pp.19-28. Available at: http://cee.technion.ac.il/eng/getfile.asp?LNGID=1&DBID=1&GID=169 [accessed 6 February 2013].

Ramming, M. (2002). *Network Knowledge and Route Choice*. Massachusetts Institute of Technology. Available at: http://pdf.aminer.org/000/304/418/proposal_of_improving_model_for_default_probability_prediction_with_logit.pdf [accessed 5 December 2012].

Robinson, C. & Pollard, M. (2009). *Land-Use and Transport Integration in Scotland (LATIS) - TMfS:07 Demand Model Development Report*. MVA Consultancy.

Samadzad, M. (2012). *Econometrics of Urban Transport Mode Choice for the Residents of the Greater Paris Region*. Université Paris-Est.

Sheffi, Y. (1985). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Massachusetts Institute of Technology, Prentice-Hall.

Significance (2012a). *Documentatie van GM 2011*.

Significance (2012b). Transportation model for STIF. Presentation at Significance, 10 May.

Significance (2013). *Etude sur la perception du confort dans les transports collectifs en Ile-de-France: Phase quantitative - preferences declares*. 4 April version.

Sobel, K. (1980). Travel Demand Forecasting by Using the Nested Multinomial Logit Model. *Transportation Research Record*, 775, pp.48-55. Available at: http://ntl.bts.gov/data/letter_nz/tdma.pdf#page=52 [accessed 21 November 2012].

Société du Grand Paris (2012). *Société du Grand Paris*. Available at: http://www.societedugrandparis.fr/wp-content/uploads/2012/08/booklet-anglais.pdf [accessed 27 May 2013].

Syndicat des transports d'Île-de-France (2006). *Fare System in Ile de France*.

Syndicat des transports d'Île-de-France (2010). *Guide tarifaire, Le guide Île-de-France, À partir du 1er juillet 2010*.

Syndicat des transports d'Île-de-France (2011). *T Zen: Un nouveau mode de transport en Île-de-France - Dossier de presse*. Available at: http://www.stif.org/IMG/pdf/DP_T_Zen_basse_def.pdf [accessed 28 May 2013].

Syndicat des transports d'Île-de-France (n.d.). *Manuel Antonin 2*.

TRAFICO et al. (2009). Beschreibung des Verkehrsmodells (Personenverkehr und Güterverkehr). In *Verkehrsprognose Österreich 2025+, Endbericht*. Ch. 3.

Train, K. (2002). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Tuinenga, J., Pieters, M. & Debrincat, L. (2006). *ANTONIN: updating and comparing a transport model for the Paris region*. RAND Europe, Syndicat des transports d'Île-de-France.

UK Department for Transport (2009). *National Transport Model High Level Overview*.

Van der Zijpp, N. & Fiorenzo-Catalano, S. (2005). Path enumeration by finding the constrained K-shortest paths. *Transportation Research Part B*, 39, pp.545-63.

Van Eck, G. (2011). Dynamische toedeling en toetsing van grootschalige multimodale vervoerssystemen. *Colloquium Vervoersplanologisch Speurwerk*, 24-25 November. Available at: http://www.cvs-congres.nl/cvspdfdocs/cvs11_049.pdf [accessed 18 January 2013].

Van Nes, R. (2011). Choice set generation. Presentation at TU Delft, 23 November.

Van Nes, R. & Bovy, P. (2008). *Course CT5802: Advanced transportation modelling and network design*. TU Delft.

Veitch, T. & Cook, J. (2010). *OtTransit: Uses and Functions*. Omnitrans International. 10 March version.

Weber, W. (1966). *Die Reisezeit der Fahrgäste öffentlicher Verkehrsmittel in Abhängigkeit von Bahnart und Raumlage*. Technische Hochschule Stuttgart.

Wen, C. & Koppelman, F. (2001). The generalized nested logit model. *Transportation Research Part B*, 35, pp.627-41.

Werkmaatschappij Amsterdam-Almere (2012). *Het IJmeeralternatief - Eindrapport Werkmaatschappij Amsterdam-Almere*. Rijk-regioprogramma Amsterdam-Almere-Markermeer. Available at: http://www.rraam.nl/Publicaties/Documenten+en+Publicaties+RRAAM/downloads_getfilem.aspx?id=205375 [accessed 18 June 2013].

Wikimedia (2012a). Block matrix. *Wikipedia, The Free Encyclopedia*. Available at: http://en.wikipedia.org/w/index.php?title=Block_matrix&oldid=523595900 [accessed 24 January 2013].

Wikimedia (2012b). Heaviside step function. *Wikipedia, The Free Encyclopedia*. Available at: http://en.wikipedia.org/w/index.php?title=Heaviside_step_function&oldid=530662224 [accessed 3 January 2013].

Wikimedia (2012c). Rank (linear algebra). *Wikipedia, The Free Encyclopedia*. Available at: http://en.wikipedia.org/w/index.php?title=Rank_%28linear_algebra%29&oldid=530415861 [accessed 23 January 2013].

Wikimedia (2013a). A* search algorithm. *Wikipedia, The Free Encyclopedia*. Available at: http://en.wikipedia.org/w/index.php?title=A*_search_algorithm&oldid=557032807 [accessed 30 May 2013].

Wikimedia (2013b). Linear least squares (mathematics). *Wikipedia, The Free Encyclopedia*. Available at: http://en.wikipedia.org/w/index.php?title=Linear_least_squares_%28mathematics%29&oldid=533669445 [accessed 23 January 2013].

Wikimedia (2013c). Paris aire urbaine. *Wikipedia, The Free Encyclopedia*. Available at: http://en.wikipedia.org/w/index.php?title=Paris_aire_urbaine&oldid=549141186 [accessed 17 April 2013].

Willigers, J. (2007). *Memo NHB trips v6*. Significance.

Willigers, J. & Tuinenga, J. (2007). *Re-estimation of mode-destination choice models STIF Antonin 2: technical memorandum*. Significance.

Yao, E. & Morikawa, T. (2003). A Study on Integrated Intercity Travel Demand Model. *International Conference on Travel Behaviour Research*, 10, 10-15 August. Available at: http://wiki.umd.edu/transportation/images/2/27/A_Study_of_an_Integrated_Intercity_Travel_Demand_Model.pdf [accessed 21 January 2013].

Zhang, L. & Xiong, C. (2011). Multi-modal Inter-regional Travel Demand Estimation. *Transportation Research Board Annual Meeting*, 90.

Appendix A. Overview of GroeiModel and ANTONIN main choice models

For the problem definition presented in Chapter 2, it may be useful to understand the inner workings of the GroeiModel and ANTONIN, such that one can understand what the exact cause of a particular problem is. This appendix therefore sheds some light on that. A subset of the information presented here is also included in Chapter 4.

The GroeiModel (Significance, 2012a) is used as the Dutch national and regional transport models, which only differ in input data, not in model structure. ANTONIN (Syndicat des transports d'Île-de-France, n.d.; Willigers & Tuinenga, 2007) is the French transport model for the Île-de-France region containing Paris.

Both models are tour-based (i.e. simultaneously model outward and return trips, see Section 7.4) and use a nested logit construction (see Appendix C.3) for the main part of the choice model, which is displayed below[59]:

[59] The stop/repeat models for tour generation are not displayed here (see Section 7.3). Also, for the GroeiModel, the nests for car driver, car passenger, cycle and walk are not displayed.

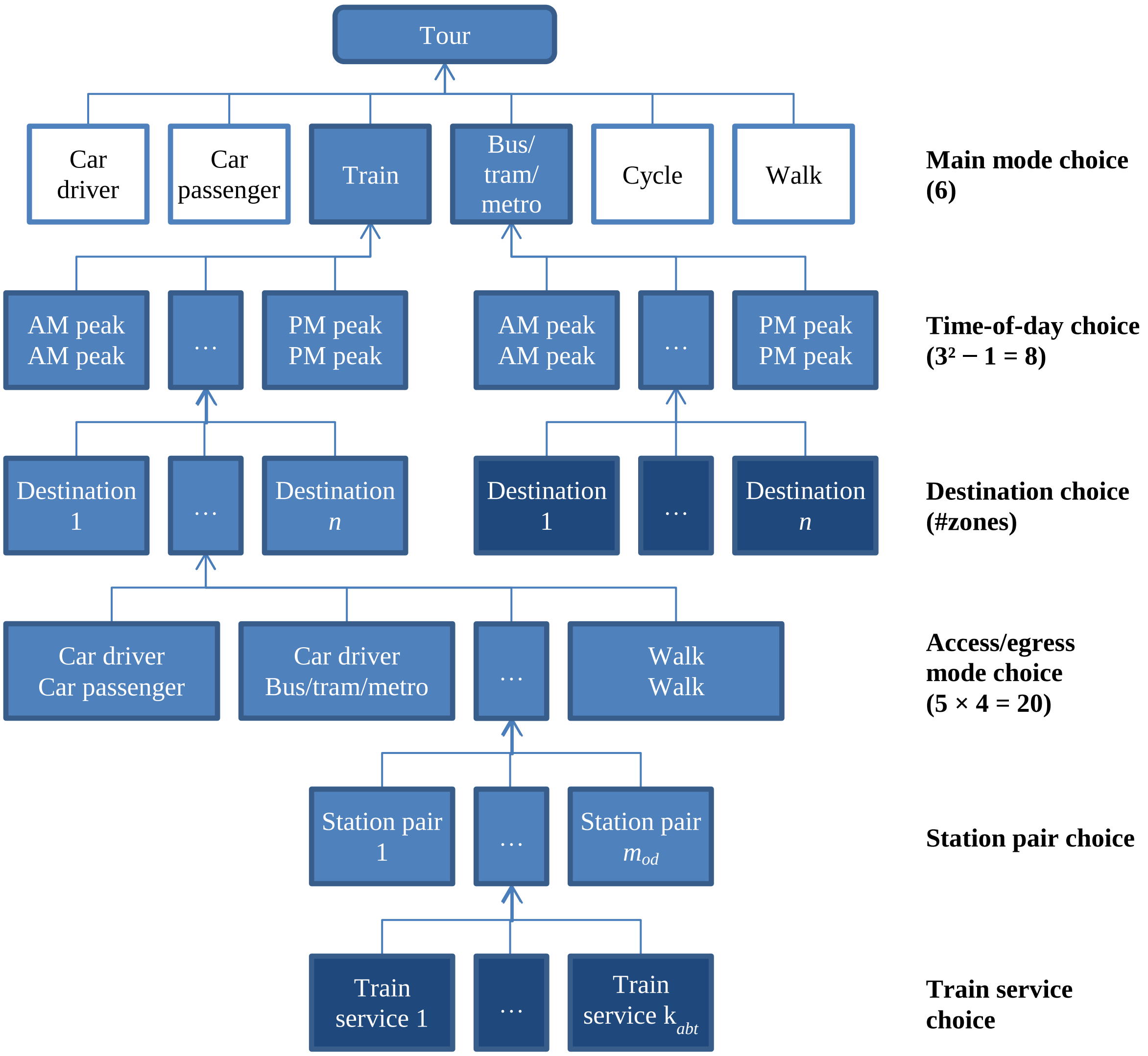


Figure 44: GroeiModel 2011 choice model structure

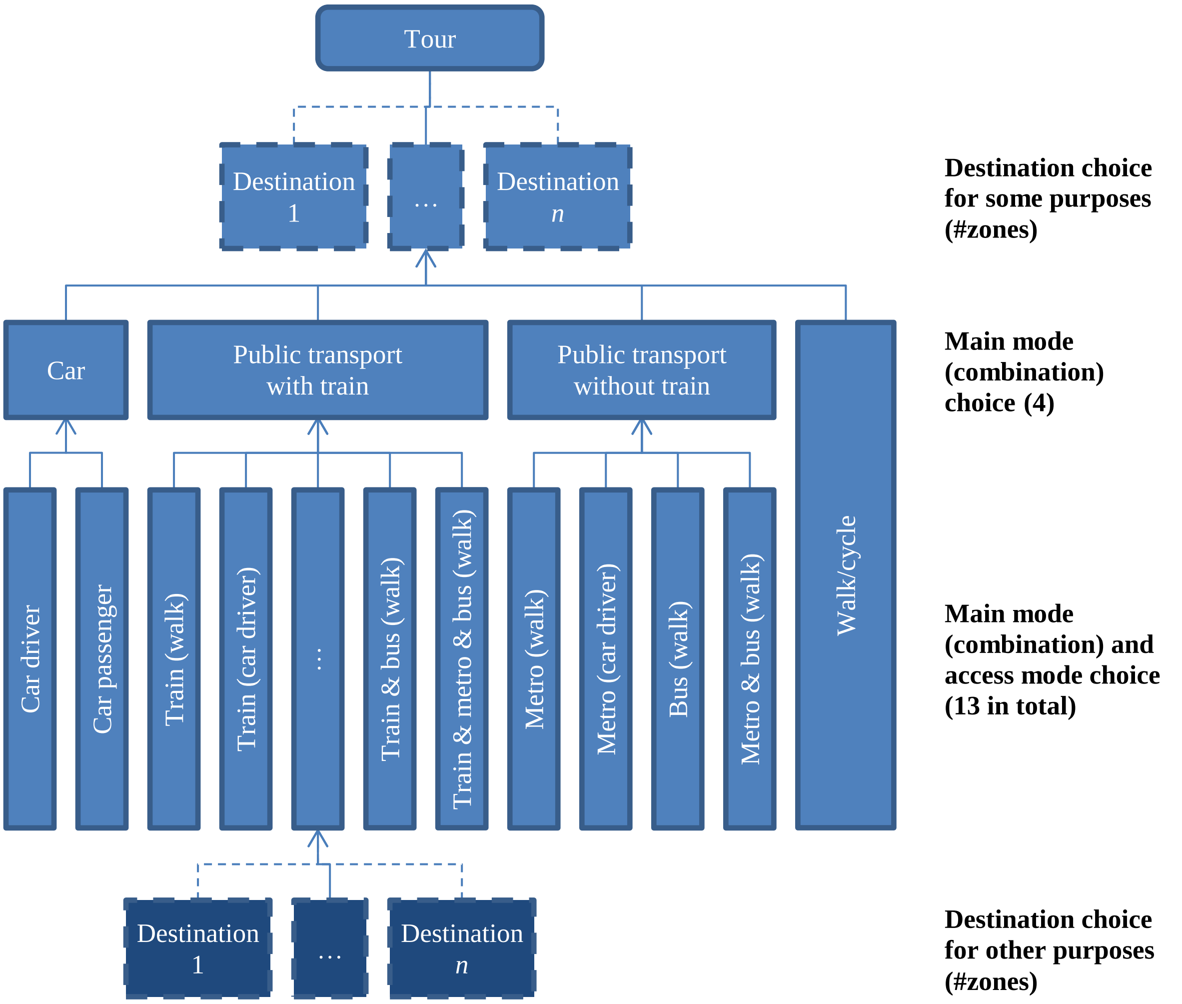


Figure 45: ANTONIN 2 choice model structure; the positioning of destination choice depends on tour purpose

Contrary to the GroeiModel, ANTONIN has no time-of-day choice (ANTONIN uses fixed proportions per purpose) and the logsum of the choice structure as a whole is not propagated to the tour generation model. However, there are more interesting observations relating to public transport and thus to this research project:

- Contrary to the GroeiModel, the nesting structure of ANTONIN is not the same for all travel purposes. The destination choice may be positioned above or below the mode choice. Additionally, for some purposes the mode choice nest or even all nests have been removed, which is equivalent to setting the nest coefficients to one (Sobel, 1980). The structure that is chosen depends on the optimal fit for the survey data for each purpose.

- In ANTONIN, to keep the model computation time down, car driver as access mode is not available for all public transport mode combinations. In the GroeiModel, access modes other than walking are only available for train tours.

- In ANTONIN, car passenger is not available as access or egress mode for public transport, contrary to the GroeiModel. In both the GroeiModel and ANTONIN, due to car availability considerations, car driver is not available as egress mode, but as an access mode only. So for ANTONIN, the egress mode is always walk (hence it is not included in the choice tree).

- While the GroeiModel includes assignment of train traffic to specific train services within the choice model[60], it has no assignment at all for bus/tram/metro traffic. ANTONIN uses an assignment phase separate from the choice model, but with the chosen combination of modes as a constraint. It generates a number of routes and uses a multinomial logit model for route choice.

Other notable differences between both models include:

- Although both models contain driving licence and car ownership models, ANTONIN, contrary to the GroeiModel, also contains a PT pass ownership model.

- ANTONIN contains a pivot-point procedure for the modes car driver, car passenger, PT and walk/cycle. The GroeiModel contains such a procedure for the modes car driver and train.

- The GroeiModel, contrary to ANTONIN, has a feedback loop from the car assignment to the choice model, such that the road congestion resulting from the model influences the choices of travellers in the next iteration of the model, until equilibrium is reached.

[60] In this case, the train service represents the route part in the train network (see Appendix D for definitions of services and routes). Note that although the train service choice occurs in GroeiModel component SES, the train service set generation occurs in an external program TRANS. The train service set is stored in the TPI files.

Appendix B. List of consulted existing travel demand models

This appendix lists the consulted existing models that mainly serve as examples in Chapter 4, as well as literature sources supporting the statements about the model structure in that chapter.

The following official national/regional models have been consulted:

- the Dutch *GroeiModel* (Significance, 2012a);
- the French *ANTONIN* (Syndicat des transports d'Île-de-France, n.d.; Willigers & Tuinenga, 2007; Citilabs, 2008) for Île-de-France;
- the Danish *PETRA* (Zhang & Xiong, 2011; Jovicic, 2001);
- the British *National Transport Model (NTM)* (UK Department for Transport, 2009);
- the Scottish *Transport Model for Scotland (TMfS)* (Johansson, 2009; Robinson & Pollard, 2009);
- the Swiss *Nationales Personenverkehrsmodell (NVPM)* (Bundesamt für Raumentwicklung ARE, 2006);
- the Swedish *SAMPERS* (Algers et al., 2000), distinguishing the model components for domestic traffic and international traffic;
- the Austrian *Verkehrsmodell Österreich (VMÖ)* (TRAFICO et al., 2009).

Additionally, the following model structures have been consulted:

- *OmniTRANS* (Veitch & Cook, 2010), which can be used to build many different models;
- *Albatross*[61] (Arentze & Timmermans, 2004);
- Benjamins (2001);
- Fiorenzo-Catalano (2007);
- Integrated intercity travel demand model (Yao & Morikawa, 2003) for Japan;
- High-speed rail (HSR) model (Cascetta & Coppola, 2012) for Italy;
- Samadzad (2012) for Île-de-France (France).

Finally, the following models have been consulted that do not intend to cover the full mode choice:

- Hoogendoorn-Lanser (2005), distinguishing the cross-nested logit model, the multinomial logit model and the path size logit model, all of which require train as the main mode;
- *INDY* (Bliemer et al., 2004), which, on its own, requires car-driver as the only mode.

[61] Albatross does not include an assignment component. However, a level-of-service matrix (see Section 5.1) is used. This means Albatross uses an implicit all-or-nothing assignment where all travellers are assigned to the routes that were used to construct the level-of-service matrix. An implicit interface between mode and route choice thus does exist, such that this model can be included in the comparison of Chapter 4.

Appendix C. Discrete choice models

Once a choice set has been established, travellers still have to decide which multi-modal route to take. For this, a choice model is needed. The choice model should use a set of alternatives and measurable attributes of these alternatives to determine the probability that each of the alternatives is selected.

Although most models presented here focus on route choice, some of them are also applicable to other choice situations, for example mode choice.

C.1. Multinomial probit

The multinomial probit model is based directly on stochastic link impedance z_m for each link $m \in M$. The expected value of link impedance z_m^0 is determined by attributes $\mathbf{X_m}$ and a standard-normal error ε_m. These values directly determine[62] the normal-distributed correlated utilities U_i of routes $i \in I$ (Bovy et al., 2006):

$$
\begin{aligned}
z_m^0 &= \boldsymbol{\beta}^{\mathrm{T}} \mathbf{X_m} \quad \forall m \in M \\
z_m &= z_m^0 + \varepsilon_m \sqrt{\theta \hat{z}_m^0} \quad \forall m \in M \\
U_i &= -\sum_{m \in R_i} z_m \quad \forall i \in I
\end{aligned}
$$

Practical considerations

The multinomial probit model requires Monte Carlo simulation to estimate the choice probabilities of routes. Algebraic approximations do exist, but are not accurate enough (Benjamins, 2001). On the other hand, a priori choice set generation is not necessary[63], which means that no route can be incorrectly excluded from the choice set. The Monte Carlo simulation makes the model considerably slower than the logit-family of choice models, which is described below.

C.2. Multinomial logit

The multinomial logit model is the best-known choice model (Van Nes & Bovy, 2008). It assumes that alternatives $i \in I$ have measurable attributes $\mathbf{X_i}$ and that its stochastic utilities U_i can be defined as deterministic linear combinations V_i of those attributes plus independent and identically Gumbel-distributed error terms ε_i:

$$
\begin{aligned}
V_i &= \boldsymbol{\beta}_\mathbf{i}^{\mathrm{T}} \mathbf{X_i} \quad \forall i \in I \\
U_i &= V_i + \varepsilon_i \quad \forall i \in I
\end{aligned}
$$

The following formula provides the probability p_i that the alternative is chosen:

---

[62] Note that $\hat{z}_m^0$ does not have to be equal to z_m^0, but that $\hat{z}_m^0$ is not allowed to depend on congestion in congested assignments.

[63] The Dijkstra algorithm may be repeatedly executed with sampled link impedances z_m.

$$p_i = \frac{e^{V_i}}{\sum_{j \in I} e^{V_j}} \quad \forall i \in I$$

The parameters $\boldsymbol{\beta}_i$ can be estimated based on maximum likelihood, i.e. using a dataset of known choices and attributes of alternatives, the probability that these choices are made is maximized. It is also possible to test whether these parameters are significantly different from zero using t-tests.

Independence from irrelevant alternatives

The multinomial logit model has the independence from irrelevant alternatives (IIA) property. This means that the relative odds that an individual will select one alternative from a pair of alternatives is independent of any other alternatives (Sobel, 1980), as is proven below:

$$p_{i|\{i,j\}} = \frac{p_i}{p_{\{i,j\}}} = \frac{\dfrac{e^{V_i}}{\sum_{k \in I} e^{V_k}}}{\dfrac{e^{V_i}}{\sum_{k \in I} e^{V_k}} + \dfrac{e^{V_j}}{\sum_{k \in I} e^{V_k}}} = \frac{e^{V_i}}{e^{V_i} + e^{V_j}} \quad \forall i, j \in I$$

This is not always realistic, especially not in route choices. For example, imagine a situation with three routes, with two routes partially overlapping, as indicated below:

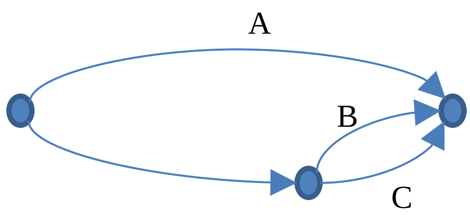


Figure 46: overlapping routes

In this situation, the utilities of route B and C are positively correlated. This implies that the relative odds of choosing route A instead of route B must also depend on the utility of route C.

Various modifications to the multinomial logit model have been proposed to deal with choices from correlated alternatives. These will be discussed below.

C.3. Nested logit

The nested logit model[64] is a very common solution to the problem of correlated alternatives. The basic idea is to group several alternatives into nest. The utility of such a nest is based on the utilities of the members of the nest. This creates multiple interacting multinomial logit models at different levels. Each of these logit models can contain two types of alternatives in

[64] To be precise, the utility maximising nested logit model.

its choice set: a direct alternative or a nest that represents a group of alternatives from a lower-level logit model. This formulation allows multiple levels of nesting.

For direct alternatives, the utility formula is identical to that of the multinomial logit model. For nests, the formula is different: the utility U_i of a nest $i \in I$ with direct children C_i, optionally with common attributes $\mathbf{X_i}$[65], is defined as follows (Sobel, 1980; Koppelman & Wen, 1998):

$$V_i = \boldsymbol{\beta}_\mathbf{i}^\mathrm{T} \mathbf{X_i} + \theta_i \, \mathrm{E}\left[\max_{j \in I_i} U_j \right] = \boldsymbol{\beta}_\mathbf{i}^\mathrm{T} \mathbf{X_i} + \theta_i \ln \sum_{j \in I_i} e^{V_j}$$

$$U_i = V_i + \varepsilon_i$$

This formula contains the expected maximum utility of the children of the nest. This is called the logsum. It is larger than the maximum expected utility of the children, which in turn is larger than the average expected utility of the children. The difference is "diversity benefit" (Daly, 2012), because each individual can choose his personally optimal child alternative.

The formulas for direct alternatives (no children) and nests (with children) may be combined:

$$V_i = \begin{cases} \boldsymbol{\beta}_\mathbf{i}^\mathrm{T} \mathbf{X_i} & \text{if } C_i = \varnothing \\ \boldsymbol{\beta}_\mathbf{i}^\mathrm{T} \mathbf{X_i} + \theta_i \ln \sum_{j \in C_i} e^{V_j} & \text{otherwise} \end{cases} \quad \forall i \in I$$

$$U_i = V_i + \varepsilon_i \quad \forall i \in I$$

The probabilities of alternatives (including nests themselves) are calculated by applying the multinomial logit formula at the level of considered alternative and multiplying the result with the probability of the parent nest, if any:

$$p_i = \begin{cases} p_a \dfrac{e^{V_i}}{\sum_{j \in C_a} e^{V_j}} & \text{if } \exists a \in I : i \in C_a \\ \dfrac{e^{V_i}}{\sum_{j \in I \setminus \bigcup_{a \in I} C_a} e^{V_j}} & \text{otherwise} \end{cases} \quad \forall i \in I$$

Nest coefficients

For each nest, the nest coefficient θ_i is determined by the correlation ρ_i between pairs of alternatives within the nest (Koppelman & Wen, 1998):

$$\rho_i = 1 - \theta_i^2 \Leftrightarrow \theta_i = \sqrt{1 - \rho_i}$$

For consistency reasons, the nest coefficient must satisfy $0 < \theta_i \leq 1$ (Sobel, 1980), implying that the higher-level choices are more stochastic than the lower-level choices:

[65] Common attributes can also be placed in the utility functions of the nested alternatives themselves. The coefficients then should be divided by θ_i.

- If $\theta_i < 0$, an increase in utility of an alternative decreases the probability its parent nest is selected.
- If $\theta_i = 0$, a change in utility of an alternative does not affect the probability its parent nest is selected.
- If $\theta_i = 1$, the results are identical to the multinomial logit model.
- If $\theta_i > 1$, an increase in utility of an alternative may increase the probabilities of other alternatives in the same nest as well. The lower the probability of the nest $\mathrm{P}(i)$, the lower the upper bound for θ_i must be to prevent this (Herriges & Kling, 1996). To have consistency under all circumstances, θ_i must not be larger than one.

The nest coefficients θ_i can be included in the maximum likelihood estimation process. If a nest coefficient is larger than one or not significantly smaller than one, the nesting can be removed (by setting the coefficient to one).

Practical considerations

Due to its hierarchical structure, the nested logit model is hard to apply to route choice, because in any but the simplest network there exist routes that share links with more than one other possible route: such routes would need multiple parents in the nested logit tree. Here is such an example where route B correlates with both routes A and C:

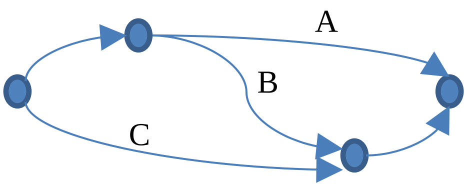


Figure 47: route choice problem that cannot be solved by a nested logit model

The nested logit model is however a very common model to link different choice processes together. Both the GroeiModel (Significance, 2012a) and ANTONIN (Willigers & Tuinenga, 2007) use it to integrate mode and destination choice. This means that tours either using the same mode or having the same destination are correlated tour alternatives in the combined mode and destination choice. For the GroeiModel, this nesting structure also contains time-of-day choice, train access/egress mode choice and train station choice, and the logsum of the whole choice structure is also used in the tour generation model.

C.4. Cross-nested logit

The cross-nested logit model[66] allows alternatives to be members of multiple nests. For the formulas, a distinction is necessary between the final alternatives $C \subset I$ and the nests $M = I \setminus C$. Each alternative $i \in C$ has a utility U_i like in the simpler logit models:

$$V_i = \boldsymbol{\beta}_\mathbf{i}^{\mathrm{T}} \mathbf{X}_\mathbf{i} \quad \forall i \in C$$
$$U_i = V_i + \varepsilon_i \quad \forall i \in C$$

Each alternative $i \in C$ has for each nest $m \in M$ a degree of membership $\alpha_{im} \geq 0$. The following formula exists for the probabilities of nests and alternatives (Bierlaire, 2006):

$$p_m = \frac{\left(\sum_{i \in C} \left(\alpha_{im} e^{V_i} \right)^{1/\theta_m} \right)^{\theta_m}}{\sum_{n \in M} \left(\sum_{i \in C} \left(\alpha_{in} e^{V_i} \right)^{1/\theta_n} \right)^{\theta_n}} \quad \forall m \in M$$

$$p_{i|m} = \frac{\left(\alpha_{im} e^{V_i} \right)^{1/\theta_m}}{\sum_{j \in C} \left(\alpha_{jm} e^{V_j} \right)^{1/\theta_n}} \quad \forall i \in C, m \in M$$

$$p_i = \sum_{m \in M} p_m p_{i|m} \quad \forall i \in C$$

These equations can be rewritten to:

$$p_m = \frac{e^{\theta_m \ln \sum_{i \in C} \sqrt[\theta_m]{\alpha_{im} e^{V_i}}}}{\sum_{n \in M} e^{\theta_n \ln \sum_{i \in C} \sqrt[\theta_n]{\alpha_{in} e^{V_i}}}} \quad \forall m \in M$$

$$p_{i|m} = \frac{\sqrt[\theta_m]{\alpha_{im} e^{V_i}}}{\sum_{j \in C} \sqrt[\theta_m]{\alpha_{jm} e^{V_j}}} \quad \forall i \in C, m \in M$$

A utility of the nest U_m can now be recognized, which looks very similar to the nested logit model:

$$V_m = \theta_m \ln \sum_{i \in C} \sqrt[\theta_m]{\alpha_{im} e^{V_i}} \quad \forall m \in M$$
$$U_m = V_m + \varepsilon_m \quad \forall m \in M$$

The whole model can now be rewritten to:

66 In literature, 'generalized nested logit model' is sometimes used instead of 'cross-nested logit model' if the nest coefficients θ_m are allowed to differ per nest $m \in M$ (Wen & Koppelman, 2001; Benjamins, 2001). In route choice applications 'link-nested logit model' may also be used.

$$V_i = \begin{cases} \boldsymbol{\beta}_\mathbf{i}^\mathrm{T} \mathbf{X}_\mathbf{i} & \text{if } i \in C \\ \theta_i \ln \sum_{j \in C} \sqrt[\theta_i]{\alpha_{ji} e^{V_j}} & \text{otherwise} \end{cases} \quad \forall i \in I$$

$$p_i = \begin{cases} \sum_{m \in M} p_m \dfrac{\sqrt[\theta_m]{\alpha_{im} e^{V_i}}}{\sum_{j \in C} \sqrt[\theta_m]{\alpha_{jm} e^{V_j}}} & \text{if } i \in C \\ \dfrac{e^{V_i}}{\sum_{j \in M} e^{V_j}} & \text{otherwise} \end{cases} \quad \forall i \in I$$

This proves the cross-nested logit model is a generalization of the two-level nested logit model. The cross-nested logit model can be written as a two-level nested logit model where each alternative $i \in C$ may be present in multiple nests $m \in M$ with linearly transformed utilities U_{im}[67]:

$$\begin{aligned} V_i &= \boldsymbol{\beta}_\mathbf{i}^\mathrm{T} \mathbf{X}_\mathbf{i} \quad \forall i \in C \\ V_{im} &= V_i / \theta_m + \upsilon_{im} \quad \forall m \in M, i \in C_m \\ U_{im} &= V_{im} + \varepsilon_{im} \quad \forall m \in M, i \in C_m \\ V_m &= \theta_m \ln \sum_{i \in C_m} e^{V_{im}} \quad \forall m \in M \\ U_m &= V_m + \varepsilon_m \quad \forall m \in M \\ p_i &= \begin{cases} \sum_{m \in M : i \in C_m} p_m \dfrac{e^{V_{im}}}{\sum_{j \in C_m} e^{V_{jm}}} & \text{if } i \in C \\ \dfrac{e^{V_i}}{\sum_{j \in M} e^{V_j}} & \text{otherwise} \end{cases} \quad \forall i \in I \end{aligned}$$

By comparing this result with the nested logit model equations, the following conditions are found under which the cross-nested logit model equals a nested logit model:

$$\begin{cases} \upsilon_{im} = (1 - 1/\theta_m) V_i & \forall m \in M, i \in C_m \\ C_m \cap C_n = \varnothing & \forall m, n \in M : m \neq n \end{cases}$$

Nest memberships

The nest memberships should satisfy the following equation:

$$\sum_{m \in M} \alpha_{im} = 1 \quad \forall i \in C$$

[67] $\ln \sqrt[\theta_m]{\alpha_{im}} = \upsilon_{im}$ has been substituted.

This both provides a useful interpretation (Wen & Koppelman, 2001) – each alternative is allocated once in total – and makes sure the model is unbiased, i.e. the expected utilities $\mathrm{E}\left[U_i\right]$ are not distorted (Abbé et al., 2007).

Practical and theoretical considerations

The cross-nested logit model is suitable for route choice problems. In this case, the alternatives are routes and the nests are links, and for each route $i \in C$ with links $R_i = \{m \in M : i \in C_m\}$ the nest membership is determined by link lengths z_m divided by the total route lengths $Z_{ii} = \sum_{n \in R_i} z_n$ (Hoogendoorn-Lanser et al., 2005; Ramming, 2002; Bekhor et al., 2006):

$$\alpha_{im} = \begin{cases} \dfrac{z_m}{Z_{ii}} & \text{if } m \in R_i \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in C, m \in M$$

The logsum parameters θ_m can be estimated in the maximum likelihood procedure, usually assuming that they are all equal (Wen & Koppelman, 2001), or may be derived from the network structure (Bekhor et al., 2006):

$$\theta_m = 1 - \frac{\sum\limits_{i \in C} \alpha_{im}}{\left|\{i \in C : m \in R_i\}\right|} = 1 - \frac{1}{\left|C_m\right|} \sum_{i \in C_m} \frac{z_m}{Z_{ii}}$$

Alternatively, for the nested logit model interpretation the nest membership can be written as:

$$\upsilon_{im} = \theta_m \ln \frac{z_m}{Z_{ii}} \qquad \forall m \in M, i \in C_m$$

The route choice model can now be seen as a model in which at the top level, a link is chosen. At the bottom level, a route passing the particular link is chosen, thereby penalizing routes for length on other links ($\upsilon_{im} \leq 0$). Like in a nested logit model, the attractiveness of the routes at the bottom level influences the link choice at the top level.

From a behavioural perspective, the theoretical foundation of the cross-nested logit model for route choice can be criticized in that people chose a route directly instead of simultaneously choosing a link and a route passing that link (Hoogendoorn-Lanser et al., 2005).

The model has also been applied to mode choice (Bierlaire, 2006). In this case there can for example be a nest for public transport and a nest for car transport that both partially contain the alternative of using public transport with car as access mode. A similar structure could be used for different PT systems (Hoogendoorn-Lanser, 2005).

Because the cross-nested logit model can be interpreted as a special two-level nested logit model, it is possible to define a logsum of the cross-nested logit model as a whole. Also, the lower-level utility functions may contain logsums of other nested logit models as a whole. This means the cross-nested logit model can theoretically be embedded anywhere inside a larger nested logit structure, although no practical examples of such applications were found in literature (see also Section C.6).

C.5. Paired combinatorial logit

The paired combinatorial logit model is a special case of the cross-nested logit model in which nests are created for all pairs of alternatives $M = \{i, j \in C : i \neq j\}$ (Koppelman & Wen, 2000; Wen & Koppelman, 2001). Each of these nests contains two members $i, j \in C$ with $\alpha_{im} = \alpha_{jm} = 1$ (or $\upsilon_{im} = \upsilon_{jm} = 0$). Correlations in pairs of alternatives are specified by the values of θ_m.

Theoretical considerations

From a behavioural perspective, it is disputed whether travellers systematically evaluate all alternatives in pairs, like the paired combinatorial logit model assumes, in particular in case of a large number of alternatives (Hoogendoorn-Lanser et al., 2005).

Practical considerations

The paired combinatorial logit model has been applied to multimodal route choice in a supernetwork (Benjamins, 2001). In this case however not enough attention seems to have been given to the parameters θ_m: these are correctly based on the theoretical correlation using the overlapping length of two routes $Z_{ij} = \sum_{m \in R_i \cap R_j} z_m$ (Hoogendoorn-Lanser et al., 2005), but these correlations cannot be converted directly into θ_m values (Koppelman & Wen, 2000) like in the following formula:

$$\theta_{(ij)} \overset{\text{incorrect}}{=} 1 - \rho_{(ij)} = 1 - \frac{Z_{ij}}{\sqrt{Z_{ii} Z_{jj}}}$$

If the correlations are known, the parameters θ_m can only be computed numerically and depend on the number of alternatives (Koppelman & Wen, 2000).

The performance seems similar to the multinomial probit model (Benjamins, 2001).

Like the cross-nested logit model, the paired combinatorial logit model has also been applied to mode choice; this allows to explicitly model correlations, and thus the amount of competition, between all modes (Wen & Koppelman, 2001).

C.6. Network GEV

The concept of the cross-nested logit model may be generalised further (i.e. allowing more nesting levels) to the network generalised extreme value (GEV) model proposed by Daly & Bierlaire (2006). It allows a large amount of freedom in the shape of the choice tree (the network). Multinomial logit, nested logit with an arbitrary number of levels, cross-nested logit and paired combinatorial logit are all special cases of the network GEV model.

Using the same notation as for the previous logit models, the utilities of alternatives and nests $i \in I$ are defined as (Newman, 2008):

$$V_i = \begin{cases} \boldsymbol{\beta}_\mathbf{i}^\mathrm{T} \mathbf{X}_\mathbf{i} & \text{if } i \in C \\ \theta_i \ln \sum_{j \in C_i} \sqrt[\theta_i]{\alpha_{ji} e^{V_j}} & \text{otherwise} \end{cases} \quad \forall i \in I$$

The probabilities of alternatives and nests $i \in I$ are calculated as follows, with the root nest denoted r:

$$p_{i|a} = \frac{\sqrt[\theta_a]{\alpha_{ia} e^{V_i}}}{\sum_{j \in C_a} \sqrt[\theta_a]{\alpha_{ia} e^{V_i}}} \quad \forall a \in I, i \in C_a$$

$$p_i = \begin{cases} 1 & \text{if } i = r \\ \sum_{a \in I : i \in C_a} p_a p_{i|a} & \text{otherwise} \end{cases} \quad \forall i \in I$$

Nest memberships

The choice tree needs to be finite, circuit-free and contain a single root nest (Daly & Bierlaire, 2006). How the nest memberships should be normalised, depends on the shape of the choice tree (Newman, 2008).

For a so-called 'crash free' tree in which all choice sequences leading to the same final choice split at the first choice of the sequence, the following normalisation is unbiased:

$$\sum_{a \in I : i \in C_a} \sqrt[\theta_r]{\alpha_{ia}} = 1 \quad \forall i \in I \setminus \{r\}$$

Unbiased here means that the expected utilities of the alternatives $\mathrm{E}[U_i]$ are not modified in a way that would affect the choice probabilities p_i.

Newman (2008) also found an unbiased normalisation for so-called 'crash safe' trees in which all choice sequences leading to the same final choice join at the last choice of the sequence. This one is not listed here due to its complexity. In case of a single level of nests, the tree is both 'crash free' and 'crash safe' and both normalisation procedures lead to the same result identical to the cross-nested logit model with $\theta_r = 1$.

Nest coefficients

Due to the root $\sqrt[\theta_a]{\ldots} \equiv \ldots^{1/\theta_a}$ in the choice probability formula, the nest coefficients are effectively equal to θ_i / θ_a when $i \in C_a$ is a sub-nest of nest $a \in I$. The requirement that nest coefficients should lie between zero and one is then, identical to other nested logit models (Newman, 2008):

$$0 < \theta_i \leq \theta_a \quad \forall a \in I, i \in C_a$$

For the alternatives $i \in C$ themselves and for any degenerate nests $i \in I \setminus C : |C_i| = 1$ (i.e. nests with a single child), the parameters θ_i are undetermined (Newman, 2008). For the root nest r , the parameter θ_r is typically set to one (Newman, 2008):

$$\theta_r = 1$$

Practical and theoretical considerations

The network GEV model provides large flexibility in defining the choice process. Like with the previous logit models, the modeller is responsible for using a meaningful choice tree. If parameter estimation results in nest memberships α_{ia} equal to zero, this may suggest the choice tree should be adapted (Daly & Bierlaire, 2006).

Like the cross-nested logit model, the nest memberships may be specified a-priori based on the choice set or estimated. However, their normalisation should be taken into account, which also puts restrictions on the choice tree structure.

C.7. Joint network GEV

The joint network GEV model is a modification of the network GEV model in which not the elemental alternatives at the bottom of the tree form the choice set, but the different paths through the choice tree form the choice set (Papola & Marzano, 2013). In other words, the choice is defined as the sequence of all intermediate decisions, instead of a final outcome.

Because a full choice tree path is now chosen, additional utility may be collected along this path, unlike the network GEV model:

$$V_i = \begin{cases} \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X_i} & \text{if } i \in C \\ \boldsymbol{\beta}^{\mathrm{T}}\mathbf{X_i} + \theta_i \ln \sum_{j \in C_i} \sqrt[\theta_i]{\alpha_{ji} e^{V_j}} & \text{otherwise} \end{cases} \quad \forall i \in I$$

The conditional choice probability formula is the same as for the network GEV model, but one now calculates the probability p_r for a particular path $r \in \hat{C}$ through the choice tree, by simply multiplying all conditional probabilities along the path, rather than the probability p_i for each bottom-end of the choice tree $i \in C$:

$$p_{i|a} = \frac{\sqrt[\theta_a]{\alpha_{ia} e^{V_i}}}{\sum_{j \in C_a} \sqrt[\theta_a]{\alpha_{ia} e^{V_i}}} \quad \forall a \in I, i \in C_a$$

$$p_r = \prod_{a \in R_r : \exists! i \in C_a \cap R_r} p_{i|a} \quad \forall r \in \hat{C}$$

Practical and theoretical considerations

Papola & Marzano (2013) apply this model to route choice by defining this as a sequence of link choices, as a natural way to take route overlap into account, and proposing formulas for nest coefficients θ_a and nest memberships α_{ia}. The flexibility of the network GEV tree effectively circumvents the problems of the nested logit model for route choice described in Section C.3.

Due to its structure, the model looks interesting for ad hoc route choice behaviour, and, unlike other route choice models, the joint network GEV model indeed prefers route overlap at the beginning of the route over route overlap at the end of the route. However, logit error terms are traditionally based on choice aspects unknown to the researcher (Train, 2002), while ad hoc route choice suggests it includes choice aspects unknown to the decision-maker, while the uncertainties even reduce during the decision process.

The model requires to assume linear valuation of route attributes, since the route utility must be the sum of all link utilities. Furthermore, the required independence of link utilities seems problematic for multi-modal contexts, since for example the utilities of links of the same mode cannot be positively correlated. Also, the suggested nest coefficient and nest membership specification may be hard to apply to multi-modal contexts, since the best path from a node to the destination depends on the valuation of route attributes. Furthermore, since there may be considerable differences in valuation of paths, the number of possible paths may not be a good basis for defining nest memberships.

In a comparison with the multinomial probit model for a road network, the joint network GEV model slightly outperformed the probit model. It is noteworthy that congestion could be neglected in this application, such that ad hoc route choice seems unlikely to have played a role in this test case.

C.8. C-logit and path size logit

The C-logit model and the path size logit model do not use nesting structures, but only modify the utilities of the alternatives to incorporate the effect of correlations (Hoogendoorn-Lanser et al., 2005)[68]:

$$V_i = \boldsymbol{\beta}_\mathbf{i}^\mathrm{T}\mathbf{X}_\mathbf{i} + \upsilon_i = \boldsymbol{\beta}_\mathbf{i}^\mathrm{T}\mathbf{X}_\mathbf{i} + \gamma\xi_i \quad \forall i \in I$$

$$U_i = V_i + \varepsilon_i \quad \forall i \in I$$

$$p_i = \frac{e^{V_i}}{\sum_{j \in I} e^{V_j}} \quad \forall i \in I$$

The following table lists the utility modifications that have been proposed (Hoogendoorn-Lanser et al., 2005; Frejinger et al., 2009):

[68] In C-logit, υ_i is called 'commonality factor'. In path size logit, e^{ξ_i} is called 'path size factor'.

Formula for ξ_i	Model name	Basic element	Symmetric penalties
$-\ln \sum_{j \in I} \rho_{ij}^{\chi}$	C-logit	Route	Yes
$-\ln \sum_{m \in R_i} \frac{z_m}{Z_{ii}} \left\| \left\{ j \in I : m \in R_j \right\} \right\|$	C-logit	Link	No
$\sum_{m \in R_i} \frac{z_m}{Z_{ii}} \ln \frac{1}{\left\| \left\{ j \in I : m \in R_j \right\} \right\|}$	C-logit	Link	No
$-\ln \left(1 + \sum_{j \in I \setminus \{i\}} \rho_{ij} \frac{Z_{ii} - Z_{ij}}{Z_{jj} - Z_{ij}} \right)$	C-logit	Route	No
$\ln \sum_{m \in R_i} \frac{z_m}{Z_{ii}} \frac{1}{\left\| \left\{ j \in I : m \in R_j \right\} \right\|}$	Path size logit	Link	No
$\ln \sum_{m \in R_i} \frac{z_m}{Z_{ii}} \frac{1}{\sum_{j \in I : m \in R_j} \left(\frac{Z_{ii}}{Z_{jj}} \right)^{\chi}}$	Extended path size logit	Link	No
$\ln \sum_{m \in R_i} \frac{z_m}{Z_{ii}} \frac{1}{1 + \sum_{j \in I \setminus \{j\} : m \in R_j} \max \left(\frac{1}{\mathrm{E}[f_j]}, 1 \right)}$	Expanded path size logit	Link	No

Table 30: variants of the C-logit and path size logit models

Only the first formulation is symmetric, i.e. for an overlapping link between two alternatives, both alternatives get the same penalty. Only the first and fourth formulations are route-based and could theoretically be generalized to non-route choice situations[69]; the others are link-based.

In the path size logit model, for each link in a route, the route gains utility if the link is not used by other routes. Note that all other routes using the link are counted equally, regardless of their attractiveness or length. To remove this ‘bias’ towards long routes, the extended path size logit model was introduced.[70] In the extended path size logit model, for each link in a route, the route gains utility if it is the shortest route using that link. The extended path size logit model equals the normal path size logit model for $\chi = 0$.

The expanded path size logit model is another modification to the path size logit model, taking into account the expected number of times $\mathrm{E}\left[f_j \right]$ an alternative is generated by the choice set generation procedure[71] (Frejinger et al., 2009). Compared to the original path size logit model, if a likely generated route and a unlikely generated route overlap, the likely route gets a larger penalty for the overlap.

[69] Note that $\frac{Z_{ii} - Z_{ij}}{Z_{jj} - Z_{ij}} = \frac{\mathrm{Var}(i) - \mathrm{Cov}(i,j)}{\mathrm{Var}(j) - \mathrm{Cov}(i,j)}$ (Van Nes & Bovy, 2008).

[70] When comparing the formulas, note that $\left| \left\{ j \in I : m \in R_j \right\} \right| \equiv \sum_{j \in I : m \in R_j} 1$.

[71] The probability distribution of f_j is to be derived from the choice set generation algorithm.

Theoretical considerations

The C-logit model lacks theoretical foundation for the definition of ξ_i; because of this, it is not clear what formula to use (Ramming, 2002). For this reason, the path size logit model was developed based on choice theory for aggregate alternatives.

Contrary to the other logit models, the extended path size logit model cannot be seen as an approximation of the covariance matrix, due to the large penalties on long routes (Hoogendoorn-Lanser, 2005). The original path size logit model can be used if the route alternatives in the choice set are comparable in length; if this is not the case, the quality of the choice set may be disputed (Hoogendoorn-Lanser et al., 2005).

The expanded path size logit model is questionable from a behavioural perspective, since it assumes that people chose directly from universal choice set instead of a subset of considered alternatives (Frejinger et al., 2009). Since the model also has not been applied yet to real-life data, it is probably wise to stick with the original path size logit model.

Path size coefficients

For the path size and extended path size logit models, the weighted path size factor $e^{\gamma\xi_i}$ follows the following formula for a set of fully overlapping alternatives:

$$e^{\gamma\xi_i} = \left(e^{\xi_i}\right)^{\gamma} = \left(\frac{1}{\sum_{j\in I} 1^{\chi}}\right)^{\gamma} = \frac{1}{|I|^{\gamma}}$$

From this equation, one can derive that the path size coefficient γ must satisfy $0 < \gamma \leq 1$:

- If $\gamma < 0$, the existence of overlapping alternatives increases the probability of the alternative being selected (since $e^{\gamma\xi_i} > 1$).
- If $\gamma = 0$, overlap is not taken into account (since $e^{\gamma\xi_i} = 1$).
- If $\gamma = 1$, a set of identical alternatives is valued identically to a single one of these alternatives (since $e^{\gamma\xi_i} = 1/|I|$).
- If $\gamma > 1$, a set of identical alternatives is valued less than a single one of these alternatives (since $e^{\gamma\xi_i} < 1/|I|$).

Practical considerations

C-logit and path size logit models are frequently used for route choice due to their simple structure (Frejinger et al., 2009). The C-logit model has been applied in practice to multimodal route choice in a supernetwork by Carlier et al. (2003), although they do not explain which formulation of the commonality factor was used.

Research suggests that $\chi \to \infty$ yields the best results for the extended path size logit model (Hoogendoorn-Lanser et al., 2005). In this case, the extended path size logit model outperforms the cross-nested logit model mentioned earlier[72] (Bekhor et al., 2006). Both outperform the normal path size logit model with $\gamma = 1$ which only yields very small improvements compared to the multinomial logit model (Hoogendoorn-Lanser et al., 2005; Bekhor et al., 2006).

C.9. Logit-kernel

The logit-kernel model[73] uses a full covariance matrix to account for overlap instead of a nesting structure or an additional utility component. The utilities U_i are similar to a multinomial logit model, but the error terms ε_i are no longer independent (Hoogendoorn-Lanser et al., 2005; Ramming, 2002):

$$\begin{gathered} U_i = \boldsymbol{\beta}^{\mathrm{T}} \mathbf{X_i} + \varepsilon_i \quad \forall i \in I \\ \boldsymbol{\varepsilon} = \mathbf{FT}\boldsymbol{\zeta} + \boldsymbol{\upsilon} \end{gathered}$$

Here, $\boldsymbol{\zeta}$ are independent and identically distributed random variables with $\mathrm{E}\left[\zeta_i\right] = 0$ and $\mathrm{Var}\left(\zeta_i\right) = 1$, and $\boldsymbol{\upsilon}$ are independent and identically distributed Gumbel variables. Furthermore, $\mathbf{F}$ is the link-route incidence matrix and $\mathbf{T}$ is a lower-triangular matrix based on link lengths z_m.

The covariance matrix then follows the following formula (Ramming, 2002; Hoogendoorn-Lanser et al., 2005):

$$\mathrm{Cov}\left(\boldsymbol{\varepsilon}\right) = \mathbf{FTT}^{\mathrm{T}}\mathbf{F}^{\mathrm{T}} + \mathrm{Var}\left(\upsilon\right)\mathbf{I}$$

Like the multinomial probit model, the logit-kernel model requires Monte Carlo simulation (for $\boldsymbol{\xi}$) (Ramming, 2002), making it slow compared to the other models.

Practical considerations

The logit-kernel model is very successful in improving likelihood, but other model parameters tend to become less significant than in other logit models. Also, the parameter estimation process is quite slow and there are stability issues with the solution (Hoogendoorn-Lanser et al., 2005).

[72] A combination of the cross-nested logit model and the extended path size factor (both modelling route overlap, not to be confused with Section E.1), performs even better, although no theoretical interpretation of this model is given.
[73] Also known in literature as mixed logit model or hybrid logit model (Fiorenzo-Catalano, 2007).

Appendix D. Time-dependent public transport choice model

This appendix derives a time-dependent path size logit model for the choice of a service in a timetable-based public transport network with a periodic timetable. The time-dependency is incorporated because the reasonable service alternatives and waiting times depend on the moment the traveller arrives at the initial station. Although incorporating time-dependencies, the model is designed as a static assignment tool.

This appendix also investigates some special cases and extensions of this model, e.g. showing that the frequency-based public transport network is a special case of multiple connected timetable-based networks. Finally, route choice in case of multiple independent timetables is discussed.

D.1. Assumptions and definitions

In all cases, it is assumed that:

- all timetables are periodic;
- the arrivals of travellers at the origin station is uniformly distributed;
- the relation between waiting time and utility is linear;
- choices within timetables are based on the current values of the time-dependent utilities upon entering the system governed by the timetable;
- higher-level choices are based on the expected values of the logsums of the timetables.

A service is defined as either a single line or a single permutation of lines with identical, known transfer locations and transfer waiting times. Only the waiting time at the boarding of the first line is a stochastic variable.

A route is defined as either a single service or a single permutation of services with known transfer locations but unknown transfer waiting times that gets the traveller from his origin station to his destination station. The waiting time at the boarding of every service is a stochastic variable.

A run is defined as an instance of a service at a single known departure time. A service with multiple departure times in the period consists of multiple runs.

The continuous and discrete Heaviside step functions are respectively defined as (Wikimedia, 2012b):

$$H(x)=\begin{cases}0 & \text{if } x<0\\ \text{undefined} & \text{if } x=0\\ 1 & \text{if } x>0\end{cases}\quad \forall x\in\mathbb{R}$$

$$H[n]=\begin{cases}0 & \text{if } n<0\\ 1 & \text{if } n\geq 0\end{cases}\quad \forall n\in\mathbb{Z}$$

D.2. Multiple regular services timetable

A periodic timetable with period T is considered. Each run $i\in C$ has an intrinsic utility V_i and a departure time $t_{d,i}$. Travellers arriving at moment t within the period have a waiting time $t_{w,i}(t)$, determining the run utility $\hat{V}_i(t)$:

$$t_{w,i}(t)=\begin{cases}t_{d,i}-t & \text{if } t\leq t_{d,i}\\ t_{d,i}+T-t & \text{otherwise}\end{cases}$$

$$\begin{aligned}\hat{V}_i(t)&=V_i+\beta_{t_w}t_{w,i}(t)\\ &=\begin{cases}V_i+\beta_{t_w}\left(t_{d,i}-t\right) & \text{if } t\leq t_{d,i}\\ V_i+\beta_{t_w}\left(t_{d,i}-t+T\right) & \text{otherwise}\end{cases}\\ &=V_i+\beta_{t_w}\left(t_{d,i}-t+H\left(t-t_{d,i}\right)T\right)\end{aligned}$$

The probability $p_i(t)$ of a run being chosen is determined by a time-dependent path size logit model:

$$p_i(t)=\frac{\alpha_i(t)e^{\hat{V}_i(t)+\gamma\xi_i(\boldsymbol{\alpha}(t))}}{\sum_{j\in C}\alpha_j(t)e^{\hat{V}_j(t)+\gamma\xi_j(\boldsymbol{\alpha}(t))}}$$

The logical function $\alpha_i(t)\in\{0,1\}$ may be used to exclude run i from the choice set at certain times because at these times they are dominated by other runs with better characteristics (including initial waiting time). To include all runs, use $\alpha_i(t)=1\ \ \forall i\in C$. Note that the same run a period T later is automatically excluded from the choice set.

The model includes a path size factor $e^{\xi_i(\boldsymbol{\alpha}(t))}$ that obviously depends on the choice set composition $\boldsymbol{\alpha}(t)$. The model can work with any of the C-logit or path size logit formulations. Note that the expression of route overlap cannot include initial waiting time; other measures like number of legs are fine.

The choice probability $p_i(t)$ is still time-dependent. However, the time-independent expected choice probability $\mathrm{E}[p_i]$ may be calculated by integrating over all possible times:

$$\mathrm{E}\left[p_i\right]=\frac{1}{T}\int_0^T \frac{\alpha_i\left(t\right)e^{\hat{V}_i(t)+\gamma\xi_i(\boldsymbol{\alpha}(t))}}{\sum_{j\in C}\alpha_j\left(t\right)e^{\hat{V}_j(t)+\gamma\xi_j(\boldsymbol{\alpha}(t))}}dt=\frac{1}{T}\int_0^T \frac{\alpha_i\left(t\right)e^{V_i+\beta_{t_w}\left(t_{d,i}-t+H\left(t-t_{d,i}\right)T\right)+\gamma\xi_i(\boldsymbol{\alpha}(t))}}{\sum_{j\in C}\alpha_j\left(t\right)e^{V_j+\beta_{t_w}\left(t_{d,j}-t+H\left(t-t_{d,j}\right)T\right)+\gamma\xi_j(\boldsymbol{\alpha}(t))}}dt$$

$$=\frac{1}{T}\int_0^T \frac{\alpha_i\left(t\right)e^{V_i+\beta_{t_w}\left(t_{d,i}+H\left(t-t_{d,i}\right)T\right)+\gamma\xi_i(\boldsymbol{\alpha}(t))}}{\sum_{j\in C}\alpha_j\left(t\right)e^{V_j+\beta_{t_w}\left(t_{d,j}+H\left(t-t_{d,j}\right)T\right)+\gamma\xi_j(\boldsymbol{\alpha}(t))}}dt$$

This integral can be solved by ordering the runs by departure time $t_{d,i}$ and noting that, between two subsequent departures, the multinomial logit model is constant[74]:

$$\mathrm{E}\left[p_i\right]=\frac{1}{T}\left(\begin{array}{l}\int_0^{t_{d,1}} \frac{\alpha_{i0}e^{V_i+\beta_{t_w}t_{d,i}+\gamma\xi_i(\boldsymbol{\alpha}(t))}}{\sum_{j\in C}\alpha_{j0}e^{V_j+\beta_{t_w}t_{d,j}+\gamma\xi_j(\boldsymbol{\alpha}(t))}}dt+\sum_{k=1}^{|C|-1}\int_{t_{d,k}}^{t_{d,k+1}} \frac{\alpha_i\left(t\right)e^{V_i+\beta_{t_w}\left(t_{d,i}+H\left(t-t_{d,i}\right)T\right)+\gamma\xi_i(\boldsymbol{\alpha}(t))}}{\sum_{j\in C}\alpha_j\left(t\right)e^{V_j+\beta_{t_w}\left(t_{d,j}+H\left(t-t_{d,j}\right)T\right)+\gamma\xi_j(\boldsymbol{\alpha}(t))}}dt\\ +\int_{t_{d,|C|}}^{T} \frac{\alpha_{i0}e^{V_i+\beta_{t_w}\left(t_{d,i}+T\right)+\gamma\xi_i(\boldsymbol{\alpha}(t))}}{\sum_{j\in C}\alpha_{j0}e^{V_j-\beta_{t_w}\left(t_{d,j}+T\right)+\gamma\xi_j(\boldsymbol{\alpha}(t))}}dt\end{array}\right)$$

$$=\frac{1}{T}\left(\frac{\alpha_{i0}e^{V_i+\beta_{t_w}t_{d,i}+\gamma\xi_i(\boldsymbol{\alpha_0})}}{\sum_{j\in C}\alpha_{j0}e^{V_j+\beta_{t_w}t_{d,j}+\gamma\xi_j(\boldsymbol{\alpha_0})}}\left(t_{d,1}+T-t_{d,|C|}\right)+\sum_{k=1}^{|C|-1}\int_{t_{d,k}}^{t_{d,k+1}}\frac{\alpha_{ik}e^{V_i+\beta_{t_w}\left(t_{d,i}+H[k-i]T\right)+\gamma\xi_i(\boldsymbol{\alpha_k})}}{\sum_{j\in C}\alpha_{jk}e^{V_j+\beta_{t_w}\left(t_{d,j}+H[k-j]T\right)+\gamma\xi_j(\boldsymbol{\alpha_k})}}dt\right)$$

$$=\frac{t_{d,1}+T-t_{d,|C|}}{T}\frac{\alpha_{i0}e^{V_i+\beta_{t_w}t_{d,i}+\gamma\xi_i(\boldsymbol{\alpha_0})}}{\sum_{j\in C}\alpha_{j0}e^{V_j+\beta_{t_w}t_{d,j}+\gamma\xi_j(\boldsymbol{\alpha_0})}}+\sum_{k=1}^{|C|-1}\frac{t_{d,k+1}-t_{d,k}}{T}\frac{\alpha_{ik}e^{V_i+\beta_{t_w}\left(t_{d,i}+H[k-i]T\right)+\gamma\xi_i(\boldsymbol{\alpha_k})}}{\sum_{j\in C}\alpha_{jk}e^{V_j+\beta_{t_w}\left(t_{d,j}+H[k-j]T\right)+\gamma\xi_j(\boldsymbol{\alpha_k})}}$$

This formula gives the probabilities of runs being chosen. To find the probabilities of services being chosen, one can simply sum the probabilities of the runs of that service.

The logsum $\bar{V}\left(t\right)$ of all runs (i.e. the timetable utility) can be calculated to embed this run choice model in a larger nested logit model:

$$\bar{V}\left(t\right)=\ln\sum_{i\in C}\alpha_i\left(t\right)e^{\hat{V}_i(t)+\gamma\xi_i(\boldsymbol{\alpha}(t))}$$

Again, the time dependency is removed by taking the expected value $\mathrm{E}\left[\bar{V}\right]$:

[74] Note that $\boldsymbol{\alpha}\left(t\right)=\begin{cases}\boldsymbol{\alpha_0} & \text{if } 0<t<t_{d,1}\vee t_{d,|C|}<t<T\\ \boldsymbol{\alpha_k} & \text{if } t_{d,k}<t<t_{d,k+1}\end{cases}$ or $\alpha_i\left(t\right)=\begin{cases}\alpha_{i0} & \text{if } 0<t<t_{d,1}\vee t_{d,|C|}<t<T\\ \alpha_{ik} & \text{if } t_{d,k}<t<t_{d,k+1}\end{cases}$.

$$
\begin{aligned}
\mathrm{E}\left[\bar{V}\right] &= \frac{1}{T}\int_0^T \ln\sum_{i\in C}\alpha_i(t)e^{\hat{V}_i(t)+\gamma\xi_i(\boldsymbol{\alpha}(t))}dt = \frac{1}{T}\int_0^T \ln\sum_{i\in C}\alpha_i(t)e^{V_i+\beta_{t_w}\left(t_{d,i}-t+H\left(t-t_{d,i}\right)T\right)+\gamma\xi_i(\boldsymbol{\alpha}(t))}dt \\
&= \frac{1}{T}\int_0^T \ln\left(e^{-\beta_{t_w}t}\sum_{i\in C}\alpha_i(t)e^{V_i+\beta_{t_w}\left(t_{d,i}+H\left(t-t_{d,i}\right)T\right)+\gamma\xi_i(\boldsymbol{\alpha}(t))}\right)dt \\
&= \frac{1}{T}\int_0^T \left(\ln e^{-\beta_{t_w}t}+\ln\sum_{i\in C}\alpha_i(t)e^{V_i+\beta_{t_w}\left(t_{d,i}+H\left(t-t_{d,i}\right)T\right)+\gamma\xi_i(\boldsymbol{\alpha}(t))}\right)dt \\
&= \frac{1}{T}\left(\int_0^T \ln\sum_{i\in C}\alpha_i(t)e^{V_i+\beta_{t_w}\left(t_{d,i}+H\left(t-t_{d,i}\right)T\right)+\gamma\xi_i(\boldsymbol{\alpha}(t))}dt-\int_0^T\beta_{t_w}t\,dt\right) \\
&= \frac{1}{T}\left(\begin{aligned}&\int_0^{t_{d,1}}\ln\sum_{i\in C}\alpha_{i0}e^{V_i+\beta_{t_w}t_{d,i}+\gamma\xi_i(\boldsymbol{\alpha}_0)}dt+\sum_{k=1}^{|C|-1}\int_{t_{d,k}}^{t_{d,k+1}}\ln\sum_{i\in C}\alpha_i(t)e^{V_i+\beta_{t_w}\left(t_{d,i}+H\left(t-t_{d,i}\right)T\right)+\gamma\xi_i(\boldsymbol{\alpha}(t))}dt \\ &+\int_{t_{d,|C|}}^{T}\ln\sum_{i\in C}\alpha_{i0}e^{V_i+\beta_{t_w}\left(t_{d,i}+T\right)+\gamma\xi_i(\boldsymbol{\alpha}_0)}dt\end{aligned}\right)-\frac{\beta_{t_w}}{T}\int_0^T t\,dt \\
&= \frac{1}{T}\left(\begin{aligned}&t_{d,1}\ln\sum_{i\in C}\alpha_{i0}e^{V_i+\beta_{t_w}t_{d,i}+\gamma\xi_i(\boldsymbol{\alpha}_0)}+\sum_{k=1}^{|C|-1}\int_{t_{d,k}}^{t_{d,k+1}}\ln\sum_{i\in C}\alpha_{ik}e^{V_i+\beta_{t_w}\left(t_{d,i}+H[k-i]T\right)+\gamma\xi_i(\boldsymbol{\alpha}_k)}dt \\ &+\left(T-t_{d,|C|}\right)\ln\sum_{i\in C}\alpha_{i0}e^{V_i+\beta_{t_w}\left(t_{d,i}+T\right)+\gamma\xi_i(\boldsymbol{\alpha}_0)}\end{aligned}\right)-\frac{\beta_{t_w}}{T}\frac{T^2}{2} \\
&= \frac{t_{d,1}}{T}\ln\sum_{i\in C}\alpha_{i0}e^{V_i+\beta_{t_w}t_{d,i}+\gamma\xi_i(\boldsymbol{\alpha}_0)}+\sum_{k=1}^{|C|-1}\frac{t_{d,k+1}-t_{d,k}}{T}\ln\sum_{i\in C}\alpha_{ik}e^{V_i+\beta_{t_w}\left(t_{d,i}+H[k-i]T\right)+\gamma\xi_i(\boldsymbol{\alpha}_k)} \\
&\quad+\frac{T-t_{d,|C|}}{T}\ln\sum_{i\in C}\alpha_{i0}e^{V_i+\beta_{t_w}\left(t_{d,i}+T\right)+\gamma\xi_i(\boldsymbol{\alpha}_0)}-\beta_{t_w}\frac{T}{2}
\end{aligned}
$$

Note the structure of this formula. The last term, the expected arrival time of travellers, is subtracted from a weighted average of inter-departure utilities including waiting times with respect to the start of the period.

D.3. Regular single service timetable

If all runs belong to a single service with intrinsic utility V, all runs are dominated by the next run to depart. Hence, at each time instant the choice set has size one and no overlap. The probabilities of runs being chosen then reduce to:

$$
\mathrm{E}\left[p_i\right]=\begin{cases}\dfrac{t_{d,1}+T-t_{d,|C|}}{T} & \text{if } i=1 \\[2ex] \dfrac{t_{d,i}-t_{d,i-1}}{T} & \text{otherwise}\end{cases}
$$

The expected service utility reduces to[75]:

[75] In the resulting formula, the expected waiting time $\mathrm{E}\left[t_w\right]$ can be identified because of the expectation operator is linear.

$$
\begin{aligned}
\mathrm{E}\left[\bar{V}\right] &= \frac{t_{d,1}}{T}\left(V+\beta_{t_w}t_{d,1}\right)+\sum_{k=1}^{|C|-1}\frac{t_{d,k+1}-t_{d,k}}{T}\left(V+\beta_{t_w}t_{d,k+1}\right)+\frac{T-t_{d,|C|}}{T}\left(V+\beta_{t_w}\left(t_{d,1}+T\right)\right)-\beta_{t_w}\frac{T}{2} \\
&= V+\frac{t_{d,1}}{T}\beta_{t_w}t_{d,1}+\sum_{k=1}^{|C|-1}\frac{t_{d,k+1}-t_{d,k}}{T}\beta_{t_w}t_{d,k+1}+\frac{T-t_{d,|C|}}{T}\beta_{t_w}\left(t_{d,1}+T\right)-\beta_{t_w}\frac{T}{2} \\
&= V+\beta_{t_w}\underbrace{\left(\frac{t_{d,1}}{T}t_{d,1}+\sum_{k=1}^{|C|-1}\frac{t_{d,k+1}-t_{d,k}}{T}t_{d,k+1}+\frac{T-t_{d,|C|}}{T}\left(t_{d,1}+T\right)-\frac{T}{2}\right)}_{\mathrm{E}[t_w]}
\end{aligned}
$$

D.4. Fixed-interval single service timetable

In case of fixed-interval runs at frequency $f=n/T$, the probabilities of runs being chosen simply equal:

$$\mathrm{E}\left[p_i\right]=\frac{1}{n}$$

The expected service utility is simplified by entering $|C|=n$ and $t_{d,i}=\left(i/n\right)T$:

$$
\begin{aligned}
\mathrm{E}\left[\bar{V}\right] &= V+\beta_{t_w}\left(\frac{1}{n}\frac{1}{n}T+\sum_{k=1}^{n-1}\frac{1}{n}\frac{k+1}{n}T-\frac{T}{2}\right) \\
&= V+\beta_{t_w}\left(\sum_{k=0}^{n-1}\frac{1}{n}\frac{k+1}{n}T-\frac{T}{2}\right) \\
&= V+\beta_{t_w}\left(\sum_{k=1}^{n}\frac{k}{n^2}T-\frac{T}{2}\right) \\
&= V+\beta_{t_w}\left(\frac{T}{n^2}\sum_{k=1}^{n}k-\frac{T}{2}\right) \\
&= V+\beta_{t_w}\left(\frac{T}{n^2}\left(\tfrac{1}{2}n\left(1+n\right)\right)-\frac{T}{2}\right) \\
&= V+\beta_{t_w}\left(\frac{T}{2n}+\frac{T}{2}-\frac{T}{2}\right) \\
&= V+\beta_{t_w}\frac{T}{2n}=V+\beta_{t_w}\underset{\mathrm{E}[t_w]}{\frac{1}{2f}}
\end{aligned}
$$

D.5. Periodic-irregular single service timetable

In the periodic-irregular single service timetable, all n runs are randomly distributed over period T, but this random pattern is the same in each period.

Each run i now runs according to a local time t_i independent of those of other runs in the same period. Without loss of generality, the departure times $t_{d,i}$ according to the local vehicle time can be set to zero. The domination indicator α_i now depends on local times of all vehicles $\mathbf{t}$. This leads to the following equations[76]:

$$t_{w,i}\left(t_i\right) = T - t_i$$
$$\hat{V}_i\left(t_i\right) = V + \beta_{t_w}\left(T - t_i\right)$$
$$\bar{V}\left(\mathbf{t}\right) = \ln \sum_{i \in |C|} \alpha_i\left(\mathbf{t}\right) e^{\hat{V}_i\left(t_i\right)} = \sum_{i \in |C|} \alpha_i\left(\mathbf{t}\right)\hat{V}_i\left(t_i\right)$$

Because all vehicles are equal, the following holds:

$$\mathrm{E}\left[p_i\right] = \frac{1}{n}$$
$$\mathrm{E}\left[\bar{V}\right] = \mathrm{E}\left[\sum_{i \in |C|} \alpha_i \hat{V}_i\right] = \sum_{i \in |C|} \mathrm{E}\left[\alpha_i \hat{V}_i\right] = n\,\mathrm{E}\left[\alpha_i \hat{V}_i\right]$$

The probability that another vehicle is behind vehicle i (i.e. its next departure will be later) equals t_i / T if the local times of the other vehicles are unknown. This leads to:

$$\mathrm{P}\left(\alpha_i\left(t_i\right) = 1\right) = \left(\frac{t_i}{T}\right)^{n-1}$$

The contribution of each vehicle to the logsum then is:

$$\begin{aligned}
\mathrm{E}\left[\alpha_i \hat{V}_i\right] &= \int_{\Omega} \alpha_i\left(\mathbf{t}\right)\hat{V}_i\left(t_i\right) P\left(d\mathbf{t}\right) \\
&= \frac{1}{T}\int_0^T \mathrm{P}\left(\alpha_i\left(t_i\right) = 1\right)\hat{V}_i\left(t_i\right) dt_i \\
&= \frac{1}{T}\int_0^T \left(\frac{t_i}{T}\right)^{n-1}\left(V + \beta_{t_w}\left(T - t_i\right)\right) dt_i \\
&= \frac{1}{T}\int_0^T \left(\frac{t_i}{T}\right)^{n-1}\left(V + \beta_{t_w} T - \beta_{t_w} t_i\right) dt_i \\
&= \frac{1}{T}\int_0^T \left(\frac{V + \beta_{t_w} T}{T^{n-1}} t_i^{n-1} - \frac{\beta_{t_w}}{T^{n-1}} t_i^{n}\right) dt_i \\
&= \frac{1}{T}\left[\frac{V + \beta_{t_w} T}{nT^{n-1}} t_i^{n} - \frac{\beta_{t_w}}{\left(n+1\right)T^{n-1}} t_i^{n+1}\right]_{t_i=0}^{t_i=T} \\
&= \frac{V + \beta_{t_w} T}{nT^{n}} T^{n} - \frac{\beta_{t_w}}{\left(n+1\right)T^{n}} T^{n+1} \\
&= \frac{V + \beta_{t_w} T}{n} - \frac{\beta_{t_w} T}{n+1}
\end{aligned}$$

[76] Because in each situation $\mathbf{t}$, the choice set has size one, the logsum can be replaced by a normal sum and the overlap correction can be removed. Note however that the element inside the choice set does depend on $\mathbf{t}$.

Hence, the logsum is given by:

$$\mathrm{E}\left[\bar{V}\right] = V + \beta_{t_w} \underbrace{\left(1 - \frac{n}{n+1}\right) T}_{\mathrm{E}[t_w]}$$

D.6. Aperiodic-irregular single service timetable

The case of aperiodic-irregularities is retrieved by extending the period of the case of periodic-irregularities to infinity:

$$\begin{aligned}
\mathrm{E}\left[p_i\right] &= \lim_{k \to \infty} \frac{1}{kn} = 0 \\
\mathrm{E}\left[\bar{V}\right] &= V + \beta_{t_w} \lim_{k \to \infty} \left(1 - \frac{kn}{kn+1}\right) kT \\
&= V + \beta_{t_w} \lim_{k \to \infty} \left(k - \frac{k^2 n}{kn+1}\right) T \\
&= V + \beta_{t_w} \lim_{k \to \infty} \left(\frac{k^2 n + k}{kn+1} - \frac{k^2 n}{kn+1}\right) T \\
&= V + \beta_{t_w} \lim_{k \to \infty} \frac{k}{kn+1} T \\
&= V + \beta_{t_w} \frac{1}{n + \lim\limits_{k \to \infty} \frac{1}{k}} T \\
&= V + \beta_{t_w} \frac{T}{n} = V + \beta_{t_w} \underbrace{\frac{1}{f}}_{\mathrm{E}[t_w]}
\end{aligned}$$

Note that the expected waiting time is twice as high as for a fixed-interval service; this result is confirmed by Goudappel Coffeng (2013) and Lam et al. (2002).

D.7. Overview

The following table summarizes the results:

Timetable	Run probabilities $\mathrm{E}\left[p_i\right]$	Logsum $\mathrm{E}\left[\bar{V}\right]$
Multiple regular services	$\frac{t_{d,1}+T-t_{d,\lvert C\rvert}}{T}\frac{\alpha_{i0}e^{V_i+\beta_{t_w}t_{d,i}+\gamma\xi_i(\boldsymbol{\alpha}_\mathbf{0})}}{\sum_{j\in C}\alpha_{j0}e^{V_j+\beta_{t_w}t_{d,j}+\gamma\xi_j(\boldsymbol{\alpha}_\mathbf{0})}}+$ $\sum_{k=1}^{\lvert C\rvert-1}\frac{t_{d,k+1}-t_{d,k}}{T}\frac{\alpha_{ik}e^{V_i+\beta_{t_w}\left(t_{d,i}+H[k-i]T\right)+\gamma\xi_i(\boldsymbol{\alpha}_\mathbf{k})}}{\sum_{j\in C}\alpha_{jk}e^{V_j+\beta_{t_w}\left(t_{d,j}+H[k-j]T\right)+\gamma\xi_j(\boldsymbol{\alpha}_\mathbf{k})}}$	$\frac{t_{d,1}}{T}\ln\sum_{i\in C}\alpha_{i0}e^{V_i+\beta_{t_w}t_{d,i}+\gamma\xi_i(\boldsymbol{\alpha}_\mathbf{0})}+$ $\sum_{k=1}^{\lvert C\rvert-1}\frac{t_{d,k+1}-t_{d,k}}{T}\ln\sum_{i\in C}\alpha_{ik}e^{V_i+\beta_{t_w}\left(t_{d,i}+H[k-i]T\right)+\gamma\xi_i(\boldsymbol{\alpha}_\mathbf{k})}+$ $\frac{T-t_{d,\lvert C\rvert}}{T}\ln\sum_{i\in C}\alpha_{i0}e^{V_i+\beta_{t_w}\left(t_{d,i}+T\right)+\gamma\xi_i(\boldsymbol{\alpha}_\mathbf{0})}-\beta_{t_w}\frac{T}{2}$
Regular single service	$\frac{t_{d,1}+T-t_{d,\lvert C\rvert}}{T}$ if $i=1$; $\frac{t_{d,i}-t_{d,i-1}}{T}$ otherwise	$V+\beta_{t_w}\left(\frac{t_{d,1}}{T}t_{d,1}+\sum_{k=1}^{\lvert C\rvert-1}\frac{t_{d,k+1}-t_{d,k}}{T}t_{d,k+1}+\frac{T-t_{d,\lvert C\rvert}}{T}\left(t_{d,1}+T\right)-\frac{T}{2}\right)$
Fixed-interval single service	$\frac{1}{n}$	$V+\beta_{t_w}\frac{1}{2f}$
Periodic-irregular single service	$\frac{1}{n}$	$V+\beta_{t_w}\left(1-\frac{n}{n+1}\right)T$
Aperiodic-irregular single service	0 [77]	$V+\beta_{t_w}\frac{1}{f}$

Table 31: expected run probabilities and logsums of several independent timetables

D.8. Route choice

For extending the service choice model above into a route choice model, a distinction must be made between a situation with a single timetable for all public transport lines or multiple independently-functioning timetables.

Single timetable

If the model only contains a single timetable, the service choice model in itself already is the route choice model. However, this can be further extended with private modes.

The public transport routes may be extended with separately generated[78] private access and egress legs by modifying the utilities V_i and shifting the departure times $t_{d,i}$ backwards with the travel time of the access legs. Fully private routes may be added to the choice set with

[77] The number of runs in an infinite period is infinite.
[78] By splitting the route generation problem as in Section 6.2.

constant availability $\alpha_i(t)=1$ with the waiting time component removed from the utility function.

Concatenated services from multiple timetables

If the synchronisation between lines is not present or not known, different lines are considered to belong to different independent timetables, but may be used together in a single route. The route then spans multiple services.

Theoretically, the previously mentioned method of shifting departure times could be applied with expected departure times $\mathrm{E}\left[t_{d,i}\right]$ instead of exact departure times $t_{d,i}$. However, besides creating a false sense of precision, the concatenation of services is not straightforward anymore, particularly if there are many lines with independent timetables.

Multiple timetables in hierarchy

Alternatively, the route choice problem may be reformulated as a bi-level problem. On the upper level, a permutation of timetables including transfer stations is chosen using timetable logsums $\mathrm{E}\left[\bar{V}\right]$. On the lower level, within each timetable a service is chosen using run probabilities $\mathrm{E}\left[p_i\right]$. Each level has its own route set generation procedure: the upper level in a supernetwork of level-of-service matrices[79] and the lower level in a timetable-based public transport network.

Optionally, within a route, different β_{t_w} parameters may be applied to the expected initial waiting time (first service) and the expected transfer waiting times (all subsequent services).

However, note that the upper level works with expected utilities while the lower level works with actual utilities; this is a theoretical weakness of this model regarding its choice modelling consistency. Because of this, one may be better of neglecting synchronisation effects at the origin station by not using a time-dependent choice model as in this appendix.

---

[79] This requires the timetable logsums and thus the service utility components to be additive.

Appendix E. Proposed new advanced logit models

This appendix complements Appendix C with additional advanced logit models for route choice integrated with other choice processes. Contrary to the models from Appendix C, the models in this appendix are newly proposed here; no previous literature has been found about them. However, these models are built by combining concepts of existing models described in Appendix C.

E.1. Cross-nested path size logit

The cross-nested path size logit model, a hybrid form of cross-nested logit and path size logit, is here proposed to integrate mode and route choice into a single logit model in case routes may use multiple modes. Each nest $m \in M$ represents a modal constraint and each alternative $i \in C$ a route. Since routes may satisfy multiple modal constraints, one gets a cross-nested logit construction. For each modal constraint, overlap between routes is modelled using a path size factor.

Initially, the nest memberships α_{im} are specified by the generalised lengths Z_{im} of the route $i \in C$ that satisfy each modal constraint $m \in M$. The measure of Z_{im} can be the same as the measure of overlap that will be used for calculating the path size factors. The nest memberships are equal to:

$$\alpha_{im} = \frac{Z_{im}}{\sum_{n \in M} Z_{in}} \quad \forall i \in C, m \in M$$

The insertion of the (extended) path size factor $e^{\gamma_m \xi_{im}}$ into the nest memberships α_{im} leads to modified nest memberships $\hat{\alpha}_{im}$[80]:

$$\hat{\alpha}_{im} = \alpha_{im} e^{\gamma_m \xi_{im}} \quad \forall i \in C, m \in M$$

$$e^{\xi_{im}} = \sum_{l \in R_i} \frac{z_l}{Z_{ii}} \frac{\alpha_{im}}{\sum_{j \in C: l \in R_j} \alpha_{jm} \left(\frac{Z_{ii}}{Z_{jj}} \right)^{\chi}} \quad \forall i \in C, m \in M$$

Note that if $\sum_{m \in M} \alpha_{im} = 1$, then $0 < \sum_{m \in M} \hat{\alpha}_{im} \leq 1$ in general[81]. Such a relaxation of the nest memberships does not have impact on the validity of the cross-nested logit model (Bierlaire, 2006). From the cross-nested logit perspective, the path size factors introduce bias, i.e. a

---

[80] ξ_{im} has been defined such that $\xi_{im} = 0$ if alternatives in a nest do not overlap, like in the original (extended) path size logit model, and such that the model reduces to the original (extended) path size logit model if $\alpha_{im} = 1$.

[81] Because $\hat{\alpha}_{im} \leq \alpha_{im} \quad \forall i \in C, m \in M$.

distortion of the expected utilities $\mathrm{E}\left[U_i\right]$ (Abbé et al., 2007). This bias is in favour of non-overlapping routes. Note that such bias is inherent to the way the path size model works: the non-nested path size logit model (see Appendix C.8) also modifies the expected utilities $\mathrm{E}\left[U_i\right]$ instead of correlating the error terms ε_i.

The choice probabilities become:

$$p_m = \frac{e^{\theta_m \ln \sum_{i \in C} \sqrt[\theta_m]{\alpha_{im} e^{V_i + \gamma_m \xi_{im}}}}}{\sum_{n \in M} e^{\theta_n \ln \sum_{i \in C} \sqrt[\theta_n]{\alpha_{in} e^{V_i + \gamma_n \xi_{in}}}}} \quad \forall m \in M$$

$$p_{i|m} = \frac{\sqrt[\theta_m]{\alpha_{im} e^{V_i + \gamma_m \xi_{im}}}}{\sum_{j \in C} \sqrt[\theta_m]{\alpha_{jm} e^{V_j + \gamma_m \xi_{jm}}}} \quad \forall i \in C, m \in M$$

$$p_i = \sum_{m \in M} p_m p_{i|m} \quad \forall i \in C$$

Relations with other logit models

This model combines properties of the path size logit model and the cross-nested logit model. Like the path size logit model for route choice, the probability p_i of a route $i \in C$ depends on the overlap with all other routes. Like the cross-nested logit model for mode choice, alternatives matching the same modal constraint have positively correlated utilities.

With a normally nested logit model with a path size logit model at the bottom level, the modal constraints cannot overlap, because the different nests are independent. Using the cross-nested logit model instead of the nested logit model overcomes this limitation. By setting $\alpha_{im} \in \{0,1\}$, one can show that the cross-nested path size logit model is indeed an extension of the nested path size logit model:

$$p_m = \frac{e^{\theta_m \ln \sum_{i \in I_m} e^{\frac{1}{\theta_m} V_i + \frac{\gamma_m}{\theta_m} \xi_{im}}}}{\sum_{n \in M} e^{\theta_n \ln \sum_{i \in I_m} e^{\frac{1}{\theta_n} V_i + \frac{\gamma_n}{\theta_n} \xi_{in}}}} = \frac{e^{\theta_m \ln \sum_{i \in I_m} e^{\tilde{V}_i + \tilde{\gamma}_m \xi_i}}}{\sum_{n \in M} e^{\theta_n \ln \sum_{i \in I_m} e^{\tilde{V}_i + \tilde{\gamma}_n \xi_i}}} \quad \forall m \in M$$

$$p_{i|m} = \frac{e^{\frac{1}{\theta_m} V_i + \frac{\gamma_m}{\theta_m} \xi_{im}}}{\sum_{j \in I_m} e^{\frac{1}{\theta_n} V_i + \frac{\gamma_n}{\theta_n} \xi_{in}}} = \frac{e^{\tilde{V}_i + \tilde{\gamma}_m \xi_i}}{\sum_{j \in I_m} e^{\tilde{V}_i + \tilde{\gamma}_n \xi_i}} \quad \forall i \in I_m, m \in M$$

$$p_i = p_m p_{i|m} \quad \forall i \in I_m, m \in M$$

However, note that the nested path size logit model is not a generalisation of the path size logit model: even if all nest coefficients would be set to one, the path size factors are still calculated at nest level. The following figure illustrates the position of the cross-nested path size logit model in the logit family:

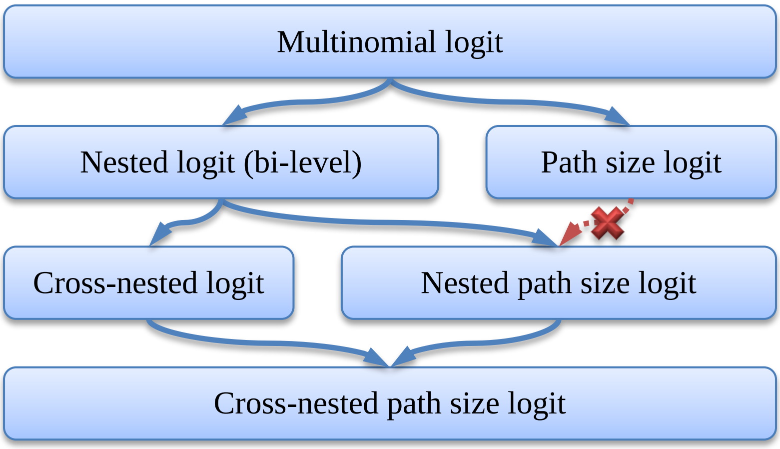


Figure 48: ancestors of the cross-nested path size logit model; each arrow indicates a generalisation/extension

Path size coefficients

For a set of fully overlapping alternatives in the same nest, the weighted path size factor $\sqrt[\theta_m]{e^{\gamma_m \xi_{im}}}$ equals:

$$\sqrt[\theta_m]{e^{\gamma_m \xi_{im}}} = \left(e^{\xi_{im}}\right)^{\gamma_m/\theta_m} = \left(\frac{\alpha_{im}}{\sum_{j\in C}\alpha_{jm}1^{\chi}}\right)^{\gamma_m/\theta_m} = \left(\frac{\alpha_{im}}{\sum_{j\in C}\alpha_{jm}}\right)^{\gamma_m/\theta_m}$$

If the alternatives have equal nest memberships, this reduces to:

$$\sqrt[\theta_m]{e^{\gamma_m \xi_{im}}} = \frac{1}{|C|^{\gamma_m/\theta_m}}$$

Using similar reasoning as in Appendix C.8, this leads to the condition $0 < \gamma_m / \theta_m \le 1$, or $0 < \gamma_m \le \theta_m$, for the path size coefficient γ_m.

E.2. Network GEV path size logit

Like the cross-nested logit model, the cross-nested path size logit model can be extended to allow multiple levels of nesting. This results in the network GEV path size logit model proposed by this section. Section 7.1 provides examples.

In order to prevent bias other than the bias created by the path size factors, the choice tree will be defined to be 'crash safe' (see Appendix C.6). This means that except for the bottom-level of the tree, all nest memberships are zero or one:

$$\alpha_{im} \in \{0,1\} \quad \forall i,m \in I \setminus C$$

These values will be assumed to be predefined (e.g. capturing correlations between similar modes), such that only at the bottom-level of the tree the nest memberships still need to be determined. Because in this situation, all nest memberships need to specified based on route characteristics, similar to Section E.1, rather than estimated, like Newman (2008) proposed, an alternate version of Newman's 'crash safe' normalisation will be given below in order to get a formula for the nest memberships that does not create bias.

Initial nest membership specification

First, define the indirect nest membership A_{im} as a result of nest memberships α_{in} of an alternative $i \in I$ in nest $m \in M$ or any direct or indirect children n of this nest m (so $A_{im} \geq \alpha_{im}$), according to the following formula (Newman, 2008):

$$A_{im} = \begin{cases} \alpha_{im} & \text{if } i \in C \\ \left(\sum_{j \in C_m} \sqrt[\theta_m]{A_{ij}} \right)^{\theta_m} & \text{otherwise} \end{cases}$$

Now, the following normalisation can be applied on the root nest r [82]:

$$A_{ir} = 1 \quad \forall i \in C$$
$$\theta_r = 1$$

Assuming that no single nest is allowed to have both alternatives as well as other nests as direct children[83], the nest memberships of alternatives α_{im} can now be specified by walking down the choice tree:

$$A_{in} = \left(\frac{Z_{in}}{\sum_{j \in C_n} Z_{jn}} \sqrt[\theta_m]{A_{im}} \right)^{\theta_m} \quad \forall i \in C, m \in I, n \in C_m$$
$$\alpha_{im} = A_{im} \quad \forall m \in I, i \in C_m$$

Here, Z_{im} it the generalised length of route $i \in C$ satisfying the modal constraint represented by nest $m \in I \setminus C$. This procedure is consistent with the definition of A_{im} and with the cross-nested path size logit model defined in Section E.1, which is a special case of the network GEV path size logit model defined here.

Path size factor inclusion

Again, the nest memberships need to be modified to include the path size factor. The same formula for the modified nest memberships $\hat{\alpha}_{im}$ as in Section E.1 can be used, using the original nest memberships from this section as input.

Also, for the path size coefficient γ_m, the same condition $0 < \gamma_m \leq \theta_m$ holds using identical reasoning.

---

[82] The mutually correlated error terms ε_i of the utilities are Gumbel-distributed with location parameter $\ln A_{ir}$ and scale parameter 1 (Newman, 2008). Because the location parameters are equal for all alternatives (to $\ln 1 = 0$), the model is unbiased.

[83] This is an additional restriction on the choice tree structure compared to the definition of a 'crash safe' choice tree of Newman (2008).

Appendix F. Access-/egress-based classification

This appendix describes an alternative classification method for models with structure 1B from Chapter 4, based on access to and egress from public transport.

The access-/egress-based classification is inspired by choice structure 3B from Chapter 4. However, the access/egress stop choice has been removed to allow route choice without nests such that utilities are defined at route level (e.g. allowing non-linear costs) and overlap may be handled with path size factors. Furthermore, the nests classifying public transport routes with identical access/egress mode pairs have been removed such that route overlap with routes of other mode permutations may be taken into account. The resulting choice tree is displayed below:

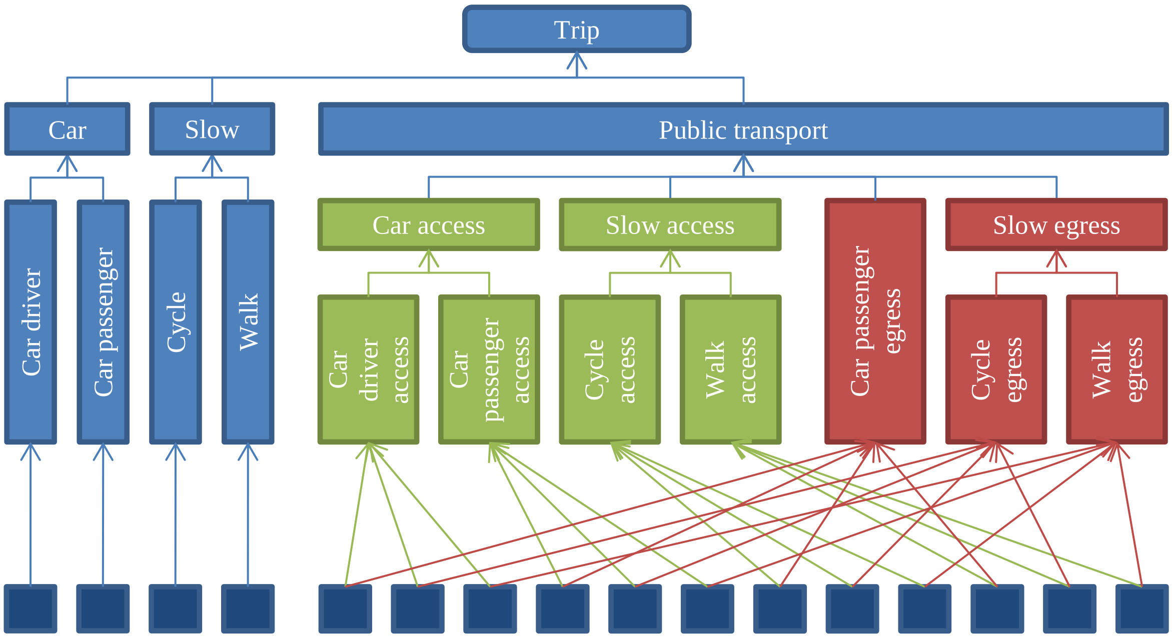


Figure 49: choice tree in case of access/egress classification; each dark box represents an arbitrary number of routes

From a behavioural perspective, the traveller chooses the main network, one private access or egress mode and then the full route.

If the path size factors and the nests for mode similarities are removed from this model the structure of the public transport nest is equal to the cross-nested logit model of Hoogendoorn-Lanser (2005) (also described by Hoogendoorn-Lanser et al. (2006)).

For the private modes, the measure of route overlap to be used to calculate path size factors, is travel time, while for the public transport nest, the number of legs is used.

F.1. Required parameters

In addition to valuation parameters for network attributes, this model requires the following parameters to be specified:

- path size coefficients for car and public transport;
- access/egress mode-specific constants for all private modes[84];
- access and egress nest memberships (the sum of these should be fixed at one);
- nest coefficients for all nests – the nesting structure itself also has to be specified;
- mode-specific constants for all private modes (one of these should be fixed at zero).

Note that within public transport, while PT modes do not have separate nests in the choice tree, mode-specific boarding penalties can be embedded in the network (see Section 7.2).

F.2. Comparison with mode-based classification

Section 7.1 already mentions disadvantages of the access-/egress-based classification compared to the adopted mode-based classification. The following table may help further in understanding the differences:

Classification	Route overlap	Modal overlap and mode similarities for private modes	Modal overlap and mode similarities for PT modes	Basis for nest memberships of routes
No classification (pure supernetwork approach)	✓	✗	✗	N/A
Access-/egress-based classification	ignored between PT and non-PT contexts	ignored across contexts	✗	context
Mode-based classification	✓	✓	✓	travel time and number of legs

Table 32: comparison of classifications

Here, a context is either the access, egress or only part of a route.

[84] According to Hoogendoorn-Lanser (2005), no significantly different values are found for the access modes and the egress modes.

Appendix G. Base matrices and route usage

Assuming that a pivot-point procedure corrects OD matrices of certain networks using base matrices, this appendix proposes an extension of the pivot-point procedure to handle multiple routes potentially traversing multiple networks (i.e. multi-modal routes). This is achieved by defining how to switch between route level and origin/destination/transfer node level. This results in a pivot-point procedure using base matrices for networks to correct route flows in a supernetwork.

A base matrix prescribes the amount of users of a network, entering at one node and leaving at another node[85]. These may be original origins or final destinations, but also transfers to other network parts. Each base matrix cell thus represents a network entry/exit node pair $t \in T$ with base patronage B_t. Note that, for this appendix, the base matrix $\mathbf{B}$ is actually a (column) vector in mathematical sense[86]; the same holds for the prediction matrix $\mathbf{P}$ and the synthetic matrices.

The extended pivot-point procedure has three steps. First, the route usage needs to be aggregated to network entry/exit node pair usage (Section G.1). Second, the traditional pivot-point procedure at network entry/exit node pair level is carried out (Section G.2). Third, the outcome needs to be disaggregated again to route usage (Section G.3). This final route usage information should then be more accurate than the original information.

G.1. Calculation of synthetic matrices

Each origin zone[87] $o \in O$ contains Q_o decision-makers with chances p_i of choosing alternative[88] $i \in C_o$, leading to each alternative having $q_i = Q_o p_i$ users. For each route, it is trivial to count the number of times α_{ti} the route passes a node pair $t \in T$ and uses the corresponding network in between ($\alpha_{ti} \in \{0,1\}$ except for some cyclic routes). If this represents the base situation, the cell B'_t of the synthetic base matrix $\mathbf{B}'$ may be calculated by the following formula:

$$B'_t = \sum_{o \in O} Q_o \sum_{i \in C_o} \alpha_{ti} p_i = \sum_{o \in O} \sum_{i \in C_o} \alpha_{ti} q_i \quad \forall t \in T$$

Of course, a similar formula holds for the synthetic prediction matrix $\mathbf{P}'$:

[85] The derivation in this appendix also holds if the concept of entry/exit node is more abstract, e.g. like aggregated origin/destination zones as used in the current pivot-point procedure of ANTONIN (Syndicat des transports d'Île-de-France, n.d.).

[86] Note that this allows multiple base matrices to be used for multiple parts of the supernetwork by simply appending their vectors.

[87] If there are multiple user classes or travel purposes, these can simply be considered as separate origin zones in this appendix.

[88] Alternatives generally are multi-modal routes to destinations at particular times of day and may also include not making a trip.

$$P'_t = \sum_{o \in \bar{O}} \sum_{i \in \bar{C}_o} \bar{\alpha}_{ti} \bar{q}_i \quad \forall t \in T$$

G.2. Application of traditional pivot-point procedure

Now that the synthetic matrices have been prepared at the same level of detail as the base matrix, the traditional pivot-point procedure is applied to calculate the prediction matrix at the same level. Generally, the following formula is used (Daly et al., 2005):

$$P_t = \frac{P'_t}{B'_t} B_t \quad \forall t \in T$$

Usually other formulas are applied in case of extreme changes between the base demand and the predicted demand (Daly et al., 2005; Significance, 2012a). However, the precise formulation does not influence the rest of this appendix and is therefore outside its scope.

G.3. Recalculation of route usage

Now, the route usage $\hat{q}_i$ needs to be recalculated based on the prediction matrix P_t, satisfying the following equation:

$$P_t = \sum_{o \in O} \sum_{i \in C_o} \alpha_{ti} \hat{q}_i \quad \forall t \in T$$

This problem can be rewritten to the following linear algebra problem, where C_t is the set of routes that use the node pair $t \in T$ and its corresponding network between these nodes (assuming routes are acyclic):

$$\mathbf{A}\hat{\mathbf{q}} = \mathbf{P}$$

$$\mathrm{A}_{ti} = \begin{cases} 1 & \text{if } i \in C_t \\ 0 & \text{otherwise} \end{cases} \quad \forall t \in T, i \in \bigcup_{o \in O} C_o$$

The route flows $\hat{\mathbf{q}}$ can be solved only if each used route traverses a non-empty set of node pairs $\tau \subset T$, i.e. traverses at least one node pair $t \in T$. Therefore, routes that do not traverse node pairs should be removed from the least-squares estimation, substituting the synthetic data $\mathbf{q}$ for them.

To handle potential overdetermined systems (i.e. due to a comprehensive base matrix), the least squares solution to the equation can be used (Wikimedia, 2013b):

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\hat{\mathbf{q}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}$$

$$\hat{\mathbf{q}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1} \mathbf{A}^{\mathrm{T}}\mathbf{P}$$

The least squares solution exists if and only if $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is invertible. Because $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ must have full rank and $\mathrm{rank}\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right) = \mathrm{rank}\left(\mathbf{A}\right)$ (Wikimedia, 2012c), the rank of $\mathbf{A}$ must be $\mathrm{rank}\left(\mathbf{A}\right) = \sum_{o \in O} |C_o|$. This is the case if and only if $\mathbf{A}$ is injective (Wikimedia, 2012c), i.e. if different route usage $\hat{\mathbf{q}}$ always leads to different node pair counts $\mathbf{P}$. Unfortunately, this is generally not the case, i.e. the problem in general may be underdetermined.

Incorporating a-priori route shares

The above derivation shows that an additional assumption is necessary. By assuming that the relative contributions of routes $\bar{q}_i / \sum_{j \in C_t} \bar{q}_j$ to a pair count P_t remain the same as in the synthetic prediction matrix, the system becomes generally solvable as shown below.

One starts with empty matrix $\tilde{\mathbf{A}}$ and empty vector $\tilde{\mathbf{P}}$. For each node pair $t \in T$, for each route using the node pairs $i \in C_t$, new rows $\underset{\leftrightarrow}{\mathbf{A}}_{\mathbf{ti}}$ and $\underset{\leftrightarrow}{P}_{ti}$ are inserted at the bottoms of $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{P}}$ respectively, according to the following formulas:

$$\underset{\leftrightarrow}{\mathbf{A}}_{\mathbf{ti}} = \left[\delta(i,1) \quad \cdots \quad \delta\left(i, \sum_{o \in O} |C_o| \right) \right] \quad \forall t \in T, i \in C_t$$

$$\delta(i,j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

$$\underset{\leftrightarrow}{P}_{ti} = \frac{\bar{q}_i}{\sum_{j \in C_t} \bar{q}_j} P_t \quad \forall t \in T, i \in C_t$$

Again, for routes that are not counted by one or more node pairs, the synthetic usage prediction $\bar{\mathbf{q}}$ must be used. For these routes, the following additional rows must be inserted:

$$\underset{\leftrightarrow}{\mathbf{A}}_{\mathbf{i}} = \left[\delta(i,1) \quad \cdots \quad \delta\left(i, \sum_{o \in O} |C_o| \right) \right]$$

$$\underset{\leftrightarrow}{P}_{i} = \bar{q}_i$$

The structure of the rows guarantee that $\operatorname{rank}(\tilde{\mathbf{A}}) = \sum_{o \in O} |C_o|$, implying that the following least-squares solution $\tilde{\mathbf{q}}$ always exists:

$$\tilde{\mathbf{A}}^{\mathrm{T}} \tilde{\mathbf{A}} \tilde{\mathbf{q}} = \tilde{\mathbf{A}}^{\mathrm{T}} \tilde{\mathbf{P}}$$

$$\tilde{\mathbf{q}} = \left(\tilde{\mathbf{A}}^{\mathrm{T}} \tilde{\mathbf{A}} \right)^{-1} \tilde{\mathbf{A}}^{\mathrm{T}} \tilde{\mathbf{P}}$$

Solution algorithm

Creating the matrix $\tilde{\mathbf{A}}$ in computer memory is impossible due to its extreme size ($|T| \times \sum_{o \in O} |C_o|$), so a smart algorithm is needed to solve this least-squares problem. By reordering the rows (i.e. sorting on routes instead of node pairs), $\tilde{\mathbf{A}}$ can be made block-diagonal:

$$\tilde{\mathbf{A}} = \left[\begin{array}{c|c|c} 1 & & \\ \vdots & & \\ 1 & & \\ \hline & \ddots & \\ \hline & & 1 \\ & & \vdots \\ & & 1 \end{array}\right] \qquad \tilde{\mathbf{P}} = \left[\begin{array}{c} P_{\underleftrightarrow{t}i} \\ \vdots \\ P_{\underleftrightarrow{t}i} \\ \hline \vdots \\ \hline P_{\underleftrightarrow{t}i} \\ \vdots \\ P_{\underleftrightarrow{t}i} \end{array}\right]$$

Using block matrix multiplication (Wikimedia, 2012a), left-multiplying both sides of the equation with $\tilde{\mathbf{A}}^{\mathrm{T}}$ results in:

$$\tilde{\mathbf{A}}^{\mathrm{T}}\tilde{\mathbf{A}} = \begin{bmatrix} \max\left(\left|\{t \in T : i \in C_t\}\right|, 1\right) & & \\ & \ddots & \\ & & \max\left(\left|\{t \in T : i \in C_t\}\right|, 1\right) \end{bmatrix}$$

$$\tilde{\mathbf{A}}^{\mathrm{T}}\tilde{\mathbf{P}} = \begin{bmatrix} \begin{cases} \sum_{t \in T : i \in C_t} P_{\underleftrightarrow{t}i} & \text{if } \exists t \in T : i \in C_t \\ \bar{q}_i & \text{otherwise} \end{cases} \\ \vdots \\ \begin{cases} \sum_{t \in T : i \in C_t} P_{\underleftrightarrow{t}i} & \text{if } \exists t \in T : i \in C_t \\ \bar{q}_i & \text{otherwise} \end{cases} \end{bmatrix}$$

This leads to the following solution of the problem:

$$\tilde{q}_i = \begin{cases} \dfrac{1}{\left|\{t \in T : i \in C_t\}\right|} \sum_{t \in T : i \in C_t} \dfrac{\bar{q}_i}{\sum_{j \in C_t} \bar{q}_j} P_t & \text{if } \exists t \in T : i \in C_t \\ \bar{q}_i & \text{otherwise} \end{cases}$$

$$= \begin{cases} \dfrac{\bar{q}_i}{\left|\{t \in T : i \in C_t\}\right|} \sum_{t \in T : i \in C_t} \dfrac{P_t}{\sum_{j \in C_t} \bar{q}_j} & \text{if } \exists t \in T : i \in C_t \\ \bar{q}_i & \text{otherwise} \end{cases}$$

Or, alternatively, written as a correction factor:

$$\frac{\tilde{q}_i}{\bar{q}_i} = \begin{cases} \dfrac{1}{\left|\{t \in T : i \in C_t\}\right|} \sum_{t \in T : i \in C_t} \dfrac{P_t}{\sum_{j \in C_t} \bar{q}_j} & \text{if } \exists t \in T : i \in C_t \\ 1 & \text{otherwise} \end{cases}$$

The solution can be calculated efficiently by first calculating and storing $P_t / \sum_{i \in C_t} \bar{q}_i$ for each node pair $t \in T$ and then looping over all routes $i \in \bigcup_{o \in O} C_o$ calculating and applying correction factors, which simply is averaging these values over the traversed node pairs $t \in T : i \in C_t$.

Appendix H. Observed mode choices in Île-de-France

This appendix lists some general statistics about mode usage in the case study described in Chapter 8.

H.1. Converted observations

The analysis in this appendix initially uses all successfully converted observations[89] (see Section 8.2), which is a superset of the observations used for model estimation (see Table 9 and Table 14). Note that trips with an origin and/or destination in Grande Couronne are not included in this initial analysis, since they are not used for model estimation.

The tables below list the statistics, where red cells indicate mode permutations that have been excluded from the case study model for theoretical reasons and orange cells indicate mode permutations that have been excluded because the low number of observations led to problems during model estimation:

Mode	Any leg		Multiple legs	
Walk[90]	1870	(67.3%)	1613	(58.1%)
Transilien	146	(5.3%)	6	(0.2%)
RER	582	(21.0%)	109	(3.9%)
Metro	1046	(37.7%)	456	(16.4%)
Tram	107	(3.9%)	3	(0.1%)
RATP Paris bus	169	(6.1%)	7	(0.3%)
RATP banlieue bus	396	(14.3%)	40	(1.4%)
Optile bus	34	(1.2%)	2	(0.1%)
Car driver	739	(26.6%)	0	(0.0%)
Motor driver	108	(3.9%)	0	(0.0%)
Bicycle	78	(2.8%)	1	(0.0%)
Car/motor passenger	62	(2.2%)	0	(0.0%)
All trips	2778	(100.0%)	1638	(59.0%)

Table 33: number of observed trips containing any leg or multiple legs of a particular mode

[89] About the totals, note that one short PT leg elimination warning changed a PT trip into a private mode trip. In the main text this trip remains counted as a PT trip but in the analysis in this attachment it shows up as a non-PT trip.

[90] Each PT-to-PT transfer is counted as a walk leg here, even if no walking is required.

Private mode	Direct leg		Access leg		Egress leg	
Walk	232	(20.4%)	1564	(95.5%)	1632	(99.6%)
Car driver	691	(60.6%)	47	(2.9%)	1	(0.1%)
Motor driver	107	(9.4%)	1	(0.1%)	0	(0.0%)
Bicycle	72	(6.3%)	4	(0.2%)	3	(0.2%)
Car/motor passenger	38	(3.3%)	22	(1.3%)	2	(0.1%)
Sum	1140	(100.0%)	1638	(100.0%)	1638	(100.0%)

Table 34: number of observed non-PT trips per mode and PT trips per access and egress mode

PT mode nest	Trips	
Train	697	(42.6%)
Metro/tram	752	(45.9%)
Bus	189	(11.5%)
Sum	1638	(100.0%)

Table 35: number of observed trips using train, trips using metro/tram but not train and trips using bus but not train/metro/tram

H.2. All original observations

For a more complete description of mode usage in Île-de-France, this appendix continues with the following tables analysing the same statistics for all inter-zonal morning-peak home-work trip observations, without the conversion step that turns survey legs into ANTONIN network legs, and including Grande Couronne[91]:

Mode	Any leg		Multiple legs	
Walk	3026	(52.5%)	2662	(46.2%)
Transilien	579	(10.0%)	37	(0.6%)
RER	1287	(22.3%)	256	(4.4%)
Metro	1511	(26.2%)	596	(10.3%)
Tram	146	(2.5%)	7	(0.1%)
RATP Paris bus	243	(4.2%)	12	(0.2%)
RATP banlieue bus	573	(9.9%)	73	(1.3%)
Optile bus	253	(4.4%)	27	(0.5%)
Car driver	2574	(44.6%)	0	(0.0%)
Motor driver	183	(3.2%)	0	(0.0%)
Bicycle	114	(2.0%)	6	(0.1%)
Car/motor passenger	207	(3.6%)	0	(0.0%)
All trips	5765	(100.0%)	2757	(47.8%)

Table 36: mode usage for all original observations

[91] Due to the incomplete PT network description in Grande Couronne, applying route conversion here would not be likely to lead to good results.

Private mode	Direct leg		Access leg		Egress leg	
Walk	274	(9.1%)	2435	(88.3%)	2729	(90.0%)
Car driver	2338	(77.7%)	230	(8.3%)	6	(0.2%)
Motor driver	182	(6.1%)	1	(0.0%)	0	(0.0%)
Bicycle	97	(3.2%)	15	(0.5%)	8	(0.3%)
Car/motor passenger	117	(3.9%)	76	(2.8%)	14	(0.5%)
Sum	3008	(100.0%)	2757	(100.0%)	2757	(100.0%)

Table 37: private mode usage for all original observations

PT mode nest	Trips	
Train	1715	(62.2%)
Metro/tram	773	(28.0%)
Bus	269	(9.8%)
Sum	2757	(100.0%)

Table 38: PT mode nest usage for all original observations

Appendix I. Traveller characteristics in Île-de-France

This appendix lists some general statistics about the availability of private vehicles and PT discounts for the case study in Chapter 8. The former influences the availability of alternatives, while the latter influences the utility of available alternatives as described in Appendix K. All statistics in this appendix apply to the data set that is used in Section 8.3.

I.1. Private vehicle availability

The following table lists the availability of vehicles as defined in Section 8.3:

Criterion	Travellers		Other travellers	
Car available	1678	(66.3%)	854	(33.7%)
Motor available	300	(11.8%)	2232	(88.2%)
Car and/or motor available	1736	(68.6%)	796	(31.4%)
Bicycle available	1257	(49.6%)	1275	(50.4%)
Car, motor and/or bicycle available	1993	(78.7%)	539	(21.3%)

Table 39: travellers with car, motor and bicycle availability

I.2. PT discount availability

The following table lists the availability of PT discounts as defined in Appendix K.2:

Cost category	No discount (i.e. pay 100%)		50% off (i.e. pay 50%)		Free travel (i.e. pay 0%)	
SNCF/RATP "regional"	1100	(43.4%)	16	(0.6%)	1416	(55.9%)
SNCF/RATP "urban"	1099	(43.4%)	16	(0.6%)	1417	(56.0%)
Optile	1121	(44.3%)	0	(0.0%)	1411	(55.7%)
Minimum discount (i.e. the worst PT legs)	1121	(44.3%)	0	(0.0%)	1411	(55.7%)
Maximum discount (i.e. the best PT legs)	1099	(43.4%)	16	(0.6%)	1417	(56.0%)

Table 40: travellers with PT discounts

For example, 55.7% of the travellers always travel for free (minimum discount is free travel), while 43.4% always need to pay the full price (maximum discount is no discount). Note that this table includes all travellers, not just the PT users.

Appendix J. Limitations of model estimation software

This appendix lists some of the limitations of choice model estimation software that negatively affected this research project. The software packages ALOGIT and Biogeme are considered here. Eventually, only the first of these has been used. Note that the choice models to be estimated in this research project are characterised by a large number (i.e. ten thousands) of route alternatives, of which only a few are available to a decision-maker: the routes for the specific OD pair of the traveller.

J.1. Control and data files

Unlike ALOGIT, a considerable problematic aspect of Biogeme is that all attribute values and alternative availability indicators must be specified in either a data file or in the *[Expressions]* section of the control file (Bierlaire, n.d.). For example, the travel time of a route cannot be entered directly into its utility function: instead, one must create a variable like *TravelTimeRoute2* which is then defined elsewhere, and define expressions like *Origin32Destination54* indicating whether a routes belong to the current OD pair. If car, motor and bicycle ownership is to be taken into account, this last expression for alternative availability must become something like *Origin32Destination54Vehicles101*, multiplying the number of availability expressions with eight.

To keep the size of the control file reasonable, route attributes can be placed in the data file. However, the data file is required to have a fixed number of columns. It is not feasible to create columns for all routes for all attributes. One can however re-use columns across OD pairs, provided that a maximum number of routes per OD pair is defined, to reduce the size of the data file.

A trick to reduce the number of columns in the data file is to define all possible values of integer attributes, like the numbers of legs per mode, in the *[Expressions]* section, and refer to these variables in the utility functions.

Further efficiency gains could be achieved by re-using alternatives for routes of different OD pairs, similar to re-using data file columns as described above, but this removes the possibility of nested logit models, as each alternative must have a fixed position in the choice tree (Bierlaire, n.d.).

Biogeme does not appear to work efficiently with a large number of alternatives like in revealed preference route choice applications. A simple multinomial logit model takes more than three hours to read the input data, even if the maximum number of routes is constraint to fifty (compare this with Figure 36).

ALOGIT also consumes a lot of time reading input data, but it is considerably faster than Biogeme and does not need a maximum number of routes. Where traveller-dependent route attributes like costs and path size would be specified in the data file for Biogeme, these can be included in the ALOGIT control file using simple *recode()* commands and expressions with help variables. Because the input files for ALOGIT are a lot simpler in route choice situations like this, ALOGIT is more flexible to use. For these reasons, ALOGIT has been chosen to be used for the case study in Chapter 8.

J.2. Network GEV

Network GEV models cannot be estimated by ALOGIT. However, with Biogeme also several specific network GEV problems are encountered, in addition to the problems mentioned above, which is why it was not applied in this research project, despite it being the theoretically ideal model in Chapter 7.

First, there are considerable technical problems. On Windows PCs, the programme crashes after a couple of iterations by running out of computer memory. This problem was successfully circumvented by switching to a Linux PC. However, a new problem now appears: the computation speed of the iterations is extremely low. A test run for this project has been aborted after almost four days, after about 150 iterations.

Second, the definition of the network GEV model in Biogeme is quite impractical when the nest memberships need to be fixed and the nest coefficients need to be estimated, as necessary according to Appendix E.2. Due to notational differences, in Biogeme, the nest coefficients need to be specified as $\sqrt[\theta_a]{\alpha_{ia}}$ (Daly & Bierlaire, 2006). However, according to the formulas in Appendix E.2, these values would depend on the nest coefficients of all second and higher level ancestor nests of the alternative. Hence the values $\sqrt[\theta_a]{\alpha_{ia}}$ should not be fixed in Biogeme, which is necessary since it represents which modes occur in which routes, which is of course not to be estimated.

This second problem disappears if only a single level of nesting is used, but then the model is identical to a cross-nested logit model. For this project this would mean that similarities between modes must be neglected, while Chapter 8 demonstrates that these exist among PT modes.

Note that if path size factors are also included to account for route overlap, these would have to be included in the fixed nest memberships as well. This would make it impossible to estimate the path size coefficient, but this could in theory be solved by simply repeating the estimation with several possible values.

J.3. Cross-nested logit

In theory, the cross-nested logit model could be used to approximate the correlation structure of a network GEV model by creating a single level of nests containing all sources of correlation. While ALOGIT cannot estimate cross-nested logit models either, this could at least circumvent the problems of Biogeme with network GEV models. For example, the following cross-nested logit model can include all correlations included in the ideal network GEV model of Section 7.1:

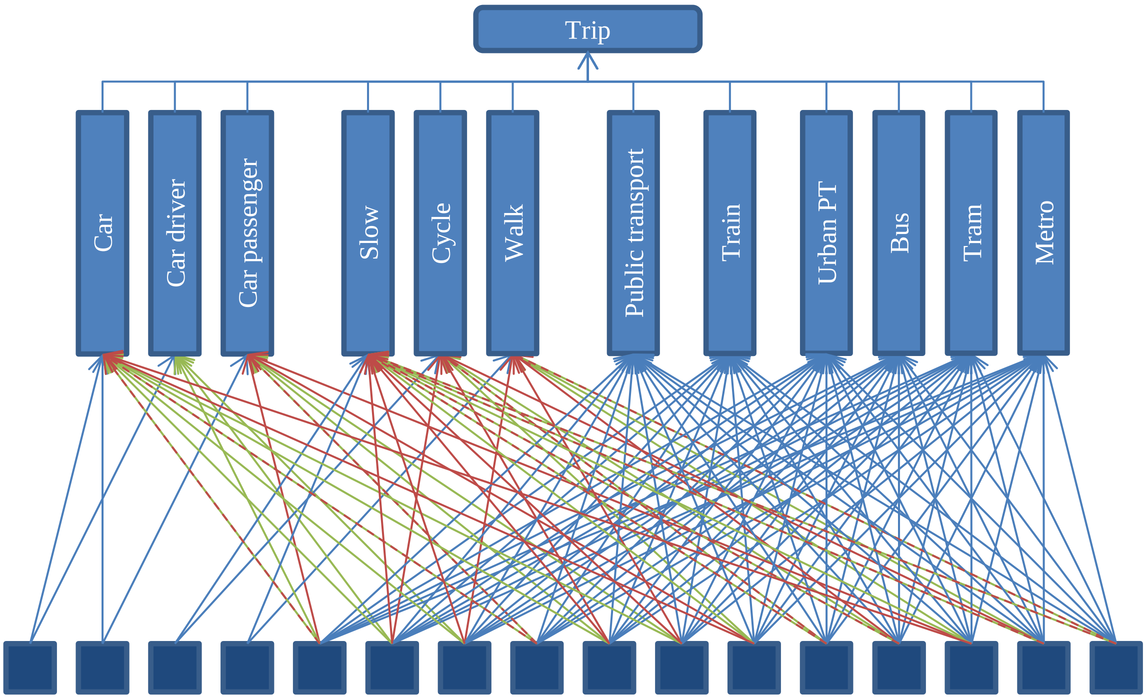


Figure 50: a cross-nested logit model approximating the ideal network GEV model

While this at first sight seems to avoid the need for a network GEV model, in reality it does not. The reason lies in the estimation of the model. In the cross-nested logit model, the nest memberships serve two purposes: indicate to which modes a route belongs and indicate to what extent each of the corresponding sources of correlation play a role. For the first purpose, the nest memberships need to be fixed, since it is a fundamental characteristic of the route. For the second purpose however, the nest memberships need to be estimated, since it is not known beforehand at which level of the original mode choice tree the most important correlations between routes occur.

Like with the network GEV model, path size factors could only be included in the fixed nest memberships, even if mode similarities are neglected by removing the nesting of modes from the network GEV model.

J.4. Nested path size logit

Because of the reasons above, a nested path size logit model is deemed more appropriate for the case study in Section 8.3. Since the alternatives then only belong to a single nest in a nested logit model, the path size factors can simply be included into the utility function such that its coefficient can be estimated.

This means that all three types of correlations between utilities – route overlap, modal overlap and mode similarities – can all be included in the same model, although there are some limitations on each of these: mode overlap and mode similarities have to be restricted to a main mode and route overlap can only be taken into account among routes sharing the same main mode.

Appendix K. Travel costs in Île-de-France

This appendix describes how the monetary costs of a generated or observed route are calculated in the case study of Chapter 8.

K.1. Normal costs

The case study model defines the monetary costs of a route as the sum of the costs of each leg. For a leg, the costs consist of two parts: boarding costs and additional distance-related costs. Here, the boarding costs may depend on the mode of the previous leg. For strategy segments in the PT network, the distance-related costs are a frequency-weighted average of the distance costs of the corresponding route segments.

For the private modes, costs are only calculated for the modes car driver and motor driver (not for car/motor passenger). Car driver has a distance cost of 0.092 €/km (price level 2001) (Willigers & Tuinenga, 2007). For simplicity, motor driver costs are assumed to be the same and parking costs are neglected.

The costs for PT legs are mostly based on the ticket t price, which equals 0.96 € (price level 2003) (Syndicat des transports d'Île-de-France, 2006)[92]. The costs of the PT legs are approximated as follows:

- 1 ticket t for metro, unless the previous leg is Transilien, RER or metro;
- 1 ticket t for bus and tram, plus 1 additional ticket t for each block of 2.5 km above 12.5 km for Optile bus legs;
- 1 ticket t for Transilien and RER if the boarding, alighting and intermediate nodes are inside Paris, unless the previous leg is Transilien, RER or metro;
- 0.09 €/km for Transilien and RER if the leg is partially or completely outside Paris, plus 1.48 € if the leg is partially in Paris or 0.68 € otherwise.

Whether a node is inside Paris or not, is determined by finding the closest zonal centroid based on crow fly distance. Since ANTONIN contains a department number for each zone, it can then be checked whether the corresponding department number equals 75.

K.2. Reduced prices

In order to take the influence of reduced PT fares into account during model estimation, the case study model divides travel costs of a particular route into the following four cost categories:

- non-PT costs (i.e. car driver and motor driver);

[92] A ticket t costs 1.20 € at price level 2010 (Syndicat des transports d'Île-de-France, 2010). However, (approximation) formulas for the other costs at a comparable price level are not available, so the 2003 value has been used.

- SNCF/RATP “regional” service costs (i.e. banlieue bus segments with either the boarding node, the alighting node or any intermediate node outside Paris, and Transilien);
- SNCF/RATP “urban” service costs (i.e. banlieue bus segments inside Paris, Paris bus, RER, metro and tram);
- Optile (bus) service costs.

Each possible PT reduction then is schematised such that for one or more of these cost categories, only a percentage of the full ticket price has to be paid. Obviously, none of these PT reduction schemes affect the non-PT cost category.

Note that this is an approximation. Due to possible interactions between costs of consecutive PT legs, the division of costs over the categories may be imperfect. Furthermore, PT passes that are valid for only a part of a single leg do not work in the model.

Reduced PT fares can be activated for an observation in the EGT by the age and PT pass ownership of the observed person that are specified in the survey data. Note that no PT pass ownership choice component is included in the simple case study model (nor an age choice component); these are assumed to be exogenous variables. The case study model can be extended to include PT pass ownership choices according to the principles of Section 7.3.

Note that ANTONIN2 uses a different approach: except for Carte Orange passes for home-work tours, it inserts the costs of a PT pass into the costs of a PT tour (Syndicat des transports d'Île-de-France, n.d.; Willigers & Tuinenga, 2007), using the average number of trips per pass to scale it to a single tour. On the other hand, its PT pass ownership choice model does not take pass costs into account, since no specific destination is known at that time. This design violates the proposal in Section 7.3.

Unfortunately, the EGT could not record multiple PT passes for the same person. Therefore, fore some people, the calculated travel costs may be too high. However, since most PT passes offer free travel for all modes, this is a minor problem.

The following table lists the assumed effects of each age and PT subscription in the EGT on the travel costs in each of the categories, as the *percentage of the cost that has to be paid*. Since the case study uses the home-work purpose only, it is generally assumed that the PT pass is valid between this origin and destination; see the footnotes for more information. The majority of the information is based on Syndicat des transports d’Île-de-France (2010).

Condition	SNCF/RATP "regional" price	SNCF/RATP "urban" price	Optile price
N/A	100%	100%	100%
Younger than 4 years	0%	0%	0%
Younger than 10 years[93]	100%	50%	50%
65 years or older[94]	50%	50%	100%
Carte Orange[95] **(weekly or monthly)**	0%	0%	0%
Carte Intégrale	0%	0%	0%
Carte Imagine'R (scholar or student)[96]	100%	100%	100%
Carte Solidarité (weekly or monthly)	0%	0%	0%
Carte Hebdo de Travail[97]	100%	100%	100%
Carte Optile Scolaire[98]	100%	100%	100%
Carte Améthyste[99]	0%	0%	100%
Carte Émeraude	100%	0%	100%
Carte Rubis	100%	100%	0%
Forfait Gratuité Transport	0%	0%	0%
Other (unknown) subscription	100%	100%	100%

Table 41: percentage of the PT price that has to be paid under various personal conditions that can be specified in the EGT survey; if a person satisfies multiple conditions, the cheapest option is chosen

[93] People under the age of 10 can get tickets t for half the price (Syndicat des transports d'Île-de-France, 2010).
[94] Everybody of 65 years and older without a Carte Améthyste Gratuité is assumed to have a Carte Améthyste ½ Tarif (Syndicat des transports d'Île-de-France, n.d.; Willigers & Tuinenga, 2007).
[95] The Carte Orange is bought for a specific pair of tariff zones (Syndicat des transports d'Île-de-France, 2010), so this should be adapted for other travel purposes. The Carte Orange is now superseded by a similar subscription on the Navigo pass.
[96] The Carte Imagine'R is bought for a specific pair of tariff zones (Syndicat des transports d'Île-de-France, 2010) and is rarely used for home-work trips (Syndicat des transports d'Île-de-France, n.d.; Willigers & Tuinenga, 2007). This should be adapted for other travel purposes.
[97] Detailed information about this pass and its usage conditions could not be found, but it seems to be a card for 12 days other than Sundays.
[98] The Abonnement Scolaire offers free transport, but is only valid for Optile buses for the home-school travel purpose (Courriers de l'Ile de France, n.d.).
[99] EGT respondents selecting Carte Améthyste are assumed to mean the Carte Améthyste Gratuite.

Appendix L. Stated preference mode choice estimation for Île-de-France

This appendix estimates a mode choice model based on stated preference research in Île-de-France described by Significance (2013). In this experiment, each question asked the respondent for a choice among with three PT modes, based on photos of the exterior and interior of the vehicles (without passengers), the in-vehicle travel time, the vehicle headway and the access time towards the PT mode. For some modes, both an old and a new vehicle design were presented.

Since no situations with transfers were used, the boarding penalties become simple mode-specific constants. The (new) Transilien has been chosen as the reference mode with a mode-specific constant of zero. Based on the stated preference data, the following multinomial logit model SP-1 can be estimated with the parameters resembling the models in Section 8.3 as much as possible; SP-2 is the same model with the old vehicle designs eliminated from the choice sets:

	SP-1		SP-2	
Log-likelihood	-9851.4		-5150.5	
ρ^2	0.241		0.257	
Observations	13295		8113	
Free coefficients	15		11	
Private mode time	-8.11 h^{-1}	(-41.0)	-8.49 h^{-1}	(-30.4)
PT in-vehicle time	-11.68 h^{-1}	(-11.8↲)	-12.32 h^{-1}	(-9.1↲)
Non-long waiting time	-22.37 h^{-1}	(-25.7↲)	-22.42 h^{-1}	(-18.5↲)
Transilien	0		0	
Old Transilien	-0.64	(-9.1)	N/A	
RER	0.09	(3.0)	0.11	(3.2)
Old RER	-0.35	(-6.3)	N/A	
Metro	-0.13	(-2.6)	-0.01	(-0.1)
Old metro	-0.22	(-3.2)	N/A	
Tram	0.21	(2.1)	0.47	(3.4)
Bus	-0.18	(-3.6)	-0.05	(-0.7)
Old bus	-0.10	(-1.4)	N/A	
Express bus	-0.37	(-6.1)	-0.37	(-5.2)
T Zen	0.21	(2.1)	0.50	(3.5)
1st alternative bias	0		0	
2nd alternative bias	-0.21	(-7.5)	-0.25	(-6.2)
3rd alternative bias	-0.29	(-10.3)	-0.29	(-7.3)

Table 42: multinomial logit model estimation for the stated preference research

The estimation includes bias dummies to correct for how the alternatives were presented to the respondent. Remarkably, the private mode time, i.e. the access time, appears to be valued lower than PT in-vehicle time, which does not match the findings from Section 8.3 even if the access mode would be car or motor. Also, the differences in mode-specific constants are low compared to the time coefficients.

Some noticeable differences are found in the mode-specific constants compared to the boarding penalties in Section 8.3. The bus appears to be much more competitive here, valued very similar to the metro, and the tram really stands out positively. The lack of a significant difference between metro and bus seems quite unrealistic compared to reality. Also, the design of the vehicle appears to be just as important as the mode, particularly for the Transilien.

These last results could be explained by the selection criteria for participants of the survey – the usually used modes are intentionally very equally distributed over all existing modes[100], except for the underrepresented express bus (Significance, 2013) which is also the worst mode in SP-2 – which limits the reliability of the estimated mode-specific constants. The stated preference questionnaire was also not limited to the home-work travel purpose.

Note that participants with more than 25 minutes current access time have been excluded from the estimation, because these participants were offered alternatives with more than 30 minutes access time. Unless the origin zone is sufficiently large, such large access times are not permitted by the route set generator in Section 8.1, and since non-linear effects were found in the valuation of access time, it has been decided to exclude these observations to get estimates that are more comparable with Section 8.3.

[100] The T Zen did not yet exist at the time of the questionnaire.