

Accounting for similarities between alternatives in discrete choice models based on high-resolution observations of transport behaviour

Report

Author(s):

Schüssler, Nadine

Publication date:

2010

Permanent link:

<https://doi.org/10.3929/ethz-b-000263353>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

DISS. ETH NO. 19093

**ACCOUNTING FOR SIMILARITIES BETWEEN
ALTERNATIVES IN DISCRETE CHOICE MODELS
BASED ON HIGH-RESOLUTION OBSERVATIONS
OF TRANSPORT BEHAVIOUR**

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

NADINE SCHÜSSLER

Diplom Wirtschaftsingenieurin

born 10.05.1979

citizen of
Germany

accepted on the recommendation of

Prof. Dr. Kay W. Axhausen, examiner
Prof. Dr. Michel Bierlaire, co-examiner
Dr. Stephane Hess, co-examiner

2010

Contents

Abstract	ix
Zusammenfassung	xi
1 Introduction	1
1.1 Motivation	1
1.2 Structure of the dissertation	4
1.3 Contributions	6
2 Recent Developments Regarding Similarities in Transport Modelling	7
2.1 The MNL model and its IIA property	8
2.2 Nesting structures	12
2.3 Multivariate error terms	16
2.4 Adjustment terms	18
2.4.1 C-Logit	21
2.4.2 Path Size Logit	22
2.4.3 The Independence of a Connection	25
2.4.4 Competing Destinations	26
2.4.5 Prospective Utility	28
2.4.6 The Sequence Alignment method	28
2.4.7 Spatial learning	29
2.4.8 Dependencies between decision-makers	30
2.4.9 Attribute derived similarity	31
2.4.10 The Concept of Dominance	32
2.4.11 Approaches beyond random utility	32
2.5 Conclusion and outlook	34
3 Processing GPS Raw Data Without Additional Information	37
3.1 Introduction and related work	38
3.2 Data cleaning and data smoothing	40
3.3 Trip and activity detection	44

3.4	Mode detection	45
3.5	Results of the GPS postprocessing	50
3.6	Computational performance	54
3.7	Conclusion and outlook	54
4	Map-Matching of GPS Points on High-Resolution Navigation Networks Using Multiple Hypothesis Technique	57
4.1	Introduction	58
4.2	Related work	59
4.3	Implementation	63
4.4	Evaluation of the resulting routes and the computational efficiency	72
4.5	Conclusion and outlook	76
5	Route Choice Sets for Very High-Resolution Data	79
5.1	Introduction and related work	80
5.2	Generating route sets	82
5.2.1	Stochastic Choice Set Generation (SCSG)	85
5.2.2	Route set generation with Breadth First Search on Link Elimination (basic BFS-LE algorithm)	85
5.2.3	Performance optimisation 1: BFS-LE(PO1)	88
5.2.4	Performance optimisation 2: BFS-LE(P02)	89
5.3	Computational performance	90
5.4	Structure of the derived route sets	94
5.5	Conclusion and outlook	97
6	Accounting for Route Overlap in Urban and Suburban Route Choice Decisions Derived from GPS Observations	99
6.1	Introduction and related work	100
6.2	Related work	101
6.3	Modelling approach	103
6.4	Derivation of the choice sets	106
6.5	Modelling results	111
6.6	Application of the modelling results	123
6.7	Conclusion and outlook	125
7	Accounting for Similarities in Destination Choice Modelling	127
7.1	Introduction	128
7.2	A general framework of similarity treatment in destination choice modelling	129
7.3	Conclusion and outlook	134

8	Summary and Outlook	137
8.1	Deriving observations for choice modelling from GPS raw data	137
8.2	Choice set generation	140
8.3	Choice modelling based on high-resolution observations	142
	Dank	145
	Bibliography	165

List of Figures

3.1	Raw vs. smoothed speeds and accelerations	43
3.2	Raw vs. smoothed points	43
3.3	Example of speed and acceleration distribution	48
3.4	Membership functions of the fuzzy variables	49
3.5	Trips and stages per day compared to the Swiss Microcensus 2005	52
3.6	Trip distance and duration distributions compared to the Swiss Microcensus 2005	53
3.7	Distribution of stage modes compared to the Swiss Microcensus 2005	54
3.8	Distance distributions per mode compared to Swiss Microcensus 2005	55
4.1	Derivation of the perpendicular distance	60
4.2	Derivation of the initial route candidates	64
4.3	Flow chart of the route development for a trip segment	66
4.4	The end of a link is reached if...	67
4.5	Cumulative distributions of data quality measures	69
4.6	Joining trip segments to complete trips	71
4.7	Distribution of computation time per GPS point depending on the maximum number of candidate paths	72
4.8	Relationship between the number of successfully matched trips and the average computation time	73
4.9	Distribution of the runs with the lowest number of candidate paths that produced the best results	75
4.10	Score per GPS point - averaged per route	75
5.1	Average computation time of the algorithms for 20 alternatives and 90 minutes time abort threshold	84
5.2	Basic BFS-LE tree	87
5.3	Examples of merging edges for producing a topologically equivalent network	90

5.4	Average computation time depending on route set size and time abort threshold	91
5.5	Computation time depending on the least cost path distance	92
5.6	Computation time depending on the number of non-pass nodes on the least cost path	93
5.7	Path size distributions for the resulting route sets	95
5.8	Number of motorway segments with less than 3 links	96
6.1	Cumulative distributions of choice set sizes resulting from the rule-based reduction	108
6.2	Travel time distributions for the different choice sets	110
6.3	Path Size distributions for the different choice sets	111
6.4	Travel time parameters for models estimated on unreduced choice sets	114
6.5	Ratios of the travel time parameters relative to extra-urban roads for unreduced choice sets with 100 alternatives	115
6.6	Utility correction for route overlap per road type depending on the degree of overlap	117
6.7	Travel time parameters for models estimated on randomly reduced choice sets	118
6.8	Travel time parameters for models estimated on similarity distribution-based reduced choice sets	119
6.9	Travel time parameters for models estimated on similarity-based reduced choice sets	120
6.10	Travel time parameters for models estimated on rule-based reduced choice sets	122

List of Tables

3.1	Fuzzy rules for mode detection	50
3.2	Overall statistics of the GPS study compared to the Swiss Microcensus 2005	51
6.1	Choice sets used in the model estimation	109
6.2	Models tested for each choice set type	112
6.3	Parameters of the basic model for unreduced SCSG choice sets with and without Sampling Correction	113
6.4	Adjusted rho squares for the unreduced choice sets . . .	115
6.5	Model results for the BoxCox(PSRT1) model and the most appropriate choice sets	123
6.6	Percentage of ODs for which the chosen route obtained the highest probability	124
6.7	Initial and final log-likelihoods of the forecasting models	125

Abstract

In transport research, discrete choice models are used to represent many aspects of travel behaviour. They model the choice of a decision-maker who is confronted with a set of discrete alternatives and chooses the alternative that maximises his or her utility. Discrete choice models give analysts and political decision-makers the opportunity to assess the impacts of infrastructure schemes and transportation policies in different scenarios with relatively low effort. Considering the scale of these projects and their significance for society, a realistic representation of travel behaviour is crucial.

To achieve more realistic travel behaviour models, researchers recently focussed – amongst others – on two major issues: a more accurate representation of the actual behaviour and the available alternatives by the means of high-resolution data and the treatment of similarities between alternatives. The high-resolution observations are obtained from GPS surveys which offer several advantages. Modellers benefit from more accurate and reliable information while the participants' burden is reduced substantially as long as the GPS recording is not combined with time-consuming questioning.

However, the increasing spatial resolution leads to new challenges, in particular regarding the treatment of similarities. Choice models estimated on high-resolution data are characterised by a large number of alternatives and complex similarity structures due to an increasing number of alternatives that differ only slightly from each other. Thus, state-of-the-art models have to be able to handle large choice sets and at the same time be flexible and able to accommodate various similarity structures.

In this dissertation, GPS observations for 32,000 person-days recorded by inhabitants of the Swiss cities of Zurich, Winterthur, and Geneva are processed with the aim to estimate car route choice models with an appropriate treatment of the similarities between the alternatives. The processing included filtering and cleaning the GPS records, deriving trips, activities and modes, and identifying the chosen routes by matching the car trips to a high-resolution network. Then, the choice sets were gen-

erated. The subsequently estimated route choice models tested different adjustment terms to account for route overlap. During the entire process, interesting lessons were learned about the new challenges introduced by the high level of spatial detail. These lessons are of vital importance for all analysts who want to model travel behaviour with this new generation of individual travel behaviour observations.

Zusammenfassung

In der Verkehrsplanung werden mit diskreten Entscheidungsmodellen vielfältige Aspekte des Verkehrsverhaltens abgebildet. Modelliert wird die Entscheidung einer Person, die aus einem Satz von Alternativen diejenige die mit dem grössten Nutzen auswählt. Diskrete Entscheidungsmodelle bieten Planern und politischen Entscheidungsträgern die Möglichkeit, die Auswirkungen von Infrastrukturveränderungen und verkehrspolitischen Massnahmen für verschiedene Szenarien mit relativ geringem Aufwand zu bestimmen und zu vergleichen. Angesichts der Grössenordnung dieser Projekte und ihrer Bedeutung für die Gesellschaft, ist eine realistische Abbildung des Verkehrsverhaltens in den Modellen äusserst wichtig.

Um eine realistischere Abbildung des Verkehrsverhaltens zu erreichen, hat sich die Forschung in letzter Zeit unter anderem auf zwei Themengebiete konzentriert: die genauere Abbildung des tatsächlichen Verhaltens und der Auswahlalternativen durch den Einsatz von hochaufgelösten Daten und die Behandlung von Ähnlichkeiten zwischen den Alternativen. Hochaufgelöste Beobachtungen des Verkehrsverhaltens stammen aus GPS Studien, die zahlreiche Vorteile bieten. Verkehrsplaner profitieren von genaueren und zuverlässigeren Informationen über das Verkehrsverhalten während der Aufwand für die Befragten gering ist, sofern die GPS Aufzeichnungen nicht von zeitaufwändigen Befragungen begleitet werden.

Andererseits birgt die grössere räumliche Auflösung auch neue Herausforderungen für die eigentliche Modellierung, besonders im Hinblick auf die Behandlung von Ähnlichkeiten zwischen den Alternativen. Entscheidungsmodelle, die auf hochaufgelösten Daten geschätzt werden, sind durch eine grosse Anzahl an Alternativen und komplexe Ähnlichkeitsstrukturen gekennzeichnet, da es viel mehr Alternativen gibt, die sich in den wesentlichen Merkmalen kaum von anderen Alternativen unterscheiden. Daher müssen moderne Entscheidungsmodelle grosse Alternativensätze verarbeiten können und gleichzeitig flexibel genug sein um vielfältige Ähnlichkeitsstrukturen abzubilden.

In Rahmen dieser Dissertation wurden GPS Beobachtungen ausgewertet, die 32'000 Personen-Tage umfassen und von Personen wohnhaft in Zürich, Winterthur und Genf stammen. Das Ziel war, aus diesen Daten Routenwahlmodelle für PW-Fahrten zu schätzen, die in geeigneter Weise die Ähnlichkeiten zwischen den Routen berücksichtigen. Dazu wurden die GPS Punkte filtert und geglättet, zwischen Fahrten und Aktivitäten unterschieden, für die Fahrten die verwendeten Verkehrsmittel identifiziert und die gewählten Routen innerhalb eines hochaufgelösten Netzwerks bestimmt. Anschliessend wurden die Alternativensätze generiert. Die daraus geschätzten Routenwahlmodelle testeten verschiedene Korrekturfaktoren für die Behandlung der Ähnlichkeiten. Im Verlaufe aller Arbeitsschritte wurden interessante Erkenntnisse über die neuen Herausforderungen gewonnen, die die Verwendung von hochaufgelösten Daten mit sich bringt. Diese Erkenntnisse werden eine entscheidende Hilfe für andere Verkehrsplaner sein, die diese noch relativ neue Möglichkeit der Beobachtung des individuellen Verkehrsverhaltens nutzen möchten.

Chapter 1

Introduction

1.1 Motivation

Discrete choice models are standard for modelling consumer behaviour. In transport research, they are applied to all aspects of travel behaviour, including, but not limited to, household activity scheduling, destination choice, route choice and mode choice. They model the choice of a decision-maker who is confronted with a set of discrete alternatives and chooses the one that maximises his or her utility. The utility depends on the decision-maker's individual preferences, the choice situation, the characteristics of the alternative and its similarities with the other available alternatives. Employing discrete choice models, analysts and political decision-makers have the opportunity to assess the impacts of planning alternatives and infrastructure schemes in different scenarios with relatively low effort. Therefore, discrete choice models are of special importance for the evaluation of transport policies, such as infrastructure investments or the introduction of tolls. Considering the scale of these projects and their significance for society, a realistic representation of travel behaviour is crucial.

To achieve more realistic travel behaviour models, researchers recently focussed – amongst others – on two major issues: a more accurate representation of actual behaviour and available alternatives by means of high-resolution data and treatment of similarities between alternatives. Since the two issues are inherently linked with each other, they are both addressed in this dissertation. The overall aim is to estimate discrete choice models on high-resolution data with an appropriate treatment of similarities between alternatives.

The increasing use of new survey technologies, in particular GPS, makes a completely new form of individual travel behaviour observations available to transport modellers. Instead of asking the survey participant

to recall and explain the comprehensive schedule of the previous day(s) – including his activities, trips, modes, routes and departure times – his or her daily routine and movements can be recorded second-by-second and with an accuracy of, at least under ideal conditions, 5-10 metres. Moreover, inherent problems of recollection-based surveys originating from people’s inability to judge distances and travel times and their tendency to forget short trips or activities, are avoided. Thus, the modeller benefits from more accurate and reliable information about times, geographic locations, and routes while the participants’ burden is reduced substantially as long as the GPS recording is not combined with extensive questioning to derive additional information, such as trip purposes and transport modes.

Yet, the new survey technologies also bring new challenges. The first challenge is the postprocessing of the data that is required to make the observations usable for model estimation. As stated above, one of the aims associated with the use of GPS in transport surveys is to reduce the burden for the participants. This can only be achieved if the GPS data collection does not involve time-consuming questioning to derive additional information. However, without additional information, i.e. modes and trip purposes, extensive post-processing is required.

The second challenge is the choice set generation procedure. To exploit the advantages of the high level of spatial detail in the observations, a high-resolution representation of the infrastructure, i.e. the road and public transport network, the activity locations, etc., is essential. This, however, substantially raises the requirements for the choice set generation, particularly in terms of computation time but also regarding the choice set composition. A lot of the recently published behaviourally advanced choice set generation procedures are basically not computable in reasonable time for the high level of spatial resolution. Thus, faster algorithms are required that produce credible choice sets.

The third challenge is the management of the large number of alternatives available in a high-resolution infrastructure representation. The high level of detail considerably amplifies the difference between the number of objectively available and subjectively considered alternatives because many alternatives in the large objective choice set deviate only slightly from each other. Thus, the choice set generation needs to explore many more alternatives to find all alternatives relevant for the choice observed. However, since the individually considered choice set is small and larger choice sets lead to longer model estimation times, the modeller should reduce the choice set size using appropriate reduction procedures that increase the behavioural realism of the choice set.

The treatment of similarities between alternatives is also affected by the high spatial resolution and the large number of alternatives. The large number of alternatives necessitates the use of simple model structures such as the Multinomial Logit (MNL) model. However, one major shortcoming of the MNL model is its inability to adequately account for the similarities between alternatives due its independence of irrelevant alternatives (IIA) property. This shortcoming is even more problematic since the overall similarity within the choice sets increases due to the high level of spatial detail. Suitable approaches to overcome the IIA property have to accommodate various and complex similarity structures and at the same time be computationally efficient. In principle, there are three groups of approaches: Allowing for non-zero off-diagonal elements in the variance-covariance matrix of the errors, employing factorial error components in addition to i.i.d. Gumbel errors, and adjustment terms in the systematic part of the utility function. The approach of using adjustment terms is especially appealing because of its simplicity and elegance. Instead of structuring the choice set a priori or taking the chance of misleading assumptions about correlations, only the type of similarities is specified. That way, the individual characteristics of the alternatives are accounted for and a value is assigned to the impact of specific interdependencies. Moreover, adjustment terms are calculated prior to model estimation and treated like other attributes in the deterministic part of the utility. Thus, they only lead to a small increase in computation time and are especially suitable for the modelling of choices with many alternatives. However, adjustment terms strongly depend on the choice situation and are usually not offhand transferable. The analyst has to determine for each choice situation which adjustment term is most appropriate. This applies especially if new types of data or new choice problems are involved.

In order to address the issues raised above, this dissertation describes the steps that were undertaken for the analysis of a large scale GPS data set. The data set contains about 32,000 person-days recorded by inhabitants of the Swiss cities of Zurich, Winterthur, and Geneva. 4,882 participants were asked to carry an on-person GPS logger for 6.65 days on average. No additional information, such as socio-demographic attributes, modes or trip purposes, was available for this study. The GPS observations were processed with the aim to estimate car route choice models. The processing included filtering and cleaning the GPS records, deriving trips, activities and modes, and the identification of the chosen routes by matching the car trips to a high-resolution network. Then, the choice sets were generated. The subsequently estimated route choice models tested

different adjustment terms to account for route overlap. During the entire process, interesting lessons were learned about the new challenges introduced by the high level of spatial detail. These lessons are of importance for all analysts who want to model travel behaviour with this new generation of individual travel behaviour observations.

One example for this new kind of models are destination choice models estimated using GPS observations and individual facilities instead of zones. Therefore, the thesis concludes with the conceptual considerations for these high-resolution destination choice models that base on the experiences gained with the route choice models. Similarities originating from the route and mode to reach the destination are accounted for along with several other aspects of similarities present in a destination choice context.

1.2 Structure of the dissertation

The individual parts of this dissertation have been presented at conferences or published in journals. Some of them were joined and updated into the chapters of this dissertation.

Chapter 2 summarises and evaluates different approaches to account for similarities between alternatives that have been proposed in the literature. After a detailed introduction of the MNL model and its IIA property, three different types of approaches are presented with a special focus on their appropriateness for high resolution applications.

Chapter 3 describes a GPS post-processing procedure needing no input other than the most basic GPS raw data: three-dimensional positions and timestamps. The procedure comprises several modules. First, the data is thoroughly cleaned and smoothed. Second, trips and activities are determined. Third, the trips are segmented into single-mode stages and the transport mode for each stage is identified. The procedure is applied to the GPS records collected in Zurich, Winterthur and Geneva. The results are compared to the Swiss Microcensus 2005 to demonstrate that the data obtained is ready for further applications, such as discrete choice model estimations.

Before the results of the GPS processing procedure described in Chapter 3 can be employed for route choice modelling, the chosen routes have to be extracted from a network. This is done with the map-matching algorithm presented in Chapter 4. The algorithm is designed to match large-scale GPS data sets on a high-resolution navigation network in acceptable computation time. The chapter describes the implementation of

the algorithm and evaluates its performance both in terms of accuracy and computational efficiency.

The high level of spatial detail offered by the GPS observations raises the requirements for the choice set generation procedure, particularly in terms of computation time but also regarding the choice set composition. Only algorithms based on shortest path search are computationally feasible for the high-resolution network used in this work. Thus, Chapter 5 presents a new route set generation algorithm based on shortest path search with link elimination. The proposed procedure ensures high diversity between the routes as well as computational feasibility for large-scale problems. To demonstrate the usability of the algorithm, its performance and the resulting route sets are compared to those of a Stochastic Choice Set Generation algorithm.

The route choice models based on the data derived in the Chapters 3, 4 and 5 are described in Chapter 6. Different choice set generation procedures as well as choice set sizes are evaluated regarding their effect on the choice set composition and the resulting route choice models. In addition, the impact of different route attributes is investigated. The focus, however, is put on the analysis of the adjustment terms, that account for route overlap. Different formulations are tested in order to evaluate which mechanisms are at work in car route choice in an urban or suburban context.

Chapter 7 presents a general framework for the treatment of similarities in a discrete choice model for destination choice of secondary activities. The framework combines several aspects of similarity derived from spatial location, the journey to and from the destination, trip chaining restrictions, and the attributes of the alternatives themselves. Moreover, it is applicable to a simultaneous route, mode and destination choice model.

The dissertation closes with a summary of the results and an outlook towards open research questions in Chapter 8.

1.3 Contributions

In this thesis, several contributions to the literature are made. The main contributions in order of appearance in the thesis are:

- The development of advanced GPS processing procedures for data cleaning, trip and activity detection, mode identification and map-matching in Chapters 3 and 4.
- The implementation of an efficient choice set generation procedure for routes in a high-resolution network in Chapter 5.
- The evaluation of different choice set reduction procedures that allow the management of the large number of alternatives in route choice sets while increasing the realism of the choice sets in Chapter 6.
- The use of a road type specific Path Size factor that delivered more stable results and more insights into behavioural mechanisms at work in car route choice in Chapter 6.

Chapter 2

Recent Developments Regarding Similarities in Transport Modelling

This chapter is based on the papers:

Schüssler, N. and K. W. Axhausen (2007) Recent developments regarding similarities in transport modelling, paper presented at the *7th Swiss Transport Research Conference*, Ascona, September 2007.

Axhausen, K. W. and N. Schüssler (2007) Similarity, visibility and similar concepts: A note, *Working Paper*, **458**, IVT, ETH Zurich, Zurich.

Discrete choice models have manifold applications in the representation of consumer behaviour. They model the choice of a decision-maker who is confronted with a set of discrete alternatives and has to choose one of them based on its utility. The utility depends on the decision-maker's individual preferences, the choice situation, the characteristics of the alternative and its similarities with the other available alternatives. The most prominent discrete choice model is the Multinomial Logit (MNL) model. One of its major shortcomings is its inability to adequately account for the unmodelled similarities between alternatives due to its independence of irrelevant alternatives (IIA) property. Overcoming the IIA property is therefore a major research issue for discrete choice modelling.

This chapter summarises and evaluates different approaches to overcome the IIA property that have been proposed in the literature. After a more detailed introduction of the MNL model and its IIA property, three different types of approaches are presented with a special focus on their appropriateness for large scale applications. The chapter concludes with a discussion of the different approaches and an outlook towards future research.

2.1 The MNL model and its IIA property

Discrete choice models are standard for modelling consumer behaviour. In transport research, they are applied to all aspects of travel behaviour, including, but not limited to household activity scheduling, destination choice, route choice and mode choice. Therefore, discrete choice models are of special importance for the evaluation of transport policies, such as infrastructure investments or the introduction of tolls.

They are based on the idea that an individual - the decision-maker - is confronted with a set of discrete alternatives, the choice set, and that he or she has to choose one of them. As a decision rule, it is assumed that the decision-maker seeks to maximise his or her personal utility. The utility of each alternative is characterised by its measurable attributes in interaction with unobserved sensitivities captured by the deterministic component V_{in} of the utility function. Beyond that, there are utility components that cannot be measured directly due to several reasons. First, there is heterogeneity of preferences across decision-makers. Second, the knowledge and the information processing abilities of decision-makers are limited. Third, there are further uncertainties regarding the choice process, including attributes which the analyst is not able or not resourced to measure. These elements are usually represented by the ran-

dom term ε_{in} of the utility function. Thus, the following utility function is postulated:

$$U_{in} = V_{in} + \varepsilon_{in} \quad (2.1)$$

with V_{in} being defined as $V_{in} = f(\beta, x_{in})$, where β is a vector of taste coefficients, and x_{in} a vector of the attributes of alternative i as faced by respondent n in the specific choice situation. In addition, socio-demographic attributes of respondent n can be included in the systematic part of the utility function.

The discrete choice model itself estimates for each alternative the probability of being chosen from a given choice set:

$$P(i|C_n) = P[U_{in} \geq U_{jn}, \forall j \in C_n] \quad (2.2)$$

The most commonly used discrete choice model is the Multinomial Logit Model (MNL) proposed by McFadden (1974). It is based on the assumption that the random terms, often called error terms, are identically and independently (i.i.d.) Gumbel distributed. The choice probability of each alternative i can then be derived as:

$$P(i|C_n) = \frac{e^{\mu V_{in}}}{\sum_j e^{\mu V_{jn}}} \quad (2.3)$$

Thereby, μ is a positive scale parameter and related to the variance of the Gumbel variable $Var(\varepsilon) = \frac{\pi^2}{6\mu^2}$. In the absence of a heterogeneous population μ is generally constrained to a value of 1.

The advantages of the MNL model are its flexibility in terms of its deterrence sensitivity, and the ease of the parameter estimation (Ben-Akiva and Lerman, 1985). On the other hand, the MNL model has several disadvantages, the most prominent being the *Independence of Irrelevant Alternatives* (IIA) property: The ratio of the choice probabilities of two alternatives does not depend on the existence or the characteristics of other choice alternatives.

$$\frac{P(i|C_n)}{P(k|C_n)} = \frac{\frac{e^{\mu V_{in}}}{\sum_j e^{\mu V_{jn}}}}{\frac{e^{\mu V_{kn}}}{\sum_j e^{\mu V_{jn}}}} = e^{\mu(V_{in}-V_{kn})} \quad (2.4)$$

An illustration for this problem is the well-known red bus/blue bus paradox (Debreu, 1960), which describes two mode choice situations. First, the decision-maker is facing two alternatives: taking the car or a red bus. It is assumed, that each alternative has a choice probability of

50%. In the second scenario, a blue bus with the same attributes relevant for the decision as the red bus is added to the choice set. Because the new alternative is just another option for using public transport, one would expect, that the share of the additional alternative comes completely at the expense of the red bus and the resulting choice probabilities should be: $P_{Car} = 50\%$, $P_{redBus} = 25\%$ and $P_{blueBus} = 25\%$ – ignoring for now the potential mode shift because of increased frequencies on the bus network. However, because of the IIA property, the MNL returns the same choice probability for each alternative ($P_{Car} = 33\frac{1}{3}\%$, $P_{redBus} = 33\frac{1}{3}\%$ and $P_{blueBus} = 33\frac{1}{3}\%$) to guarantee that the ratio between the probabilities for the car and the red bus stays equal to one.

Though the red bus and the blue bus obviously share a lot of characteristics and are therefore similar, the MNL model ignores this completely. The same applies for any other choice context, as demonstrated for example by Daganzo and Sheffi (1977) for private transport route choice, where the similarity between routes is derived from their overlap. Thus, one possible interpretation of the IIA property of the MNL model is its failure account for similarities between alternatives. Mathematically, similarities can be represented by correlations. Since the error terms in the MNL model are independently distributed, no unobserved correlations are included in the model as can be seen from the variance-covariance matrix for 5 alternatives depicted below, where because of its immanent symmetry, only the upper triangle of the matrix is shown. The matrix consists only of the variances of the alternatives' utilities. The covariances are assumed to be equal to 0.

$$\begin{pmatrix} \sigma^{in} & 0 & 0 & 0 & 0 \\ & \sigma^{in} & 0 & 0 & 0 \\ & & \sigma^{in} & 0 & 0 \\ & & & \sigma^{in} & 0 \\ & & & & \sigma^{in} \end{pmatrix}$$

This property leads to biased parameter estimates and behaviour forecasts, if the assumption of uncorrelated alternatives is not justified. Furthermore, the model can miss an important aspect of the actual choice behaviour. Solving this issue is still an ongoing research topic as is the question whether similarities between alternatives have positive or negative effects on their choice probabilities. One derivation postulates that similarities reduce the probability to be chosen. However, recent studies such as Weis (2006), van Eggermond (2007), Hoogendoorn-Lanser and Bovy (2007) or Frejinger and Bierlaire (2007) suggest that this assump-

tion does not hold for all choice contexts. A positive influence of similarities can for example be derived from the possibility to switch routes or connections while the passenger is traveling or to a strong preference for certain alternative attributes that are also present in the chosen alternative such as a specific departure time, travel time or fare.

There are different ways of describing similarities between alternatives. The most general way would be the Universal Logit (or Mother Logit) model introduced by McFadden (1975). In the Universal Logit model the utility of an alternative does not only depend on the attributes of this alternatives but also on the attributes of all other alternatives. However, since McFadden *et al.* (1977) expressed doubts about the model's consistency with utility maximisation – which could be dispelled by Hettinger (2008) – and the difficulty of defining an appropriate utility specification (Ben-Akiva, 1974a), the number of its applications are limited. Instead, three different approaches have been used that will be described in the following sections:

- imposing a nesting structure,
- explicitly modelling the correlation using multivariate error terms, and
- introducing adjustment terms in the deterministic part of the utility function.

The first group of models is presented in Section 2.2 and contains most Multivariate Extreme Value (MEV) models other than the MNL model. The alternatives are subdivided into groups, called nests. The nests describe the correlation structure in the way that alternatives belonging to the same nest are correlated with each other. The nests can be disjoint or alternatives can belong to more than one nest.

A prominent representative of the second group of models is the Error-Components Logit (ECL) formulation of the Mixed Multinomial Logit (MMNL) model. In addition to the i.i.d. Gumbel distributed error term a multivariate randomly distributed error term is introduced in the utility function that captures the correlation between alternatives. Section 2.3 discusses the MMNL and the Multinomial Probit and some of their applications to transportation problems.

The models of the third group aim to capture similarities by correcting the systematic component of the utility function. They rest upon the assumption that the utility of an alternative is influenced by its degree of similarity with other alternatives. Thus, they add a deterministic adjustment term that measures the similarity to the utility function. Adjustment term models can also be seen as specific cases of the Universal

Logit model where the utility is separated into a part that depends only on the attributes of this alternatives and a second part that depends on the attributes of other alternatives. The crucial issue for this approach is the appropriate choice of the adjustment term. Section 2.4 examines several adjustment terms that have been previously proposed in the literature.

2.2 Nesting structures

The family of Multivariate Extreme Value (MEV) models was introduced by McFadden (1978) under the name of Generalised Extreme Value (GEV) models. McFadden (1978) also demonstrated that the MNL model is a MEV model. The most popular model of this family, apart from the MNL, is the Nested Logit model, first presented by Ben-Akiva (1973, 1974b). Since the Nested Logit model is not able to capture all kinds of correlations, the Cross Nested Logit model was introduced by McFadden (1978). Bierlaire (2006) proved that the Cross-Nested Logit model is also a member of the MEV family of models, that the original formulation by Ben-Akiva and Bierlaire (1999) is equivalent to the formulations by Papola (2004) and Wen and Koppelman (2001) and that the formulations by Small (1987) and Vovsha (1997) are special cases of this formulation. The Cross-Nested Logit model was further generalised to the Network MEV model (Daly and Bierlaire, 2006). All these models are briefly discussed in this section.

The basic idea of the **Nested Logit (NL) model** is to divide all alternatives of a choice set into disjoint nests. Correlations may remain within the nests, but between the nests they are eliminated. Thus, the entire utility function for alternative i belonging to nest C_{mn} has to be reformulated. The systematic component incorporates the alternative specific effects V'_{in} as well as the impacts associated with nest m , $V_{C_{mn}}$:

$$U_{in} = V'_{in} + \varepsilon_{in} + V_{C_{mn}} + \varepsilon_{C_{mn}} \quad (2.5)$$

where ε_{in} and $\varepsilon_{C_{mn}}$ are independent. The distribution of the error-term ε_{in} remains i.i.d. Gumbel with a scale parameter $\mu_m \geq \mu$, while the error-terms $\varepsilon_{C_{mn}}$ jointly follow a distribution in a way that the random variable $\max_{j \in C_{mn}} U_{jn}$ is Gumbel distributed with scale parameter μ .

Each nest C_{mn} has a composite utility $V'_{C_{mn}}$, also called *expected maximum utility* or *Logsum*:

$$V'_{C_{mn}} = V_{C_{mn}} + \frac{1}{\mu_m} \ln \sum_{j \in C_{mn}} e^{\mu_m V_{jn}} \quad (2.6)$$

where $V_{C_{mn}}$ is the utility common to all alternatives in nest C_{mn} . Thus, the probability of choosing alternative i that is part of nest C_{mn} from the individual choice set C_n can be calculated as the product of the probability, that nest C_{mn} is chosen from the set of all nests and the probability that alternative i is chosen from the alternatives belonging to nest m :

$$P(i|C_n) = P(C_{mn}|C_n) \cdot P(i|C_{mn}) \quad (2.7)$$

where

$$P(C_{mn}|C_n) = \frac{e^{\mu V_{C_{mn}}'}}{\sum_{l=1}^M e^{\mu V_{C_{ln}}'}} \quad (2.8)$$

and

$$P(i|C_{mn}) = \frac{e^{\mu_m V_{in}}}{\sum_{j \in C_{mn}} e^{\mu_m V_{jn}}} \quad (2.9)$$

For $\frac{\mu}{\mu_m} = 1 \forall k$ the NL model collapses to the MNL model.

As such, correlation between the error-terms of alternatives nested together is introduced. However, the model does not capture potential correlation between nests. This can be illustrated by the variance-covariance matrix. The example here shows the covariances for 5 alternatives, of which alternatives 1 and 2 belong to the first nest, alternative 3 to a second one and alternatives 4 and 5 to the third one:

$$\begin{pmatrix} \sigma^{in} & \sigma_{12}^{in} & 0 & 0 & 0 \\ & \sigma^{in} & 0 & 0 & 0 \\ & & \sigma^{in} & 0 & 0 \\ & & & \sigma^{in} & \sigma_{45}^{in} \\ & & & & \sigma^{in} \end{pmatrix}$$

A solution to the problem of missed correlations between alternatives that do not belong to the same nest is the **Cross-Nested Logit (CNL) model**. In the CNL model, each alternative can belong to more than one nest. To represent the degree of membership to a nest, an allocation parameter $0 \leq \alpha_{im} \leq 1$ is introduced. Ben-Akiva and Bierlaire (1999) formulate the utility function of the CNL as:

$$U_{in} = V_{in}' + \varepsilon_{in} + V_{C_{nm}} + \varepsilon_{C_{nm}} + \ln \alpha_{im} \quad (2.10)$$

where ε_{in} and $\varepsilon_{C_{nm}}$ are defined as in Equation 2.5. The choice probability

for an alternative i has then to be calculated over all nests m it partially belongs to:

$$P(i|C_n) = \sum_{m=1}^M P(C_{mn}|C_n) \cdot P(i|C_{mn}) \quad (2.11)$$

The Logsum for each nest C_{mn} of the CNL model is calculated by

$$V'_{C_{mn}} = V_{C_{mn}} + \ln \sum_{j \in C_{mn}} \alpha_{jm} e^{V_{jn}} \quad (2.12)$$

and $P(i|C_{mn})$ is reformulated to

$$P(i|C_{mn}) = \frac{\alpha_{im} e^{V_{in}}}{\sum_{j \in C_{mn}} \alpha_{jm} e^{V_{jn}}} \quad (2.13)$$

It is important to note, that any functional relationship can be defined for the allocation parameter α_{im} depending on the choice context, though often, simple point-estimates are used. Abbe *et al.* (2007) derived a normalisation for α_{im} which is:

$$\sum_m \alpha_{im}^{\frac{\mu}{m}} = c, \forall j \in C_n \quad (2.14)$$

where c is a constant that does not depend on i and is usually assumed to equal 1.

Thus, the CNL model is theoretically able to depict all kinds of correlation structures by allowing the error-terms of alternatives that are somehow nested together to be correlated. The following example shows the matrix for a five alternatives belonging to three nests example with the following membership structure: alternative 1 belongs to nest a, alternative 2 to nests a and b, alternative 3 to nest b, alternative 4 to nests b and c and alternative 5 to nest c.

$$\begin{pmatrix} \sigma^{in} & \sigma_{12}^{in} & 0 & 0 & 0 \\ & \sigma^{in} & \sigma_{23}^{in} & \sigma_{24}^{in} & 0 \\ & & \sigma^{in} & \sigma_{34}^{in} & 0 \\ & & & \sigma^{in} & \sigma_{45}^{in} \\ & & & & \sigma^{in} \end{pmatrix}$$

Due to its flexibility, the CNL model has been applied to various small and medium sized transport problems. Departure time choice mod-

els employing the CNL model were estimated by Ramming (2002) and Small (1987) whose Ordered Generalised Extreme Value model is mathematically identical to the CNL model. Vovsha and Bekhor (1998) presented the Link Nested Logit model, a CNL model for route choice, that was also applied by Ramming (2002) and Bekhor *et al.* (2007). Mode choice CNL models were estimated by Bierlaire *et al.* (2001) and combined mode and destination choice models by Bekhor (2007) and Hess *et al.* (2005) whereas Hess and Polak (2006) and Hess (2005) achieved a joint treatment of correlations between airport, airline and access-mode through the use of a CNL model. However, in large scale applications the CNL model soon leads to highly complex structures, which make it difficult to specify and computationally hard to estimate.

Another relaxation of the NL model's restrictions is the **Paired Combinatorial Logit (PCL) model** by Chu (1989); Koppelman and Wen (2000). The PCL model allows for correlation between every two alternatives by creating a nest for each pair of alternatives and estimating a dissimilarity parameter for each nest. The nests themselves are independent from each other. Thus, the similarity relationship between each pair of alternatives is independent of the similarity relationship between other pairs of alternatives. The PCL model is also a member of the MEV family as shown by Wen and Koppelman (2001). Koppelman and Wen (2000) and Wen and Koppelman (2001) applied the PCL model to mode choice data. Bhat and Guo (2004) and Bekhor and Prashker (2008) employed it in a residential location choice and destination choice context, respectively. In the models by Bhat and Guo (2004) and Bekhor and Prashker (2008), only spatially adjacent zones are nested together. Bhat and Guo (2004) combine the PCL model with a MMNL model to account for unobserved taste heterogeneity while Bekhor and Prashker (2008) add an additional nesting structure to allow for non-spatial attribute driven correlation between alternatives.

To derive a more general formulation for Nested Logit models, Wen and Koppelman (2001) proposed the **Generalised Nested Logit (GNL) model**. It summarises the NL, CNL, PCL models and other NL derivatives through normalisation of the CNL model structure. The GNL model fractionally assigns each alternative to a nest and different Logsums can be calculated for each nest. The normalisation of Wen and Koppelman (2001) was formally proved by Abbe *et al.* (2007). Furthermore, the latter gave guidance to derive a CNL model from any arbitrary variance-covariance structure.

Another approach of generalisation was proposed by Daly and Bierlaire (2006) with the **Network MEV model**. The authors showed that

any correlation structure represented by a network with certain properties can be modelled with a NGEV model and, furthermore, that every such model is a MEV model. The properties for the network are straightforward: The network is not allowed to include circuits, it has to have one root node without predecessors, the alternatives have to be represented by leafs without successors, and each node in the network has to be part of a continuous path between the root and at least one alternative. This model formulation is especially appealing because of its intuitive way of capturing even complex correlation structures. It eases the formulation of a complex model by its recursive definition.

2.3 Multivariate error terms

The most exhaustive way to account for correlation between alternatives is to use a **Multinomial Probit model**. In a Probit model, as discussed for example by Daganzo (1979), multivariate Normal distributed error terms replace the i.i.d. Gumbel distributed ones of the MNL resulting in the most general variance-covariance structure:

$$\begin{pmatrix} \sigma^{in} & & & & \\ & \sigma_{12}^{in} & & & \\ & \sigma^{in} & & & \\ & & \sigma_{23}^{in} & & \\ & & \sigma^{in} & & \\ & & & \sigma_{34}^{in} & \\ & & & \sigma^{in} & \\ & & & & \sigma_{45}^{in} \\ & & & & \sigma^{in} \end{pmatrix}$$

Thus, any variance-covariance structure can be specified and all kinds of correlation structures between the alternatives of the choice set can be depicted. Probit model have, for example, been applied by Yai *et al.* (1997) or Daganzo and Sheffi (1977) to car route choice problems. The respective authors specified models, in which the covariances of the route utilities are proportional to the length of link overlaps. However, the formulation of the Probit model is complex and its choice probabilities do not have a closed form. It can not be solved analytically and requires simulation for estimation as well as application. Thus, it is only applicable if the number of parameters and alternatives is small.

The **Mixed Multinomial Logit (MMNL) or Logit Kernel (LK) model** (Ben-Akiva and Bolduc, 1996; McFadden and Train, 2000) was introduced with the aim to combine the advantages of a Probit model with those of a Logit model. The i.i.d. Gumbel distributed error term ε_{in} is maintained and a multivariate randomly distributed error term η_{in} with

zero mean is added to the utility function:

$$U_{in} = V_{in} + \eta_{in} + \varepsilon_{in} \quad (2.15)$$

There is no a priori constraint on the distribution of η_{in} though most applications assume a multivariate Normal distribution.

Two conceptually different approaches of the MMNL model have been developed: the *Error-Components Logit (ECL)* model and the *Random-Coefficients Logit (RCL)* model. In the ECL model, correlation between alternatives is accounted for by letting them share the same error component η_{in} . In the RCL model unobserved taste heterogeneity across individuals is accommodated by specifying some entries of the vector β in the equation $V_{in} = f(\beta, x_{in})$ to be random variables for which the error term η_{in} represents the deviation from the mean.

In the present context, the ECL approach is of special interest. In theory, an ECL model can approximate any correlation structure, including heteroscedastic ones, arbitrarily closely. As such, the model can also replicate the variance-covariance matrix of the general Probit model. Like the Probit model, the ECL model has the disadvantage that simulation is required in estimation and application. In addition, imposing the right identification restrictions, so that a unique solution can be obtained from the infinite set of optimal solutions of the unconstrained model, is a difficult and time-consuming task and an often overlooked one as argued by Walker (2002). If the MMNL model additionally allows for random taste variation (e.g. in an RCL framework), these problems go much further because before identification issues can be solved, the appropriate distribution function for the random parameters has to be determined. This altogether makes the model difficult to be applied in large-scale forecasting systems. For further discussion see Walker (2002), Walker *et al.* (forthcoming) and Ben-Akiva and Bolduc (1996).

Several studies have successfully applied ECL models to transport related choice problems. Bekhor *et al.* (2002) and Ramming (2002) estimated ECL route choice models. In their models, the utility function of each route contains one error term per link in the route. If a link is used by more than one route, the according error term is shared among the routes. However, due to the computational complexity of the model, it was only applied to a medium scale scenario.

The only MMNL model that can reasonably be estimated for large scale route choice scenarios is the **Subnetwork model** by Frejinger and Bierlaire (2007). Correlation is not established using link overlap but so-called subnetwork components. A subnetwork component is a continu-

ous subsection of the network that is easily identifiable and behaviourally relevant. Subnetwork components can either be derived from the network hierarchy or from route descriptions in personal interviews. Routes using the same subnetwork component are assumed to be correlated even if they are not physically overlapping. For them a joined error component is estimated. The authors tested different model specifications with subnetworks based on a data set containing 2978 observations private transport route choice in the city of Borlaenge, Sweden. The choice set size ranges from 2 to 43 alternative paths with a majority of choice sets containing less than 15 paths.

The **Mixed Spatially Correlated Logit (MSCL) model** suggested by Guo and Bhat (2007) combines an MMNL model with a Paired Generalised Nested Logit model. It has been developed for residential location choice. The PGNL structure accounts for correlations between adjacent spatial units whereas the mixing Normal distribution captures unobserved taste heterogeneity. The approach was used to model the residential location choice of 236 households within Dallas County for zones of different sizes and characteristics. The authors found that combining a closed-form correlation structure with an open-form account for taste variations resulted in a good model-fit and was computationally efficient compared to a pure MMNL model.

A Mixed MEV model was also applied by Hess *et al.* (2005) to model data from a Stated Preference long-distance mode choice survey in Switzerland. The aim of the survey was to estimate the hypothetical demand for a new transport system in Switzerland, the so-called Swiss Metro (Abay, 1999; Bierlaire *et al.*, 2001). Nested Logit and Cross Nested Logit models are combined with Normal distributed random terms to capture taste heterogeneity. Their results emphasise the significant risk of confounding effects of taste heterogeneity and correlation since these two phenomena are not necessarily clearly distinguishable. This is especially pointed out by the difficulties the authors experienced with the estimation of the Mixed CNL model which were partly due to the model's complexity and partly to the data that was not rich enough for the Mixed CNL model.

2.4 Adjustment terms

Adjustment terms are based on the assumption that one can measure the similarity of an alternative with other alternatives and correct its utility accordingly. Therefore, a similarity attribute A_{in} is calculated for each

alternative i and choice set C_n . Then, as shown in Equation 2.16, this attribute is added to the deterministic part of the utility function as an adjustment term. The error terms remain i.i.d. Gumbel distributed and the variance-covariance matrix is the same as for the MNL model:

$$U_{in} = V_{in} + f(A_{in}) + \varepsilon_{in} \quad (2.16)$$

where A_{in} is the adjustment term accounting for the similarity between alternative i and all other alternatives $j \neq i$ and $f()$ is the transformation of A_{in} , ensuring $f(A_{in}) \leq 0$ for the possible range of A_{in} .

The formulation follows the generally made assumption (Fotheringham, 1983; Cascetta *et al.*, 1996; Ben-Akiva and Bierlaire, 1999) that the similarity of an alternative with other, competing alternatives decreases its utility and, thus, its probability to be chosen. A_{in} has no behavioural interpretation. It is just a computational convenience obviating the need for the more complex formulations, such as Nested, Cross Nested, tree Logit or Probit models. It will ensure, that the share of the alternative will be reduced.

Still, if the underlying similarity is perceived by the decision maker, then the question arises what form the similarity has and what behavioural reaction can be expected. Four generic mechanisms come to mind:

- **Loosing visibility** as being undistinguishable from the other alternatives, which translates into a lower chance of inclusion into the choice set and therefore to be chosen.
- **Joint risks** through common elements, in the spatial domain joint bottlenecks (links, transfer locations, or facilities), or chances of shared shocks (strikes, accidents, timetable dependence, etc.). Again, for the usually risk averse decision maker this will lead to a reduction of the attractiveness of the alternative.
- Becoming a **super-alternative**, as similar alternatives provide joint opportunities and therefore a higher chance of achieving one's goals. An example is a shopping center providing redundancy by offering in each price band a range of essentially similar goods distinguished by their branding only. The super-alternative effect should increase the choice of any of the constituent alternatives.
- Gaining **super-visibility** by being the best of a class of essentially similar alternatives, even if only by a marginal amount. This should increase the chances of being chosen both through more

frequent inclusion in the choice set, but also through the prestige of consuming/employing the best alternative conferring status on the chooser.

As the relative effect of the four different mechanisms in any choice situation is unclear a priori it is impossible to impose the rigid constraints formulated in Equation 2.16. Weis (2006), Hoogendoorn-Lanser *et al.* (2005a) and Hoogendoorn-Lanser and Bovy (2007) document such seemingly counter-intuitive results. The modeller should therefore test:

$$U_{in} = V_{in} + \alpha f(A_{in}) + \varepsilon_{in} \quad (2.17)$$

where α is the parameter for the adjustment term A_{in} that has to be estimated.

One can also note, that the four mechanisms affect visibility, i.e. the probability of inclusion in the choice set, and the attribute values, i.e. utility level, either positively or negatively. Both paths to derive the specific adjustment term can be found in the literature.

The range of attributes which can be involved in the mechanisms outlined above is large in the scheduling choices travel behaviour research is interested in. For the choice of activity participation, sequence, timing, duration, location and group size and the associated travel choices of staging, mode, connection and route, including parking and access point, the following sources of similarities come to mind:

- Temporal vicinity of the time-space paths
- Spatial overlap of routes
- Overlap in type and price of the facilities used
- Similar brand, comfort and other attributes
- Price, tolls, fares on certain parts of the transport network
- Density of supply in the “vicinity” of the alternative
- Type and quality of the supply in the “vicinity” of the alternative

This makes clear, that scheduling choices will always need corrections in areas where the IIA assumption is not justified due, for example, to spatial and temporal constraints that produce similarities. The basic question is: How to measure these similarities in the particular choice context? The literature knows several answers to this question for various choice problems, some of which are discussed in the following.

2.4.1 C-Logit

One of the earliest adjustment terms proposed in the literature is the C-Logit model developed by Cascetta *et al.* (1996) for car route choice models. The so-called Commonality Factor CF_{in} indicates the percentage of route length that route i shares with other routes by comparing the total length of route i with the length of the overlapping links. Cascetta *et al.* (1996) proposed three different formulations for CF_{in} :

$$CF_{in} = \sum_{j \in C_n} \left(\frac{L_{ij}}{\sqrt{L_i \cdot L_j}} \right)^\gamma \quad (2.18)$$

$$CF_{in} = \sum_{a \in \Gamma_i} \frac{l_a}{L_i} N_{an} \quad (2.19)$$

$$CF_{in} = \sum_{a \in \Gamma_i} \frac{l_a}{L_i} \ln N_{an} \quad (2.20)$$

where γ is a coefficient to be estimated, L_i and L_j are the lengths of routes i and j , L_{ij} is the length of links shared by routes i and j , Γ_i the set of links of route i , l_a the length of link a , and N_{an} number of routes using link a .

Cascetta *et al.* (1996) tested all three specifications of CF_{in} concluding that, compared to the Probit model, all three specifications yield similar choice probabilities, though the C-Logit model consistently assigns slightly lower choice probabilities to independent alternatives. Regarding the question which of the three specifications should be used they state that Equations 2.18 and 2.19 deliver better results for alternatives that have similar generalised costs whereas Equation 2.20 works better for alternatives with varying overall generalised costs. Cascetta *et al.* (1996) themselves applied Equation 2.19 whereas Equation 2.18 is used by Ramming (2002) and Cascetta *et al.* (2002). Prato and Bekhor (2007a) and Ramming (2002) worked with an additional specification of CF_{in} and compared it to other approaches such as the Path Size Logit model:

$$CF_{in} = \left[1 + \sum_{j \in C_n, i \neq j} \left(\frac{L_{ij}}{\sqrt{L_i \cdot L_j}} \right) \left(\frac{L_i - L_{ij}}{L_j - L_{ij}} \right) \right] \quad (2.21)$$

Vrtic (2003) used the formulation in Equation 2.18 and combined it with a Nested Logit model to the **Nested C-Logit (NCL) model**. The NCL model was developed for a simultaneous route and mode choice model. The nesting structure accounts for unobserved similarities between private transport and public transport alternatives respectively. Deterministic correlations within the nests are captured by the Commonality Factor.

2.4.2 Path Size Logit

The Path Size (PS) Logit model of Ben-Akiva and Bierlaire (1999) was also developed for route choice problems. The length of each route is corrected by the so-called Path Size PS_{in} . Only a distinct route, i.e. a route with no overlaps with other routes, can get the maximum path size of 1. Path Sizes different from 1 are calculated based on the length of the links within the route i and the length of the routes that share a link with it relative to the length of the shortest route using the link. Ben-Akiva and Bierlaire (1999) propose two different formulation for PS_{in} , the first one being

$$PS_{in} = \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \frac{1}{\sum_{j \in C_n} \delta_{aj}} \quad (2.22)$$

where Γ_i is the set of all links of route i , l_a is the length of link a , and L_i the length of route i . δ_{aj} equals 1 if link a is on route i and 0 otherwise. The second formulation additionally accounts for the relative ratio between the length of the shortest route $L_{C_n}^*$ in C_n using link a and the length of each route j using link a .

$$PS_{in} = \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \frac{1}{\sum_{j \in C_n} \delta_{aj} \frac{L_{C_n}^*}{L_j}} \quad (2.23)$$

Ramming (2002) states that this model formulation has a major shortcoming: Its second term is not affected by the length of other than the shortest route if a link is used by more than one route. Thus, he derived a **General Path Size (GPS) factor**. He reformulates the second part of Ben-Akiva and Bierlaire's Path Size factor to account for the contribution of the individual links. The basic idea is to give each link the size 1 and to allocate this size among the routes using that link. The size of a route is then calculated as the sum of its link sizes weighted according to the length of the route compared to the length of other routes using that link.

The influence of this weighting is given by the size allocation parameter γ .

$$GPS_{in} = \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \frac{1}{\sum_{j \in C_n} \left(\frac{L_i}{L_j} \right)^\gamma \delta_{aj}} \quad (2.24)$$

Especially for large γ , Ramming (2002) achieved the best model results for $\gamma \rightarrow \infty$, this formulation assigns the size of a shared link primarily to the shortest route using that link.

However, Hoogendoorn-Lanser *et al.* (2005a), who applied the PS factor and the GPS factor to multi-modal route choice, as well as Frejinger and Bierlaire (2007) found the interpretation of this approach difficult. In contrast to the original PS factor that can be interpreted as an approximation of the variance-covariance matrix, the GPS factor introduces asymmetry into the model by explicitly favouring the shortest route. In addition, the empirical analysis of the GPS factor showed that it captures part of the explanatory power of the variables related to the units the GPS factor is measured in. Furthermore, Hoogendoorn-Lanser *et al.* (2005a) expressed the need to have a close look at the value of γ before applying it and to explicitly estimate β_{PS} , which had been fixed to 1 by Ramming (2002) and Ben-Akiva and Bierlaire (1999).

Recently, Bovy *et al.* (2008) proposed another approach to derive the Path Size factor, resulting in the so-called Path Size Correction (PSC) term. Bovy *et al.* (2008) argue that their PSC term has a clear theoretical derivation from the notion of aggregate alternatives and a nested logic structure based on common links between route alternatives and offers a more intuitive interpretation of the role of correlation due to spatial overlap. Moreover, it outperforms the classic PS factor in the application to an synthetic network and two empirical datasets. The PSC term is defined as follows:

$$PSC_{in} = - \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \ln \sum_{j \in C_n} \delta_{aj} \quad (2.25)$$

Thus, analogously to the classic PS_{in} , the PSC_{in} depends on the number of shared links, the lengths of these common links and the numbers of distinct routes using each common link and a completely independent route obtains a PSC_{in} of 0. However, different from the PS_{in} , there is no upper bound for the absolute value of PSC_{in} . The utility reduction increases with an increasing number of common links in a route, increasing lengths of the common links, and increasing number of other routes

from the choice set that overlap with one or more links of the route.

A complete different variant of the Path Size factor was developed by Hoogendoorn-Lanser and Bovy (2007) for route choice modelling in multi-modal networks. Their **trip part specific Path Size factor** enables the modeller to account for varying valuations of overlap between different parts of the trip, namely the the train part of the trip and the access and egress part to and from the train station. In contrast to the classic Path Size formulation, the trip part specific Path Size is based on *stages* and not links. A stage is defined as a part of a trip between two nodes with a single mode, vehicle, or service type without transferring. This accounts for the particular perception of overlap in multi-modal trips. As Hoogendoorn-Lanser *et al.* (2005b) and Hoogendoorn-Lanser *et al.* (2005a) have demonstrated, overlap in multi-modal trips is not perceived in terms of time or distance but in terms of the number of shared stages. Thus, the length of a trip or sub-route in the trip part specific Path Size does not equal the distance travelled but the number of stages in the trip or sub-route. Hoogendoorn-Lanser and Bovy (2007) tested three different formulations of the trip part specific Path Size factor:

$$pS_{ixn} = \frac{1}{L_{ix}} \sum_{a \in \Gamma_{ix}} \frac{l_a}{N_{na}} \quad (2.26)$$

$$Ps_{ixn} = \frac{1}{L_{ix}} \sum_{a \in \Gamma_{ix}} \frac{l_a}{n_{nax}} \quad (2.27)$$

$$PS_{ixn} = \frac{1}{L_i} \sum_{a \in \Gamma_{ix}} \frac{l_a}{N_{na}} \quad (2.28)$$

where L_i is the length of the of full route i , L_{ix} is the length of sub-route s_{ix} of trip part x and route i , Γ_{ix} is the set of all stages in trip part x of route i , l_a is the length of stage a , N_{na} is the number of unique *full routes* using stage a , and n_{nax} is the number of unique sub-routes for trip part x using stage a .

After estimating models with all three trip part specific Path Size formulations, Hoogendoorn-Lanser and Bovy (2007) found that overlap in the home end and the activity end of a trip was valued negatively, whereas overlap in the train part was valued positively, implying that redundancy in the train part increases the attractiveness of a route. This might be caused by a hierarchical choice process in multi-modal trips. Travellers

probably first decide about the trip part with the lowest frequency, which is usually the train-part. Thus, this part becomes a super-alternative and all alternatives containing this specific train connection obtain a higher choice probability. Since this result was stable over all models and the performance of all models was improved significantly compared to classic Path Size models, they concluded that they captured a new behavioural effect. The formulation that worked best in their experiments was the one in Equation 2.28. In closing, they suggest to test this kind of Path Size also for single mode trips using, for example, a road type specific Path Size.

2.4.3 The Independence of a Connection

The Independence of a Connection (IND) factor, an adjustment term specifically designed for public transport connection choice, was presented by Friedrich *et al.* (2001). It is based on the assumption that the similarity in the spatial dimension is less decisive for public transport connection choice and mainly restricted to shared transfer points. Instead, temporal similarity aspects are highly relevant as is the similarity with respect to price. The Independence of a Connection factor is defined as the reciprocal of the sum of similarities of alternative i with all other alternatives j in the choice set:

$$IND_{in} = \frac{1}{\sum_j f_i(j)} \quad (2.29)$$

The similarity itself is measured considering the time gap between corresponding departure (DEP) and arrival (ARR) times and the differences in perceived journey times (PJT) and prices.

$$f_i(j) = \left(1 - \frac{x_i(j)}{s_x}\right) \cdot \left(1 - \gamma \cdot \min \left\{1, \frac{s_z \cdot |y_i(j)| + s_y \cdot |z_i(j)|}{s_y \cdot s_z}\right\}\right) \quad (2.30)$$

where $x_i(j) = \frac{|DEP(j)-DEP(i)|+|ARR(j)-ARR(i)|}{2}$, $y_i(j) = PJT(j) - PJT(i)$, and $z_i(j) = price(j) - price(i)$. s_x , s_y and s_z set the range of influence of $x_i(j)$, $y_i(j)$ and $z_i(j)$ respectively. s_y and s_z depend on the sign of $y_i(j)$ and $z_i(j)$ in order to model the asymmetry between connections. If there is a difference in terms of perceived journey time, the superior connection will exert a stronger influence on the inferior one, the same applies for the price.

Weis (2006) and van Eggermond (2007) applied the IND factor. Weis (2006) examined ground-based public transport and had no price data available. He found, that similarities between alternatives have a negative influence on their choice probabilities, which complies with the assumptions by Friedrich *et al.* (2001). This finding was confirmed by van Eggermond (2007), who analysed air transport choice and had price data available. However, for a different formulation of the utility function, he found a positive influence of similarities on the choice probabilities. As a conclusion he pointed out that it is very important for the analyst to consider, what effects are actually captured by which part of the utility function as the IND factor interacts with other decision-attributes.

2.4.4 Competing Destinations

For alternatives in a destination or location choice context, the most obvious aspect of similarity is spatial proximity. One of the first researchers who accounted for spatial proximity in destination choice modelling was Fotheringham (1983) with his Competing Destinations (CD) model. The underlying assumption is a two step decision process: The decision-maker first chooses a broader region and second an alternative within that region. Therefore, the utility of each alternative is affected by the number of alternatives in the same region. With an increasing number of alternatives within the same region the probability for each alternative to be recognised, and, thus, to be chosen, decreases. Two formulations for the adjustment term have been presented so far: Fotheringham (1983) suggested to sum up the reciprocal distances d_{ij} from a store i to all $I - 1$ other stores j in the universal choice set and to weight the distances relative to the utility of the corresponding store.

$$CD_{in} = \left(\frac{1}{I - 1} \sum_{j, j \neq i} \frac{V_{jn}}{d_{ij}} \right)^{\theta} \quad (2.31)$$

A second formulation has been proposed by Borgers and Timmermans (1987). It simply takes into account the average distance from store i to all other stores.

$$CD_{in} = \left(\frac{1}{I - 1} \sum_{j, j \neq i} d_{ij} \right)^{\theta} \quad (2.32)$$

In both formulations, θ is a parameter to be estimated. Following Fotheringham (1983)'s assumption that the main similarity mechanism at work is the one of loosing visibility, the adjustment term enters the utility function in transformed and without an additional parameter that would allow for a positive impact of the similarity.

According to Bernardin *et al.* (2009), the main weakness of the CD model is that it only measures the net effect of spatial proximity between destination choice alternatives while, in reality, there are two opposing forces at work: *spatial competition* and *agglomeration effects*. Spatial competition derives from similar alternatives located nearby. Because of them the alternative loses visibility and its choice probability decreases. Agglomeration effects, on the other hand, arise from alternatives nearby that offer complementary goods or activities. Due to the presence of complementary alternatives the alternative becomes a super-alternative with increased choice probability. Therefore, Bernardin *et al.* (2009) state that two adjustment terms have to be included in the utility function, each with its own parameter to be estimated. The two adjustment terms represent the accessibility of substitutes AS_{in} or complements AC_{in} from alternative i . In case the alternatives are facilities, AS_{in} and AC_{in} can be derived straightforwardly. For the more complex case of zones, Bernardin *et al.* (2009) suggest the following formulations :

$$AC_{in} = \ln \sum_j D_{ij} F_j e^{\alpha_C c_{ij}} \quad (2.33)$$

$$AS_{in} = \ln \sum_j (2 - D_{ij}) F_j e^{\alpha_S c_{ij}} \quad (2.34)$$

where F_j is the total number of facilities in zone j , c_{ij} is the travel cost to get from zone i to zone j and α_C and α_S are parameters to be estimated. Even though α_C and α_S lead to an substantial increase in estimation time, Bernardin *et al.* (2009) prefer this accessibility formulation since the impact of the travel cost on the accessibility is determined by the data and not defined a priori by the analyst. The most crucial variable, however, is the degree of dissimilarity D_{ij} between alternatives i and j . It is calculated by:

$$D_{ij} = 1 - \sum_c w_c \frac{F_{ic} F_{jc}}{F_i F_j} \quad (2.35)$$

where F_{ic} is the number of facilities of category c in zone i , F_i is the total

number of facilities in zone i and w_c is a weighting function representing the number of times a facility of category c is visited in the study area. Bernardin *et al.* (2009) show that their model (slightly) outperforms the MNL and CD model and reacts behaviourally reasonable when a new alternative n is added to the choice set.

2.4.5 Prospective Utility

Closely related to the aspect of spatial proximity is the similarity derived from trip chaining. As Bernardin *et al.* (2009) discussed, the choice probability of an alternative can increase if it is surrounded by complementary activities because the decision-maker is then able to execute several activities in one trip. Based on the same assumption, Kitamura (1984) developed a destination choice model that explicitly accounts for trip chaining effects by introducing an adjustment term called Prospective Utility (PU) which recursively integrates the utility that can be derived from subsequent activities into the utility of the destination under consideration:

$$PU_{in} = \sum_j q_{jn}(U_{jn}\theta d_{ij}) \quad (2.36)$$

where q_{jn} is the probability that decision-maker n carries out an activity at location j after his activity at location i , U_{jn} is the utility of said activity at location j , d_{ij} is the spatial distance between i and j , and θ is the disutility parameter for d_{ij} . PU_{in} can be interpreted as a measure of perceived accessibility of zone i . It can be modified to account for different trip purposes and due to its recursiveness also for longer trip chains.

2.4.6 The Sequence Alignment method

A different approach that can be used to account for trip chaining effects is the Sequence Alignment Method (SAM). The SAM allows to determine the degree of similarity between alternatives, that compromise multiple characteristics, which themselves can have a multivariate description. A transport example are trip and activity chains. Trip and activity chains consist of multiple activities, that each have several properties such as type, location, timing and duration as well as trips with characteristics such as mode, duration and distance. The SAM originates from molecular biology and was introduced into the field of travel behaviour research by Wilson (1998). Joh *et al.* (2001) and Wilson (2008)

provided important enhancements. The SAM employs the concept of biological distance rather than geometrical distance. Biological distance is defined as the smallest number of attribute changes (mutations) that is necessary to equalise two sequences. Thus, the SAM allows to measure (dis)similarity regarding different attributes as well as the sequential order of activities. It is very flexible and allows to determine a simple measure of similarity even for alternatives with different types of attributes and complex interdependencies. Feil *et al.* (2009) used the SAM to integrate a similarity measure in the utility function of MATSim MATSim-T (2008). It considers the similarity between activity schedules in terms of activity chain sequence, mode choice and activity location choice.

2.4.7 Spatial learning

Another similarity aspect related to spatial proximity is spatial learning. Since a decision-maker can only make a choice between alternatives he knows, it is an ongoing research issue how a decision-maker gets to know new destination choice alternatives. He or she might have been told about it by a friend or colleague, found it on the internet or, most importantly regarding the treatment of similarities, discovered it while travelling. One way to depict the spatial learning process and the resulting spatial knowledge of a decision-maker is to draw his or her **Mental Map**. Several studies (e.g. Chorus and Timmermans, 2009; Hannes *et al.*, 2008; Mondschein *et al.*, 2008) have been conducted to explore the relationships between mental maps on the one side and socio-demographic characteristics and travel behaviour on the other side. Their main findings are:

- Living in an area for a longer time improves the quality of a mental map.
- The regular use of modes that require active navigation (e.g. walk, bike, car) improves the quality of a mental map significantly.
- Most activities have a standard mode-destination setting that is only changed if necessary.
- When deviating from their default option, decision-makers choose from a repertoire of standard alternative destinations that are often spatially linked to the default option.

This illustrates how much a destination choice set, and the choice probability of an alternative, depend on the places we have already visited. However, this issue has obtained little attention in the literature

so far (with few exceptions, e.g. Schlich, 2001; Schönfelder and Samaga, 2003). One reason might be the lack of longitudinal survey data in the past. With the advent of more and more longitudinal diary and GPS studies, this obstacle should be overcome soon. As a first approach to account for the effect of repeatedly visited destinations, Sivakumar and Bhat (2007) introduced a variable in the utility function which indicates whether the destination, in their case the zone, was chosen in the previous time period or not.

2.4.8 Dependencies between decision-makers

The focus of the work presented by Mohammadian *et al.* (2005) was on the introduction of spatial dependencies between decision-makers instead of alternatives. They developed a Mixed Logit model for new housing projects that accounts for taste heterogeneity and correlations between alternatives. In addition, a **Spatial Dependency Parameter** SP_{in} is introduced into the systematic part of the utility function to account for spatial correlation between the decision-makers.

$$SP_{in} = \sum_{s=1}^S \rho_{nsi} y_{si} \quad (2.37)$$

where $s = 1, \dots, S$ are the decision-makers whose choice influences the choice of decision-maker n while evaluating alternative i and y_{in} is equal to 1 if decision-maker s has chosen alternative i and 0 otherwise. The parameter ρ stands for a matrix of coefficients representing the influence that the choice of one decision-maker has on another decision-maker while he chooses alternative i . Mohammadian *et al.* (2005) define

$$\rho_{nsi} = \lambda e^{-\frac{D_{ns}}{\gamma}} \quad (2.38)$$

where D_{ns} is the spatial distance separating decision-makers n and s , and λ and γ are parameters to be estimated.

Dugundji and Walker (2005) also focussed on the explicit account for dependencies between decision-makers. They employed a field effect variable in the deterministic part of the utility function. This variable represents the dependency of a decision-maker's choice on the overall share of connected decision-makers that choose the alternative in question. However, instead of capturing only spatial dependencies, they suggest a network structure to represent any kind of dependencies between decision-makers, especially social ones. In the dependency network,

each decision-makers is symbolised by a node and his or her dependencies by links. Other correlations between alternatives in this model have been captured by a CNL model.

2.4.9 Attribute derived similarity

Other adjustment terms, that have originally been proposed in the context of destination choice but that are equally applicable in any choice between alternatives, are directly derived from the attributes of the alternatives. Amongst them are the similarity measures of: Batsell (1982)

$$A_{in} = \exp\left(\frac{1}{I-1} \sum_j \sum_k \theta |x_{ik} - x_{jk}|\right), \quad (2.39)$$

Borgers and Timmermans (1987)

$$A_{in} = \prod_k \left[\frac{1}{I-1} \sum_j |x_{ik} - x_{jk}| \right]^{\theta/K}, \quad (2.40)$$

and Meyer and Eagle (1982)

$$A_{in} = \left[\frac{1}{I-1} \sum_{j \neq i} 0.5|r_{ij} - 1| \right]^{\theta}. \quad (2.41)$$

where x_{ik} is the value of attribute k for alternative i , r_{ij} is the observed Pearson product moment correlation between alternatives i and j across their attributes, θ a parameter to be estimated and I the total number of alternatives.

All these measures could be interpreted as mean field effects. In the spatial choice context it would be obvious to weight their contributions by an appropriate spatial weight matrix. The wider econometric literature on mean field effects has generally not been integrated into the discrete choice literature, with exceptions such as Dugundji and Walker (2005). Employing an exponential distance (generalised cost) weighting of the other alternatives would result in the inclusion of accessibility terms, as for example proposed by Kitamura (1984) to capture the super-alternative mechanism discussed above. See Train (1986) on the necessary logarithmic transformation of such a super-alternative term, which he derived in the context of the aggregation of individual facilities into

zones.

2.4.10 The Concept of Dominance

A more sophisticated measure to account for different attributes of the alternatives is the Concept of Dominance introduced by Cascetta and Papola (2009). It is based on the assumption that an alternative is less likely to be taken into account if it is dominated by other alternatives. Alternative j dominates alternative i , if the utility of all attributes of j is higher than (or equal to) the utility of the equivalent attributes of i . Following that concept, a **Dominance Factor** DF_{in} is calculated for each alternative i , indicating the number of alternatives dominating i . Cascetta and Papola (2009) and Cascetta *et al.* (2007) use two specifications for the Dominance Factor. In the first specification, they assume that alternative j dominates alternative i if the utility of j is greater than that of i for all attributes of i and j while at the same time the generalised costs c_{oj} of getting from origin o to destination j are smaller than those of getting from o to i . The second dominance measure originates from the concept of intervening opportunities (Stouffer, 1940). In order to dominate i , destination j has to fulfil the conditions formulated above and, in addition, has to be situated on the route from origin o to destination i . In this case, j is an intervening opportunity on the route to i .

Martínez *et al.* (2008) used the Dominance Factor as cut-off value for their Constrained Multinomial Logit Model which models the probability of an alternative to be included in the individual choice set of the decision-maker with a binomial logit function. They detected that the dominance affects the utility in a non-linear way. Accordingly, further research is advised concerning the way the Dominance Factor should enter the utility function. Non-linear transformations should be tested as well as minimum or maximum thresholds.

2.4.11 Approaches beyond random utility

In the light of this discussion, some authors have raised the question, whether Random Utility Maximisation is the right framework to model choice behaviour. Completely new approaches have been proposed that do not have the same shortcomings due to their underlying assumptions. Example for such new frameworks are *Prospect Theory* and *Random Regret Minimisation*. Prospect Theory was derived from the observation that decision makers frame their choice relative to the status quo and that gains and losses relative to this status quo are treated differently (Kahne-

man and Tversky, 1979). Regret minimisation builds upon the assumption that decision-makers do not seek to maximise their utility but rather aim to minimise their regret. Both theories can be, from the perspective taken here, also be seen as similarity variables resulting in adjustment terms. The formulations of the respective adjustment terms are presented in the following.

Kahneman and Tversky (1979) developed **Prospect Theory** to model decision making under risk because they found that choices made among risky prospects are inconsistent with utility maximisation. People underweight outcomes that are merely probable in comparison with outcomes that are obtained with certainty leading to risk aversion in choices involving sure gains and to risk seeking in choices involving sure losses. People also discard components that are shared by all prospects under consideration leading to inconsistent preferences when the same choice is presented in different forms. To model these phenomena, Kahneman and Tversky (1979) recommend to compare the outcomes to a reference point (usually the status quo) and assign values to gains and losses rather than to the final asset. The value function should be concave for gains, convex for losses, and steeper for losses than for gains. Instead of probabilities decision weights are used that are generally lower than the corresponding probabilities, except in the range of low probabilities.

In the context of the discussion in this chapter, the adjustment term derived from Prospect Theory is the result of a differential, but complete comparison across all attributes of the alternatives:

$$A_{in} = \sum_{k \in K} [\gamma_k^+ f_k^+ (\max \{0, x_{ik} - x_{0k}\}) + \gamma_k^- f_k^- (\min \{0, x_{ik} - x_{0k}\})] \quad (2.42)$$

where A_{in} is the adjustment term for alternative i and choice set C_n , x_{ik} is the value of attribute k for alternative i , $f(\cdot)$ is an appropriate transformation of the attribute difference, K is the set of attributes considered and γ_k is a parameter to be estimated. Different parameters and transformations can apply for gains $+$ and losses $-$. The shortcoming of this adjustment term is, that there is only one reference point for comparison and that similarity to non-reference alternatives is ignored.

Regret Theory was originally suggested independently by Bell (1982), Fishburn (1982) and Loomes and Sugden (1982) for pairwise choices between lotteries. Recently, Chorus *et al.* (2008) extended it to make it applicable to multinomial and multi-attribute choices, such as travel choices. Their **Random Regret-Minimisation** framework pro-

posed is built on the assumption that regret arises if a non-chosen alternative turns out to be more attractive than the chosen one. It is calculated by comparing the utility of each attribute of an alternative to the best utility of the same attribute of all other alternatives. Thereby, the framework also takes into account that decision-making is not fully compensatory. In addition, it is able to model risky choices and the postponement of choices due to information limitations. The resulting adjustment term can then be defined as:

$$A_{in} = \max_{j \in C_n} \left\{ \sum_{k \in K} \max \{0, \gamma_k f(\delta_k [x_{ik} - x_{jk}])\} \right\} \quad (2.43)$$

where A_{in} is the adjustment term for alternative i and choice set C_n , x_{ik} is the value of attribute k for alternative i , $f()$ is an appropriate transformation of the attribute difference, K is the set of attributes considered, γ_k is a parameter to be estimated and δ_k is an indicator which equals 1, if the worse value is the larger one, and -1, if it is smaller one.

Chorus *et al.* (2008) go even one step further by suggesting that the systematic utility V_{in} consists only of the adjustment term. Similar formulations could also be based on rank and dominance variables. In line with the discussion above, regret can be seen as a negative super-visibility. Moreover, the regret formulation is a loop-sided version of the more general Prospect Theory.

2.5 Conclusion and outlook

Many modern choice models are characterised by a large number of alternatives in the choice set and a complex structure of similarities between these alternatives. This is caused by two trends: the increasing use of high-resolution data originating for example from GPS studies and the growing effort to model several steps of the classic four step approach simultaneously. Hence, models are needed that are able to handle large choice sets and do not require too much effort for computation, specification and identification. On the other hand, suitable approaches have to be flexible and able to accommodate various and complex similarity structures.

This applies especially to the models described in Section 2.4. The inclusion of adjustment terms in the deterministic part of the utility function is very appealing because of its simplicity and elegance. Instead of structuring the choice set a priori, the type of similarities and their functional form is specified. This type accounts for the individual char-

acteristics of the alternatives in the choice set and imposes a value to the impact of specific interdependencies. Practical applications of the models described in this chapter demonstrated, that the IIA property has been well accounted for and that the models could be estimated with relatively low computational costs even for large sets of alternatives.

However, these models also suffer from some shortcomings. They are designed with respect to a specific choice context and usually miss some aspects of the correlation between alternatives. While adjustment terms for some choice situations have been extensively investigated and appropriate factors have been well established, similarities in other choice situations have hardly been tackled by the means of adjustment terms. Particularly public transport connection choice and destination choice need further investigation.

In addition, more empirical work is needed to determine which of the four mechanisms identified in Section 2.4 is dominant for which choice dimension. This problem particularly important for multi-dimensional choices, in which the effects overlay each other. This requires the systematic analysis or reanalysis of relevant examples. Recent work on public transport connection choice (Weis, 2006; van Eggermond *et al.*, 2007) has shown that for the best model fit the sign and transformation of the adjustment terms does not necessarily match the theoretical expectations of their developers. They have overlooked some of the mechanisms at work.

Chapter 3

Processing GPS Raw Data Without Additional Information

This chapter combines the papers:

Schüssler, N. and K. W. Axhausen (2008) Identifying trips and activities and their characteristics from GPS raw data without further information, paper presented at *8th International Conference on Survey Methods in Transport*, Annecy, May 2008.

Schüssler, N. and K. W. Axhausen (2009) Processing GPS raw data without additional information, *Transportation Research Record*, **2105**, 28–36.

In recent years, studies surveying individual travel behaviour based on GPS records have become more and more important due to their manifold advantages compared to classic travel survey methods. Researchers benefit from more accurate and reliable information while participants' burden can be reduced substantially if the GPS data collection does not involve time-consuming questioning. However, without additional information, i.e. modes and trip purposes, extensive post-processing is required to derive data that can be used for analysis and model estimation.

This chapter describes a post-processing procedure needing no input other than the most basic GPS raw data: three-dimensional positions and timestamps. First, the data is thoroughly cleaned and smoothed. Second, trips and activities are determined. Third, the trips are segmented into single-mode stages and the transport mode for each of the stages is identified. The procedure is applied to GPS records collected in the Swiss cities of Zurich, Winterthur and Geneva. 4882 participants carried an on-person GPS-receiver for 6.65 days on average. The resulting trip and activity rates, distance and duration distributions as well as mode shares are compared to the Swiss Microcensus 2005 to demonstrate that derived data is ready for further applications, such as discrete choice model estimation.

3.1 Introduction and related work

Since the first GPS studies in the mid-1990s (e.g. Wagner, 1997; Casas and Arce, 1999; Yalamanchili *et al.*, 1999; Draijer *et al.*, 2000; Wolf, 2000; Pearson, 2001), this new method of surveying individual travel behaviour has gained increasing attention in transport research. Compared to previous travel survey methods researchers benefit from more accurate and reliable information about times, geographic locations, and routes. At the same time, participants' burden is reduced substantially if the GPS data collection does not involve time-consuming questions to derive additional information, such as trip purposes and transport modes. However, without additional information, extensive data processing is required to derive data that can be used for analysis and model estimation. Accordingly, current research focuses on the development of GPS post-processing procedures that allow the researcher to derive all necessary information, such as start and ending time, mode, and trip purpose directly from the GPS records. Moreover, continuously growing sample sizes lead to an increasing demand for automated procedures with low computational cost.

The choice of approach strongly depends on whether the collected GPS data is vehicle-based or person-based. In vehicle-based studies (e.g. Du and Aultman-Hall, 2007; Schönfelder *et al.*, 2006; Ogle *et al.*, 2002; Biding and Lind, 2002), the participants' vehicles are usually equipped with GPS loggers that record only, when the engine of the vehicle is running. Accordingly, recognition of individual trips is relatively easy using time differences between recorded points. In addition, short stops can be quite reliably determined by identifying times when the vehicle's speed registers zero. However, there are also some shortcomings related to vehicle-based data. First and foremost, all other modes are omitted, even though they are essential for the analysis of transport behaviour in an urban environment. Second, the real trip origins and destinations must be estimated since only vehicle movements are recorded.

Thus, person-based GPS studies have recently become more popular, although they raise the requirements for the post-processing procedures considerably. In addition to data filtering, trip detection and map-matching, the analyst has to detect the modes used by the participant. Moreover, the method for trip detection needs refinement and the map-matching has to be done either on a multi-modal network or, if that is not available, multiple networks.

As summarised in Stopher (2009) several authors have started to address these problems (e.g. Bohte and Maat, 2008; Stopher *et al.*, 2007; Flamm and Kaufmann, 2007; Tsui and Shalaby, 2006; Chung and Shalaby, 2005; de Jong and Mensonides, 2003). Basically, all approaches contain individual modules for:

- Data filtering
- Detection of trips and activities
- Determination of single-mode stages
- Mode identification
- Map-matching

Some authors include additional features such as merging of stages after the mode detection (de Jong and Mensonides, 2003) or feedback between map-matching and mode detection (Tsui and Shalaby, 2006). However, until now, all of these methods have been tested only on small samples or test scenarios and most still require manual intervention, particularly during mode detection.

Reliance on manual interaction is not feasible for the data set at hand or any other large GPS data set. The data set at hand contains about 32,000 person-days recorded in the Swiss cities of Zurich, Winterthur,

and Geneva. The original study was conducted by a private sector company trying to determine whether or not participants noticed certain billboards (Pasquier *et al.*, 2008). We obtained the data, but without the socio-economic details of the respondents, from one of the sponsors of the original data collection effort as part of a joint project. 4882 participants were asked to carry an on-person GPS logger for 6.65 days on average. No additional information, such as modes or trip purposes, was collected. The network used in the map-matching was the Navteq network, a high resolution navigation network covering all regions of Switzerland and containing 408,636 nodes and 882,120 unidirectional links. Due to the large amount of data and the high resolution of the network, it was necessary to work not within a GIS environment but to implement new procedures in JAVA. This chapter describes the design of these procedures. Their advantages and shortcomings will be discussed with respect to the overall aim: deriving a data set ready for further applications, i.e. discrete choice model estimations, from GPS raw data, with no further information.

The remainder of this chapter is structured as follows: the next section focuses on data cleaning and smoothing. Subsequently, it is described how GPS records are subdivided into trips and activities. Since each trip can still contain more than one mode, the mode detection presented afterwards starts with segmenting trips into single-mode stages. The actual mode detection is executed using a fuzzy logic approach based on speed and acceleration characteristics of the stages. The functioning of the map-matching is subsequently illustrated on the sub-sample of car stages. However, as it was demonstrated in Menghini *et al.* (2009) it can also be applied to other modes. Before presenting the conclusions and the outlook on future work, the results of the proposed approach are compared to the Swiss Microcensus 2005, since no validation data was available from the study itself.

3.2 Data cleaning and data smoothing

The positioning accuracy of GPS receivers under ideal conditions lies between five and ten metres (Wolf, 2006). In reality, however, it is usually much worse due to several error sources. For instance, there might be less than the four satellites in view that are required to precisely calculate a three-dimensional position. Even if there are enough satellites in view, they might not be ideally positioned, evidenced by a high position dilution of precision (PDOP) value (Wolf *et al.*, 1999). While this

leads to GPS positions that are completely different from the actual position of the receiver, the so-called *warm start/cold start problem* results in missing GPS points at the beginning of the trip due to the time the GPS receiver needs to acquire the position of at least four satellites in view (Stopher *et al.*, 2005).

In addition, there are random errors caused for example by satellite or receiver issues, atmospheric and ionospheric disturbances, multi-path signal reflection or signal blocking (Jun *et al.*, 2007). Multi-path errors (also called *urban canyoning errors* because they typically appear in urban canyons) are especially troublesome. The GPS signal is reflected by buildings, walls or surfaces and the corresponding GPS positions jump and are often scattered around the actual position of the receiver. Signal blocking, however, leads to missing GPS points and is of special importance for person-based GPS surveys since its frequency varies for the different means of transport. While GPS reception is generally good when the participant is walking, cycling and or travelling by car, it fluctuates considerably for public transport journeys, depending on the proximity of the person to the nearest window (Draijer *et al.*, 2000; de Jong and Mensonides, 2003).

There are several ways to overcome the problems caused by the GPS errors described above. Data filtering, for instance, takes care of systematic errors while data smoothing removes random errors. All these approaches, however, depend on the information available in the study. Previous studies (e.g. Wolf *et al.*, 1999; Ogle *et al.*, 2002) showed that the number of satellites in view and the PDOP value are fairly efficient in determining systematic errors. Unfortunately, they are not accessible here. Therefore, other criteria to identify erroneous data points had to be developed. One example is the altitude value. Considering the clear variation in Swiss topology, all points with an altitude value of less than 200 and more than 4200 metres above sea level are removed.

Another criterion are sudden jumps in position. Position jumps are detected by comparing the distance between two consecutive GPS points with the distance the person could have travelled in the time interval assuming a maximum speed of 50 *m/s* and a random error buffer of 30 metres. The GPS points are split into so-called quality segments each of which comprises all points between two consecutive position jumps. Subsequently, every two adjacent quality segments are compared and the GPS points of the shorter one are deleted until the end of the quality segment is reached or the distance between every two consecutive GPS points is smaller than the threshold defined above. The whole procedure is repeated until all positional jumps are removed. It is important to note

that for this filter the three-dimensional distance is used, because errors in longitude and latitude are often accompanied by fairly strong jumps in altitude, while in the subsequent trip and mode detection only the two-dimensional positions are considered.

Concerning random errors, several approaches were implemented and tested to select the optimal smoothing technique. Since no speeds from Doppler measurements are available, speed and acceleration have to be calculated directly from the position and timestamp of the GPS points. Hence, the position of the GPS points is smoothed rather than the speed. A Gauss kernel smoothing approach was used. For each coordinate dimension $c \in x, y, z$ the smoothed value $\tilde{c}(t)$ at time t is individually calculated as

$$\tilde{c}(t) = \frac{\sum_j (w(t_j) \cdot c(t_j))}{\sum_j w(t_j)} \quad (3.1)$$

with $c(t_j)$ being the raw value of the coordinate c at time t_j and $w(t_j)$ the Gaussian Kernel function computed for each point of time t_j by

$$w(t_j) = \exp -\frac{(t - t_j)^2}{2\sigma^2} \quad (3.2)$$

The Kernel bandwidth, represented by σ , is set to 10 seconds, which results in a 15 second smoothing range because this is a reasonable time frame for real behavioural changes as opposed to signal jumps. Accordingly, the directional speed for each coordinate c is the first derivative with respect to t of the smoothed position and the acceleration the second derivative with respect to t .

The effects of the data cleaning and smoothing are depicted in Figures 3.1 and 3.2. Figure 3.1 shows the development of speed and acceleration over time for a sample individual. In the left column, speed and acceleration are derived from the raw data, whereas in the right column, they are calculated after filtering for unrealistic altitudes and smoothing, but before the final filtering for unrealistic speeds and accelerations. As can be seen, much of the noise in the raw data could be removed. In particular, the completely unrealistic speed jumps which result in acceleration values higher than 10 m/s^2 are excluded without being explicitly filtered. The resulting progression of speeds and accelerations provides reasonable patterns, especially for trips in an urban environment.

Spatial effects of the filtering and smoothing are illustrated in Figure 3.2. First and foremost, it can be seen that most of the outliers, here especially noticeable at activity locations, are attenuated. Second, the

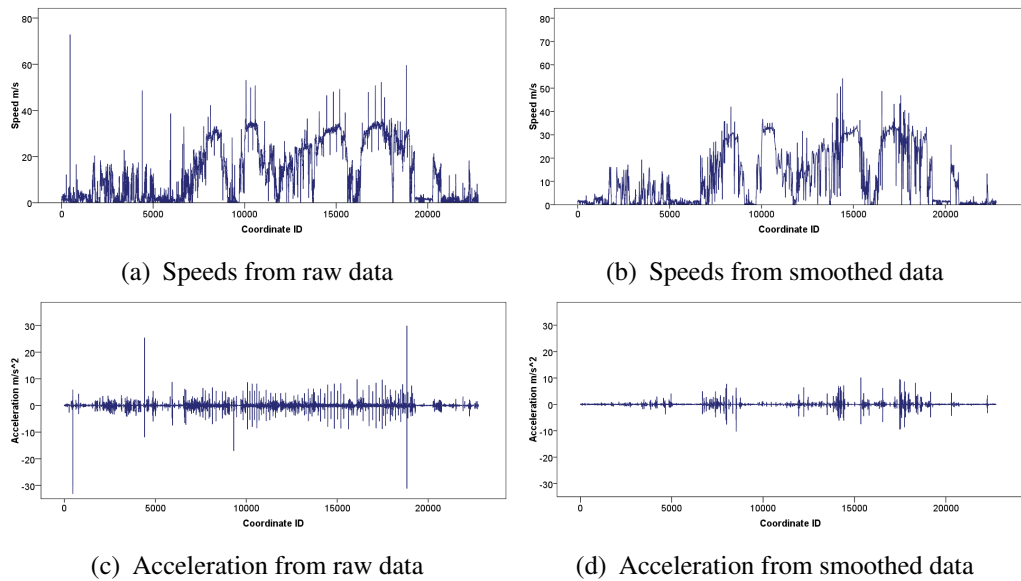


Figure 3.1: Comparison of speeds and accelerations from raw vs. smoothed data - Person 1548

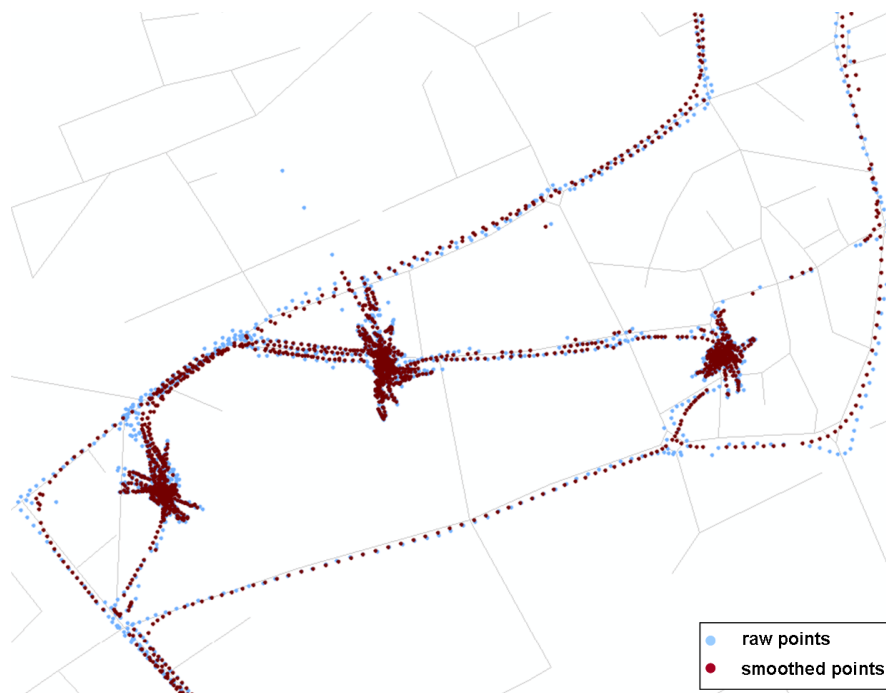


Figure 3.2: Comparison of point positions raw vs. smoothed - Person 1548

overall movement trajectories can be identified more clearly, since small deviations from the general path are reduced. The downside of this is that some corners are cut stronger than in reality due to the 15-second smoothing interval. This is, however, not an important shortcoming be-

cause, first, it is not the spatial positions of the GPS points but the speeds and accelerations resulting from the overall movement trajectories that are taken into account in the trip, activity and mode detection, and, second, the original coordinates are stored along with the smoothed coordinates. Thus, both types of coordinates can be used for the map-matching.

3.3 Trip and activity detection

The filtered and smoothed GPS points then have to be subdivided into trips and activities. Since the GPS points have been collected person-based, two basic types of activities are considered: activities with ongoing GPS recording and activities with signal loss. Activities with ongoing GPS recording are either characterised by speeds that are close to zero (e.g. Schönfelder *et al.*, 2006; Tsui and Shalaby, 2006) or by so-called *bundles of GPS points* (e.g. Doherty *et al.*, 2001; Stopher *et al.*, 2005). A bundle is a sequence of GPS points positioned very close to each other, i.e. within a diameter of about 30 metres that equates to approximately three times the standard deviation of the measurement accuracy (Stopher *et al.*, 2005). Therefore, two criteria to detect activities with ongoing recording were established. Analogously to Tsui and Shalaby (2006), the first criteria flags an activity, when the speed is lower than 0.01 m/s for at least 120 seconds. The second criteria to detect activities is based on *point density*. In this study, the point density is defined by the number of points that are within a 15 metres radius around the GPS point in question. The 15 metre radius corresponds to the definition of point bundles by Stopher *et al.* (2005). If the point density exceeds 15 for at least 10 points or 300 seconds an activity is flagged. The time threshold was chosen considerably higher than the point threshold since activities in buildings are often linked with signal loss. The density measure of 15 was derived by analysing the different point density patterns for trips by different modes and activities. Car trips usually result in point densities of 2-3 whereas the point density for walk trips lies between 7 and 10. Thus, point densities of 15 most likely occur only if the participant stayed at the same place.

Activities with signal loss are detected by means of the time difference between two consecutive GPS points. The threshold beyond which it is assumed that an activity took place varies in the literature between 45 (Pearson, 2001) and 300 seconds (e.g. Wolf *et al.*, 2004; Doherty *et al.*, 2001), whereas most studies apply 120 seconds. In this study, however, a 900 second threshold is used - a high value compared to former stud-

ies. But examination of the GPS points showed that a shorter dwell time would lead to too many wrongly detected activities due to bad reception during trips. As these reception losses will be handled by the map-matching algorithm, they do not need to be considered here.

Each of the three criteria is used individually to determine *potential activity start* and *potential activity end* points. It is important to note, that each activity can be detected by more than one criterion. Consequently, the potential activity start and end points are joined in a way that the outermost potential activity start and end points are considered to be the true activity start and end points regardless of the criterion they belong to. Moreover, if a new activity starts shortly (maximum 15 GPS points, i.e. about 15 seconds) after the last one has ended, the two activities are joined. This rule, on the one hand, accounts for measurement errors and, on the other hand, considers that trips of less than 15 GPS points cannot be reasonably used for route choice modelling. After finding all activity start and end points, activity and trip objects are generated and stored in separate lists. Only the trip objects are used in the subsequent analysis. The activity objects will be used later to analyse trip purposes as well as trip and activity chains.

Since no information about the actual activities was available to validate the procedure, an analyst was asked to divide the records of 25 persons, randomly chosen from the Zurich sample and containing 152 person days, into trips and activities based on spatial and temporal patterns and human intuition. The comparison of this classification with the outcome of the trip and activity detection revealed that 97% of the activities could be detected by at least one of the criteria and none was falsely detected.

3.4 Mode detection

Determining the modes used by the participants is one of the major research issues for person-based GPS studies. It is the crucial step to make them usable for large-scale applications. However, few approaches for an automated mode detection have been published to date. As most approaches presented so far, the mode detection approach implemented in this dissertation comprises two steps: segmentation into single-mode stages and then mode assignment for each of these stages. Based on the circumstances in the study area, five modes are distinguished:

- walk
- cycle

- car
- urban public transport (i.e. bus and tram)
- rail

Since the GPS records only show the spatial and temporal movement and nothing about the circumstances or accompanying persons, the car modes comprises all trips travelled by car regardless if the person was driving, a passenger or riding a taxi. Analogously, the mode detection cannot distinguish between tram and bus trips because they have very similar movement patterns. The mode detection does, however, detect the transfer between different public transport vehicles and treats the according trip parts as individual stages. Other modes are neglected because their cumulative share amounts only to 0.8% of all stages in the Swiss Microcensus on Travel Behaviour 2005 made by people living in the study area.

The segmentation of trips into single-mode stages implements the definition that walking is required for every mode change and for every transfer between public transport vehicles. The procedure, which follows the mode detection method presented by Chung and Shalaby (2005) and Tsui and Shalaby (2006), exploits the uniqueness of the walk mode with consistently low speeds and accelerations. In addition, it considers that a mode change can occur during signal loss. Therefore, a new stage is created if a signal loss of more than a threshold occurs. Three types of *potential mode transfer points (MTP)* are detected: *end of walk (EOW)*, *start of walk (SOW)*, and *end of gap (EOG)* points. Thereby the start of gap point is implicitly defined as the point before the EOG point. Since Tsui and Shalaby (2006) showed that their thresholds for signal loss duration and walking speed and acceleration delivered reliable results, the same thresholds were applied in this study. The threshold for signal loss duration is 120 s and the speed and acceleration thresholds for walking are 2.78 m/s and 0.1 m/s², respectively. Afterwards, the potential MTPs are aligned, to ensure that each walking stage is enclosed by exactly one SOW (or EOG) and one EOW (or EOG) point, that the speed in a walking stage never exceeds 2.78 m/s, and that the derived stages are sufficiently long. For a walking stage the minimal duration is 60 s and for all other modes 120 s (Tsui and Shalaby, 2006).

Subsequently, the mode for each stage has to be derived. Former studies have applied different approaches to achieve this. Bohte and Maat (2008), Stopher *et al.* (2005), Chung and Shalaby (2005), and de Jong and Mensonides (2003) use rules based, for instance, on average and maximum speed, proximity to certain network elements (e.g. bus stops

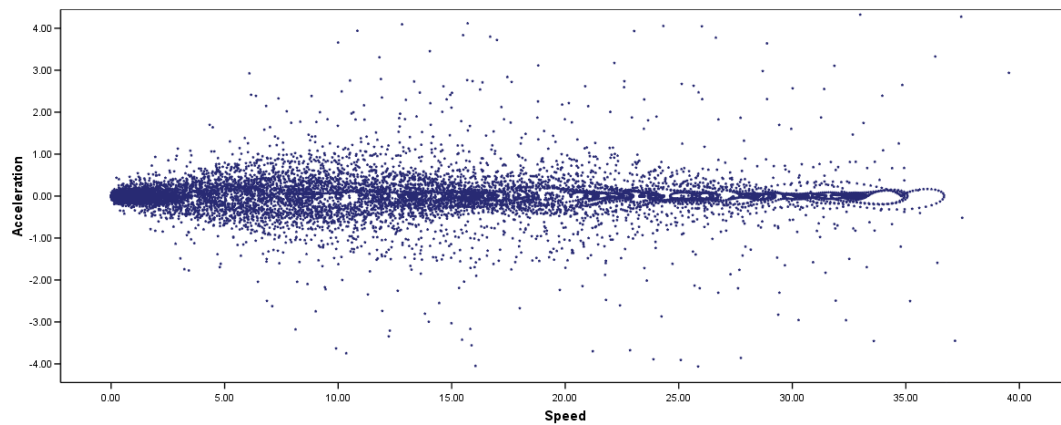
or train stations), or the deviation from the street network. Zheng *et al.* (2008) evaluate four inference models (Decision Trees, Bayesian Networks, Support Vector Machines and Conditional Random Fields) and obtain best results for the Decision Tree model. In this dissertation, however, a fuzzy logic approach based on speed and acceleration characteristics is employed, as introduced by Tsui and Shalaby (2006), since it best accounts for ambiguity in the allocation of modes to observed characteristics such as speed or acceleration.

An open source fuzzy engine (Sazonov *et al.*, 2002) is used for the implementation of the fuzzy logic component of the mode detection. The crucial elements are the fuzzy variables, the fuzzy rules describing the relationship between the modes and the fuzzy variables, and the membership functions representing the different levels of the fuzzy variables. Three fuzzy variables were chosen, each with three membership functions: the median of speed, and the ninety-fifth percentiles of the speed and acceleration distributions. These statistical location parameters were deliberately chosen over the average speed or the maximum speed and acceleration to make the algorithm more robust against outliers.

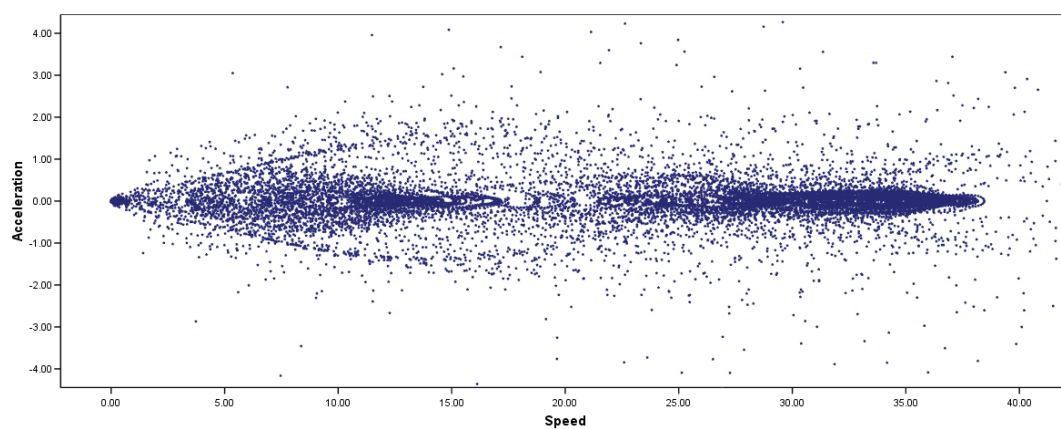
The trapezoidal membership functions are described by four key points: Start point, left top corner, right top corner, and end point. They were chosen after an analysis of the available modes and the speed and acceleration characteristics in the GPS data. Figure 3.3 depicts an example of a distribution of speed and acceleration combinations for two persons with obviously different travel behaviour. Overall, five more or less well-defined clusters can be distinguished:

1. speed $< 2 \text{ m/s}$ and acceleration $< 0.15 \text{ m/s}^2$
2. speed $4\text{-}8 \text{ m/s}$ and acceleration $< 0.2 \text{ m/s}^2$
3. speed $14\text{-}17 \text{ m/s}$ and acceleration $< 0.3 \text{ m/s}^2$, with accelerations up to 1 m/s^2 for points leading to the cluster
4. speed $20\text{-}28 \text{ m/s}$ and acceleration $< 0.3 \text{ m/s}^2$, with accelerations up to 1.6 m/s^2 for points leading to the cluster
5. speed $> 30 \text{ m/s}$ and acceleration $< 0.4 \text{ m/s}^2$, with accelerations up to 4 m/s^2 for points leading to the cluster

Although these clusters cannot be directly assigned to specific modes, they give an idea about the distribution of speed and acceleration of the modes. The first cluster, for example, characterises walking with low speeds and very low accelerations, whereas the fifth cluster contains



(a) Person 1548



(b) Person 16048

Figure 3.3: Example of speed and acceleration distribution

mainly trips travelled with high speeds by car on the motorway or by high-speed train. The fourth cluster describes speed and acceleration patterns that are typical for travel on country roads or on the InterRegio or rapid-transit railway system. Clusters 2 and 3 on the other hand cannot be so easily matched to individual modes, since they can be caused by the whole range of urban means of transport, such as cycle, urban public transport and car. They are, however, used to determine the key points of the membership functions, as they are presented in Figure 3.4.

Having established the membership functions, fuzzy rules are derived. They characterise the modes with regard to the fuzzy variables. As depicted in Table 3.1, each mode is described by at least one rule. Since the ranges of the membership functions overlap, more than one rule can apply to the same stage and multiple modes can be assigned to that stage. The defuzzify method combines the membership values for

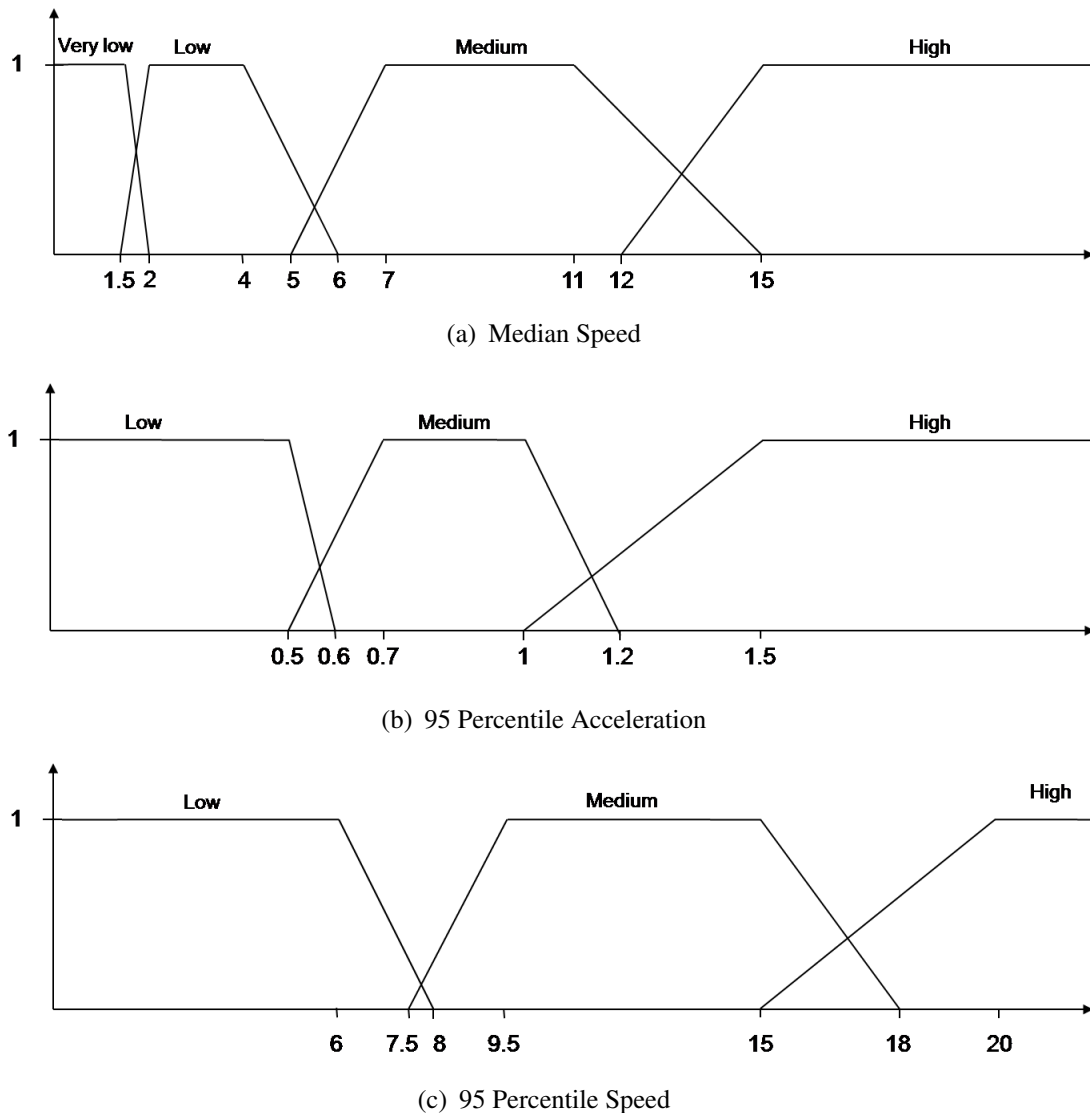


Figure 3.4: Membership functions of the fuzzy variables

each mode using the OR operator, meaning that the final score for each mode equals the maximum membership value amongst all its rules. Subsequently, the defuzzify method calculates the likelihood for each mode by dividing the fuzzy score of each mode by the sum of the fuzzy scores of all modes.

In a third step, the reasonability of the derived mode chains is investigated. Due to the design of the stage generation process described above, every two non-walking stages are separated either by a walking stage or a time gap of at least 120 seconds. The latter accounts for the fact that the analyst cannot confirm whether a mode change has or has not occurred during that time. This, however, sometimes leads to relatively unlikely mode transitions, e.g. a direct transfer from car to train.

Median speed	95 perc. acceleration	95 perc. speed	Mode
very low	low	—	Walk
very low	medium	—	Cycle
very low	high	—	Cycle
low	low	low	Cycle
low	low	medium	UrbanPuT
low	low	high	Car
low	medium	—	UrbanPuT
low	high	low	UrbanPuT
low	high	medium	Car
low	high	high	Car
medium	low	—	UrbanPuT
medium	medium	—	Car
medium	high	—	Car
high	low	—	Rail
high	medium	—	Car
high	high	—	Car

Table 3.1: Fuzzy rules for mode detection

Therefore, the characteristics of the time gap are scrutinised, especially the average speed during the time gap. If the average speed is higher than walking speed and the two neighbouring stages are both non-walk stages, it is assumed that no mode transfer happened during the time gap. The stages are joined and the mode detection is repeated for the joined stage. In addition, two adjacent walk stages are joined if the average speed in the time gap is below the walk speed threshold. The threshold for walking speed is thereby set to 2 m/s . Thus, more realistic mode chains were derived.

3.5 Results of the GPS postprocessing

Since no information about actual trips and activities of the participants is available and the GPS sample is assumed to be representative for the population of the study areas, the Swiss Microcensus on Travel Behaviour 2005 (MZ 2005) (Swiss Federal Statistical Office, 2006) is used as the basis to validate the post-processing procedure. The Swiss Microcensus on Travel Behaviour (MZ) is conducted every five years and delivers a representative and detailed insight into the travel patterns of the Swiss population. In 2005, 33,390 individuals reported on their socio-economic background, their mobility tools, and their stage-based, recorded trips and activities on the reporting day in a computer-assisted telephone interview (CATI). In the following, results of the post-processing proce-

	Zurich	Winterthur	Geneva	MZ 2005
Number of persons	2 435	1 086	1 361	3199
Number of days per person	6.99	5.96	6.51	1.00
Number of trips per day	4.50	3.40	4.26	3.65
Average trip distance [km]	7.72	7.37	7.19	8.79
Average daily mileage [km]	34.74	23.20	29.25	32.13
Average trip duration [min]	15.17	13.71	15.05	26.21
Average number of stages per trip	1.40	1.31	1.47	1.68

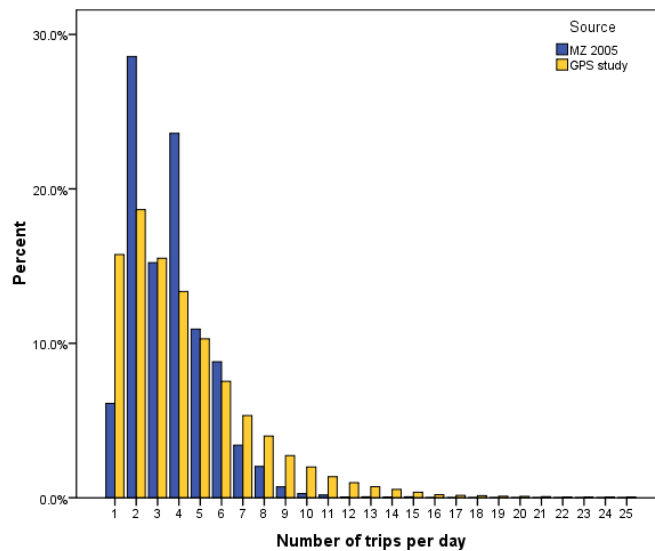
Table 3.2: Overall statistics of the GPS study compared to the Swiss Microcensus 2005

dure are compared to a sub-sample of the Microcensus 2005, made up of respondents living in Zurich, Winterthur or Geneva.

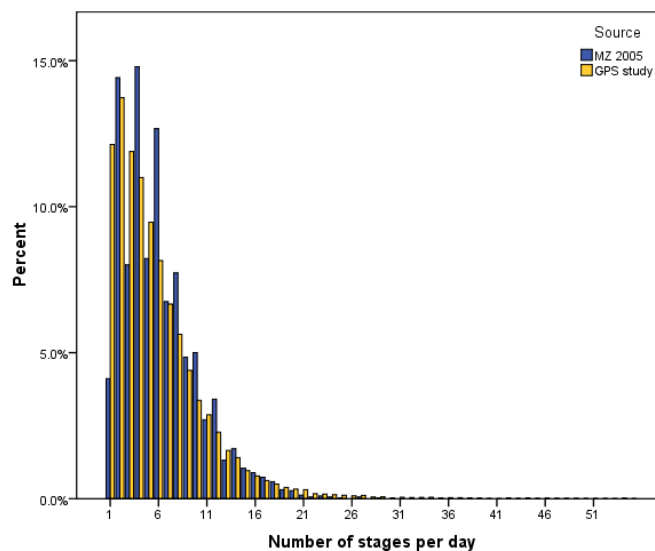
The overall statistics for the three sub-studies for Zurich, Winterthur and Geneva, and the corresponding sample in the MZ 2005 are represented in Table 3.2. Values vary between the three cities. The number of days observed per participant, for example, is about seven in Zurich compared to approximately six days in Winterthur. Overall, however, the numbers reveal the same trend. Except for Winterthur, the number of trips per day is higher than the number of trips reported in the MZ 2005 and the average trip distance and duration are smaller. This effect was expected because several previous studies (e.g. Du and Aultman-Hall, 2007; Wolf *et al.*, 2003; Yalamanchili *et al.*, 1999; Bricka and Bhat, 2006) demonstrated that shorter trips and activities are especially under-reported in recall-based surveys.

The distributions of number of trips per day and the number of stages per day and the respective distributions of the Swiss Microcensus 2005 are depicted in Figure 3.5. The trip distributions as well as the stage distributions are similarly skewed positive, although the tails of the ones derived from the GPS data are slightly longer. However, the distinct peaks for even number of trips or stages present in the Swiss Microcensus are not replicated in the GPS data. This was expected since the omission shorter trips and intermediate activities, e.g. shopping on the way home from work, often occurring in recall-based surveys leads to fewer and shorter trip chains.

The distributions of trip lengths and durations are presented in Figure 3.6. All trips longer than 20 km or 2 hours have been summarised in the last category of the respective distributions. The distributions reveal similar patterns in the MZ 2005 and the GPS data. This confirms that the trip and activity detection works properly. Moreover, the graphs reveal that the rounding of times and distances is an important issue in



(a) Trips per day



(b) Stages per day

Figure 3.5: Trips and stages per day compared to the Swiss Microcensus 2005

recall-based surveys. Estimating distances and times accurately is a real challenge for the respondents. In the Swiss Microcensus, distances are frequently rounded to km values and durations to quarter hours. This underlines once more the advantage of GPS over recall-based surveys with respect to temporal and spatial accuracy.

In Figure 3.7, the distribution of mode per stage is compared to the MZ 2005. Walk stages are excluded because they are only partially comparable. On the one hand, around 15% of mode transitions in the MZ 2005 do not meet the basic assumption that every two non-walking

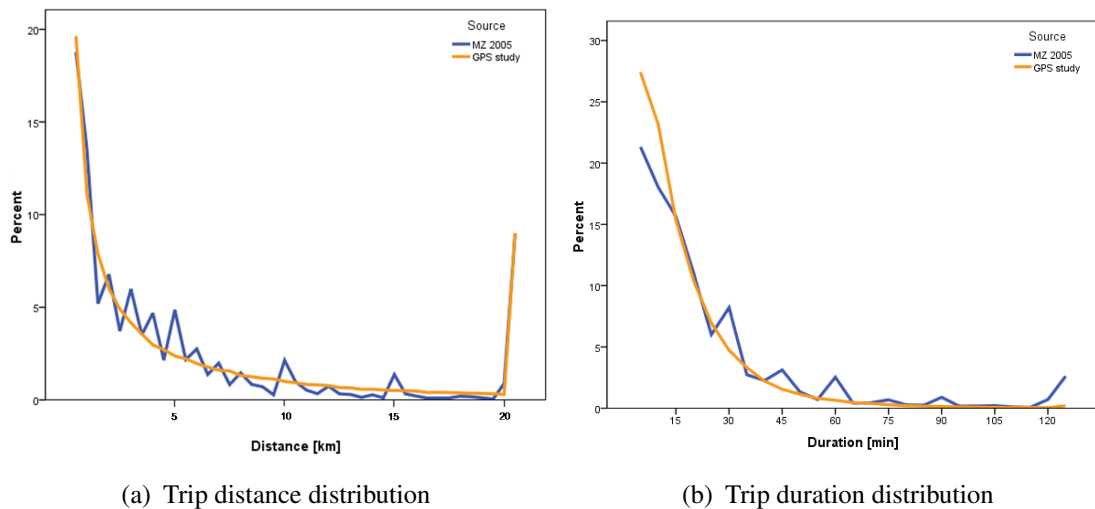


Figure 3.6: Trip distance and duration distributions compared to the Swiss Microcensus 2005

stages should be separated by a walk stage. On the other hand, trips derived from GPS data suffer from the warm start/cold start problem. Accordingly, 43% of the GPS trips start with a car, rail or urban public transport stage as opposed to 31% in the MZ 2005. Figure 3.7 reveals that rail stages are quite under-represented at this phase of the analysis. This might be due to the design of the study which focussed on passing urban bill-boards and not on inter-urban travel. Another problem is the generally poor GPS signal reception in trains. However, some rail trips are probably misclassified as car stages, as are some of the missing urban public transport stages. Two approaches to correct this are envisaged in the future. First, the mode detection parameters will be further optimised. Second, a feedback-loop between map-matching and mode-detection will be implemented.

Even closer to the MZ 2005 is the distribution of stage distance per mode, as depicted in Figure 3.8. Again the last category contains all trips that are at least 20 km long. In both surveys, the distribution follows reasonable patterns, e.g. walk trips are rather short, while car and rail are predominantly used for longer trips. Overall, the distance distributions of the individual modes derived from the GPS data are very similar to those reported in the MZ 2005. This underlines the quality of the mode detection procedure, especially since stage distance is not used in the procedure. Only the detected bike trips are quite short compared to the MZ 2005. This might indicate that some of the bike trips are misclassified walk trips. This requires further analysis in the future.

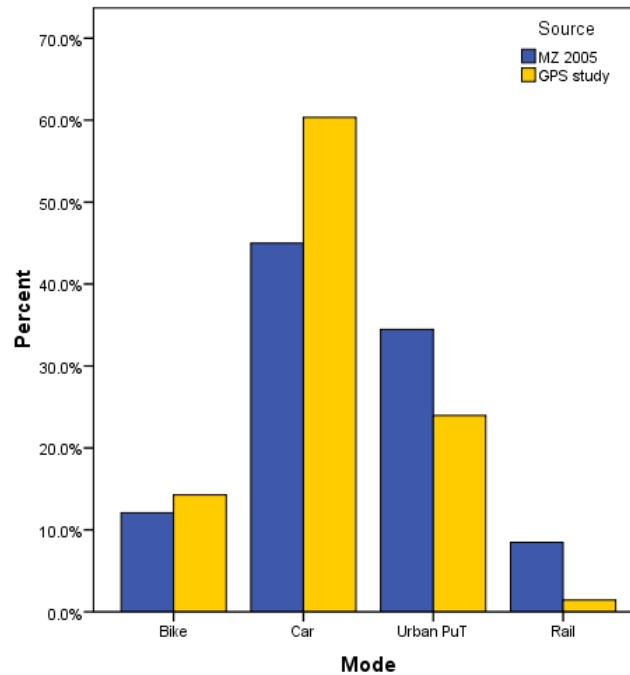


Figure 3.7: Distribution of stage modes compared to the Swiss Micro-census 2005

3.6 Computational performance

One of the main reasons for implementing this GPS post-processing procedure in Java rather than a GIS environment was the necessity of handling the study's large amount of data in a reasonable computation time. This objective was fully achieved. The computation was performed on a machine with 4 dual-core AMD Opteron 2 GHz CPUs, 4 GB RAM and a Debian Linux version 2.6.18-5-amd64. The Java 1.5 program processed the complete data set containing about 64.5 million GPS points using one CPU and 1 GB allocated memory in 7091 seconds resulting in an average speed of about 9100 GPS points per second.

3.7 Conclusion and outlook

The GPS post-processing procedures presented in this chapter allow the exploitation of a very rich data set for transport behaviour modelling. The person-based GPS data set at hand covers complete trip and activity chains over several days, thus describing the participants travel behaviour in the most comprehensive way. This makes it ideal for the employment in choice models. In addition, the sample size is large enough to obtain stable choice models and its analysis is now feasible since the procedure

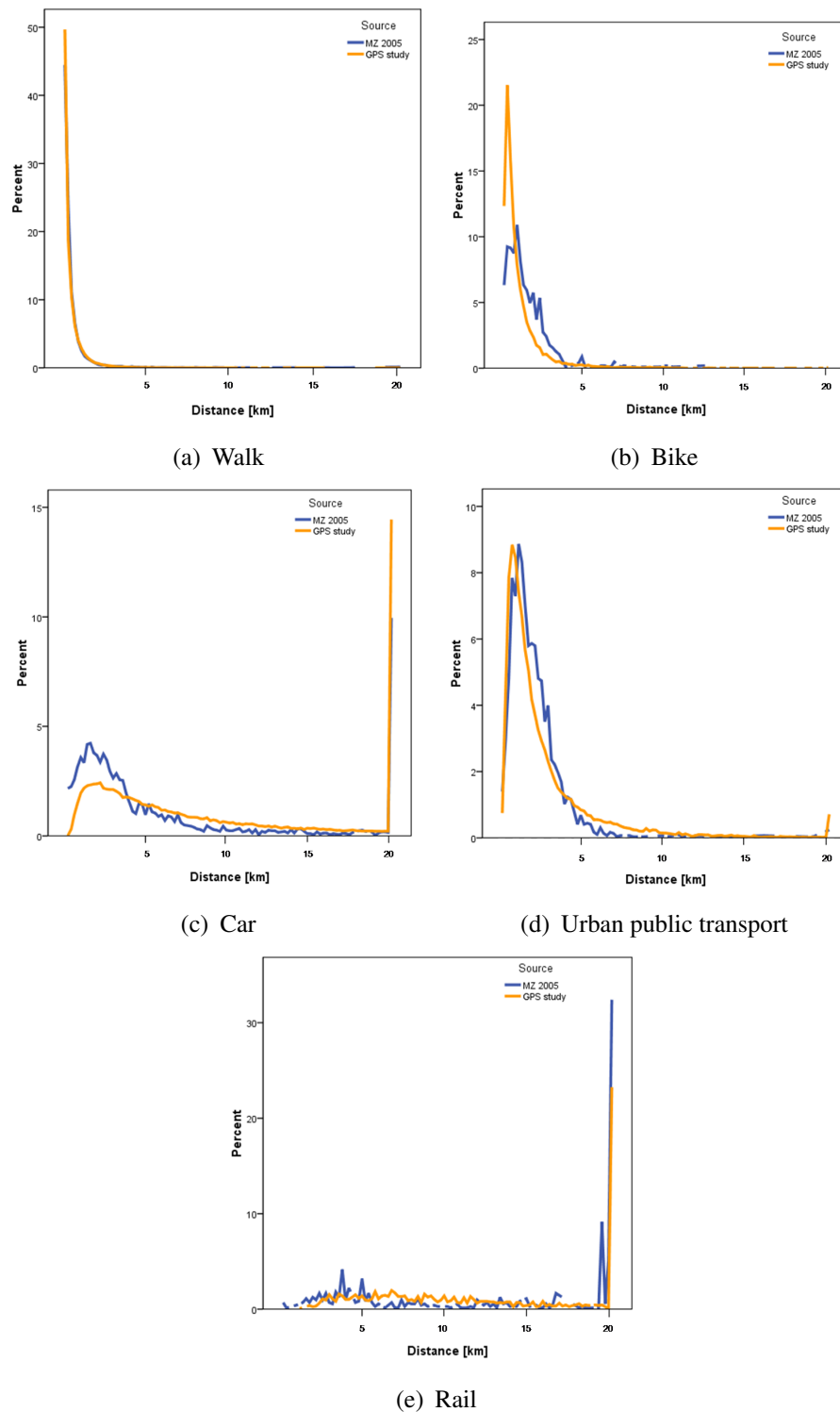


Figure 3.8: Distance distributions per mode compared to Swiss Micro-census 2005

presented in this chapter delivers sound results in a reasonable computation time even with only the most basic information.

A key success factor for the post-processing procedure described in

this chapter is the implementation of an appropriate filtering and smoothing mechanism. Finding the right approach was essential because no information about the numbers of satellites in view or their positioning was available. Though filtering based on altitude levels and unrealistic positional jump was necessary, it was not sufficient. Therefore, a Gauss kernel smoothing had to be applied as well.

The activity dwell time of 900 seconds employed in the trip and activity detection is rather high. Shorter dwell times, however, led to too many false trip ends. Instead, the detection of shorter activities is based on the identification of point bundles and zero speeds. With these criteria, the trip and activity detection delivers results very similar to those derived manually, which had to be the frame of reference since no information about the actual trips and activities was available.

For the same reason, the results of the fuzzy logic mode detection could only be compared to the MZ 2005. The comparison showed that mode detection yields realistic results. Notably, the distance distributions per mode are very satisfactory. However, in the future a more detailed validation of the mode detection should be carried out using GPS data that contains the modes actually chosen.

Chapter 4

Map-Matching of GPS Points on High-Resolution Navigation Networks Using Multiple Hypothesis Technique

This chapter consists of the paper:

Schüssler, N. and K. W. Axhausen (2009) Map-matching of GPS traces on high-resolution navigation networks using the Multiple Hypothesis Technique (MHT), *Working Paper*, **568**, IVT, ETH Zurich, Zurich.

Before the results of the GPS processing procedure described in Chapter 3 can be employed for route choice modelling, the chosen routes have to be identified in a network. This is done with a procedure called map-matching. State of the art map-matching algorithms have to be accurate as well as computationally efficient given the increasing use of GPS devices in large-scale transport studies and the augmented use of high-resolution navigation networks. Yet, only a few authors have addressed the issue of performance in the sense of computational efficiency so far.

The algorithm proposed in this chapter is an adaptation of the algorithm proposed by Marchal *et al.* (2005). It is designed to match large-scale GPS data sets on a high-resolution navigation network in an acceptable computation time. This chapter describes the implementation of the algorithm after a short overview of the existing literature. Finally, the performance of the algorithm is evaluated both in terms of accuracy and computational efficiency.

4.1 Introduction

In recent years, data from GPS-based surveys has become increasingly important since transport modellers benefit from more accurate and reliable information about times, geographic locations, and routes. At the same time, participants' burden is reduced substantially if the GPS data collection does not involve time-consuming questions to derive additional information. However, without respondent provided information, extensive data post-processing is required to derive results that can be used for analysis and model estimation. Beginning with the first GPS survey (Wagner, 1997), one of the key post-processing steps is the map-matching, i.e. the association of the GPS points to the links of a network, in order to establish the routes travelled by the survey participants.

State of the art map-matching algorithms have to be accurate as well as computationally efficient. The first algorithms focussed more on accuracy and consistency of the derived routes, since the survey samples were still rather small. Accordingly, most reviews of map-matching algorithms, (e.g. White *et al.*, 2000; Greenfeld, 2002; Quddus *et al.*, 2007) mean accuracy in terms of percentage of correctly identified links when they talk about the performance of an algorithm. However, with the increasing use of GPS devices in large-scale transport studies, the need for computational speed grows. This need is further amplified by the augmented use of high-resolution navigation networks, which are essential

for an accurate identification of the chosen routes. Yet, only few authors (Nielsen *et al.*, 2004; Marchal *et al.*, 2005) have addressed the issue of performance in the sense of computational efficiency so far.

The algorithm proposed in this chapter is an adaptation of the algorithm proposed by Marchal *et al.* (2005). It is designed to match large-scale GPS data sets on a high-resolution navigation network within acceptable computation time. This chapter describes the implementation of the algorithm after a short overview of the existing literature. Afterwards, the performance of the algorithm is evaluated both in terms of accuracy and computational efficiency for a sample drawn from about 36,000 car trips executed by 2,434 persons living in and around Zurich, Switzerland, which are mapped on the Swiss Navteq network, a high-resolution navigation network covering all regions and all streets of Switzerland. The chapter closes with conclusions and an outlook on future work.

4.2 Related work

With regard to their underlying approach, map-matching procedures can best be classified into three categories:

- Geometric procedures
- Topological procedures, and
- Advanced procedures.

The most basic approaches are the geometric procedures, that only take into account the distance between the GPS points and certain network elements. Frequently used examples are the search for the nearest node or the nearest link. White *et al.* (2000) extended the nearest link search by comparing the heading of the GPS points with the heading of the links in question. And they also proposed the so-called curve to curve matching, where every two GPS points are connected by a curve and the distance between this curve and the surrounding links is minimised. The main shortcoming of all geometric procedures is that they ignore the sequence of the GPS points over time as well as the connectivity of the network links. Therefore, it is possible that the derived route oscillates back and forth between links. In addition, they are very dependent on a correct network coding and are rather sensitive to outliers. Regarding computational efficiency, Nielsen *et al.* (2004) demonstrated that the nearest node search is in any way fast enough for the problem at hand. Yet, the error rate of the matching of nearly 3,000 trips from the Copenhagen area on an associated 300,000 link network was with 33% too high for any real life application.

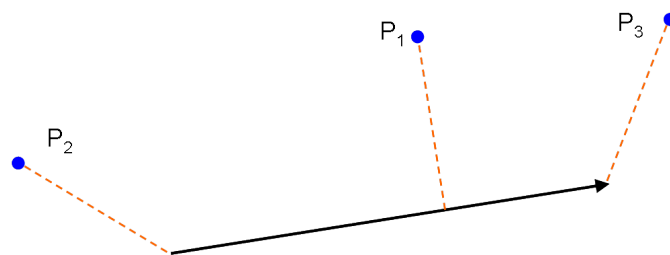


Figure 4.1: Derivation of the perpendicular distance

In contrast to the geometric procedures, topological procedures do not only account for the distance between the GPS points and network elements, but also for the sequence or history of GPS points and the connectivity of network elements. Most procedures work in two steps. First, the initial node or link is found using geometric approaches. Afterwards, the route is developed by choosing a link out of a set of candidate links. Usually, this set consists of the last matched link and the links succeeding that link. Some authors extend it to all links preceding the last matched link (Chung and Shalaby, 2005) or to the links succeeding the succeeding links (Greenfeld, 2002). For the choice of the most likely link out of the set of candidate links, different criteria can be employed. The most common one is the *perpendicular distance* between the GPS point and the link. The perpendicular distance equals the minimum of the euclidean distance between the GPS point and its orthogonal projection on the link, the euclidean distance to the start node and the euclidean distance to the end node. In the following, the distance measure that, out of these three measures, delivers the minimum value is called the *relevant perpendicular distance*. In Figure 4.1, the relevant perpendicular distance for GPS point P_1 is the distance to its orthogonal projection on the link, for P_2 the distance to the start node and for P_3 the distance to the end node.

Other criteria for choosing a link out of the set of candidate links are the heading of the GPS point compared to the one of the link (e.g. Greenfeld, 2002; Chung and Shalaby, 2005; Velaga *et al.*, 2009), the position of the point relative to the link, which is derived from the angle between the link and the line between the start node of the link and the GPS point (Quddus *et al.*, 2003) or if, and at what angle, the link and the line between the GPS point and its predecessor intersect (Greenfeld, 2002). If more than one criterion is used, they are usually weighted against each other with parameters determined using expert knowledge or a calibration procedure (e.g. Velaga *et al.*, 2009). Additional features include the treatment of outliers (Greenfeld, 2002) or a post-processing either in terms of a ex-post elimination of unlikely links based on the

percentage of the link that is covered by the points (Chung and Shalaby, 2005) or a mode specific filling of gaps between links (Tsui and Shalaby, 2006). The topological approaches outperform the geometric ones in terms of computational speed as well as route accuracy. They are faster because for each GPS point, except the first one, only a very limited number of links has to be evaluated and they are more accurate because they take into account the whole sequence of points and are less sensitive to measurement errors and outliers. However, there is still room for improvement in both respects. In particular, there is no fall-back solution in case the initial link determination failed. Moreover, there is the issue of parallel streets, that are running closely next to each other. Once the algorithm chooses the wrong route, it is unable to correct such a mistake (Quddus *et al.*, 2003; Nielsen *et al.*, 2004).

To overcome these problems, in recent years more advanced approaches have been proposed. Advanced approaches do not only take into account the whole sequence of GPS points and the network topology, but also the fact that, due to errors in the GPS measurement as well as the network coding, the nearest link or node is not necessarily the right one. A lot of different procedures have been proposed of which a small selection is presented here.

A straight forward approach to account for GPS measurement errors is the construction of error or confidence regions around the GPS points (e.g. Doherty *et al.*, 2001; Ochieng *et al.*, 2004; Velaga *et al.*, 2009). The size of the error region should be derived from the error variances (Quddus *et al.*, 2007; Ochieng *et al.*, 2004). Then all links within this error region are evaluated based on distance, heading, connectivity to the previously matched link, and sometimes speed (Ochieng *et al.*, 2004). Quddus *et al.* (2006) extended Ochieng *et al.* (2004)'s error region approach by a fuzzy logic inference systems. The fuzzy rules consider different criteria such as distance, heading, speed, the quality of the position solution based on the HDOP value, link connectivity and the position of the GPS point relative to the candidate link. Thereby, separate rules apply for the initial link search and the subsequent path development. Another example for an extension of the error region concept is the map-matching based on conditional random fields introduced by Liao *et al.* (2007). They use a reduced set of evaluation criteria: Only distance and connectivity are taken into account.

An approach without the use of error regions was proposed by Nielsen *et al.* (2004). Their algorithm resembles the Dijkstra algorithm for the single-source shortest path problem. The start node is determined in a not further described preprocessing step. Starting from there, the

route is developed by adding the end nodes of all outgoing links of the current node to the set of nodes to be evaluated. The next node to be evaluated is then the node that could be reached in the shortest amount of time starting from the last node of the route so far. The score of each node is calculated based on the perpendicular distance between the GPS points and the links they are assigned to and the distance between the GPS points and the start node of the link they are assigned to. Yet, how the GPS points are assigned to the links cannot be derived from the paper even though this is a crucial aspect of any map-matching approach since a wrong assignment can lead to greatly biased scores. Another problem with this algorithm is that it cannot not guarantee to find the optimal solution, as the Dijkstra algorithm would, because the route development criterion differs from the scoring function. Moreover, even though Nielsen *et al.* (2004) claim that the algorithm is highly efficient, the risk of evaluating the whole network should not be neglected because the stop criterion is simply that the set of nodes to be evaluated should be empty.

However, the main disadvantage of all the approaches summarised hitherto is that after the evaluation of each GPS point only one link, or route, remains. The next GPS point is then matched based on the assumption that the last GPS point was matched correctly. Accordingly, if the map-matching identifies just one link wrongly, the probability is quite high that the remaining route will be wrong as well. Considering the frequency and magnitude of GPS measurement errors, especially at the beginning of a trip, this is not recommendable. Therefore, Pyo *et al.* (2001) introduced the use of the Multiple Hypothesis Technique (MHT) for map-matching. This means that, following the sequence of GPS points, several route candidates are kept in memory, developed and assigned a score. The best candidate is usually only determined when the end of the GPS sequence is reached. Two different implementations of the MHT have been presented so far. While Pyo *et al.* (2001) employed error regions and extended the existing paths by the links within the error region, Marchal *et al.* (2005) adopt a topological search algorithm. The initial set of links is determined by searching the m nodes closest to the first GPS point and creating a single-link path for each of their incident links. Afterwards it is checked whether the current GPS point can be matched to the last link of the route or if a junction was reached. If a junction was reached, a new route candidate is created for each link succeeding the current link. The new link is added at the end of the route. Each of the new route candidates is scored and saved in the set of route candidates. Because this topological search inherently en-

asures link connectivity, the scoring function is rather lean. It sums up the perpendicular distances between the GPS points and the links the points are assigned to. In a later version of the algorithm, they extended the scoring function to account for the shape of the link by comparing the distance between two consecutive GPS points with the distance between their projections on said link. Compared to this, the list of criteria in Pyo *et al.* (2001)'s scoring function is rather extensive. They use distance, heading, and, as a measure of connectivity, the number of links necessary to get from the last matched link to the link in question. In both approaches it is necessary, to limit the number of candidates kept in memory to ensure computational feasibility. Marchal *et al.* (2005) simply define a maximum number of candidates. If the size of the candidate set exceeds this number after all candidates have been processed for the current GPS point, the route candidates with the lowest scores are removed. Pyo *et al.* (2001) do not define their pruning criteria on the absolute number of alternatives. They cut candidates based on the score of the candidates relative to the score of the best candidate so far. Accordingly, their probability to maintain a rather large set of candidates is high, especially in dense urban areas. This leads, combined with the need to establish an error region around each GPS record, to relatively high computational cost whereas Marchal *et al.* (2005) could show that their algorithm is accurate as well as fast enough for the map-matching of large amounts of GPS data on a high-resolution navigation network. Therefore, the algorithm presented by Marchal *et al.* (2005) is chosen as the basis for the algorithm proposed in this chapter.

4.3 Implementation

The map-matching presented in this chapter is the fifth step of a framework that derives mode specific route choice observations from GPS records. It follows the data filtering, detection of trips and activities, mode stage determination, and mode identification described in Chapter 3. Like its preceding steps, the map-matching is implemented in JAVA. While the first four steps did not employ any other information but the GPS points, the map-matching by nature requires the use of a network. In order to profit from existing infrastructure, the map-matching uses several elements of MATSim (MATSim-T, 2008). This includes the representation of the network as well as other helpful methods such as the calculation of distances between GPS points and certain network elements, or the search for all nodes within a certain radius.

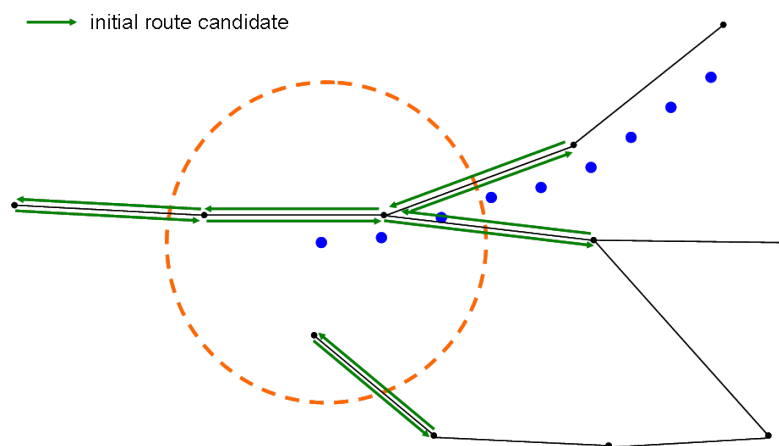


Figure 4.2: Derivation of the initial route candidates

The map-matching procedure itself consists of six consecutive steps:

1. Trip segmentation
2. Determination of initial route candidates
3. Route development
4. Selection of the most likely route candidate
5. Stage filtering
6. Treatment of the gaps between trip segments

The first four steps are based on the work of Marchal *et al.* (2005). However, several aspects have been improved to increase the accuracy of map-matching results.

First, the GPS points of each trip are subdivided into continuous segments. Therefore, spatial and temporal gaps in the sequence of GPS points are detected. If the time gap between two subsequent GPS points is longer or the distance is larger than a threshold defined by the analyst, the map-matching cannot deliver trustworthy results. Several combinations of the thresholds for time and distance gap were tested before the final thresholds of 120 seconds and 500 metres were chosen. The GPS points before and after such a gap are stored as separate trip segments. Moreover, a minimum number of points per segment was defined because the GPS points after a recording gap tend to be noisy potentially leading to unreliable map-matching results especially when they are all matched to the same link. Thus, segments containing less than 10 points were removed because a segment with 10 or more points has a high likelihood to be matched to at least two links. Subsequently, the steps two to five are executed for each segment individually, before the segments are joined to complete trips in step six.

Second, the initial set of single-link route candidates is derived. As depicted in Figure 4.2, all nodes within the a radius of 750 metres around the first GPS point are found. Afterwards, a single-link route candidate is created for each link connected to at least one of these nodes. For each route candidate, the first GPS point is assigned to the one link of the route, the score is calculated and it is stored in the list of current route candidates. In case the number of route candidates found this way is lower than 25, the search radius is increased by 100 metres and the whole process is run again. This is repeated until there are at least 25 routes candidates.

The route development process is illustrated in Figure 4.3. For each GPS point, all route candidates remaining from the previous iteration are evaluated. If the route candidate contains only one link, it is first checked if the start node of this link was reached, i.e. the distance to the start node is the relevant perpendicular distance. In this case, the route candidate is discarded because apparently the GPS points are running in the opposite direction of the link. Subsequently, it is examined if the point can be assigned to the last link of the route candidate or if the end of this link was reached. Given that the GPS point can be assigned to the last link, this assignment is stored and the score of the route candidate is recalculated. If the end of the link was reached, for each link succeeding the last link a copy of the route candidate is created, the succeeding link is appended to the route, the GPS point assigned to that link, and the score is recalculated. Then, the old route candidate is removed and the newly created route candidates are added to the set of route candidates. However, before a route candidate is added, it is verified, that it is a valid route candidate. A route is only valid if it is unique, i.e. the exact same sequence of links is not already included in the set of route candidates, if the new link does not head back to the start node of the last link, i.e. the route contains no u-turn, and if the route does not contain any link twice. Cycles in terms of using the same node twice are allowed since in one-way street systems with turn restrictions it is sometimes necessary, to pass the same crossing twice.

A crucial aspect of this procedure is the way to determine if the end of a link was reached. Even though this is an important issue for many map-matching procedures, only Ochieng *et al.* (2004) and Marchal *et al.* (2005) have discussed their way of doing this in more detail. Ochieng *et al.* (2004) use the relative position of the GPS point to the link as well as heading changes that exceed a certain threshold as end of link criteria, whereas Marchal *et al.* (2005) compare the distance travelled by the GPS points with the length of the link. If the GPS points cover at

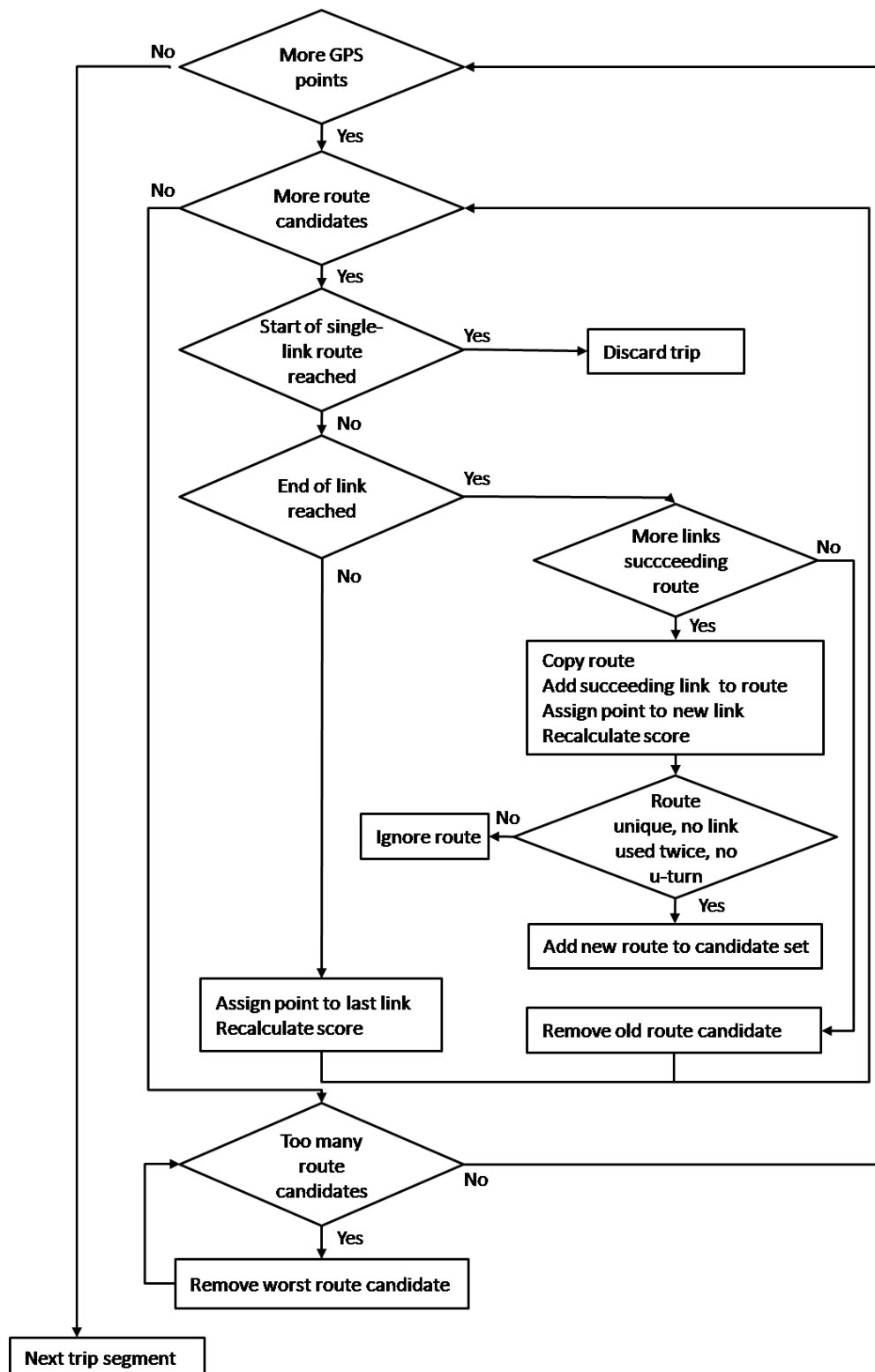


Figure 4.3: Flow chart of the route development for a trip segment

least a certain percentage of the link, the end of the link is reached. As shown in Figure 4.4, the proposed algorithm uses a combination of these criteria. The GPS point has reached the end of the current link if the

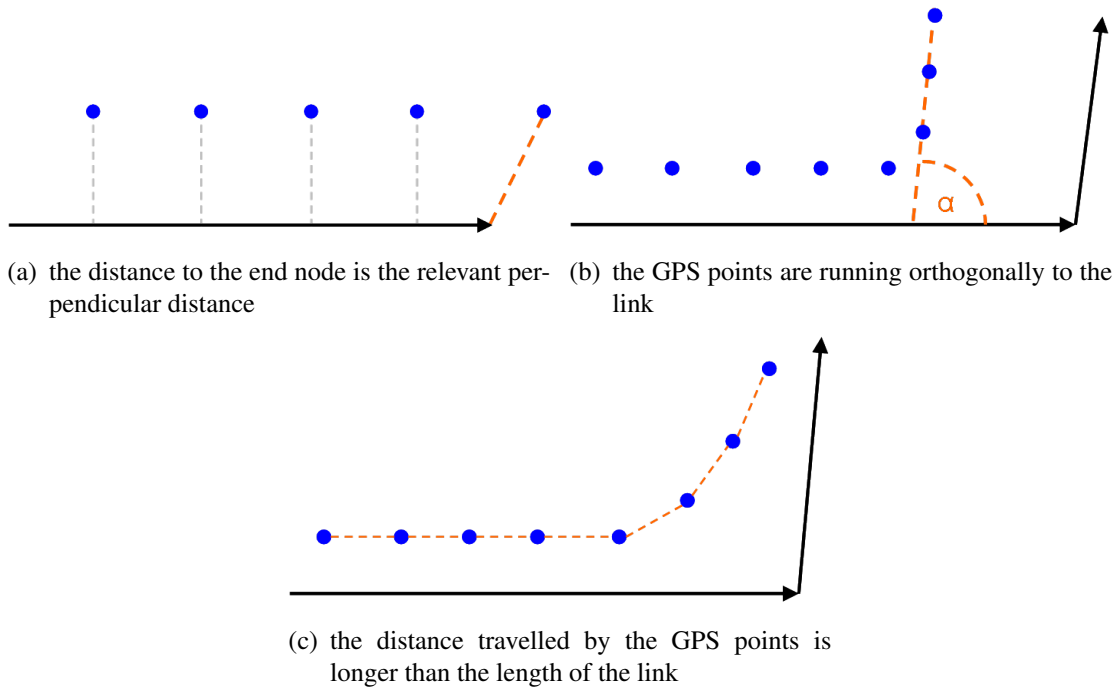


Figure 4.4: The end of a link is reached if...

distance to the end node is the relevant perpendicular distance or if the GPS points are running in an orthogonal or opposite direction, i.e. at an angle bigger than 85° , to the link or if the distance travelled by the GPS points is longer than the length of the link.

When all route candidates have been processed for the current GPS point, the number of remaining route candidates is compared to the maximum number of route candidates N_{max} defined by the analyst. If there are too many route candidates, the ones with the worst score are removed until the size of the route candidate set equals N_{max} . Thereby, N_{max} has to be chosen carefully. On the one hand, a high N_{max} ensures that the route candidate set contains the actually chosen route even if it obtains lower scores at the beginning or in between. On the other hand, a low N_{max} improves the computational performance. Several values for N_{max} have been tested in the development of the map-matching algorithm. The results of these tests are discussed in Section 4.4.

Another decisive factor for the success of a map-matching procedure is the way the score of each path candidate is calculated. As discussed in Section 4.2, several criteria have been used so far. The most popular ones were the perpendicular distance between the GPS points and the links they were assigned to, the heading of two successive GPS points compared to the heading of the link, the speed of the GPS points, and the connectivity of the link to the preceding link. In the proposed algorithm

the score is calculated based on the perpendicular distances between the GPS points and the links they are assigned to and a speed malus in case the speed of the GPS points exceeds the free-flow speed on the link in question, as specified in Equation 4.1:

$$SC_{path} = \sum_P \sum_L (\beta_1 d(p_i, l_j) \delta_{ij} + \beta_2 (v(p_i) - v_{ff}(l_j))^2 \gamma_{ij}) \quad (4.1)$$

where $L\{l_1, l_2, \dots, l_t\}$ is the set of links composing the path, $P\{p_1, p_2, \dots, p_p\}$ the set of GPS points and δ_{ij} equals 1 if p_i is assigned to l_j , and 0 otherwise. In addition, $v(p_i)$ is the GPS speed at point p_i , $v_{ff}(l_j)$ is the free-flow speed on link l_j and γ_{ij} equals 1 if $v(p_i) > v_{ff}(l_j)$ and 0 otherwise. The parameters β_1 and β_2 have two functions, analogously to those of β -parameters in a utility function. On the one hand, they allow the combination of different score contributions with different units, such as m and m/s . On the other hand, they serve as weighting factors for the different contributions of the perpendicular distance and the speed malus. As such, they can be defined by the analyst. Currently, in absence of better evidence, they are set to $1 \frac{1}{m}$ and $1 \frac{s^2}{m^2}$. Systematic testing is needed to derive the best parameter combination. Since no information about the actual travel speeds on the individual links is available, the speed malus only considers speed differences resulting from the GPS points travelling faster than the free-flow speed and not from GPS points travelling slower. Slower speeds are probably as often caused by traffic conditions as by travelling on the wrong link whereas higher speeds, especially substantially higher ones, are presumably more often caused by the examination of the wrong link than by speeding of the participants who knew that their travel profiles were recorded. In order to punish especially those substantially higher speeds the speed malus enters the score in the form of the squared speed difference. Other frequently employed criteria were not taken into account for different reasons. A score for connectivity, for example, was obsolete considering the design of the algorithm and the heading was found to be too erratic due to jumps in the GPS points.

In the end, after processing all GPS points of the trip segment, the route candidate with the lowest score is determined, since it is presumably the route actually travelled by the participant. However, before assigning the route to the trip segment, a final validity check is carried out. If the route contains only one link, it is not assigned to the trip segment and the trip segment itself is neglected in the sixth step, the joining of the trip segments to complete trips. A map-matching of GPS points to just one link can be terribly misleading, especially in directed networks. In such a network, a two-way road between two nodes is represented by

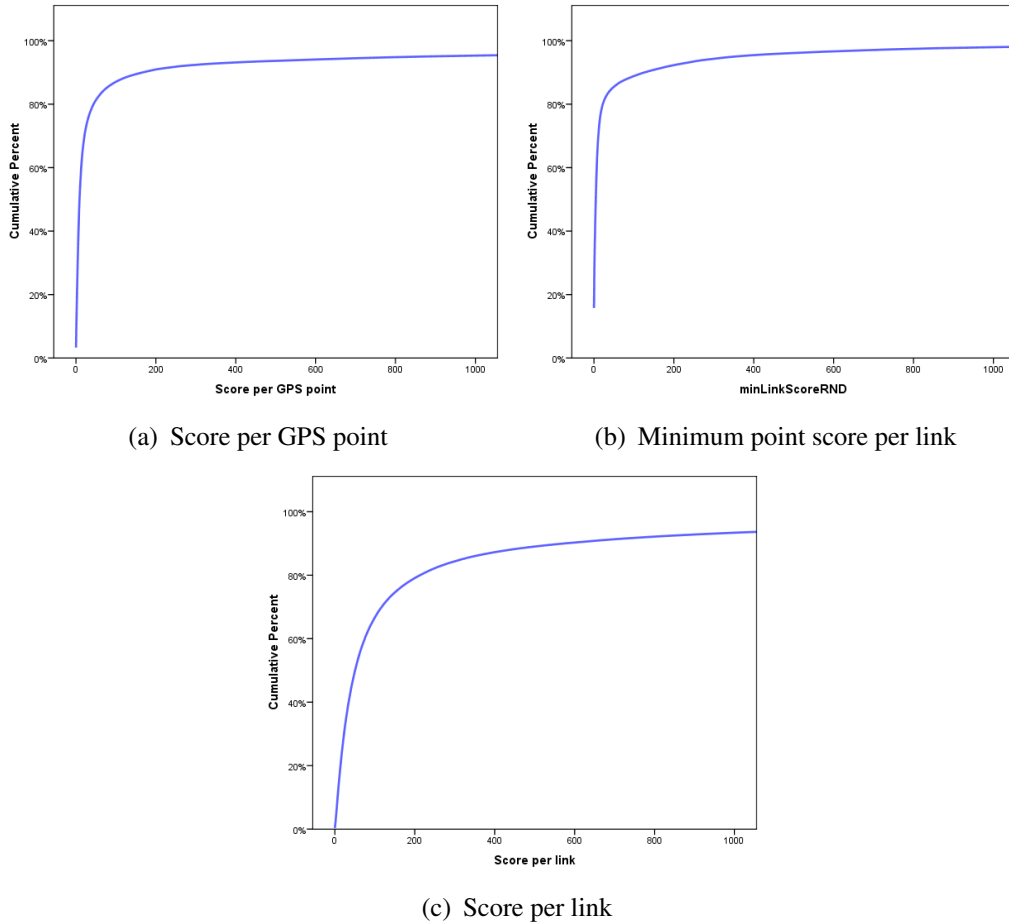


Figure 4.5: Cumulative distributions of data quality measures

two separate links, which would both obtain the same score. The selection of one of these single-link route candidates would be totally arbitrary and might have detrimental effects on the subsequent joining of the trip segments.

The next validity check is related to the quality of the network representation. An important requirement for each map-matching algorithm to work properly is a correct, consistent and complete representation of the real transport network by the network used for the map-matching. Unfortunately, hardly any network currently available can guarantee this requirement. Especially missing network links lead to problems for map-matching algorithms because the real route taken by the traveller can not be reproduced. Instead, either no route is found or a route that works its way around the missing link. This route does not reflect the actual behaviour and it is up to the researcher to decide how to deal with it. Since the applications using the map-matching results require that the matched routes are reliably reflecting the actual route choice behaviour, it was decided to filter these routes based on the assumption that there

are no systematic errors in the network coding but that missing links are randomly distributed throughout the entire network. The filtering mechanisms exploit the fact that routes which include detours around missing links are characterised by high score values. Thus, the filtering was implemented in two different ways. First, the average score per GPS point is evaluated. If it exceeds a value defined by the analyst, the route is removed. Second, the minimum point score for each link is determined. The point score is the contribution of a GPS point to the score of the link it is assigned to. If the minimum point score for a link is higher than a predefined value, the link is marked as *odd*. If a route contains more than 3 odd links, it is discarded.

The filtering thresholds were derived from the distributions of the score per GPS point and of the minimum point score per link that are depicted in Figure 4.5 (a) and (b). The 85 percentiles of the distributions were used, resulting in a threshold of 75 for the average score per GPS point in a segment and a threshold of 100 for the minimum point score per link. This ensures a high quality and reliability of the remaining routes. However, the characteristics of the trips for which the routes are discarded are stored in a separate file to allow for a check of the representativeness of the resulting routes.

In order to obtain complete trips, the gaps between the trip segments have to be filled. Depending on the quality of the GPS points and the density of the network the length of the gaps and the number of links required to fill them can vary significantly. In addition, no information is available on where the participant actually travelled during the period of signal loss. Thus, a consistent assumption has to be made on how a gap between two subsequent trip segments should be filled. Because different assumptions will have different impacts on the resulting routes, the assumptions employed should always be considered in later analysis. Two basic assumptions have been presented in the literature. Tsui and Shalaby (2006) claim that people try to avoid turning maneuvers as much as possible. Therefore, they use the heading of the links in the gap as main criterion to close the gap. Others employ shortest path search to fill gaps in route observations (Pasquier *et al.*, 2008) or participant descriptions (Ramming, 2002).

The procedure applied here is illustrated in Figure 4.6. The procedure combines a shortest path search with a treatment of low quality map-matching results. First, *questionable* links at the end of the segment preceding the link and the start of the segment succeeding the link are identified. Then, the shortest path is searched between the last node of the segment preceding the gap and the first node of the segment suc-

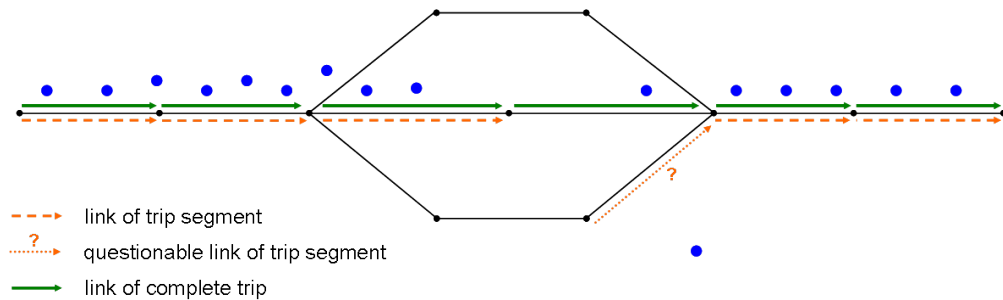


Figure 4.6: Joining trip segments to complete trips

ceeding the gap. In addition, shortest path search is executed between the start nodes of all questionable links at the end of the segment preceding the gap and the end nodes of all questionable links at the start of the segment succeeding the gap. Subsequently, the shortest of these shortest paths is determined and after two concluding quality tests, the links of the trip segments and the shortest of the shortest paths are joined to form the route which was most likely taken by the participant and can now be used for further analysis and choice modelling.

The introduction of the search for questionable links stems from the issue that the map-matching can be especially unreliable at the start and the end of a trip segment due to sparse or noisy GPS points. Therefore, the score assigned to the individual links is evaluated. If it is higher than a certain threshold, the link is marked as questionable. In this application, a rather rigorous threshold of 50 score points was chosen that corresponds to the median of the link score distribution depicted in Figure 4.5 (c). This was intended since the threshold does not automatically lead to a deletion of the trip but instead to a treatment with care. If, however, all links of a trip segment are marked as questionable, something is probably not right with the entire segment and the whole trip is discarded.

The concluding two quality tests concern the routes resulting from the combination of the map-matching and the shortest path search. On the one hand, all trips containing the same link twice are discarded. These are detected by finding trips for which at least one trip segment is entirely included in the shortest paths filling a gap. This usually occurs in case of wrong map-matching around the entry and exit ramps of motorways. On the other hand, trips which contain shortest paths that cannot be travelled in the time available according to the GPS points are dropped as well.

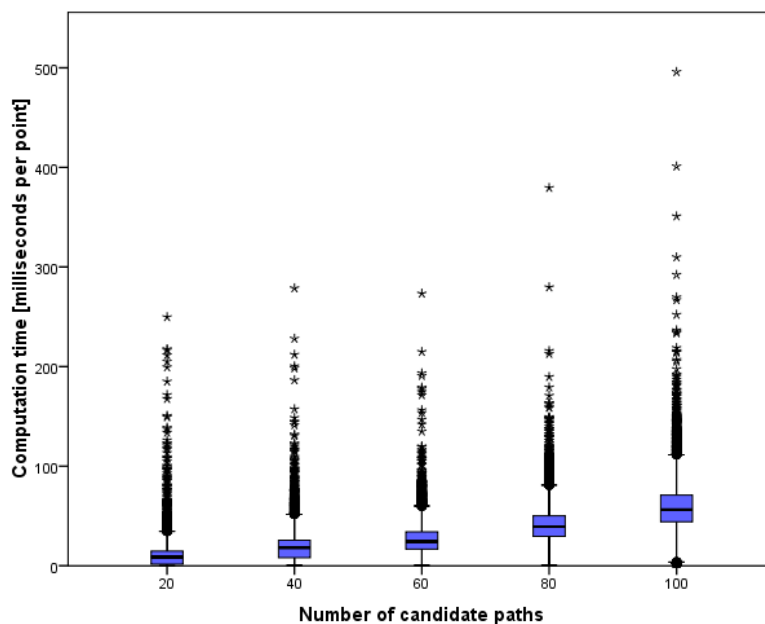


Figure 4.7: Distribution of computation time per GPS point depending on the maximum number of candidate paths

4.4 Evaluation of the resulting routes and the computational efficiency

All tests described in this section were run on systems having two Dual-Core AMD Opteron Processors 2222 running at 3 GHz. Memory was connected through a front side bus clocked at 1,000 MHz. As the code was not multi-threaded, only one of the CPU cores was actually used by the tests. The Java 1.6 program runs on 1 CPU using 2 GB allocated memory. The test data set contained about 4.1 million GPS points recorded for 250 persons and 1776 person-days. Since the map-matching is embedded in a bigger framework for the processing of GPS raw data described in Chapter 3, the 4.1 million raw data points were first filtered, subdivided into trips and activities and assigned a mode. Subsequently, 3932 car stages comprising 2.4 million GPS points were matched to the Swiss Navteq network, a high-resolution navigation network covering all regions of Switzerland and containing 408,636 nodes and 882,120 unidirectional links representing the entire Swiss street network, including minor and access roads. Thus, the Navteq network contains 44-times more links than the planning network for the same area (Vrtic *et al.*, 2005). In order to evaluate the computation time of the map-matching alone, the computation time for the GPS processing without map-matching was subtracted from the total computation time of this set of test runs.

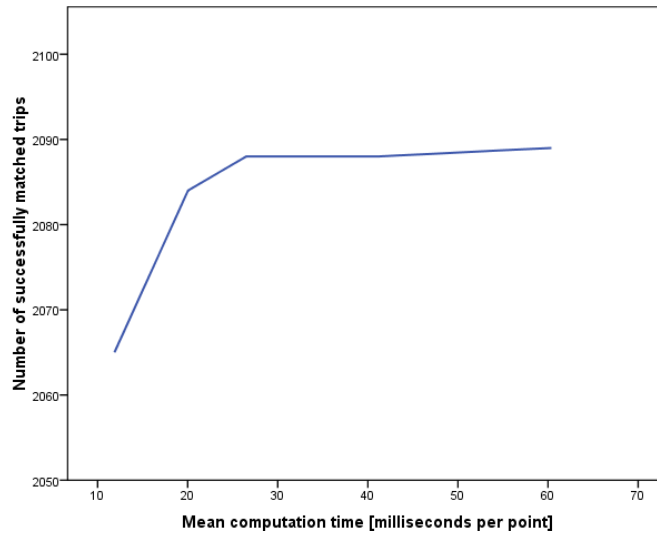


Figure 4.8: Relationship between the number of successfully matched trips and the average computation time

One of the strongest drivers for the computational performance of the map-matching is the maximum number of candidate paths that is maintained. Therefore, 5 test runs with 20, 40, 60, 80 and 100 candidates respectively were executed. In the following, the results of these test runs are evaluated with respect to computation time, number of properly matched trips and the distribution of scores. The aim is to determine the optimal number of candidates that produces reliably results within a reasonable computation time.

Figure 4.7 depicts the distribution of computation time per GPS point depending on the maximum number of candidate paths for the map-matching itself, i.e. the steps 2-4 of the procedure described in Section 4.3. As expected, the computation time increases with the number of candidate paths. The total computation time for map-matching the sample of 250 car trips including all steps outlined in Section 4.3 ranges between 393 minutes for 20 candidates and 3022 minutes for 100 candidates. This translates into a mean computation time per route between 6 and 46 seconds and a mean computation time per GPS point between 0.010 seconds and 0.075 seconds. Since the computation time increases substantially with the number of candidate paths this number should be kept as low as possible.

In this dataset, the number of routes that could be matched successfully varies only slightly with the number of candidates between 2065 and 2089 with the strongest increase between 20 and 40 candidates from 2065 to 2089 successfully matched routes. At the same time, the mean computation time per GPS point increases substantially as can be seen in

Figure 4.8. Thus, about 40 candidate paths would deliver the best trade-off between computation time and map-matching success. The more important result, however, is that compared to the total number of car routes which should have been matched, the number of successfully matched routes is extremely low. The question is why so many routes could not have been matched. A first manual check of the data for a few persons revealed that most of the problems were network related. In more than 40% of the cases the map-matching was unsuccessful because links were missing in the network. Some of these links were used by several persons or repeatedly by the same person, making a network coding error very probable. Another 40% of the unsuccessful map-matchings occurred because part of the trip or the whole trip had been travelled off-network. The complete off-network trips were usually trips outside Switzerland. Since both error sources concern the quality and the range of the network and not the map-matching algorithm, solutions have to be found on the data side. One possibility is to use the high error scores in the map-matching to impute missing network links. This is, however, a future research issue. In the meantime, the representativeness of the successfully matched routes can be derived from the characteristics of all trips stored in a separate file as described in the previous section. An issue that should be solved within the map-matching is the third most common reason for an unsuccessful map-matching: U-turns and even double u-turns, i.e. driving in circles. U-turns and circles account for a little bit less than 10% of the unsuccessful matchings. Since they are an aspect of real life travel behaviour, the map-matching algorithm should be able to handle them. This is, however, a future research issue. In addition, a more detailed analysis of the remaining unsuccessful map-matchings is necessary. For them no apparent error source could be detected in this preliminary analysis. They might reveal potential areas for improvements.

Continuing the comparison between the runs with different numbers of candidates paths, Figure 4.9 reveals that there are only few deviations in the route matchings between the different runs and that for 95.3% of the trips the best map-matching solution, i.e. the one with the lowest score, could already be found with 20 candidate paths. For further 3% the best map-matching result is derived with 40 candidate paths, meaning, that only for less than 2% a number of candidate paths higher than 40 is required to find the ideal solution. Moreover, the score differences are only marginal and the distributions of scores, as presented in Figure 4.10 for the scores per GPS point, are very similar for all runs. Considering the substantially higher computation times for runs with a higher number of candidate paths, the additional gain in accuracy for more than 40

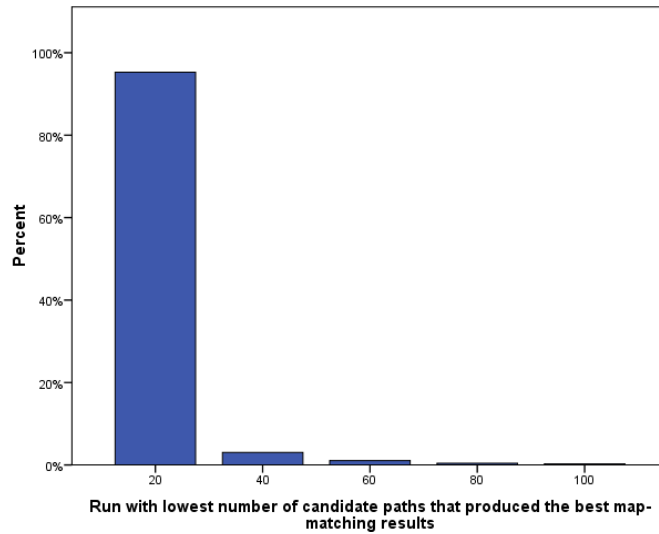


Figure 4.9: Distribution of the runs with the lowest number of candidate paths that produced the best results

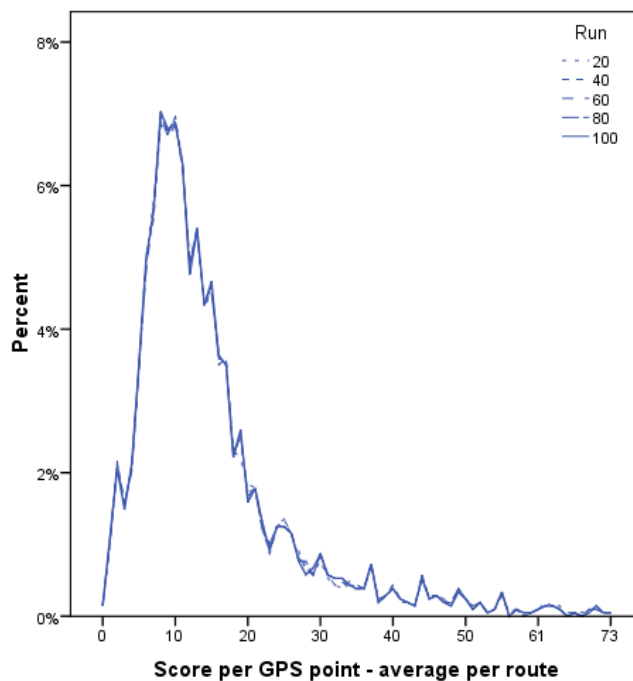


Figure 4.10: Score per GPS point - averaged per route

candidates paths is negligible.

Figure 4.10 also reveals that for the majority of routes the average score per GPS point is below 20 with an average of about 15 and a median of about 11. Considering that the distance measurement accuracy of the GPS points under ideal conditions lies between 5 and 10 metres and that especially longer links are coded in a way that their position does only

overlay at the beginning and the end with the position of the underlying road, this is a very good result for the map-matching.

Summarising, the results discussed in this section with regard to the question of which would be the ideal maximum number of candidates path, one can say that a higher number of candidate paths does only slightly increase the quality of the results while leading to a substantial increase in computation time. Thus, a maximum number of candidates paths between 20 and 40 can be recommended. This corresponds with the findings by Marchal *et al.* (2005) who advocated a maximum number of candidate paths of 30 for their algorithm.

4.5 Conclusion and outlook

The map-matching algorithm presented in this chapter is able to match a large-scale GPS data set to a high-resolution navigation network in an acceptable computation time. Manual checks of the matched roads as well as the analysis in Section 4.4 suggest that the map-matching reliably delivers the routes actually taken by the participants if the underlying network is correct, consistent and complete. Trips for which a road used by the traveller is missing in the network representation or which took place (partly) off the network are filtered because for them no accurate representation of the actual travel behaviour is possible. These trips might, however, be used in the future to correct and complement the network.

Employing the Multiple Hypothesis Technique has several advantages. The most important one is that it makes the map-matching results more robust against erroneous map-matching due to noisy GPS data or a simplified network coding where the shape of the link does not exactly follow the course of the actual road. Another advantage is that it results in more than one route candidate. These route candidates might allow to approximate the uncertainty connected with the outcomes of any map-matching algorithm. How this can be done is a future research issue. Initial ideas on how to explicitly model the likelihood of potential routes without employing a map-matching procedure have been presented by Bierlaire *et al.* (2009b). This interesting approach is, however, not yet ready for real size applications.

The most important future research issue is the treatment of u-turns in the GPS data. So far, trips that contain u-turns are also filtered by design. This is a shortcoming of the algorithm and should be resolved in the future. Currently, also trips with bad scores for which there was no apparent reason were filtered. These trips should be analysed in order

to improve the success rate of the algorithm. In the meantime, the representativeness of the successfully matched routes can be derived from the characteristics of all trips stored in a separate file as described in the previous sections.

Although the computation time of the algorithm is already acceptable, there is still some room for improvement. A first step would be to decrease number of path candidates during the development of a path depending on the differences between the scores. In the beginning, a wide exploration and a large number of candidates is necessary to ensure that the right route is contained in the candidate set. But while the path development progresses the score differences between the right route and the other candidates increases considerably so that it might be sufficient to reduce the candidate set to 10-20 alternatives.

Another worthwhile investigation is to test the computational performance and the quality performance for different network resolutions. This algorithm was specifically designed for very high-resolution networks. The application to lower resolution networks could give some more valuable insight into the workings of the algorithm especially since it will also be used for the map-matching of public transport trips to the much sparser public transport network.

In terms of embedding the map-matching in the bigger framework for processing GPS data, the next step would be a feedback of the map-matching results to the mode detection. This would improve the distinction between car trips and particularly rail-based public transport trips. So far, especially the share of rail trips resulting from the mode detection is rather low and the current hypothesis is that some of the actual rail trips are misinterpreted as car trips. A feedback from the map-matching might help to identify these trips.

Chapter 5

Route Choice Sets for Very High-Resolution Data

This chapter is based on the paper:

Schüssler, N., M. Balmer and K. W. Axhausen (2010) Route choice sets for very high-resolution data, paper presented at the *89th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2010.

The next processing step that is required to estimate route choice models that account for similarities for the car trips observed in the GPS data is the generation of the choice sets. The high level of spatial detail, that was essential for an accurate identification of chosen routes in Chapter 4, substantially raises the requirements for the choice set generation procedure, particularly in terms of computation time but also regarding the choice set composition.

This chapter addresses these issues and presents a route set generation based on shortest path search with link elimination. The proposed procedure combines a *Breadth First Search* with a *topologically equivalent network reduction* to ensure a high diversity between the routes, as well as computational feasibility for large-scale problems. To demonstrate the usability of the algorithm, its performance and the resulting route sets are compared to those of a Stochastic Choice Set Generation algorithm.

5.1 Introduction and related work

The next processing step that is required to estimate route choice models that account for similarities for the car trips observed in the GPS data is the generation of the choice sets. Evidently, the route alternatives have to be generated on the same network as the chosen routes were detected. Thus, the basis for the choice set generation is again the Swiss Navteq network, a high-resolution navigation network covering all regions and all streets of Switzerland. The high network resolution is essential for an accurate identification of chosen routes. However, it substantially increases the requirements for the choice set generation. On the one hand, the computational efficiency of choice set generation algorithms becomes a predominant issue again. The identification of a route is substantially more time consuming in a high-resolution network, while at the same time the number of possible routes increase considerable, implying that more routes must be evaluated to ensure that all relevant routes are considered. Consequently, many of the advanced algorithms that focus on the behavioural realism of choice sets will run for weeks to generate choice sets for as few as 500 OD pairs.

On the other hand, as several authors (e.g. Prato and Bekhor, 2007b; Bekhor *et al.*, 2006; Bliemer and Bovy, 2008) have demonstrated, size and composition of the choice set strongly influence the outcome of model estimation. Misspecifications of the choice set lead to biased parameter estimates and choice probabilities. As Bliemer and Bovy (2008) showed, this is especially the true when the model accounts for correla-

tion between alternatives. Ideally, all relevant and no irrelevant alternatives should be included in the choice set. To achieve this, two different approaches can be employed. The analyst can either model the membership of an alternative to the choice set (e.g. Swait, 2001; Morikawa, 1996) or generate the choice set in a step prior to the modelling. In route choice situations, the universal choice set, i.e. all possible routes between an origin and destination pair, is not known. Therefore, choice set generation procedures have to be employed that extract routes from the network. The aim is to derive as exhaustive a route set as possible in order to ensure that all relevant alternatives are detected. Afterwards, this route set can be reduced considering attractiveness, plausibility and overlap of the routes in order to obtain the choice set (Bovy, 2009).

The most common route set generation approaches can be categorised in two ways: first, by focussing on the path-establishing procedure, into approaches using repeated least cost path search and approaches employing successive path development or, second, by focussing on the output, into stochastic and deterministic procedures. Representative of the stochastic successive path development is the random walk algorithm developed by Frejinger *et al.* (2009). Starting from the origin node, the next link is chosen (based on a Kumaraswamy distribution) depending on length of the link and the shortest path distance between its end node and the destination. This is repeated until the destination node is reached. Examples for deterministic successive path development were presented by Hoogendoorn-Lanser *et al.* (2006) for multi-modal connections and by Prato and Bekhor (2006) for car trips. Both apply branch & bound technique by creating a branch at every node in their respective networks and bounding these branches using several constraints.

A prevalent version of the route set generation with repeated least cost path search is the *Stochastic Choice Set Generation*. Before each least cost path search, the link costs in the network are drawn from a probability distribution, e.g. a normal distribution (Ramming, 2002; Dugge, 2006; Bliemer *et al.*, 2007) or a truncated normal distribution (Prato and Bekhor, 2007b). In addition to the link cost, Bovy and Fiorenzo-Catalano (2007) also randomised the preference parameters for the generalised cost function. The procedure ends when a predefined number of draws or when the route set reaches a predefined size. Even though the randomisation of link costs is time-consuming for a high-resolution network, Stochastic Choice Set Generation is applicable to the problem at hand.

Deterministic route set generation approaches using repeated least cost path search are link elimination, link penalty and path labelling. Path

labelling was introduced by Ben-Akiva *et al.* (1984). The least cost path is determined according to different cost functions, called labels. Possible labels include minimum travel time, distance, number of left turns or congestion but also maximum scenery. The number of routes in the set equals the number of labels. In the link penalty approach, presented by de la Barra *et al.* (1993), the cost function remains the same. Instead, link cost on all links of the current least cost path is increased by a certain factor. Then, the new least cost path is searched. This is repeated until a predefined number of routes are found. For link elimination, one or more links of the current shortest path are eliminated before the next shortest path is searched. The elimination follows a certain order. This order can be random, duplicating the order of appearance in the route (Azevedo *et al.*, 1993) or controlled by criteria (Prato and Bekhor, 2007b). The number of links eliminated each time increases until the required number of routes is found. Some link elimination approaches ensure that the k-least cost paths are found (Lawler, 1976; van der Zijpp and Fiorenzo-Catalano, 2005), while others only accept paths within constraints such as maximum amount of overlap with other paths or a maximum detour time (van der Zijpp and Fiorenzo-Catalano, 2005).

The algorithm presented in this chapter also employs link elimination. It was designed to meet the special requirements for route set generation in a high-resolution network. The algorithm combines a Breadth First Search with topologically equivalent network reduction to ensure a significant level of diversity between the routes as well as high computational speed. The algorithm, as well as the specifications of other algorithms examined in this context, is described in the next section. Subsequently, the usability of the algorithm is demonstrated by first comparing its computational performance and resulting route sets to those of the Stochastic Choice Set Generation. The chapter closes with some conclusions and the outlook on further work.

5.2 Generating route sets

As established in the previous section, the procedures presented here focus on generation of the route set that can afterwards be reduced to the individual choice set. The goal is to find a maximum number of feasible and low cost routes in a the shortest amount of time possible. Thereby, a *feasible* route is continuous, contains no loops and is reasonable in travel cost. *Travel cost*, in this application, is defined as the free-flow travel time. In the choice set generation, reasonability with respect to travel

cost means finding the routes with the lowest travel costs. A more restrictive interpretation of reasonability is afterwards implemented in Chapter 6.

In total, four route set generation approaches were examined with regard to their applicability to high-resolution networks:

- a Branch & Bound algorithm proposed by Prato and Bekhor (2006),
- random walk introduced by Frejinger *et al.* (2009),
- a Stochastic Choice Set Generation (SCSG), and
- a Breadth First Search on Link Elimination (BFS-LE).

The branch & bound algorithm and the random walk were implemented because the respective authors showed that the resulting route sets had attractive properties. Prato and Bekhor (2006) claim that their algorithm produces realistic, heterogeneous routes allowing estimation of models with a higher prediction accuracy than models derived from other choice sets (Prato and Bekhor, 2007b). Frejinger *et al.* (2009) hypothesises that the true choice set is the universal choice set. Thus, the choice set produced by their random walk is just a sample and if each alternative in this sample is corrected with regard to its sampling probability, unbiased parameter estimates can be obtained. Both algorithms, however, proved to be inappropriate for the high-resolution network. The branch & bound algorithm terminated in reasonable computation time only for OD pairs connected by very short paths. Given the exponential increase of computation time with the number of links in the paths (Prato, 2009) an average number of 65.69 links per chosen route was too much for the branch & bound algorithm. The random walk, on the other hand, proved to be difficult to calibrate. Several parameter settings were tested but none were suitable for the entire data set with its high variation in route lengths. If the parameters were too strict, only the shortest path was found. If they were too weak, the algorithm wandered around and needed a long time to terminate i.e. find the end node of the route. The resulting route is unreasonably long, i.e. more than ten times longer than the shortest path. Thus, both approaches led to impractically long computation times, even for a small number of alternatives and short routes as can be seen in Figure 5.1.

Figure 5.1 shows a comparison of the run times of all four algorithms on a logarithmic scale for a target choice set size of 20. A *time abort threshold* was introduced to capture OD pairs for which the choice set generation could not be completed within a time interval predefined by

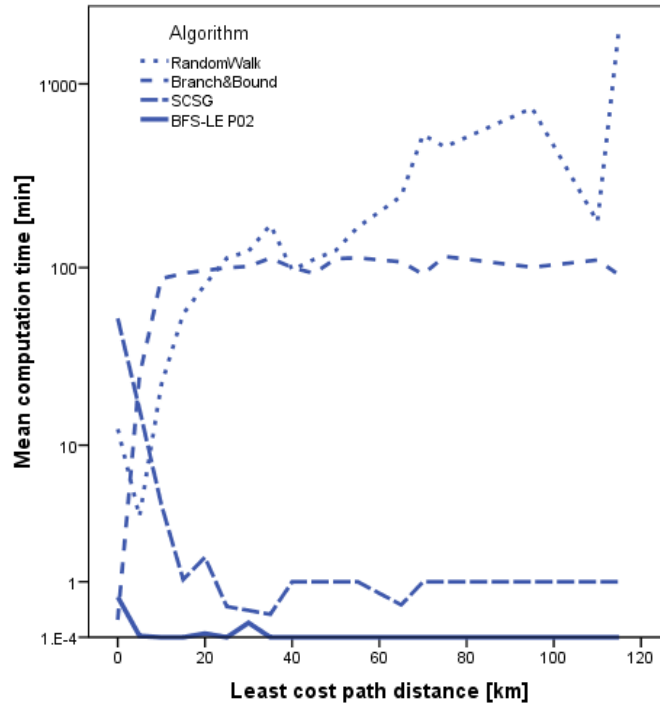


Figure 5.1: Average computation time of the algorithms for 20 alternatives and 90 minutes time abort threshold

the analyst. If the computation time for any OD pair exceeds this time abort threshold, the routes generated until then are stored as the choice set for the OD pair and the choice set generation moves on to the next OD pair. In the runs shown in Figure 5.1, this time abort threshold is set to 90 minutes per OD pair. Since the time criterion is only checked after a route has been completed, an additional abort criterion was imposed on the random walk to prevent it from exploring the network for hours trying to finish just one route and to avoid completely unrealistic routes. If the number of links in the random walk path exceeds ten times the number of links of the shortest path for the given OD pair, the random walk was stopped and restarted. If by then the time abort threshold was violated, the choice set generation for this OD pair was stopped. Thus, of the four algorithms tested in this study, only the Stochastic Choice Set Generation (SCSG) and the, performance optimised, Breadth First Search on Link Elimination (BFS-LE) were appropriate to choice set generation in the very high-resolution network at hand. Consequently, only these two algorithms are described in detail in the following and evaluated with respect to their computational performance and the structure of the resulting route sets in the remainder of this chapter.

5.2.1 Stochastic Choice Set Generation (SCSG)

In the Stochastic Choice Set Generation (SCSG) the network is changed by randomly drawing the cost of each network link from a probability distribution. The shortest path search itself is carried out using an implementation of the Dijkstra's algorithm (Dijkstra, 1959). Normal distributions – truncated at zero cost with the mean at the initial link cost and employing different multiples of the initial link cost as standard deviations – were tested. However, insufficient variation in the resulting route costs was created. Thus, the same routes were found over and over again, making it extremely time-consuming to generate route sets of sufficient size. Dugge (2006) reported similar problems for routes consisting of many small links. She reasoned that with an increasing number of links and stable travel cost, the standard deviation of a route decreases. Therefore, and because the aim was to generate as many heterogeneous routes as possible, a uniform distribution was used, ranging from zero to twice the initial link costs. The downside of this approach, however, is an increased likelihood for generating unrealistic routes, because the randomly drawn travel cost can deviate substantially from the real travel costs. This could be accounted for with an appropriate filtering such as the ones discussed in Chapter 6.

5.2.2 Route set generation with Breadth First Search on Link Elimination (basic BFS-LE algorithm)

The Breadth First Search on Link Elimination (BFS-LE) calculates repeated least cost paths of a given origin-destination (OD) pair for a given network, represented as a strongly connected, weighted, directed graph $G(V, E)$. The vertices of the graph $G(V, E)$ are geo-coded in an Euclidean space with coordinates (x_v, y_v) , while an edge $e \in E$ defines its cost as its weight (negative utility). The cost function itself can take any form and depends solely on the available network information. It does neither impair the functionality nor the computation time of the algorithm.

The least cost paths are calculated with the so-called *A-Star Landmarks* routing algorithm presented in Lefebvre and Balmer (2007). Its computational performance is at least one order of magnitude better than the Dijkstra. The SCSG method cannot take advantage of the A-Star Landmarks router's performance because it requires a preprocessing to estimate remaining travel costs. Since the SCSG changes link costs, this preprocessing would have to be re-performed after each link cost varia-

tion step.

For constructing the BFS-LE tree, some definitions have to be given first: The input network $G(V, E)$ defines the *root* of the tree and is denoted as $G^0 = G(V, E^0)$ (see Figure 5.2a). $G^d = G(V, E^d)$ is a sub-network of G^0 at depth d of the tree, while d also indicates the number of edges removed from root network G^0 . The least cost path from v_O to v_D ($v_O, v_D \in V; v_O \neq v_D$) in a network G^d is called $P^d(v_O, v_D)$ and is defined by the set of p edges $e_i^d \in P^d; i = [1..p]$. If no path exists, P^d is empty. Therefore, each *tree node* of the BFS-LE tree is defined by one sub-network G^d and its least cost path $P^d(v_O, v_D)$.

The construction of the BFS-LE tree is based on the following four rules:

- **Tree node expansion:** the creation of the *child* nodes of a BFS-LE tree node is done using the following rule: for each edge e_i^d of path P^d a sub-network $G^{d+1} = G(V, E^{d+1}) = G(V, E^d \setminus (e_i^d))$, i.e. a sub-network without edge e_i^d , is constructed and the least cost path $P^{d+1}(v_O, v_D)$ based on network G^{d+1} is calculated (see Figure 5.2 for an example).
- **Uniqueness of sub-networks:** a child node of a node at depth d will be created only if the sub-network G^{d+1} is not already created at another node of the tree at depth $d + 1$, since that node and its sub-tree do not produce new sub-networks, resp. new paths (Figure 5.2c).
- **Leaf definition:** a tree node is a *leaf* of the BFS-LE tree if the path P^d of graph G^d is empty (Figure 5.2b).
- **Breadth first:** child nodes (nodes at depth $d + 1$) will only be constructed if all parent nodes (nodes at depth d) are already created.

Last, but not least, the BFS-LE tree creates one least cost path per tree node, and at each depth d at most $b(d)$ routes are calculated. Since the routes S have to be *unique, none-null routes*, not all calculated routes in the BFS-LE tree can be part of the set. Therefore, a route P^d of a sub-network G^d of the BFS-LE tree is added to the route set S only if the route is not empty (see Figure 5.2b) and it does not already exist in S (see Figure 5.2 where the routes assigned to S are marked with a grey background). Furthermore, since the addition of a path P^d to the route set S is dependent on the paths already assigned before, and these paths are dependent on the order of parsing through the edges e_i^d of a path P^d to create the child tree nodes, it is necessary to complete the whole tree at depth d before assigning the new paths to the set S . If the set S^d keeping

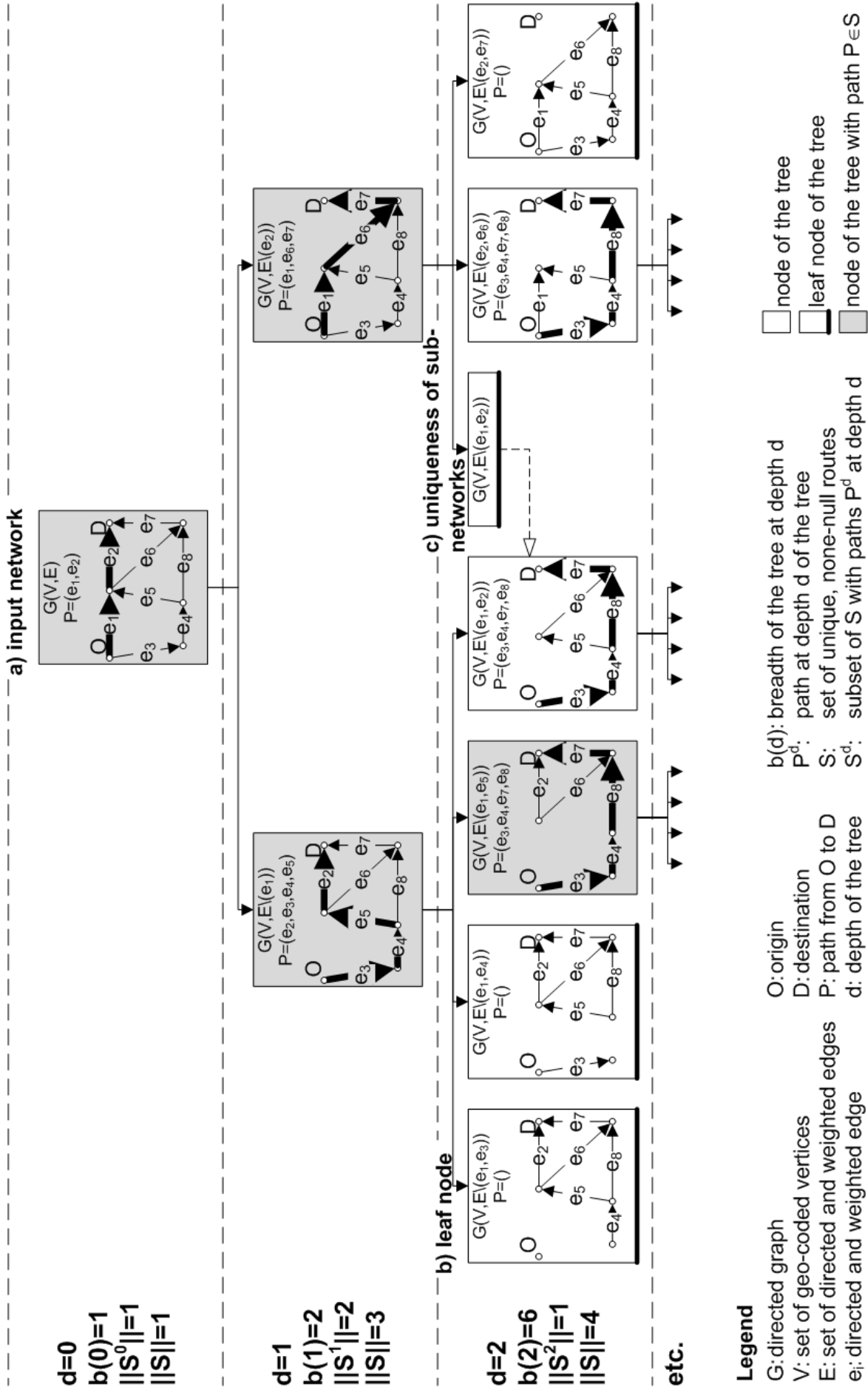


Figure 5.2: Basic BFS-LE tree: (a) is the root of the tree with the given network, (b) shows a leaf node (no path from v_O to v_D) and (c) is an example of a sub-network that is already present at depth d .

the disjoint paths P^d at depth d with $P^d \cap S = \emptyset$ contains more paths than necessary for a route set with size n then a random subset of S^d is assigned to S such that $\|S\| = n$. Otherwise the whole set S^d is added to S .

The complexity of the BFS-LE tree can be estimated via the breadth of the tree b at depth d , called $b(d)$. Since the number of child nodes of a node containing graph G^d is equal to the number of edges $e_i^d \in P^d$ while respecting *Uniqueness of sub-networks*,

$$b(d+1) \leq \sum_{j=1..b(d)} \|P_j^d\| \quad ; \quad \text{with } b(0) = 1. \quad (5.1)$$

Usually, in real, high-resolution networks a least cost path of an OD pair can easily contain many dozens of edges, which let $b(d)$ grow very fast for increasing depth d . Uniqueness of sub-networks helps to decrease that complexity, but does not typically prevent the production of a wide BFS-LE tree.

The BFS-LE route set generation method presented here produces n different paths from v_O to v_D if n paths exist. Otherwise, it will return the set of all possible paths from v_O to v_D . The algorithm guarantees the shortest paths P^0 to be part of S . Even more, each set S^d at depth d contains the $(d+1)$ th shortest path of the input network G . Therefore, assuming S^0, S^1, \dots, S^d are completely added to the resulting set S , then it contains at least the first $(d+1)$ shortest paths of network G .

While the algorithm performs well for a Manhattan-network it is also necessary to estimate *pathological cases for BFS-LE method*. For BFS-LE, the worst case happens if at each depth d only the $(d+1)$ th-shortest path is found. To create a route set S of size n for such a situation the BFS-LE algorithm needs to expand the tree until depth $n-1$. This pathological case happens only if all of the n -shortest paths of an OD pair are disjoint.

5.2.3 Performance optimisation 1: BFS-LE(PO1)

As mentioned above, the the order in which the links are processed can bias the outcome of the algorithm. Thus, it is necessary to create the whole set S^d of routes for level d before adding routes of level d to s . But this produces a fair amount of computational overhead. For example, the BFS-LE tree at depth d can already be very wide ($b(d) \gg 1$), but only one additional path of S^d needs to be added to S to reach route set size n . Furthermore, while the generation of G^d via link elimination is not costly

at all, the calculation of the shortest path – even with the *Landmarks A-Star* routing algorithm – is the most time-consuming part of the BFS-LE method. A simple, but efficient, way to reduce that computational overhead is to randomly *shuffle* the order for which tree node at depth d the shortest path is calculated. With that, it is not necessary to produce the whole breadth of the tree at depth d when the route set size is already reached and the resulting choice sets do not systematically differ from those generated with the non-optimised algorithm.

Performance gain is located only at depth d where route set size n is reached. For all previous depths, performance remains the same. That means if n is large, then the depth of the tree is large too, and the performance gain of that optimisation is negligible. But the typical route set to generate on a high-resolution network lies between 20 and 100 routes per OD pair, which typically ends up with a depth d of the BFS-LE tree between two to four. In such cases, the BFS-LE(PO1) method can efficiently reduce computational time.

5.2.4 Performance optimisation 2: BFS-LE(P02)

Based on BFS-LE(PO1), it is possible to increase computational performance even more by reducing the input network $G(V, E)$ to a *topologically equivalent network* $G' = G(V', E')$. To create G' , so-called “pass” vertices that do not model junctions, intersections or dead-ends, are removed from G and their incident edges are combined per direction. Figure 5.3 illustrates the procedure to generate G' . Therefore, $V' \subseteq V$ and $E' = E^r \cup E^m$. V' keeps the remaining vertices, called “non-pass” vertices or “non-pass” nodes in the subsequent analysis. E^r contains the set of untouched edges of E . E^m defines the set of merged edges $e_{ij..k} = \text{merge}(e_i, e_j, \dots, e_k); e_i, e_j, \dots, e_k \in E$. By assigning G' as the root of the BFS tree, the complexity of the tree is markedly reduced compared to the basic BFS-LE method, since some nodes of the tree at depth d are treated as one tree node and therefore $b(d)$ is reduced (i.e. edges e_3 and e_4 of G in Figure 5.2 will be combined to edge e_{34} and the first two tree nodes of depth $d = 2$ are treated only once which reduces $b(2)$ to 5). In order to ensure that the performance optimisation P02 does not change resulting route set, the route set still needs to be generated based on the networks G^d due to three reasons:

1. v_O and/or v_D can be part of the removed “pass” vertices and therefore are not part of V' .
2. The costs of merged edges can differ in case of non linear additive costs.

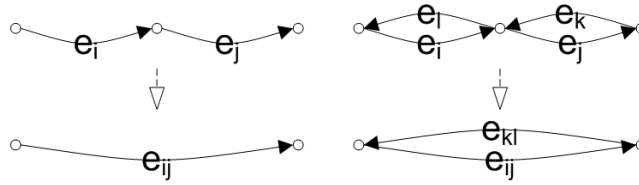


Figure 5.3: Examples of merging edges for producing a topologically equivalent network

3. The edges of the generated routes must be part of G^d .

In contrast to the BFS-LE(PO1), BFS-LE(PO2) can decrease computing time at each depth of the BFS tree. The performance gain strongly depends on (i) the number of edges that can be merged in G , (ii) the OD pair itself and (iii) the number of calculated paths of the BFS tree that contain vertices $v \notin V'$.

5.3 Computational performance

All tests described in this section were run on systems having two Dual-Core AMD Opteron Processors 2222 running at 3 GHz. The 4 GB allocated memory was connected through a front side bus clocked at 1,000 MHz. As the code was not multi-threaded, only one of the CPU cores was actually used by the tests. The performance analysis was conducted on a sub-sample of 500 OD pairs representative of the 36,000 main study OD pairs in terms of distance, main road type of the shortest path and network density at origin and destination. The main road type used by the shortest path was motorway for 9.1 %, extra-urban road for 11.6 %, urban main road for 62.2 % and local road for 7.3 % of the OD pairs. The main road type for the remaining 9.8% could not be decided because they used several road types in similar proportions. The shortest path distance distributions for each main road type showed reasonable patterns. The median of the shortest path distance was 18.93 km for motorway, 8.71 km for extra-urban road, 3.75 km for urban main road and 1.42 km for local road as the main road type. Median distance of trips that could not be categorised was 8.18 km. Concerning network density, a threshold of 15 nodes within a radius of 200 metres was chosen to distinguish between high and low density areas at the start and end points. The share for each of the resulting four density classes is about 25 %. This share, however, varies with the main road types. Extra-urban trips for instance

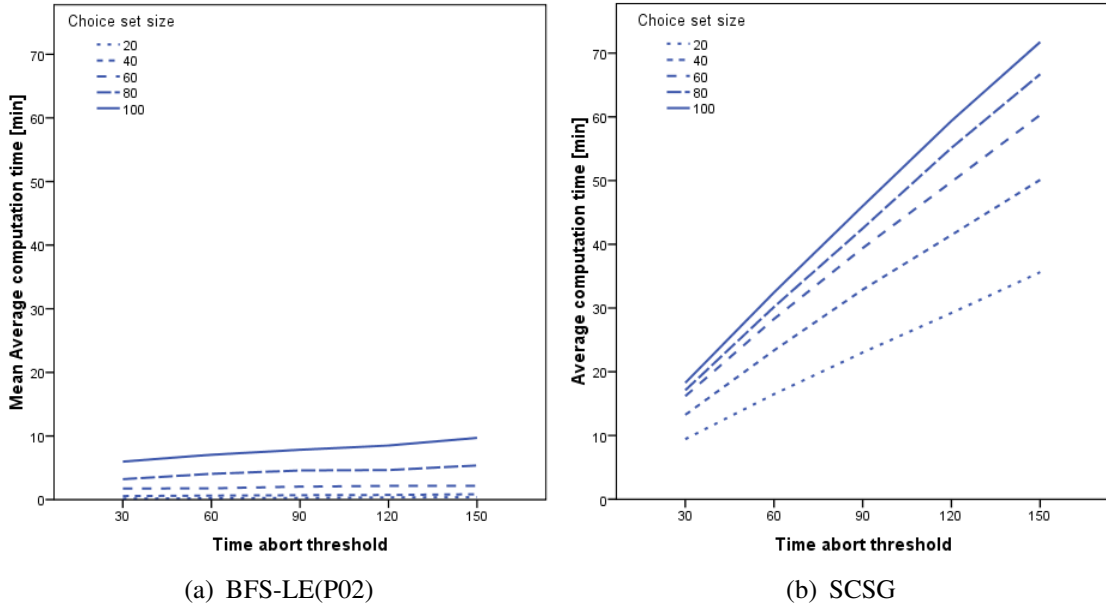


Figure 5.4: Average computation time depending on route set size and time abort threshold

tend to start and/or end in lower density areas whereas a higher share of trips using main urban main roads start and/or end in dense areas.

For performance evaluation of the BFS-LE algorithm, the BFS-LE(P02) version with both performance optimisation options (topologically equivalent network reduction and shuffling of the sub-network list at depth d) was employed in the performance analysis; whereas in the SCSG, a uniform distribution ranging from zero to twice the initial link costs was used. The computational performance was evaluated according to the following criteria:

- Route set sizes
- Time abort thresholds
- Least cost path distance
- Number of non-pass nodes on the least cost path

Figure 5.4 shows the computation time depending on the route set size and the time abort threshold. As expected, the computation time rises with increasing route set size and time abort threshold. The impact of the route set size is stronger than that of the time abort threshold. The average computation time of the SCSG is, on average, 32 times higher than those of the BFS-LE(P02) for several reasons. First, the SCSG reaches the time abort threshold much more often than the BFS-LE(P02). Second, the average time to adapt the network and determine the new

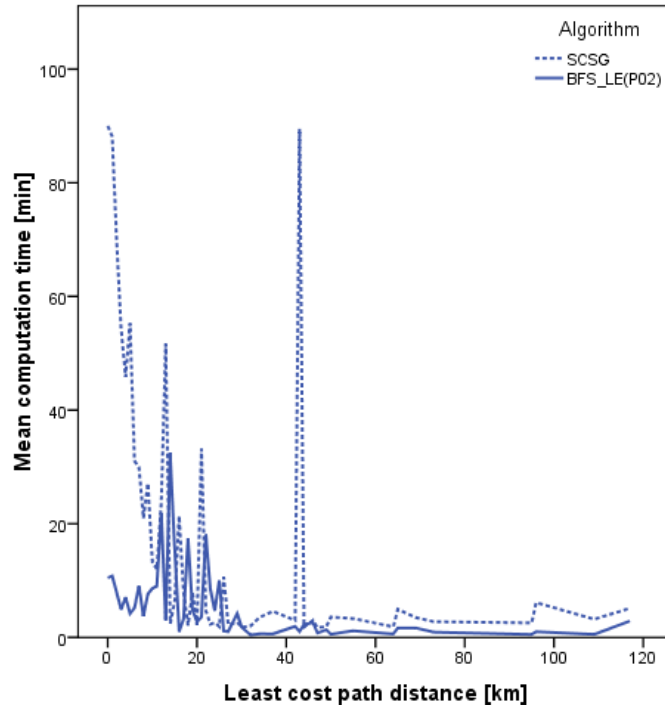


Figure 5.5: Computation time depending on the least cost path distance

shortest path is 0.45 seconds for the BFS-LE(P02) and 0.79 seconds for the SCSG. Third, the number of routes calculated to derive a route set size of 100 with a time abort threshold of 90 minutes adds up to 1039 on average for the BFS-LE(P02) and 3525 for the SCSG. However, it can also be seen in Figure 5.4 that the increase in computation time over route set size grows for the BFS-LE(P02) while it lessens for the SCSG. This originates again from the frequency the time abort threshold was reached. For the BFS-LE(P02), this number is very low in the beginning and rises strongly with increasing route set size. For the SCSG, this rise slows down for route set sizes of 80 and 100. Only a few additional OD pairs, that have not reached the abort threshold for smaller route set sizes, reach the time abort threshold for larger route set sizes.

As discussed in the previous section, the performance of both algorithms depends to a large degree on the network structure. Thus, the subsequent figures evaluate the computational performance with regard to characteristics of the OD pairs for the runs with a route set size of 100 and a time abort threshold of 90 minutes.

Figure 5.5 presents the computation time relative to the distance of the least cost path between each OD pair is shown. The BFS-LE(P02) clearly outperforms the SCSG for short trips, especially in the band under 10 km, covering 75 % of the car trips in this sample. The BFS-LE(P02) reveals unsteady behaviour only for OD pairs that are 10-30

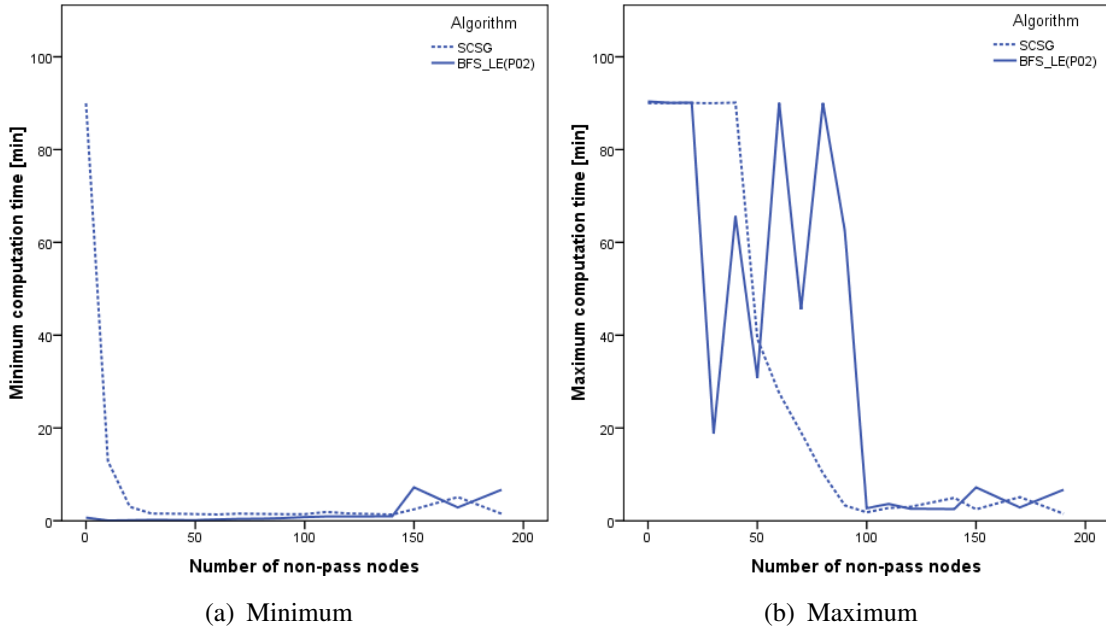


Figure 5.6: Computation time depending on the number of non-pass nodes on the least cost path

km apart. In Switzerland, this distance band typically contains trips between neighbouring agglomerations and has the highest probability of containing pathological cases as described in the previous section. For trips longer than 40 km the BFS-LE(P02) outperforms the SCSG again. However, these values have to be treated with care because they represent only few observations.

A closer look at the underlying networks structure was needed to explain the unsteady behaviour of the BFS-LE(P02) for medium-distance trips. This is done in Figure 5.6, depicting minimum and maximum computation time per OD pair depending on the number of non-pass nodes of the least cost path. The graphs reveal that, in principal, the computation time rises with an increasing number of non-pass nodes. This explains the good performance for short-distance trips. These trips take place in an urban or suburban environment where not much can and is gained by the topologically equivalent network reduction. The medium and long trips were more affected by the reduction, but each to a different extent, so that the number of non-pass nodes is only weakly correlated with distance. For those trips, the minimum computation time rises, but at the same time the probability of a pathological case, or at least reaching the time abort threshold, decreases. This leads to the unsteady performance behaviour for medium distance trips. Since this depends entirely on the network structure, there is, unfortunately, no algorithmic way to improve

the performance other than the time abort threshold used in this study.

5.4 Structure of the derived route sets

A thorough evaluation of route set generation approaches does not consider only their computational performance but also the structure and quality of resulting route sets. The guideline for this analysis were the following questions suggested by Bovy (2009):

1. Is the size of the route set sufficient?
2. How often/well is the chosen route reproduced?
3. How diverse are the routes?
4. How plausible is the hierarchical sequence?

In principal, the analyst defines for both algorithms what route set size is sufficient, and the algorithms search new routes until this route set size is reached or no further routes exist. Exceptions are OD pairs reaching the time abortion threshold. For the current settings, this occurs only for 6 % (at most) of the OD pairs with the BFS-LE(P02), but for up to 50 % of the OD pairs with the SCSG. The analyst has to decide how to treat the affected OD pairs in subsequent applications.

The reproduction of the chosen route was measured in two ways. First, the number of times the complete chosen route had been reproduced was counted. The BFS-LE(P02) achieved this for 63 % of the OD pairs with a route set size of 20 and for 73 % of the OD pairs with a route set size of 100. The respective figures of the SCSG were 64 % and 75 %. These figures are better than those reported for example by Ramming (2002) and Prato and Bekhor (2007b). With their link elimination algorithms, the chosen route was reproduced for 60 % of the OD pairs by Ramming (2002) and for 58 % of the OD pairs by Prato and Bekhor (2007b). For their stochastic choice set generation algorithms the respective figures were 38-50 % and 49-61 %. It has to be noted, though, that their choice set generation set-ups were different. Ramming (2002) performed only 48 shortest path searches for each OD pair and choice set algorithm and Prato and Bekhor (2007b) even less, namely 10.

In the second step, the route sets not containing the chosen route were examined more closely. For them, the overlap between the route that best reproduces the chosen route and the chosen route itself was calculated as percentage of the chosen route length. The resulting mean coverage of the chosen route varies between 76 % and 82 % for the BFS-LE(P02) and between 76 % and 78 % for the SCSG. Considering network resolution

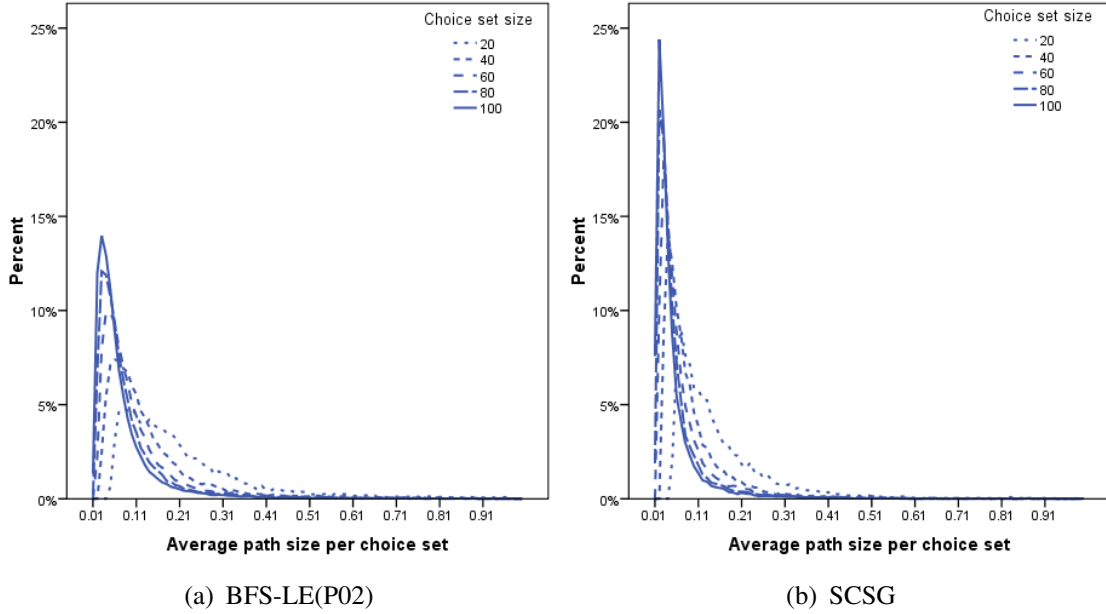


Figure 5.7: Path size distributions for resulting route sets

and number of possible paths between an OD pair, both algorithms are adequately able to reproduce the chosen route.

To investigate the diversity of the routes set, the path size PS_{in} for each route was determined using the well-known formulation of Ben-Akiva and Bierlaire (1999):

$$PS_{in} = \sum_{a \in \Gamma_i} \left(\frac{l_a}{L_i} \right) \frac{1}{\sum_{j \in C_n} \delta_{aj} \frac{L_{C_n}^*}{L_j}} \quad (5.2)$$

where Γ_i is the set of all links of path i , l_a the length of link a , L_i the length of path i and $L_{C_n}^*$ the length of the shortest path in C_n using link a . δ_{aj} equals one if link a is on path i and zero otherwise. The path size ranges between 0 and 1, with 0 indicating complete overlap and 1 no overlap at all. Figure 5.7 depicts the distribution of average path size of all routes in each route set. As expected, the overlap between the routes increases for both algorithms with route set size. For all route set sizes, however, the path size distribution indicates considerably more diversity between the routes generated with the BFS-LE(P02) than those generated with the SCSG.

The most ambiguous criterion for evaluation of the route set structure is the plausibility of the hierarchical sequence. Since this study is concerned with the generation of route sets for car route choice, this analysis focussed on the shares and sequences of road types. Four road-type categories were defined: motorways, extra-urban roads, urban main roads

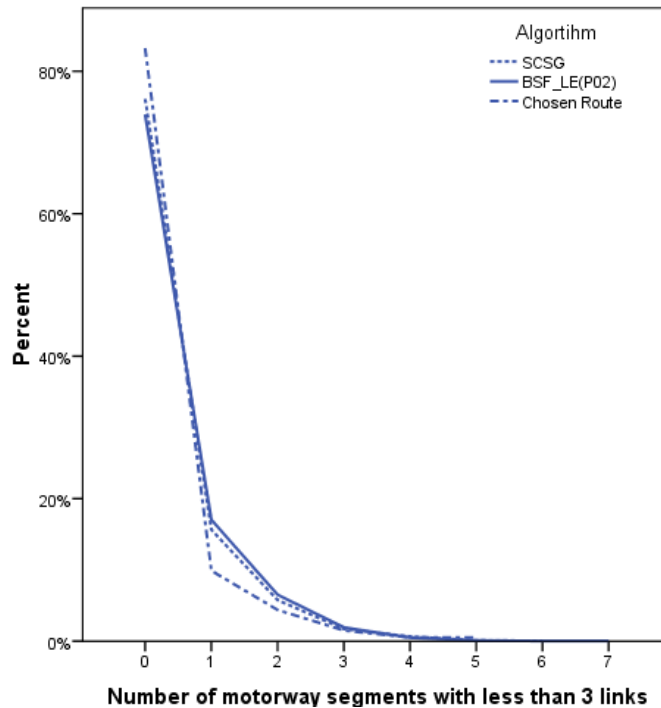


Figure 5.8: Number of motorway segments with less than 3 links

and local roads. The average share for each road type was calculated as a percentage of the total route length. The road type shares in the chosen routes were then compared to the average road type shares in the BFS-LE(P02) and the SCSG route sets. The analysis revealed, that road type shares were similar for the route sets of both algorithms and fairly stable over the different route set sizes. The chosen routes were travelled on average 12 % on motorways, 16 % on extra-urban roads, 36 % on urban-main roads and 36 % on local roads. The respective figures for the BFS-LE(P02) route sets are 16 %, 18 %, 35 % and 32 % and for the SCSG route sets 17 %, 18 %, 34 % and 31 %. This indicates that the routes generated by both algorithms reflect actual choice behaviour.

Bovy (2009)'s main concern about the plausibility of hierarchical sequences was the unrealistic shifting between different hierarchical levels. Given the structure of the Swiss road networks, repeated switching between extra-urban roads and urban main roads is inevitable for off-motorway trips between cities. Thus, unrealistic shifting between different road types is most likely to occur with motorways. Figure 5.8 shows the number of times a routes contains a motorway segment consisting of 3 or less links as percentage of the total number of routes containing at least one motorway segment. Both algorithms perform almost equally well with nearly 80 % of the trips using the motorway containing no segment consisting of 3 links or less, and over 90 % containing at most

one such motorway segment. The chosen routes, however, contain even fewer cases with repeated switching between the motorway and other road types. This implies that, in reality, repeated switching between motorways and other road types is rare but it does sometimes occur. Thus, this type of behaviour cannot be excluded a-priori as unrealistic.

5.5 Conclusion and outlook

The generation of choice sets in very high-resolution networks is a new challenge for transportation modellers. The number of routes in the universal choice set increases substantially while the size of the individual choice set, i.e. the number of relevant routes, probably remains the same. Thus, a large route set has to be extracted from the network, and afterwards reduced, to ensure that all relevant routes are included in the individual choice set. However, extracting routes from a high-resolution network is cumbersome and extremely time-consuming. Therefore, computational performance becomes a very important criterion in the evaluation of choice set generation algorithms again. Many of the advanced choice set generation algorithms presented recently are simply not applicable because they would run for weeks in order to generate choice sets for a few hundred OD pairs.

This chapter compares different route set generation algorithms and demonstrates that only approaches based on repeated shortest path search are applicable to high-resolution networks. A new algorithm is presented employing Breadth First Search on Link Elimination and includes two performance optimisation features: a randomisation of the processing order within each tree depth and a topologically equivalent network reduction. The computational performance of this BFS-LE(P02) algorithm and the quality of the resulting route sets is compared to the performance and results of a Stochastic Choice Set Generation (SCSG) algorithm.

In terms of computational efficiency, the BFS-LE(P02) clearly outperforms the SCSG. Particularly for typical urban trips under 10 km, the SCSG struggled to find enough routes to meet the required route set size, while the BFS-LE(P02) works most efficiently in this setting. Considering reproduction of the chosen route and road type composition, both algorithms perform almost equally well. The routes of the BFS-LE(P02) route sets are, however, more diverse than those of the SCSG route sets. Overall, the BFS-LE(P02) is clearly advantageous and can be recommended for generating route sets in high-resolution networks.

Another advantage of the BFS-LE(P02) is that it can use any cost

function specified by the analyst without changing the algorithm structure or computational performance. The only requirement is appropriate network information. One way to derive a more reasonable cost function would be to employ loaded dynamic travel time networks instead of free-flow times. This can be implemented straightforwardly because the whole algorithm, including the A-Star Landmarks router, is designed accordingly. Demonstrating this is a future research topic requiring additional data about network loads: for example, from a micro-simulation. Other cost functions could, for instance, be derived from the comparison between the attributes of the minimum time path and the chosen path.

The increasing use of GPS in transport surveys makes the issue of choice set generation in high-resolution networks more pressing. But it also opens up new ways to derive choice sets, namely the generation of individual choice sets from repeated GPS observations. This would allow the analyst to approximate the actual individual choice sets and allow valuable insights into the actual decision-process. Though a first attempt in this direction has been undertaken by Rich *et al.* (2007), more research is necessary before this approach can become state-of-the-art in route choice modelling.

Chapter 6

Accounting for Route Overlap in Urban and Suburban Route Choice Decisions Derived from GPS Observations

This chapter is based on the paper:

Schüssler, N. and K. W. Axhausen (2009a) Accounting for route overlap in urban and suburban route choice decisions derived from GPS observations, paper presented at the *12th International Conference on Travel Behaviour Research*, Jaipur, December 2009.

Two of the most prominent features of driver's route choice in an urban or suburban environment are the large number of alternatives and the similarity between route alternatives, both caused by the high density of urban and suburban street networks. The high density influences the way the choice set is established as well as the way the overlap between route alternatives is accounted for.

This chapter addresses both issues for the car route choice observations derived from GPS data. Different choice set generation procedures as well as choice set sizes are evaluated regarding their effect on the choice set composition and the resulting route choice models. In addition, the impact of different route attributes is investigated. The focus, however, is put on the analysis of the adjustment terms, that account for route overlap. Different formulations are tested in order to evaluate which mechanisms are at work in car route choice in an urban or suburban context.

6.1 Introduction and related work

Driver's route choice behaviour in an urban or suburban environment is shaped by a wide variety of factors. Two of the most prominent ones are the large number of alternatives available to the decision-maker and the similarity between route alternatives. Both are caused by the high density of urban and suburban street networks and both strongly influence and potentially bias the results of route choice models. Neither the decision-maker nor the analyst is able to evaluate the full set of alternatives, the universal choice set. Thus, the actual route choice and the route choice model have to be based on a subset of alternatives. The composition of the true choice set considered by the decision-maker depends on the technique he or she employed to extract routes from the network and the similarity between alternatives. However, how the decision-makers actually derive their choice set and what behavioural implications the similarity between alternatives has, is still an ongoing research issue. Previous literature often focusses either on the modelling of the choice set or on the treatment of similarities between alternatives. Only recently, some studies (e.g. Prato and Bekhor, 2007b; Bekhor *et al.*, 2006; Bliemer and Bovy, 2008) systematically investigated the interdependencies between these two aspects.

This chapter aims to continue this line of research. It examines the influence of different choice set generation approaches, similarity treatments, and their interdependencies for a car route choice model based on

very high-resolution data. The high level of detail amplifies several of the issues concerned with choice set generation (as discussed in Chapter 5) and similarity treatment since the choice set generation algorithms are more likely to produce routes with only slight deviations from each other, i.e. more overlap. In addition, it is investigated how a reduction of the choice set size, as suggested by Bovy (2009), following different paradigms influences the stability of the modelling results. Of course both questions cannot be addressed without appropriately accounting for similarities.

The rest of the chapter is structured as follows. After an overview about the related work, the modelling approach and the ways to account for similarities between route alternatives tested in this chapter are presented. Then, Section 6.4 introduces the techniques employed to reduce the choice set size and shortly characterises the choice sets used in the model estimation. Section 6.5 discusses the results of the model estimations followed by the conclusion and an outlook in Section 6.7.

6.2 Related work

Starting from the universal choice set U , the analyst has three different options to derive the individual choice set C_n of decision-maker n . The first option is to model the membership of each alternative i to C_n explicitly. Prominent examples for this approach are Swait (2001), Swait and Ben-Akiva (1987) or Morikawa (1996). They use the framework by Manski (1977) who stated that the probability that decision maker n chooses i from U depends on the probability $P(i|C_n)$ that he chooses i from C_n and the probability $P(C_n|U)$ that $C_n \subset U$ is his actual choice set. This approach, however, requires the analyst to enumerate all alternatives of U and all possible choice sets C_n . Moreover, its computational complexity is already too high for medium sized problems. Therefore, it is not applicable to route choice situation where the U is huge and unknown to the analyst.

The second option is the implicit modelling of inclusion in the choice set. Based on the assumption that the membership of an alternative to the individual choice set depends on its attributes, a continuous variable is added to the systematic part of the utility function of each alternative indicating its degree of inclusion in the choice set. This variable is either derived from assumptions about the availability or perception of an alternative (Cascetta and Papola, 2001) or from constraints which act as cutoffs and cannot be compensated by other attributes (Martínez *et al.*,

2009). These kinds of approaches are often interpreted as heuristic approximations of the explicit choice set models. However, a recent study by Bierlaire *et al.* (2009a) raised the question if the results of these approximations do indeed concur with the outcomes of an explicit choice set model or if they capture different mechanisms. Yet, the main problem for the application of the implicit choice set modelling approach to route choice modelling is again the underlying assumption that the universal choice set is known to the analyst.

Consequently, in route choice modelling mainly the third option of generating choice sets in a step prior to the modelling is used. First, a set of routes is extracted from the network with the aim to derive a master set M as exhaustive as possible in order to ensure that all relevant alternatives are detected. As discussed in Chapter 5, various approaches can be used for the route set generation. Subsequently, the individual choice set C_n can be obtained by reducing the route set M considering attractiveness, plausibility and overlap of the routes (Bovy, 2009).

Once the individual route choice set C_n is established, the choice of route i from C_n has to be modelled. This is usually done assuming Random Utility Maximisation (RUM). Various factors originating from attributes of the routes, traffic conditions, the environment around the route, the choice situation or the decision-maker can influence the route choice. One of the most prevalent factors, however, is the similarity between alternatives. Especially in urban or suburban networks the routes of a choice set can overlap extensively. As discussed in Chapter 2, there are three ways to account for similarities in route choice: imposing a nesting structure, explicitly modelling the correlation using multivariate error terms, or introducing adjustment terms in the deterministic part of the utility function.

The only nesting models that have been applied to route choice modelling are the Paired Combinatorial Logit (PCL) model (Chu, 1989; Koppelman and Wen, 2000) and the Link-Nested Logit (LNL) model (Vovsha *et al.*, 2002). Due to their computational complexity they have only been applied to small to medium size choice sets. The most extensive example documented in the literature is the application of the LNL to choice sets of up to 55 alternatives and up to about 900 links (Ramming, 2002; Bekhor *et al.*, 2006). The resulting matrix that describes the nesting structure is at least 1 order of magnitude smaller than it would be for the route choice problem at hand with up to 100 alternatives and up to 3600 links per choice set. Similar problems occur for the route choice models based on Probit (Yai *et al.*, 1997) or Mixed Multinomial Logit (MMNL) models (Ramming, 2002; Bekhor *et al.*, 2002). The only MMNL model

with a feasible computation time for a route choice model is the Subnetwork model by Frejinger and Bierlaire (2007). However, applying the Subnetwork model to very high-resolution data is computationally demanding and, thus, omitted in this study.

Thus, the only feasible way to account for similarities in route choice models based on high-resolution data is to employ adjustment terms. Since they are calculated prior to the model estimation they only lead to a small increase in estimation time. As discussed in Chapter 2, multiple adjustment term formulations have been proposed in the literature for route choice situations, several of which are evaluated in this chapter.

6.3 Modelling approach

The general model form used in this analysis is a linear in parameters Multinomial Logit (MNL) model employing the utility function proposed by McFadden (1974). For the basic model, without accounting for similarities, several variables and formulations were tested. Models with a travel impedance based on distance, free flow travel time and time-of-day dependent travel time were compared against each other. The time-of-day dependent travel time values were derived from a MATSim (MATSim-T, 2008) run of the study area. Similar to the setup reported in Balmer *et al.* (2009), the run included a demand relaxation process with routes and times employing the Charypar/Nagel utility function (Charypar and Nagel, 2005) for a synthetic population based on the Swiss Census 2000 (Swiss Federal Statistical Office, 2000) and the Swiss Navteq network. As expected, the models employing time-of-day dependent travel times resulted in the best model fit and the most precisely estimated travel time parameters.

Moreover, it was investigated if the sensitivity towards the travel time differs depending on the type of road the traveller is driving on, as previous studies (e.g. Bierlaire *et al.*, 2006) have shown. Therefore, each link of the network was assigned one of four major road types (motorway, extra-urban road, urban main road, and local road) based on the hierarchical road types coded in the Swiss Navteq network. For each route alternative, the amount of travel time spend on the respective road type was determined. A separate travel time parameter was estimated for each road type. The models using differentiated travel time parameters systematically outperformed models with only one travel time parameter and are, thus, used in the subsequent analysis.

In order to obtain unbiased travel time parameters, an additional

correction element for the perception of different travel times was introduced following the approach presented by Bierlaire *et al.* (2006). Instead of using a fixed value, the constants for each road type were weighted by the proportion of travel time spend on the respective road type relative to the total travel time of the route. The road type proportion values add up to 1. In the model estimation, the travel time proportion on extra-urban roads was defined as reference category to facilitate the interpretation of the model results.

Summarising these findings, for the basic model without treatment of similarities the deterministic part of the utility function is formulated as follows:

$$V_{in} = \beta_{ttMW} * tt_{MW} + \beta_{ttEU} * tt_{EU} + \beta_{ttUM} * tt_{UM} + \beta_{ttLR} * tt_{LR} \\ + \beta_{rtpMW} * rtp_{MW} + \beta_{rtpUM} * rtp_{UM} + \beta_{rtpLR} * rtp_{LR} \quad (6.1)$$

where tt_{MW} is the time-of-day dependent travel time on motorways, tt_{EU} is the time-of-day dependent travel time on extra-urban roads, tt_{UM} is the time-of-day dependent travel time on urban main roads, and tt_{LR} is the time-of-day dependent travel time on local roads. rtp_{MW} , rtp_{UM} , and rtp_{LR} are the proportions of travel time on motorways, urban main roads and local roads, respectively. The proportion of travel time travelled on extra-urban roads serves as reference category. The β are the parameters to be estimated.

In order to account for similarities between the routes due to route overlap, several adjustment terms have been evaluated

- two formulations of the C-Logit model introduced by Cascetta *et al.* (1996),
- two formulations of the Path Size logit model, first presented by Ben-Akiva and Bierlaire (1999),
- the Path Size Correction (PSC) term recently presented by Bovy *et al.* (2008), and
- two formulations of the trip part specific Path Size factor developed by Hoogendoorn-Lanser and Bovy (2007).

For the C-Logit model the formulations given in Equations 2.18 and 2.21 have been examined. However, the formulation in Equation 2.21 sometimes resulted in very large values and disproportionate differences in the Commonality Factor values for alternatives within the same choice

set, particularly if the choice sets were large. Consequently, the parameters for most other attributes were insignificant, suggesting that the similarity between the alternatives was the only decisive attribute. Since this was more founded on the numerical effect than on actual behaviour, the formulation was left out of the subsequent analysis.

The formulations of the Path Size factor and the Path Size Correction term are given in the Equations 2.22 and 2.23 and Equation 2.25, respectively.

The trip part specific Path Size factor was transformed into a road type specific Path Size factor. The experiments that led to formulation of the the basic model given in Equation 6.1 have demonstrated that road types are perceived differently and that the sensitivity to travel time changes depending on the road type. Thus, the hypothesis is that also the overlap on different road types has a varying influence on the choice probability of the alternative. In order to verify this hypothesis, the links belonging to a specific road type were defined as a trip part and models for the formulations given in Equations 2.26 and 2.28 were estimated.

Moreover, a sampling correction term is introduced for the choice sets generated with the Stochastic Choice Set Generation approach. According to Bovy *et al.* (2009), the Stochastic Choice Set Generation is a case of importance sampling because the probability that a route is selected in the choice set generation depends on its characteristics. Thus, routes with a higher probability to be chosen also have a higher probability to be included in the choice set. In order to account for the unequal selection probabilities Bovy *et al.* (2009) developed a Sampling Correction term SC_{in} for route choice sets originating from Stochastic Choice Set Generation. The Sampling Correction term also accounts for the influence of spatial overlap on the selection probability. SC_{in} is defined as:

$$SC_{in} = \ln \left(\frac{f_i}{Q_i} \right) \quad (6.2)$$

where f_i is the number of times, alternative i was drawn during the Stochastic Choice Set Generation and Q_i is the selection probability of i in the choice set generation. Q_i is calculated using Equation 6.3

$$Q_i = \frac{PS_i \cdot \exp(-cost_i/b)}{\sum_{j \in C_n} PS_j \cdot \exp(-cost_j/b)} \quad (6.3)$$

where $cost_i$ is the cost of route i and b is the positive variance parameter. Since a uniform distribution was employed in the Stochastic Choice Set

Generation in this study, b is determined differently than in Bovy *et al.* (2009). The variance of the link errors $V(\varepsilon_a)$ for a uniform distribution ranging from 0 to twice the costs $cost_a$ of link a equal

$$V(\varepsilon_a) = \frac{1}{12}(y - x)^2 = \frac{1}{3}cost_a^2 \quad (6.4)$$

Analogously to Bovy *et al.* (2009) it is, assumed that the route error variances are equal for all routes and can be determined from cost of the shortest path $cost_{min}$:

$$V(\varepsilon_r) = \sum_a \delta_{ar} V(\varepsilon_a) = \frac{1}{3} \sum_a \delta_{ar} cost_a^2 \quad (6.5)$$

where δ_{ar} equals 1 if link a belongs to route r and 0 otherwise. Then, b is defined as:

$$b = \sqrt{\frac{2 \cdot Cost_{min}}{\pi}} \quad (6.6)$$

6.4 Derivation of the choice sets

Since one of the aims of the modelling effort described in this chapter is to examine the influence choice set composition on route choice models, several different choice sets were created. First 1500 OD pairs were sampled from the main study OD pairs using the same sampling procedure as described in Section 5.3 that ensures the representativeness of the sampled OD pairs in terms of distance, main road type of the shortest path and network density at origin and destination. Then, with each of the two choice set generation procedure established in the last chapter, the Breadth First Search on Link Elimination (BFS-LE) and the Stochastic Choice Set Generation (SCSG), choice sets containing 20, 60 and 100 alternatives were produced. In case the chosen route was not reproduced by the choice set generation, it was added to the choice set. Third, the choice sets containing 100 alternatives are reduced following different paradigms in order to test Bovy (2009)'s recommendation of first establishing a master set that is as exhaustive as possible and then reducing this master set to the individual choice set taking into account attractiveness, plausibility and overlap of the routes. Four different choice set reduction procedures are investigated:

- Random reduction
- Similarity distribution-based reduction

- Similarity-based reduction
- Rule-based reduction

The most simple reduction procedure is the *random reduction*, where route are randomly removed from the master set until only the target number of alternatives is left. The only restriction in this procedure is that the chosen route has to remain in the choice set.

The next two reduction procedures use the overlap between the routes as criterion for the reduction. The objective of the similarity-distribution based reduction is to obtain choice sets with many different levels of overlap whereas the similarity-based reduction aims to derive choice sets that are as heterogenous as possible. Both procedures use the Path Size factor defined in Equation 2.23 as their overlap measure. For the *similarity-distribution-based reduction*, first the Path Size factor for each route in the complete choice set is calculated. Then the routes are assigned to previously defined similarity classes according to this Path Size factor. In this study, the interval of possible Path Size values between 0 and 1 was subdivided into 10 classes of equal width. Subsequently, the quota of routes for each class is calculated based on the number target routes, the number of classes and the assumption that each class should be equally represented. Then, from each class routes are randomly drawn until the class quota is met. If the number of routes in the choice set is lower than the target choice set size because some classes contain less routes than their quota, additional routes from the classes in which routes remained are drawn in random order until the target choice set size is reached.

The approach employed in the *similarity-based reduction* is based on the work by Kuby *et al.* (1997). The basic idea is to successively add routes of the master set to the choice set in a way that each new routes is least similar to the routes already contained in the choice set. First the chosen route is added to the choice set and removed from the master set. Second, the Path Size of each remaining route in the master set with respect to the routes of the choice set, and only those, is calculated. Third, the route with the highest Path Size is added to the choice set and removed from the master set. Subsequently, step two and three are repeated until the choice set has the target size.

The *rule-based reduction* evaluates certain characteristics of the alternatives and deletes all routes that violate at least one previously defined threshold. The thresholds can be defined in absolute terms or relative to the characteristics of other routes. In this chapter, four route characteristics are employed: length, travel time, plausibility, i.e. number of consecutive motorway links, and route overlap. For length and travel

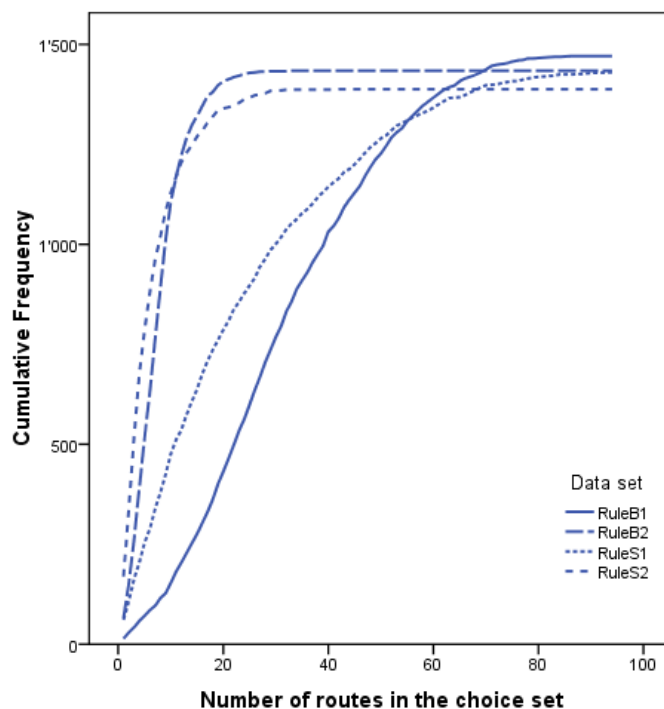


Figure 6.1: Cumulative distributions of choice set sizes resulting from the rule-based reduction

time thresholds relative to the best path according to the respective criteria are defined. If a route alternative is longer or slower than the threshold it is discarded. The thresholds for the plausibility and the overlap criterion are defined in absolute numbers. The only reasonable assumption concerning the plausibility of a route that could be made in this analysis is that people do rarely enter the motorway for only a small number of links, especially if there are other, shorter and quicker alternatives, available. Thus, all routes that contained motorway segments of less than three links were removed from the choice sets. Finally, all routes that overlapped with at least one route for a certain percentage of their own length, were discarded.

For each of the first three choice set reduction procedures, two target choice set sizes of 60 and 20 alternatives are defined. Analogously, two different levels of strictness are applied for the rule-based reduction procedure: a weak and a strong restriction level. The thresholds for the weak and strong restriction level for route length are 3.5 and 1.9 times the length of the shortest route. These values correspond to the 85 and 70 percentiles of the distribution of the ratios between the length of a route alternative and the length of the shortest path in the choice set. The respective thresholds for route travel time are 3.5 and 2 times the travel time of the quickest route. Regarding the route overlap, a maximum over-

Algorithm	Choice set size	Reduction procedure	Identification code
BFS-LE	20, 60, 100	–	B20, B60, B100
	20, 60	random	RandB20, RandB60
	20, 60	similarity distribution-based	SimDistB20, SimDistB60
	20, 60	similarity-based	SimB20, SimB60
	34, 87	rule-based	RuleB1, RuleB2
SCSG	20, 60, 100	–	S20, S60, S100
	20, 60	random	RandS20, RandS60
	20, 60	similarity distribution-based	SimDistS20, SimDistS60
	20, 60	similarity-based	SimS20, SimS60
	43, 95	rule-based	RuleS1, RuleS2

Table 6.1: Choice sets used in the model estimation

lap of 90% at the weak restriction level and 70% at the strong restriction level were tested. Other than for the first three reduction procedures, the choice set sizes resulting from the rule-based reduction cannot be predicted. They vary depending on the composition of the individual choice sets. Figure 6.1 shows the distribution of choice set sizes for the two levels of strictness and the two choice set generation procedures used to produce the master sets of 100 alternatives. It can be seen that especially the strong restrictions result in relatively small choice sets, most of them containing less than 20 alternatives, whereas the weaker restrictions lead to more variety regarding the choice set sizes.

Summarising, the overall 22 different choice set types are outlined in Table 6.1. Each choice set type is assigned a code that is used to identify the choice sets in the subsequent analyses. The choice set sizes for the rule-based reduced choice sets represent maximum choice set sizes whereas the choice set sizes for all other choice sets stand for the target choice set sizes. Moreover, the following two figures give a first insight into the compositions of the different choice sets used in the subsequent model estimations. To indicators are employed to investigate the choice set structures: the distributions of travel times and path overlap in the choice sets.

The travel time distributions for each choice set are presented in Figure 6.2. Overall, the mean travel time varies between 8.4 and 12.9 minutes and the median between 6.33 and 10.4 minutes. For both choice set generation algorithms the mean and average travel time increase with increasing choice set sizes. The choice sets originating from the BFS-LE algorithm contain longer trips and more outliers than SCSG choice sets. On the one hand, this might be caused by the removal of crucial links from the network resulting in unrealistic routes. On the other hand,

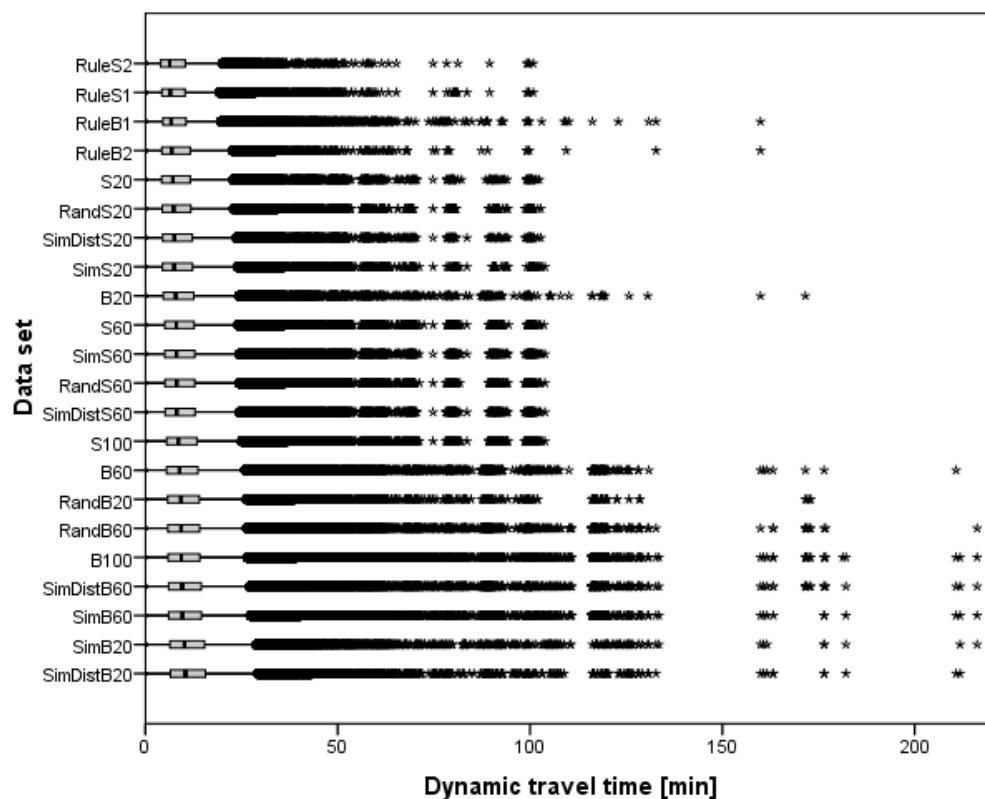


Figure 6.2: Travel time distributions for the different choice sets

the BFS-LE choice sets contain overall more routes because the algorithm reached the time abort threshold less often and additional routes tend to be longer by design. There is, however, no systematic change in the travel time distributions caused by the reduction procedures. Usually, the median travel time decreases slightly after the reduction, but there are also several cases where the median travel time increases and the travel time distribution becomes wider. Only the rule-based reduction leads to a distinct decrease in the mean travel time and a narrower travel time distribution.

In order to evaluate the route overlap, Figure 6.3 shows the distribution of the Path Size factor defined in Equation 2.23 for each choice set. It can be seen that the level of route overlap differs considerably for the different choice set generation algorithms, reduction procedures and choice set sizes. Unsurprisingly, smaller choice set sizes lead to higher Path Size factors, i.e. smaller overlap, regardless of the procedure used to derive them. The choice sets originating from the BFS-LE algorithm contain less overlap than those from the SCSG algorithm, even after the different reduction procedures were applied. Moreover, all choice set reduction procedures, except the random reduction, substantially increase the Path Size. The reduced choice sets contain even less overlap than the

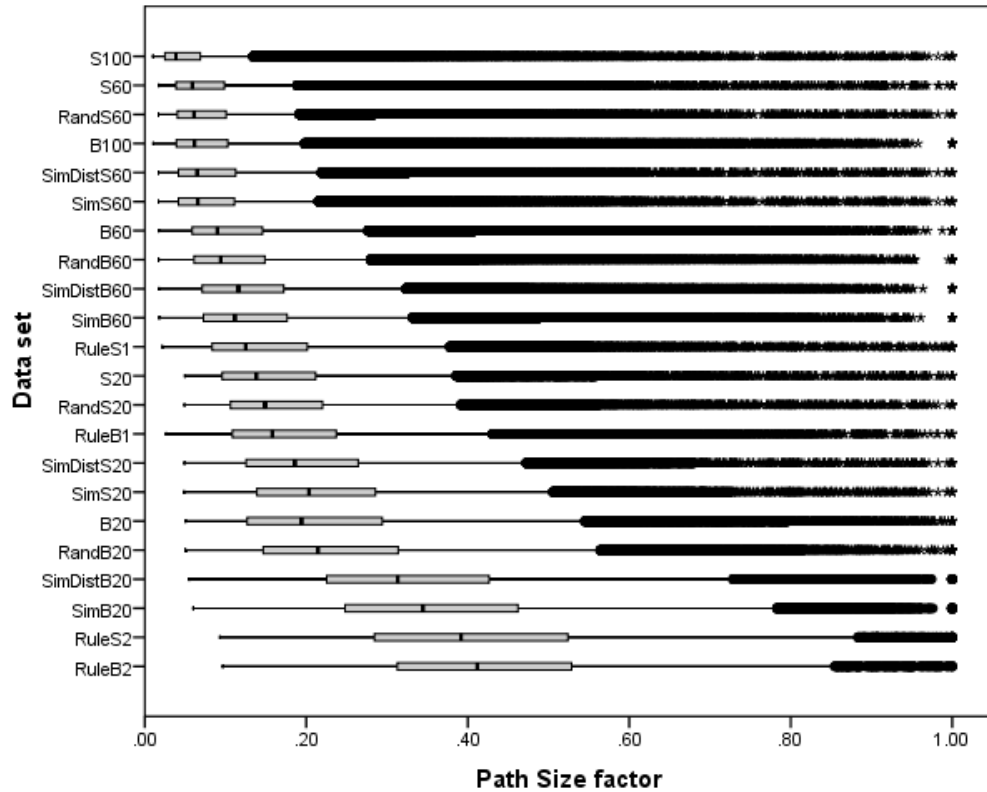


Figure 6.3: Path Size distributions for the different choice sets

choice sets of the same size derived directly from the choice set generation algorithms. This effect is strongest for the rules-based reduction and weakest for the similarity distribution-based reduction.

6.5 Modelling results

This section describes the choice models that have been estimated for different choice sets and adjustment terms. While the choice sets are outlined in Table 6.1, Table 6.2 gives an overview of the 17 models tested for each choice set type. First, a basic model without similarity treatment was estimated. Then, the six adjustment terms established in Section 6.3 were added with a logarithmic transformation, as suggested by the literature. An exception is the PSC factor by Bovy *et al.* (2008), where the logarithm is already part of the adjustment term. Afterwards, it was examined whether the natural logarithm is indeed the best transformation. Therefore, additional models were estimated with no transformation or a BoxCox transformation for all adjustment terms except the PSC factor. All models were all estimated using BIOGEME (Bierlaire, 2009).

For the choice sets derived from the Stochastic Choice Set Generation

Name	Adjustment term	Transformation
Basic model	–	–
PSC	Path Size Correction as in Equation 2.25	–
ln(PS1)	Path Size as in Equation 2.22	ln
ln(PS2)	Path Size as in Equation 2.23	ln
ln(CF1)	Commonality Factor as in Equation 2.18	ln
ln(PSRT1)	Road type specific Path Size as in Equation 2.26	ln
ln(PSRT3)	Road type specific Path Size as in Equation 2.28	ln
PS1	Path Size as in Equation 2.22	–
PS2	Path Size as in Equation 2.23	–
CF1	Commonality Factor as in Equation 2.18	–
PSRT1	Road type specific Path Size as in Equation 2.26	–
PSRT3	Road type specific Path Size as in Equation 2.28	–
BoxCox(PS1)	Path Size as in Equation 2.22	BoxCox
BoxCox(PS2)	Path Size as in Equation 2.23	BoxCox
BoxCox(CF1)	Commonality Factor as in Equation 2.18	BoxCox
BoxCox(PSRT1)	Road type specific Path Size as in Equation 2.26	BoxCox
BoxCox(PSRT3)	Road type specific Path Size as in Equation 2.28	BoxCox

Table 6.2: Models tested for each choice set type

(SCSG) additional models were tested employing the Sampling Correction (SC) term proposed by Bovy *et al.* (2009). Table 6.3 compares the results with and without the SC term for the basic model without correction for route overlap and for the choice sets directly derived from the SCSG. In these choice sets, the routes are highly correlated, especially those using mainly urban main roads. Thus, without similarity treatment, the travel time parameter for urban main roads can be positive. If, however, the SC term is added, also the travel time parameters for motorways and extra-urban roads turn positive. This effect can only be partially corrected by adding adjustment terms to the utility function and also occurs, though a little weaker, for the reduced choice sets with less correlation. Thus, the SC term is left out of the subsequent analysis.

With regard to the treatment of similarities between route alternatives, Figure 6.4 shows the impact of different adjustment terms formulations and transformations on the values of the road type specific travel time parameters for the unreduced choice sets. It can be seen that the absolute values of the travel time parameters increase with increasing choice set size and also with the inclusion of an adjustment term. While no clear trend could be detected concerning the transformation of the adjustment term, the influence of the travel time for all road types tends to be lower for models employing road type specific Path Size factor formulations.

Moreover, as can be seen in Figure 6.5, the various adjustment terms

Parameter	S20	S60	S100	S20	S60	S100
Travel time on MW	-0.13 (3.02)	-0.37 (5.44)	-0.48 (6.64)	0.14 (-0.56)	0.11 (-0.31)	0.13 (-0.59)
Travel time on EU	-0.09 (2.43)	-0.34 (6.05)	-0.46 (7.59)	0.13 (-1.38)	0.10 (-0.86)	0.10 (-0.80)
Travel time on UM	0.14 (-3.18)	0.00 (0.86)	-0.06 (2.72)	0.38 (-7.65)	0.42 (-9.32)	0.43 (-9.84)
Travel time on LR	-0.49 (6.95)	-0.95 (11.90)	-1.16 (14.16)	-0.26 (4.41)	-0.29 (4.72)	-0.32 (5.02)
Road type perc. MW	4.68 (-2.72)	4.86 (-2.72)	5.01 (-2.83)	4.31 (-2.01)	4.62 (-2.24)	4.75 (-2.48)
Road type perc. UM	3.85 (-4.36)	2.62 (-2.86)	2.19 (-2.22)	4.25 (-4.65)	3.82 (-4.23)	3.74 (-4.21)
Road type perc. LR	0.71 (0.13)	-0.25 (1.32)	-0.35 (1.44)	3.78 (-3.52)	3.34 (-3.07)	3.34 (-3.12)
Sampling Correction				0.80 (-28.61)	0.80 (-33.85)	0.80 (-35.64)
Initial LL	-4065	-5218	-5668	-4065	-5218	-5668
Final LL	-3660	-4629	-5002	-3049	-3788	-4078
Adjusted Rho square	0.10	0.11	0.12	0.25	0.27	0.28

(*) parameter not significant at 95% confidence level

MW = motorway, EU = extra-urban road, UM = urban main road, LR = local road

Road type percentage are estimated with extra-urban roads as reference category

Table 6.3: Parameters of the basic model for unreduced SCSG choice sets with and without Sampling Correction

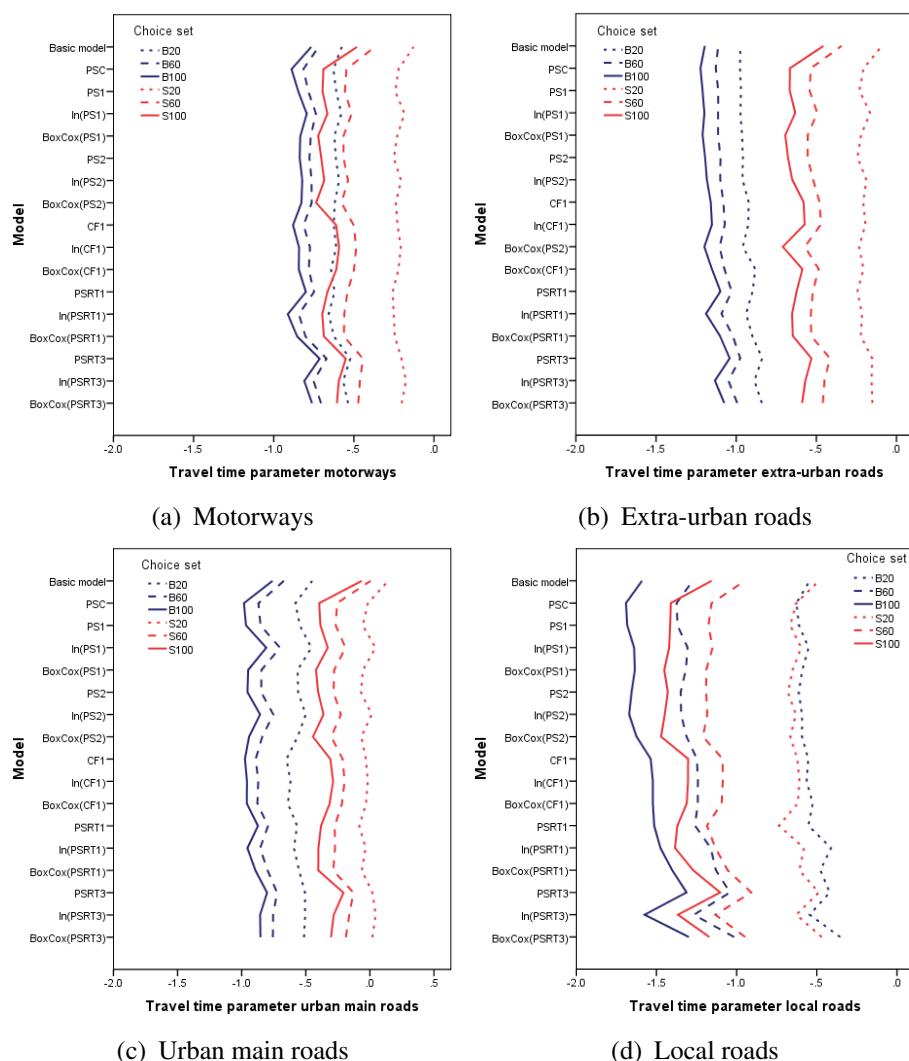


Figure 6.4: Travel time parameters for models estimated on unreduced choice sets

affect the travel time parameters for each road type differently. Most pronounced is the difference between the basic model and the similarity corrected models for both choice sets with 100 alternatives for the urban main road parameter. Relative to the extra-urban main road parameter the travel time on urban roads is punished more with similarity treatment, whereas the ratio between the travel time on motorways and the travel time parameter on extra-urban roads remains rather stable throughout all models. Furthermore, Figure 6.5 shows, that the differences between the road types are more distinct for the choice sets based on the SCSG than for the one based on the BFS-LE.

In order to determine the best transformation for each adjustment term, the adjusted rho square values for the same models are presented in Table 6.4. Note that the adjusted rho squares are only comparable for

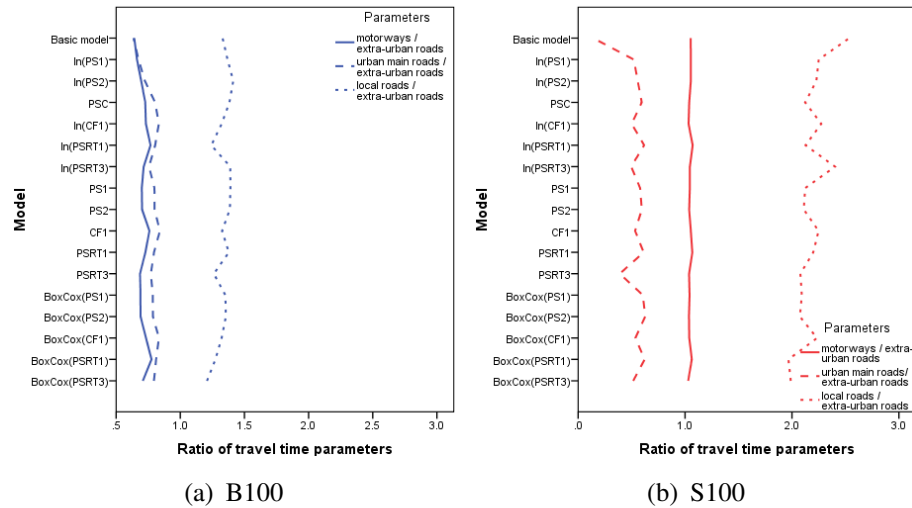


Figure 6.5: Ratios of the travel time parameters relative to extra-urban roads for unreduced choice sets with 100 alternatives

Model	B20	B60	B100	S20	S60	S100
no transformation						
Basic model	0.12	0.19	0.21	0.10	0.11	0.12
PSC	0.13	0.20	0.22	0.11	0.12	0.13
PS1	0.13	0.20	0.23	0.11	0.13	0.14
PS2	0.13	0.20	0.24	0.11	0.13	0.14
CF1	0.15	0.21	0.23	0.11	0.12	0.13
PSRT1	0.15	0.22	0.26	0.13	0.15	0.16
PSRT3	0.16	0.22	0.25	0.13	0.15	0.16
ln transformation						
ln(PS1)	0.12	0.19	0.22	0.10	0.12	0.13
ln(PS2)	0.13	0.19	0.22	0.10	0.13	0.13
ln(CF1)	0.15	0.21	0.24	0.11	0.12	0.13
ln(PSRT1)	0.15	0.21	0.24	0.11	0.13	0.14
ln(PSRT3)	0.14	0.20	0.23	0.11	0.13	0.14
BoxCox transformation						
BoxCox(PS1)	0.14	0.21	0.24	0.11	0.13	0.14
BoxCox(PS2)	0.14	0.21	0.24	0.11	0.13	0.14
BoxCox(CF1)	0.15	0.21	0.24	0.11	0.12	0.13
BoxCox(PSRT1)	0.16	0.23	0.26	0.14	0.16	0.17
BoxCox(PSRT3)	0.16	0.22	0.25	0.13	0.15	0.16

Table 6.4: Adjusted rho squares for the unreduced choice sets

models estimated from the same choice sets and cannot be used to evaluate the quality of the choice sets themselves. The models applying a BoxCox transformation performed for all choice sets and all adjustment terms at least equally well as, if not better than, the logarithmic transfor-

mation or no transformation at all. Moreover, the BoxCox transformed models, unlike the other models, delivered stable and reasonable results with respect to the travel time parameters for the individual road types over all choice sets and with regard to the overlap punishment for the different road types in the road type specific Path Size factors. Only for the models employing BoxCox transformed adjustment terms, all travel time parameters were negative for all choice sets and the relative order of the overlap punishment remained the same. Thus, only the BoxCox adjustment terms are considered in the subsequent analysis. As explained above, the PSC factor is excluded from this rule because here the transformation happens within the calculation of the adjustment term itself.

Before the impact of the different choice sets on the modelling results is scrutinised, the relative order of the overlap punishment for the different road types is examined by calculating magnitude of the utility correction for each road type depending on the degree of overlap measured by the road type specific Path Size. In Figure 6.6 the results of this calculation are exemplarily depicted for the first formulation of the road type specific Path Size (PSRT1) and the unreduced choice sets. For all these choice sets, and all other choice sets, overlap on motorways is punished the least and overlap on urban main roads the most. In between, overlap on extra-urban roads is punished considerably less than overlap on local roads. This is reasonable since the motorway network in Switzerland is not very dense. Thus, it is difficult to find alternative routes that use a different motorway (segment) for the same OD relation. A similar logic applies to extra-urban roads that connect built-up areas and where avoiding these routes would lead to large detours. In built-up areas, on the other hand, small detours are more easily available but also easily overlooked. Therefore, the difference between the large number of objectively available alternatives and the number of subjectively perceived alternatives is higher, especially concerning urban-main roads. People usually navigate using main roads as major decision-criterion and take always the same minor roads to access the main roads. Thus, alternatives using the same part of the main road network often collapse to one alternative in the perception of the traveller.

The impact of the random choice set size reduction on the travel time parameters can be seen in Figure 6.7 where the parameter values are compared to those for the unreduced choice sets with 100 alternatives. For all randomly reduced choice sets the absolute values of the travel time parameters decrease with decreasing choice set size though this decrease is less strong than for the unreduced choice sets of the same size. The same applies for the parameters of the adjustment terms. Apparently,

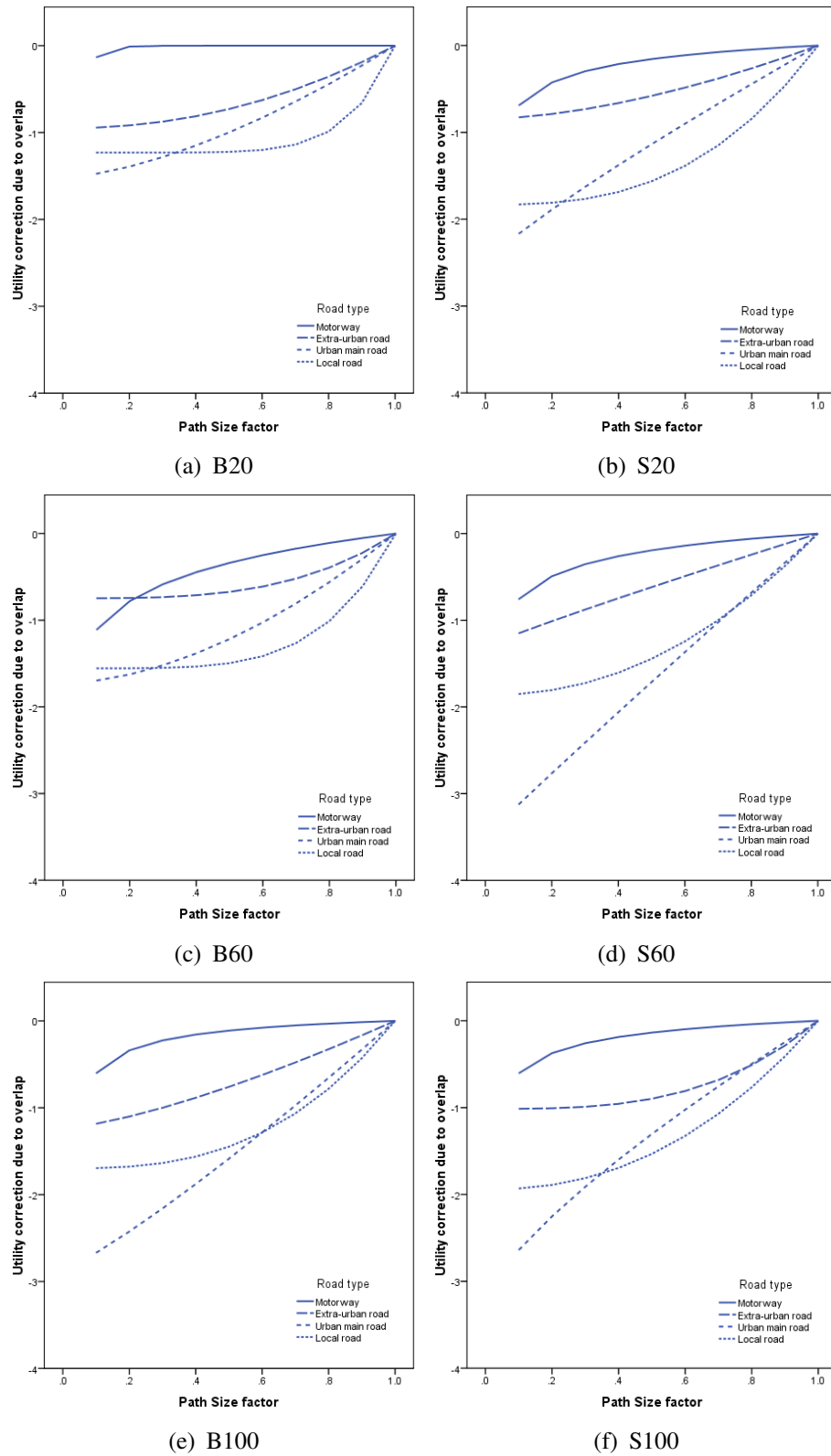


Figure 6.6: Utility correction for route overlap per road type depending on the degree of overlap

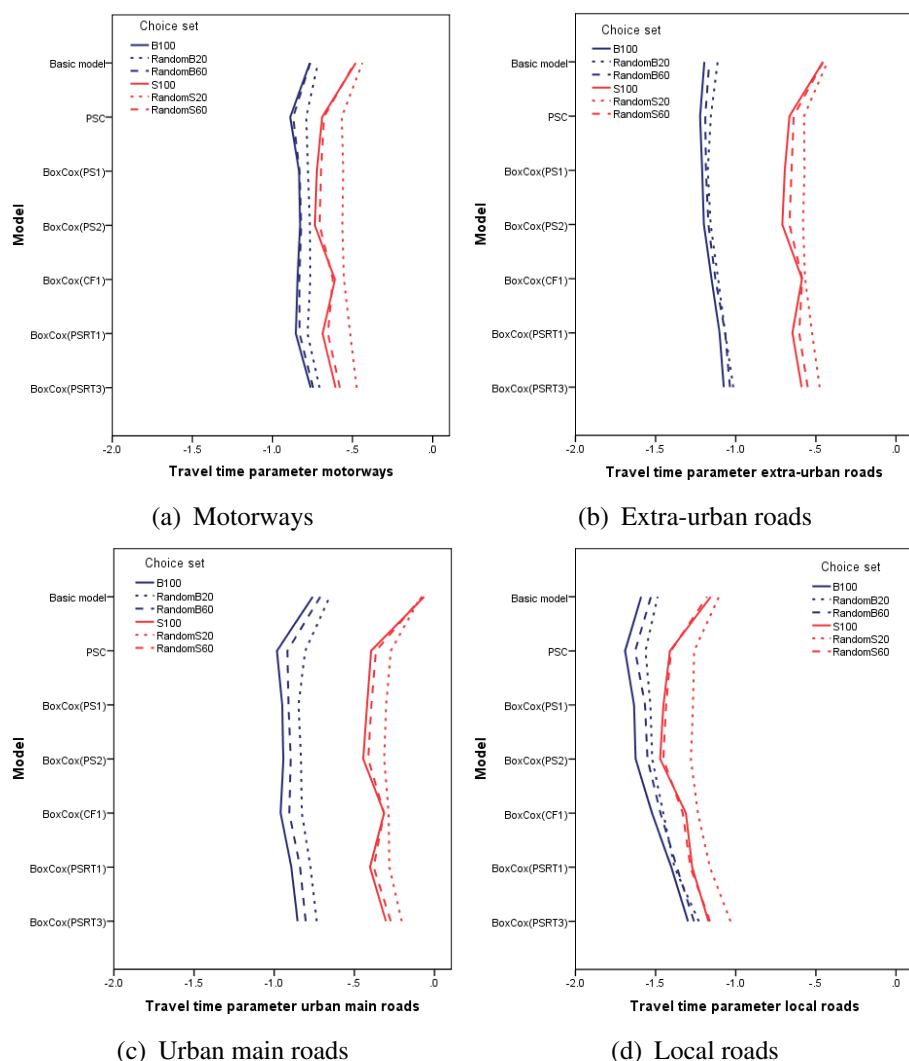


Figure 6.7: Travel time parameters for models estimated on randomly reduced choice sets

the generation of the full choice sets of 100 alternatives and the subsequent reduction causes an approximation of the parameters to those of the full choice sets. Another effect is a smoothing of the parameter values with respect to the adjustment term. The travel time parameters vary less depending on the different adjustment terms than they do it for the unreduced choice sets in Figure 6.4.

For the choice sets obtained from the similarity distribution-based reduction, the travel time parameters approximate the values for the unreduced choice sets with 100 alternatives even closer as presented in Figure 6.8. Also the smoothing effect is more distinct. Sometimes the absolute value of the travel time parameter for the reduced choice sets is even higher than the one for the full choice sets, for example for the SimDistS20 choice sets and the models employing BoxCox transformed

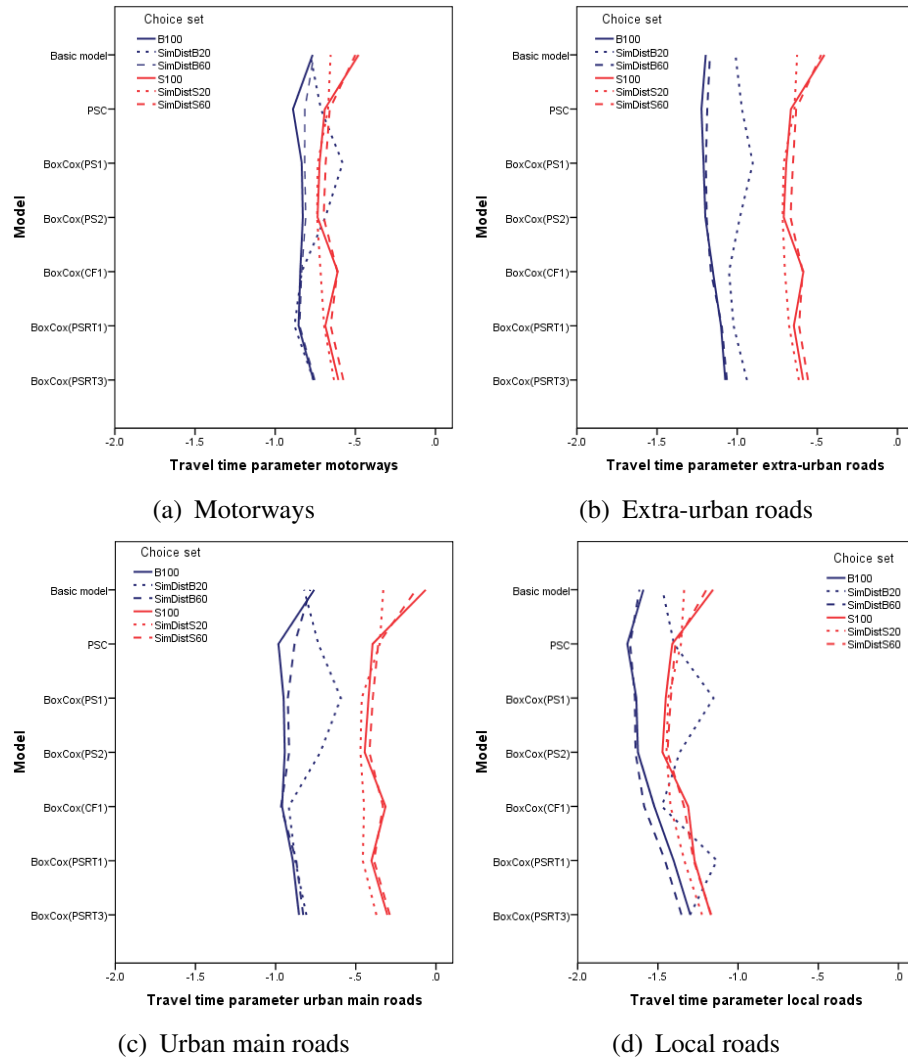


Figure 6.8: Travel time parameters for models estimated on similarity distribution-based reduced choice sets

CF1, PSRT1 or PSRT3 adjustment terms. An exception to this trend are the SimDistB20 choice sets. In these choice sets, the route overlap has been reduced substantially as shown in Figure 6.3. This leads to negative beta parameters for the adjustment terms for PS1, PS2, PSC implying that route overlap would have a positive influence on the utility of an alternative. This in turn leads to a decrease in the absolute values of the travel time parameters on the individual road types. The BoxCox(PSRT1) model, on the other hand, indicates that the positive influence of the route overlap only originates from local roads. In this model, only the parameter for the local road Path Size is negative. Accordingly, the absolute value of the travel time parameter for travel on local roads is also lower than the one estimated for the unreduced choice sets while the travel time parameters for the other road types are close to the unreduced

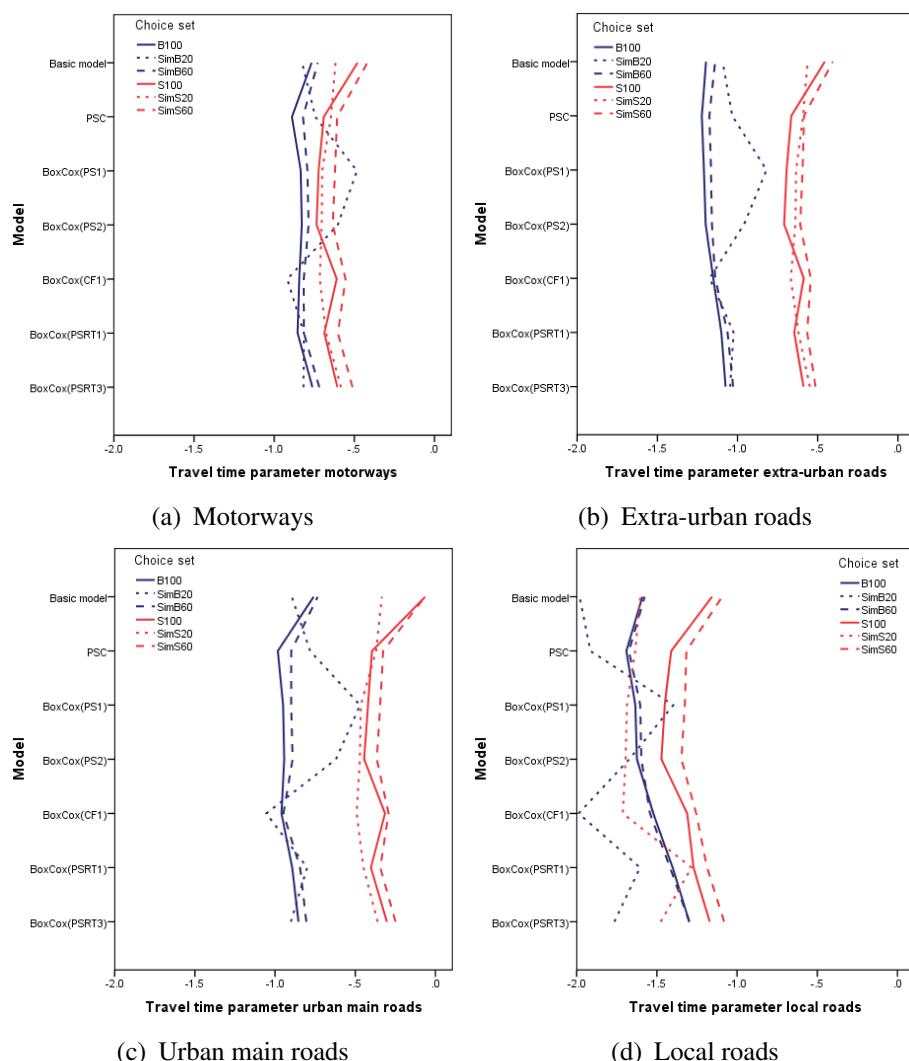


Figure 6.9: Travel time parameters for models estimated on similarity-based reduced choice sets

ones.

Figure 6.9 depicts the travel time parameters for the similarity-based reduced choice sets. The parameters for the SimB60, SimS20 and SimS60 choice sets show a similar behaviour as the corresponding choice sets derived through similarity distribution-based reduction. Though the absolute values of the parameters for the similarity-based choice sets with 60 alternatives are a little bit lower than those for the similarity-distribution based choice sets, the overall approximation and smoothing trend can be confirmed. Exceptions to this are the local road travel time parameter for the SimS20 choice sets and all parameters for the SimB20 choice sets. The local road parameter for the SimS20 choice sets reveals a higher sensitivity to travel time than in any other model estimated for on choice sets derived from the SCSG. The patterns of the travel time

parameters estimated from the SimB20 model resemble those estimated from the SimDistB20 model. They do, however, jump around even more with respect to the different forms of similarity correction, especially the local road parameter. The most stable parameters, and the ones closest to the full choice sets, deliver again the models with road type specific Path Size factors. In the PSRT1 model, not only the local road Path Size parameter is negative but also the urban main road Path Size parameter. This leads to the conclusion that the similarity-based reduction procedure, and to a lesser extent also the similarity distribution-based reduction procedure, does not remove the correlation between the routes uniformly over all route types but stronger for road types higher in the hierarchy, probably because they usually make up the longer part of a route. Moreover, correlation on local roads at the start and end of a trip is sometimes unavoidable because there are only a few local roads leading from the origin or to the destination.

As shown in Figure 6.3, the rule-based reduction procedures, especially the one employing the stricter parameters, were most efficient in the reduction of route overlap in the choice sets. Moreover, unlike the other choice set reduction procedures, they reduced the travel times considerably. This is also reflected in the travel time parameters estimated for these choice sets that are presented in Figure 6.10. Especially the absolute values of the travel time parameters for the stronger reduced choice sets RulesB2 and RulesS2 are substantially lower than those for the other choice sets because the travel time plays a less decisive role if the participant has to choose from a choice set in which the travel times are rather similar. Thus, the travel time parameters are not really comparable to the ones estimated from other choice sets. It can, however, be seen that the travel time parameters are more stable with respect to the similarity treatment than they are for other choice sets with comparably low correlation. They all follow a similar pattern that is different from the patterns observable for other choice sets. Moreover, the distance between the travel time parameters for these two choice sets is considerably smaller than it is between other choice sets with corresponding choice set size and reduction procedure but produced with different choice set generation algorithms. This implies that the choice sets somehow converge in their characteristics. An approximation in the distribution of travel times and route overlap can already be observed in Figures 6.3 and 6.2. It has to be noted, though, that the BoxCox(PSRT1) model did not converge for the RulesB2 choice sets. The values for the logarithmic transformed PSRT1 model (LN(PSRT1)) are depicted here instead because it outperformed the model without transformation.

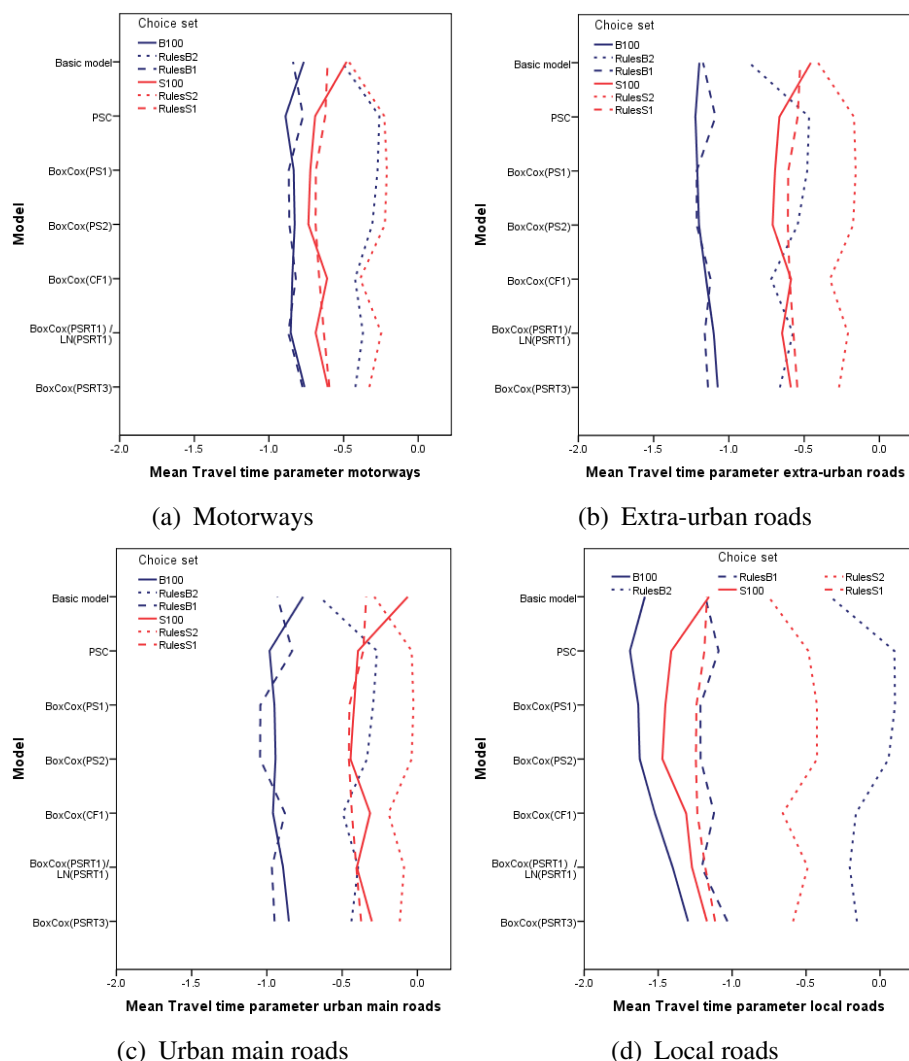


Figure 6.10: Travel time parameters for models estimated on rule-based reduced choice sets

Summing up, Table 6.5 shows the estimation results for the Box-Cox(PSRT1) model for the choice sets that turned out to be most appropriate in this study. All the travel time parameters have the right sign and lie within a similar range. Travel time on motorways is punished least and travel time on local roads most. Of the parameters indicating a preference for a particular road type only the local road parameter is in some models significant implying that the road type specific aspects are to a large extent covered in the road type specific travel times and the road type specific Path Sizes. In contrast to the results for the unreduced choice sets, the parameters for the motorway and extra-urban road Path Size factors are rarely ever significant. This confirms that in the reduction procedure the overlap on these road types has been removed to some extent while there is still substantial overlap on urban main roads

Parameter	B100	SimDistB60	RulesB1
Travel time on MW	-0.85 (11.52)	-0.85 (11.59)	-0.87 (-0.87)
Travel time on EU	-1.10 (17.31)	-1.00 (17.27)	-1.16 (18.38)
Travel time on UM	-0.89 (24.22)	-0.87 (24.24)	-0.97 (26.15)
Travel time on LR	-1.40 (17.62)	-1.46 (18.12)	-1.21 (15.52)
Road type perc. MW	0.71 (0.48)	0.62 (0.54)	0.02 (0.99)
Road type perc. UM	0.60 (0.11)	0.40 (0.42)	0.15 (0.78)
Road type perc. LR	-1.12 (2.45)	-1.23 (2.56)	-3.14 (4.90)
Beta PSRT1 MW	0.26 (-0.48)	0.30 (-0.37)	0.85 (-0.79)
Lambda PSRT1 MW	-0.20 (1.79)	-0.20 (1.63)	0.58 (-0.18)
Beta PSRT1 EU	1.21 (-1.90)	1.80 (-1.58)	2.57 (-1.13)
Lambda PSRT1 EU	0.92 (-1.78)	1.51 (-1.62)	2.56 (-1.23)
Beta PSRT1 UM	3.39 (-5.68)	4.84 (-4.86)	6.10 (-3.26)
Lambda PSRT1 UM	0.97 (-8.08)	1.58 (-7.70)	2.69 (-5.31)
Beta PSRT1 LR	4.01 (-6.20)	6.16 (-4.94)	3.77 (-4.21)
Lambda PSRT1 LR	2.15 (-7.73)	3.53 (-6.00)	2.78 (-5.61)
Initial LL	-6868	-6132	-4853
Final LL	-5081	-4366	-3490
Adjusted Rho square	0.26	0.29	0.28

(*) parameter not significant at 95% confidence level

MW = motorway, EU = extra-urban road, UM = urban main road,

LR = local road

Road type percentage are estimated with extra-urban roads
as reference category

Table 6.5: Model results for the BoxCox(PSRT1) model and the most appropriate choice sets

and local roads and only by correcting for overlap individually for each road type, the travel time parameters remained reasonable.

6.6 Application of the modelling results

Since a route choice model, or any choice model for that matter, is not estimated as an end in itself but to forecast behaviour, it is also important to evaluate its ability to predict choice probabilities. This is typically done by applying the model to a sub-sample of the data that is different from the sub-sample used for estimation. Therefore, a new sub-sample of 250 OD pairs was drawn using the same sampling procedure described in Section 6.4. For each of these OD pairs, three different choice sets were generated employing the choice set generation procedures that performed best in the analysis in the previous section. First, 100 alternatives were generated with the BFS-LE algorithm to form the B100 choice sets. Second, these 100 alternatives were reduced to the 60 alternatives of the

Choice sets	BC(PS2)	BC(PSRT1)
B100	28.5%	29.3%
SimDistB60	39.0%	35.3%
RulesB1	38.9%	40.5%

Table 6.6: Percentage of ODs sets for which the chosen route obtained the highest probability

SimDistB60 choice sets employing similarity distribution-based reduction procedure. Third, the RulesB1 choice sets were derived using the rule-based reduction with weak threshold parameters to the 100 alternatives.

Two models were applied to each choice set: the BC(PS2) model employing the BoxCox transformed Path Size Factor defined in Equation 2.23 and the BC(PSRT1) model with the BoxCox transformed road type specific Path Size Factor defined in Equation 2.26. The BC(PSRT1) model performed best in the model estimation whereas the BC(PS2) models serves as reference model. The parameters of these two models estimated on the choice sets described previous sections were applied to the corresponding new choice sets to calculate the predicted choice probabilities. The results of these choice probability forecasts are presented in the following.

To obtain a first impression of the models' prediction accuracy Table 6.6 shows how often the actual chosen route obtains the highest choice probability. The values range from 28.5% of the choice sets for the B100 choice sets and the BC(PS2) model to 40.5% for the RulesB1 choice sets and the BC(PSRT1) model. In general, the share of ODs for which the chosen route has the highest choice probability is highest for the RulesB1 choice sets and lowest for the B100 choice sets while it is higher for the BC(PSRT1) model than for the BC(PS2) model. An exception is the BC(PS2) model for the SimDistB60 choice sets. It performs better than the BC(PSRT1) for the same choice sets or the BC(PS2) for the RulesB1 choice sets.

That it is, however, misleading to conclude from this a better prediction accuracy for this model and choice set combination can be seen in Table 6.7. Table 6.7 compares the different models and choice sets using the log-likelihood of the predicted probabilities computed on the new choice sets. Rather than simply counting the number of times the chosen route had the highest choice probability, the log-likelihood accounts for the absolute value of the choice probability of the chosen route regardless of its relative value compared to the choice probabilities of other alterna-

Choice sets	Initial LL	BC(PS2)		BC(PSRT1)	
		Final LL	ρ^2	Final LL	ρ^2
B100	-1140.28	-894.06	0.216	-860.82	0.245
SimDistB60	-1015.08	-759.16	0.252	-743.40	0.268
RulesB1	-801.44	-590.52	0.263	-583.02	0.273

Table 6.7: Initial and final log-likelihoods of the forecasting models

tives. Since the different choice sets had different initial log-likelihoods those are also presented in Table 6.7 as are the ρ^2 calculated from the two log-likelihood values.

It can be seen that the predicted ρ^2 s resemble those of the estimation models. Moreover, the general trend suggested by Table 6.6 can be found again and this time without an exception. The BC(PSRT1) models outperform the BC(PS2) models for all choice sets and for both models the RulesB1 choice sets perform best and the B100 choice sets worst.

6.7 Conclusion and outlook

The aim of the research effort described in this chapter is two answer two questions: What is the most suitable choice set and what is the best way to account for similarities in a car route choice model derived from GPS observations and a very high-resolution network? Concerning the choice set it can be concluded that there is a substantial difference between model parameters estimated for a small choice set directly derived from the choice set generation algorithm and those estimated for a small choice set for which a large set of alternatives has been generated and then reduced to a smaller choice set. Since this effect even occurs if the reduction is purely random, this might be caused by missing relevant routes in the small choice sets directly derived from the choice set generation. Thus, it is advisable to first generate a large set of alternatives to ensure that all relevant routes are found and then reduce the choice set if needed. This has, however, to be balanced against the computation costs for the generation of a large set of alternatives, especially on a high-resolution network. As an example, the generation of the 1500 choice sets for this study with the BFS-LE algorithm lasted 6.6 hours for 20 alternatives, 52 hours for 60 alternatives, and 8.17 days for 100 alternatives whereas the corresponding run times for the SCSG algorithm were 23.98 days, 51.87 days, and 61.83 days. Though the actual computation time was considerably reduced by splitting up the list of OD pairs and running several jobs in parallel, it is still questionable if the SCSG algo-

rithm is appropriate for the choice set generation in very high-resolution data given that the resulting choice sets were also characterised by more overlap and a much narrower overlap distribution, even after employing reduction procedures. For the BFS-LE algorithm, on the other hand, the run times are still in an area where the generation of a large set of alternatives can be achieved in an acceptable time.

Regarding the reduction procedure, an important lesson to be learned from these experiments is that less overlap is not necessarily always better, especially if the reduction procedure does not remove the overlap uniformly over all route types. This can lead to unstable estimation results that are difficult to interpret. Thus, the analyst should either aim for a wide distribution of overlap levels or try to reduce the overlap in a meaningful way. The best way to do this is either employing the similarity distribution-based or the rule-based reduction procedure. However, the rule-based reduction is favoured because it is able to reflect behavioural assumptions about the individual choice set of a decision-maker. Moreover, it showed better prediction performance than the similarity distribution-based reduction. For both procedures additional parameter testing is required. The parameters to be tested for the similarity distribution-based reduction are the number of similarity classes and the target choice set size. The rule-based reduction has no predefined target choice set size. Instead, a systematic testing of the travel time, distance and overlap thresholds is necessary with particular attention to the effects of the individual thresholds.

The best similarity treatment in this study was the road type specific Path Size with the formulation established in Equation 2.26, the PSRT1. It outperformed the other models consistently throughout all choice sets and transformations in the estimation as well as the application. Moreover, it allows to account for the route overlap on each road type individually. This is a benefit because, apparently, the impact of route overlap on the utility of an alternative does vary with the road type. Overlap on motorways is rather negligible. Also overlap on extra-urban roads has a small impact compared to overlap on local roads or urban main roads. Concerning the transformation, the BoxCox transformation consistently outperformed the other transformations, even accounting for the additional parameter.

Chapter 7

Accounting for Similarities in Destination Choice Modelling

This chapter is based on the paper:

Schüssler, N. and K. W. Axhausen (2009b) Accounting for similarities in destination choice modelling: A concept, paper presented at the *9th Swiss Transport Research Conference*, Ascona, October 2009.

Another area of choice modelling that is highly influenced by the available alternatives by the means of high-resolution data is destination choice modelling. Destination choice decisions are influenced by several factors such as the attributes of the destination itself, its accessibility by different means of transport, the location of preceding or subsequent activities, or the attributes and accessibility of competing destinations. In the traditional four step model, a lot of these factors have been ignored or only accounted for by rough approximations.

This chapter presents a general framework for the treatment of similarities in a discrete choice model for destination choice of secondary activities. The framework combines several aspects of similarity derived from spatial location, the journey to and from the destination, trip chaining restrictions, and the attributes of the alternatives themselves. Moreover, it is applicable in a simultaneous route, mode and destination choice model.

7.1 Introduction

The car route choice models presented in Chapter 6 underline the importance of an appropriate similarity treatment especially if a large number of high-resolution choice alternatives is considered. Moreover, they show that a differentiated similarity treatment yields the best estimation and forecasting results. While the similarity treatment in the route choice models only distinguished between different road types or trip parts, a more sophisticated differentiation is necessary in a destination choice model. Similarities between destination choice alternatives originate from a variety of factors: attributes of the destination, its accessibility by different means of transport, the location of preceding or subsequent activities, or the attributes and accessibility of competing destinations and not the least similarities between the routes to and from the destination. Moreover, these different aspects of similarity interact with each other.

This chapter presents, based on the experiences gained with the route choice models, a concept to account for the different aspects of similarity between destination choice alternatives and their interactions within the framework of discrete choice modelling. Along with adjustment terms that have already been established and described in Section 2.4 some new concepts will be introduced that have not yet been formalised but provide new ideas for adjustment terms. The focus is put on the modelling of *secondary* activities, i.e. activities for which the traveller can choose a new

destination at the beginning of each trip. In contrast to that, the location of *primary* activities is determined in a long-term decision. Therefore, the location of primary activities should be modelled separately and can be treated as fixed in a model of daily travel behaviour. Classic primary activities are home, work or school, while shopping or meeting friends are typical secondary activities.

7.2 A general framework of similarity treatment in destination choice modelling

Several assumptions about the destination choice model for secondary activities are necessary to derive a framework for similarity treatment: The complete daily schedule including the order and type of activities, their timing and durations, is known. The location of the primary activities is fixed and known, as is the equipment of the household with mobility tools such as cars or public transport season tickets. The universal choice set for each activity comprises all facilities or zones in the study area where an activity of the specified type can be carried out.

Since the universal choice set is extremely large and contains many alternatives that are irrelevant for the choice of the decision-maker, it should be appropriately reduced to the individual choice set. Analogously to the rule-based reduction in the previous chapter, a first reduction should be done using the shortest path-based space-time prism approach developed by Scott (2006) who extended the work of Hägerstrand (1970). The space-time prisms approach rules out all alternatives that cannot be reached within the time-budget of the individual and, thus, obtain a very low choice probability. The space-time prism approach is better suited than the traditional way of including all alternatives within a circle around the origin because it accounts for the impact of the subsequent primary activity. Additionally, it is able to deal with sequences of secondary activities between two primary activities (i.e trip chaining effects) and can derive different choice sets for different modes. Since the number of alternatives within a space-time prism is still very large, it is advisable to use only a sub-sample of the alternatives and the version of the algorithm presented by Horni *et al.* (2009) with improved computational performance. How to draw the sub-sample is an important question but goes beyond the scope of this dissertation. In a basic version, the alternatives within the space-time prism that have been chosen by the decision-maker for the activity type in question during the survey period should be complemented by randomly drawn alternatives.

Given these prerequisites, the following aspects of similarity will be accounted for in the framework described in this section:

- similarity derived from travel mode and route
- similarity caused by spatial proximity
- similarities emerging from spatial learning and spatial repetition, and
- similarities originating from the image of the destinations.

The **travel mode and route** used to reach the destination do not only influence the set of destination alternatives known to the decision-maker, but also determine which alternatives are considered accessible. While the relationship between the destination and route choice has been hardly investigated, some authors (Hannes *et al.*, 2008; Buliung *et al.*, 2008, e.g.) showed that mode and destination choice are usually executed simultaneously. Moreover, for each regularly performed activity people have a default mode-destination combination. Only if this default alternative is not available or suitable they choose a new mode-destination combination from a set of predefined alternatives. These standard sets are derived “from habit” or because the mode-destination combination is “logical” and the effect is even stronger for public transport users than for car drivers.

Despite these findings destination choice is still usually carried out without an explicit account for mode or route choice by employing a combination of the travel costs of all modes to represent the accessibility by different transport modes. Travel costs do, however, differ significantly depending on the mode and route chosen. Some alternatives might be inaccessible by certain modes while some modes might not be available to all decision-makers. Therefore, the proper way to account for mode and route similarities in destination choice is a route, mode and destination choice model where the modes available to the decision-maker constitute the nests of a Nested Logit model and the route similarity is captured by the similarity measure that performed best in the route choice models in Chapter 6: the BoxCox-transformed trip part specific Path Size Factor $tpsPS_{ix}$ developed by Hoogendoorn-Lanser and Bovy (2007) and specified in Equation 2.26. Thereby, the different trip parts have to be defined representing the specific characteristics of each mode. Then, the destinations in each nest should be generated separately for each mode and treating only unique route and destination combinations as separate alternatives.

Following the argument of Bernardin *et al.* (2009), the effects of **spatial proximity** of substitutes and complements should be accounted for

separately. But a different measure for spatial proximity is proposed here. A buffer is created around the destination and all (complementary or substitutional) alternatives within that buffer are summed up weighted by the distance from the destination in question using a distance decay function. The form of the distance decay function and the size of the buffer should be derived from empirical data. Since not only alternatives around the destination influence the choice of the decision-maker but also alternatives on the way to and from the destination these alternatives are also incorporated in the measure of spatial proximity employing the same concept of a buffer and a weighting by distance. Thereby, the construction of the buffer along the way has to follow the characteristics of the modes. For public transport trips, it suffices to create a buffer around the public transport stops, whereas for private transport modes a buffer around the complete path is suitable.

Moreover, a distinction has to be made between models that work on the spatial resolution of travel zones and those in which facilities compose the decision entities. In case of facilities as alternatives, the spatial proximity effects arising from complements SPC_i and substitutes SPS_i for alternative i should be accounted for by:

$$SPC_i = \sum_{j \in Co} w(d(i, j)) \cdot \delta_j \quad (7.1)$$

$$SPS_i = \sum_{j \in S} w(d(i, j)) \cdot \delta_j \quad (7.2)$$

where Co is the set of alternatives that are complements to the activity in question, S is the set of alternatives that are substitutes for the activity in question, $w(d(i, j))$ is the distance decay function used as weight for alternative j and δ_j is an indicator function which equals 1 if j lies within the overall buffer area and 0 otherwise.

If the alternatives of the destination model are zones, it is necessary to determine for each zone the “size” of the complementary and substitutional facilities within each zone. The size attribute can be measured in various ways, e.g. based on the number of facilities, the number of employees in these facilities or the square footage of the facilities. Then, the size variables sc_j and ss_j have to be incorporated in the spatial proximity measures SPC_i and SPS_i . Moreover, the distance $d(i, j)$ between alternative i and alternative j is measured as distance between the zonal centroids. Finally, it has to be defined when a zone lies within the buffer, i.e. when δ_j obtains a value of 1. Several variants are possible:

The buffer has to cover the zone completely, a certain percentage of the zone or the at least the centroid. Another option would be to transform δ_j into a continuous variable ranging from 0 to 1 and representing the percentage of the zone that lies within the buffer. Merging all these considerations, the spatial proximity measures for zone i are:

$$SPC_i = \sum_{j \in C} w(d(i, j)) \cdot sc_j \cdot \delta_j \quad (7.3)$$

$$SPS_i = \sum_{j \in S} w(d(i, j)) \cdot ss_j \cdot \delta_j \quad (7.4)$$

In order to measure the similarity effects originating from spatial learning or **spatial repetition**, multiple-day observations for each survey participant are required. Following the approach suggested by Sivakumar and Bhat (2007) a adjustment term based on spatial repetition is suggested. In case of zones, the spatial repetition index should indicate the time period, e.g. the number of days, since the zone was last chosen. The spatial repetition measure SR_i is then formulated as:

$$SR_i = \frac{t_0 - t_i}{t_0} \quad (7.5)$$

where t_0 is the current period of time and t_i is the period of time when the zone i was last visited. Both times should be measured relative to the beginning of the survey period. This way, the beginning of the survey period can be defined as the default value for each t_i . Together with the normalisation by the current time this leads to a range between 0 and 1 for SR_i . 0 indicates that the zone was already visited in the same period of time and 1 that it was never visited before in the survey period.

In a model where the alternatives are elemental facilities, the definition of repetition is extended in the sense that not only a visit to the same facility is counted but also to other facilities that are close by. This can be done by either considering all facilities within a circle around the destination in question or by summarising alternatives to neighbourhoods and determining the last time the person visited the neighbourhood. The spatial repetition measure SR_i is then formulated analogously to the one for the zones:

$$SR_i = \min_{j \in N} \frac{t_0 - t_j}{t_0} \quad (7.6)$$

where N is the set of facilities that lie within the circle or neighbourhood around i , including i . For t_0 and t_j the same definitions as above apply.

Unlike the the similarity aspects discussed above that are measured on a continuous or ordinal scale, the **image of an alternative** can only be measured on a nominal scale. Moreover, the image of an alternative is difficult to define and strongly depends on the choice context, the level of spatial resolution and the characteristics of the alternatives that are available. Despite this, several authors (Huybers, 2003; Hong *et al.*, 2006) have demonstrated its importance, particularly for vacation destination choice. Labeling the alternatives according to their major attractions (e.g. mountainous, coastal, historic and exotic) or simply the name of the place results in considerably different choice behaviour and market shares. Evidently, the labels of the vacation destinations carried a lot more meaning than the attributes describing the alternatives. Analogously, the choice of a grocery shopping location can be highly affected by the chain the supermarket belongs to, the size of the shop, the range of goods or its crowdedness.

Therefore, the analyst first has to decide which characteristics are crucial for the image of a destination. Based on this, the alternatives are subdivided into categories. The categorisation can enter the utility function either by the means of dummy variables and/or in the form of a measure for image related similarity IM_i :

$$IM_i = \frac{\sum_{j \in C_n, j \neq i} \delta_{jy}}{I - 1} \quad (7.7)$$

where C_n is the choice set, I is the number of alternative in C_n , and δ_{jy} is an indicator function that takes the value of 1 if alternative j belongs to the same category y as alternative i . The values of the adjustment term range between 0 and 1, 0 indicating that no other alternatives belongs to the same category and 1 that all alternatives belong to the same category.

Finally, the similarity components have to be combined to one model that accounts for all the different aspects of similarity. The basic functional form is a Nested Logit model in which the mode alternatives comprise the nests. As in Vrtic (2003), the adjustment terms are added to deterministic parts of the utility of each destination alternative. Thereby, each adjustment term obtains its own parameter α_l . This parameter is estimated and can take positive as well as negative values because it is unclear a priori how the similarity between the alternatives will affect the choice behaviour. The utility U_{in} of an destination alternative i , mode m

and decision-maker n is then formulated as:

$$U_{in} = V_{in} + \varepsilon_{in} + V_{C_{mn}} + \varepsilon_{C_{mn}} \quad (7.8)$$

where ε_{in} and $\varepsilon_{C_{mn}}$ are the error terms of the Nested Logit model following the distributional assumptions described in Section 2.2. The deterministic part of the utility V_{in} of each destination alternative is described by:

$$V_{in} = f(\beta, x_{in}) + \alpha_1 \sum_x tpsPS_{ix} + \alpha_2 SPC_i + \alpha_3 SPS_i + \alpha_4 SR_i + \alpha_5 IM_i, \quad (7.9)$$

where $tpsPS_{ix}$ is the trip part specific Path Size Factor for trip part x . The Logsum term for each mode is calculated by:

$$V_{C_{mn}} = V'_{C_{mn}} + \frac{1}{\mu_m} \ln \sum_{j \in C_{mn}} \exp(\mu_m (f(\beta, x_{in}) + \alpha_1 \sum_x tpsPS_{ix} + \alpha_2 SPC_i + \alpha_3 SPS_i + \alpha_4 SR_i + \alpha_5 IM_i)), \quad (7.10)$$

where $V'_{C_{mn}}$ is the utility common to all alternatives in nest C_{mn} .

7.3 Conclusion and outlook

Destination choice situations are characterised by a multitude of similarities originating from different sources. This chapter offers a framework that accounts for several similarity aspects simultaneously. The framework is built in a modular way to make it adjustable in case the required data is not available. Moreover, for each similarity component, a separate coefficient has to be estimated to depict the relative influence of the similarity component on the overall choice.

It is advised to use a state of the art choice set generation procedure, the space-time prism approach, which by design takes care of the impact of trip chaining on the choice set. A Nested Logit structure accounts for the similarities derived from the access and egress modes while the similarities resulting from routes, spatial proximity, spatial learning and the image of the destination are integrated by the use of adjustment terms. Since the findings in Chapter 6 underlined the importance of a differentiated similarity treatment, the adjustment term for route similarity is specified trip part specific whereas the spatial proximity adjustment term differentiates between complementary and substitutional alternatives.

In order to estimate models based on the framework proposed in this chapter, some data requirements have to be met. Most importantly, the survey data has to have the form of a diary which records the complete daily schedule of the participant, including times of journeys and activities, the types of activities and the mode used for travelling. Additionally, multi-day records would be necessary to include similarity based on spatial repetition and observations about the routes for the route adjustment term and an accurate representation of travel costs. Concerning the infrastructure, the data should be as detailed as possible.

Regarding the choice set generation the main question is how a subsample out of all the alternatives within the space-time prism should be drawn. With respect to the spatial proximity measure the correct size of the buffer, the form of the distance decay function, the derivation of the size variables sc_j and ss_j and the definition of which zones lie within the buffer area have to be investigated. In addition, a more detailed literature research which exceeds the field of transportation research, might help to determine how the image of an alternative is formed and how it can be derived from available data.

Chapter 8

Summary and Outlook

The overall aim of this dissertation was to develop discrete choice models estimated on high-resolution data with an appropriate treatment of the similarities between the alternatives. To this end, the dissertation focussed on three different issues: the processing of GPS observations to obtain the chosen alternatives, the generation of choice sets and the model estimation to test different adjustment terms to account for similarities between alternative. Following the structure given by these three issues, this chapter gives an overview about the results of the presented work from a more global point of view. In addition, some ideas for future research directions are presented.

8.1 Deriving observations for choice modelling from GPS raw data

Since the first GPS studies were conducted in the mid-1990s, researchers are promising that this new way of observing individual travel behaviour will substantially reduce the participants' burden. However, so far this promise has not been kept. In addition to wearing a GPS recorder, the participants are asked to complete extensive questionnaires and diaries. To dispose of the time-consuming questioning post-processing procedures are required that are able to derive all relevant details of the individual travel behaviour automatically and without additional information by the participant. Moreover, the post-processing procedures have to be fast enough to process large quantities of data in acceptable time. Developing these procedures is still an ongoing research issue.

This dissertation contributes to this line of research by presenting an automatic GPS post-processing procedure needing no input other than the most basic GPS raw data, and for the map-matching the transport

network. This independence of a GIS environment or any other representation of land-use data is one of the factors making the procedure really fast and, thus, suitable for large-scale applications.

Another key success factor for the post-processing procedure is the implementation of an appropriate filtering and smoothing mechanism. Finding the right approach was essential in this study because no information about the numbers of satellites in view or their positioning was available. Though filtering based on altitude levels and unrealistic positional jump was essential, it was not sufficient. Therefore, a Gauss kernel smoothing had to be applied as well.

Different from most post-processing procedures published, the activity and trip detection described in this dissertation is based on three criteria: point bundles, dwell times and zero speed. With them, very similar results to those derived manually are obtained. The manual trip and activity derivation for a sample of the data set had to be the frame of reference since no information about the actual trips and activities was available.

The fuzzy logic mode detection also delivered reasonable results. Notably, the distance distributions per mode are very satisfactory. However, in the future a more detailed validation of the mode detection has to be carried out with data that contains the modes actually chosen. The validation will comprise the stage segregation as well as the fuzzy logic membership parameters. Possible validation data is the hand coded GPS data presented by Flamm and Kaufmann (2007). Upadhyay *et al.* (2008) already used this data to make a preliminary analysis of the fuzzy logic parameters for an earlier version of the post-processing procedure. Building on this work, the analysis will be extended using an advanced multi-dimensional Sequence Alignment Method (Wilson *et al.*, 2005).

Depending on the outcomes of this analysis, there are some options to improve the mode detection. First, there is the deviation from the mode-specific transport network. This can be used to better distinguish, for example, inter-urban rail trips from car trips. A second possible approach considers the likelihood of mode transitions and mode transition sequences, as employed by Zheng *et al.* (2008). Third, learning procedures such as presented by Auld *et al.* (2009) or Moiseeva *et al.* (2010) can be applied to refine the mode detection based on previous behaviour of the participant.

The last step of the GPS post-processing described in this dissertation is the map-matching. The map-matching reliably identifies the routes actually taken by the participants if the underlying network is correct, consistent and complete. Due to the employment of the Multiple Hypothesis

Technique, the map-matching results are robust against erroneous map-matching due to noisy GPS data or a simplified network coding where the shape of the link does not exactly follow the course of the actual road. The computation time of the map-matching, however, is only acceptable and has to be improved. A first step would be to decrease number of path candidates during the development of a path depending on the differences between the scores. In the beginning, a wide exploration and a large number of candidates is necessary to ensure that the right route is contained in the candidate set. But while the path development progresses the score differences between the right route and the other candidates increase considerably so that it might be sufficient to reduce the candidate set to 10-20 alternatives.

A second shortcoming of the map-matching is its relatively low ratio of routes that could be matched. The main reason for this are network related problems. In more than 40% of the unsuccessful map-matching cases one or more links were missing in the network. Some of these links were used by several persons or repeatedly by the same person, making a network coding error very probable. In future research, this information can be used to complete and correct the network. In return, completing the network will improve the success rate of the map-matching. Another way to improve the success rate would be to implement a handling for u-turns. Currently, all trips that contain u-turns are removed. They were left out because the map-matching results are used for route choice modelling where u-turns cannot be modelled appropriately. However, for future applications, e.g. the description of parking search behaviour at the end of a trip, an appropriate representation of u-turns will be important.

A completely different issue related to the map-matching is the identification of the public transport lines used by the participant to make the GPS observations usable for public transport route choice. To describe a public transport route, it is not enough to identify the links of the public transport network the participant travelled on. The public transport map-matching needs to identify public transport lines, the access and egress stops and the transfer points. Since no approach of doing that automatically has been published yet, this is a major research issue.

Besides delivering the chosen route description as input for choice models, the map-matching results can play another important role in the processing of GPS data. Especially for modes with similar speed profiles, such as long-distance train and car travel, the map-matching can help to refine the mode detection. Therefore, stages with ambiguous modes should be matched to all networks for whose modes they have a non-zero probability. Then, the error measure resulting from each map-matching

can be fed back into the mode detection. This does, however, require adequate networks for all modes.

Finally, an important step is still missing in the GPS processing to describe the complete daily routine of a participant: the activity purpose identification. Current procedures (e.g. Wolf *et al.*, 2001; Schönfelder and Samaga, 2003; Wolf *et al.*, 2004; Stopher *et al.*, 2005) rely on land-use and addresses reported by the participants to derive the activity purpose. This is, however, not satisfactory. On the one hand, the usage of GIS environments to obtain that land-use makes these procedures rather slow. On the other hand, the land-use does not always give sufficient information to determine the activity purpose, especially in survey areas with a high share of mixed land-use zones. Thus, it would be favourable to follow the lines of the GPS post-processing presented in this dissertation and employ as little land-use information as possible. Instead, the key criteria should be observed activity patterns such as locations that are visited multiple times, the time of day and duration of activities. A first attempt in this direction has been made by Kawasaki and Axhausen (2009). Future work should follow this approach.

8.2 Choice set generation

The generation of choice sets in very high-resolution networks is a new challenge for transportation modellers. The number of routes in the universal choice set increases substantially while the size of the individual choice set, i.e. the number of relevant routes, remains the same. In other words, many more routes have to be found and evaluated to find all routes that are relevant to the decision-maker. This reasoning is underlined by the model estimations presented in this dissertation. Substantial differences occurred between model parameters estimated for a small choice set directly derived from the choice set generation algorithm and those estimated for a small choice set for which first a large set of alternatives has been generated and then reduced to a smaller choice set. Thus, the choice set is ideally derived in a two step process. First, a large master set is extracted from the network to ensure that all relevant routes are considered. Then, the master set is reduced to the individual choice set taking into account behavioural assumptions about the decision-maker.

However, extracting routes from a high-resolution network is cumbersome and time-consuming. Moreover, the computation time increases disproportionate with the number of alternatives the procedure aims for. Therefore, computational performance is an important criterion in the

evaluation of choice set generation algorithms, even though recent research has paid little attention to this issue. This dissertation compares different choice set algorithms and demonstrates that only approaches based on repeated least cost path search are applicable to high-resolution networks. The best performance delivered a new algorithm employing Breadth First Search on Link Elimination (BFS-LE). The algorithm includes two performance optimisation features: a randomisation of the processing order within each tree depth and a topologically equivalent network reduction. Compared to the computational performance of a Stochastic Choice Set Generation (SCSG), a procedure in which the link costs are randomly drawn before each least cost path search, it performed particularly good for typical urban trips under 10 km. Considering the reproduction of chosen routes and road type compositions, the new BFS-LE algorithm and the SCSG algorithm performed almost equally well. The routes of the BFS-LE choice sets are, however, more diverse than those of the SCSG choice sets.

The issue of behavioural realism was only initially addressed in this dissertation by evaluating the impact of different choice set size reduction procedures on car route choice models. These experiments showed that less overlap is not necessarily always better, especially if the reduction procedure does not remove the overlap uniformly over all route types. Instead, these choice sets revealed unstable estimation results that are difficult to interpret. Thus, the analyst should aim for a wide distribution of overlap levels and try to reduce the overlap in a meaningful way. Of the different reduction procedures presented in this dissertation the similarity distribution-based or the rule-based reduction procedure worked best in terms of delivering stable estimation results. The rule-based reduction is favoured because it is able to reflect behavioural assumptions about the individual choice set of a decision-maker. However, additional testing of the rules is necessary before it can be stated with confidence that the resulting choice sets are behaviourally realistic. This testing includes the testing of different reduction rules and overlap thresholds and the linkage between the thresholds and socio-demographic attributes of the respondent, if those are available.

Moreover, the BFS-LE choice set generation should be run on a loaded dynamic travel time network to obtain even more realistic routes. This can be done without any increase in computation time, because the whole algorithm, including the A-Star Landmarks router, is designed accordingly. These runs do, however, require additional data about network loads. The loads can, for example, be taken from a micro-simulation.

Another future research direction is a completely new way of gener-

ating choice sets. Instead of using the network to generate alternatives, the individual choice set is derived from repeated GPS observations. This would allow the analyst to approximate the actual individual choice sets and bestow invaluable insights into the actual decision-process. Though a first attempt in this direction has been undertaken by Rich *et al.* (2007), more research is necessary before this approach can become state-of-the-art in route choice modelling.

The choice set generation procedures developed in the course of this dissertation are so far restricted to route choice for private transport modes, that use the road network infrastructure, i.e. cars and bikes. The next step will be to generate choice sets for public transport mode choice. For public transport, it is not sufficient to describe the routes in terms of network links. Instead, public transport lines and preferable even public transport connections have to be derived. Thus, for the public transport route choice set generation the BFS-LE algorithm has to be adapted in a way that it uses a network generated from the public transport schedule. The public transport router by Rieser and Nagel (2009) will be used for the least cost path search. Moreover, to fully model the daily routine of a participant, a choice set generation for destination choice alternatives has to be developed. It should be based on the computationally optimised space-time prism approach presented by Horni *et al.* (forthcoming).

8.3 Choice modelling based on high-resolution observations

Choice models estimated on high-resolution data are characterised by a large number of alternatives and complex similarity structures. Additionally, there is an growing effort to model several steps of the classic four step approach simultaneously, further increasing the number of alternatives. Thus, state-of-the-art models have to be able to handle large choice sets and should not require too much effort for computation, specification and identification. On the other hand, they have to be flexible and able to accommodate various and complex similarity structures. One way to combine these requirements is the inclusion of adjustment terms in the systematic part of the MNL utility function. This approach is very appealing because of its simplicity and elegance. Instead of structuring the choice set a priori and taking the chance of misleading assumptions about correlations, only the type of similarity is specified. This type accounts for the individual characteristics of the alternatives in the choice set and imposes a value to the impact of specific interdependencies.

Adjustment terms do, however, have a shortcoming in the sense that they are only designed with respect to a specific context and usually miss some aspects of the correlation between alternatives. While adjustment terms for some choice situations have been extensively investigated and appropriate factors have been well established, similarities in other choice situations have hardly been tackled by the means of adjustment terms. Moreover, in multi-dimensional choice situations there might be several behavioural mechanisms overlaying each other. Thus, sometimes multiple adjustment terms have to be combined and each of them needs a parameter that allows to estimate its sign. Empirical work, in this thesis but also by Weis (2006) and van Eggermond *et al.* (2007), has shown that for the best model fit the sign and transformation of the adjustment terms do not necessarily match the theoretical expectations of the developers of the adjustment terms. They have overlooked some of the mechanisms at work.

Regarding the question which adjustment term performs best in choice models estimated on high-resolution data, results have been derived for bike and car route choice so far. Surprisingly, for the bike route choice models presented by Menghini *et al.* (2009) the parameter for Path Size factor of Equation 2.23 is for no transformation significant. Apparently, similarities between alternative routes influence the route choice of cyclists considerably less than the directness of the route or even the presence of a marked bike path and the gradient of the route. For car route choice several adjustment terms were tested. The best performance yielded a road type specific Path Size employing the formulation established in Equation 2.26. It outperformed the other models consistently throughout all choice sets and transformations. Moreover, it demonstrated that the impact of route overlap on the utility of an alternative does vary with the road type. Estimating public transport route choice models that account for similarities from GPS observations is future research issue.

Another future research issue is the modelling of destination choice based on GPS observations. This dissertation describes a concept how the multitude of similarities originating from different sources can be represented simultaneously in this model. Several adjustment terms are combined with each other and for each a separate coefficient has to be estimated to depict the relative influence of the similarity component on the overall choice.

Dank

Mit den letzten Zeilen dieser Dissertation geht die lange und spannende Reise meines Doktorats zu Ende. Dabei habe ich durch viele grosser und kleine Umwege und Kurswechsel ein Forschungsfeld gefunden, das mir sehr viel Freude bereitet und dem ich mich mit grossem Vergnügen und viel Motivation in den nächsten Jahren weiter widmen werde. Möglich gemacht haben dies nicht zuletzt die Finanzierung meines Doktoratprojektes durch den Schweizerischen Nationalfonds und Stefan Muff, damals Endoxon AG, der mir die GPS-Daten zur Verfügung gestellt hat.

Natürlich wäre meine Reise, wie viele andere auch, nicht annähernd so bereichernd gewesen ohne die vielen Menschen, die mich in den letzten sechs Jahren begleitet haben. Diesen Menschen möchte ich hiermit danken. Grosser Dank gehört zunächst meinem Doktorvater Kay W. Axhausen. Er hat mir die Freiheiten gegeben, die ich zum Finden und Erforschen meines eigenen Themengebietes benötigte und hatte immer eine offene Tür und ein offenes Ohr, wenn ich Anleitung oder einen Diskussionspartner brauchte. Darüber hinaus hat er durch seinen grossen Wissens- und Erfahrungsschatz dafür gesorgt, dass mir auf meiner Reise die nötige Weitsicht und der Blick über den Tellerrand hinaus nicht verloren gingen. Nicht zuletzt hat er durch die Zusammenstellung seiner Gruppe und die Schaffung eines anregenden und kommunikativen Arbeitsklimas dafür gesorgt, dass ich auch auf steilen Abschnitten nie die Motivation verloren habe.

Milenko Vrtic danke ich für die gute Betreuung in den ersten drei Jahren meines Doktorats. Er hat mich nicht nur in die Welt der Entscheidungsmodellierung eingearbeitet, sondern mir auch vermittelt, was die wichtigen Fragen- und Problemstellungen der Verkehrsplanung in der Forschung und vor allem in der Praxis sind.

I would like to thank Michel Bierlaire and Stephane Hess for taking over the role of co-examiners of my dissertation. Their valuable comments on the previous drafts of this dissertation made the version you hold now in your hands considerably more readable and stringent and helped to improve the presentation of my research findings. More-

over, I would like to thank Michel for the productive and constructive discussions about my research and the great opportunity to meet other researchers in the area of discrete choice modelling in his yearly workshops. I am also grateful for the opportunity to stay at the University of Texas at Austin provided to me by Chandra Bhat. I met great people there and my research made some important steps forward.

Einen wichtigen Anteil an der erfolgreichen Bewältigung meiner Doktoratsreise hatten auch meine (ehemaligen) Kollegen am IVT, die die täglichen Freuden und Leiden einer Forschenden mit mir geteilt haben. Ihnen allen gilt mein Dank für die vielen fachlichen und nichtfachlichen Gespräche, die fröhlichen Kaffeerunden und anderen zusammen genossenen Getränke. Besonders danken möchte ich David Charypar. Ohne unsere stundenlangen Diskussionen über GPS-Daten und Datenbereinigung im Allgemeinen und im Besonderen wäre diese Arbeit nicht zu dem geworden, was sie heute ist. Viel mehr bedeutet mir jedoch unsere Freundschaft, unsere ausführlichen Gespräche und unsere Skiausflüge. Ebenfalls danken möchte ich Andreas Horni. Seine Eigenart alles zu hinterfragen hat auch mich dazu verleitet, an manchen Stellen mehr in die Tiefe zu gehen, und hat mir zu vielen wertvollen Erkenntnissen verholfen. Mindestens ebenso wertvoll waren und sind seine Freundschaft unsere die vielen spontanen Grill-, Wii- und Dartabende sowie unsere mal mehr oder weniger windigen Skiausflüge. Stefan Sandmeier, unser zeitweiliger “Institutshistoriker”, hat dafür gesorgt, dass wir immer mal wieder über den Tellerrand unseres Ingenieursrattatoui geschaut haben. Für diese interessanten Diskussionen danke ich ihm ebenso wie ihm und seiner Frau Nicole für ihre Freundschaft. Bei Michael Balmer möchte ich mich für die vielen offenen Ohren bedanken, die er für mich und meine L^AT_EX-, Programmier- und Choice Set Generation-Probleme hatte. Seine Beiträge und Hilfestellungen waren von unschätzbarem Wert für die Struktur meines GPS-Processing-Programms. Ausserdem wäre ohne ihn und seine ansteckende Begeisterungsfähigkeit Kapitel 5 nicht zustande gekommen. Es ist ebenso eine Freude mit ihm Paper zu schreiben wie mit ihm Bier zu trinken und über Gott, die Welt und MATSim zu philosophieren. Alex Erath danke ich für die jahrelange Komplizenschaft in diversen Projekten und Vorlesungsbetreuungen sowie für unseren lustigen Konferenztourismus. Mit Konrad Meister, der am gleichen Tag sein Doktorat begonnen hat, habe ich verschiedene Höhen und Tiefen geteilt. Es war schön, dabei einen solchen Kameraden an der Seite zu haben. Martin Tschopp und Francesco Ciari waren die besten Bürokollegen die man sich wünschen kann. Sie haben nicht nur meine stark wechselnden Bedürfnisse nach Ruhe zum Arbeiten und amüsanten

Gesprächen geteilt, sondern mir auch Einblicke in die schweizerische und italienische Lebens- und Denkweise gewährt.

Mindestens ebenso wichtig wie die gute Arbeitsatmosphäre waren für das Gelingen dieser Dissertation die Ablenkungen die mir halfen Abstand von meiner Arbeit zu bekommen. Als erstes möchte ich daher “meinen Mädels” Milena Scherer, Sigrun Beige und Anna Sundvall danken für all den Sport, die Musik, die Filme und die vielen Koch-, Back- und Mädelsabende.

Eine weitere wesentliche Bereicherung meiner Doktoratsreise war meine Mitarbeit bei der AVETH. Ich habe dadurch nicht nur neue Freunde gewonnen, sondern auch wertvolle Einblicke in die schweizerische (Hochschul-)Politik gewinnen können. Etwas durch den eigenen Einsatz verändern zu können, macht grossen Spass und viele Menschen haben dazu beigetragen, von denen ich nur einige hier erwähnen kann. Paolo Losio, meinem Mit-Co-Präsidenten in den ersten 1.5 Jahren, danke ich für die Einführung in das Gestrüpp der ETH-Administration und Politik. Anke Neumann danke ich für unsere intensive, spannende und unschätzbar bereichernde Co-Präsidentenschaft. Ihre Energie und Begeisterungsfähigkeit war ansteckend und wir haben nicht nur einiges erreicht sondern auch ein unvergessliches Jahr zusammen gehabt. Michael Waschbüsch und Axel Krauth danke ich für die langen Jahre der Freundschaft und der Vernehmlassungsbearbeitungen. Gemeinsam haben wir auch uns auch von den grössten Vernehmlassungsfluten nicht unterkriegen lassen. Martin Wieser und Michèle Grieder danke ich für die vielen schönen Abende und den Spass.

Anke Eßer, Anne Held und Kerstin Wilson haben mich nicht nur durch mein Studium begleitet, sondern auch mein Doktorat aus der Ferne tatkräftig unterstützt. Ihre jahrelange Freundschaft und ihre (teilweise) Schicksalsgemeinschaft auf der langen Reise des Doktorats bedeutet mir unbeschreiblich viel. Sie waren immer für mich da auf meinen beruflichen und privaten Achterbahnfahrten und ich hoffe, ich habe ihnen ein klein wenig davon zurückgeben können, was sie mir gegeben haben.

Das grösste “Danke Schön” von allen gilt jedoch meinen Liebsten. Meine Eltern haben von klein auf meine Neugier und meinen Wissensdurst gefördert und mich in allem unterstützt, was ich angefangen habe. Sie haben mir nicht nur das Gefühl gegeben, dass ich alles erreichen kann, wenn ich es nur will, sondern haben mich auch in die Welt hinaus ziehen lassen. Meine Schwester Vanessa war mir immer eine gute Gefährtin und meine Kindheit und Jugend wäre ohne sie nur halb so schön gewesen. Eine ebensolche Freude ist es, meine Schwester Mariam aufwachsen zu sehen. Du bist ein tolles Mädchen und wirst

sicher eine beeindruckende Frau! Marcel danke ich für seine unendliche Liebe und Geduld. Du bist mein Zuhause, mein Ruhepol und meine Kraftquelle. Du hast mir vor allem in den verrückten Monaten des Zusammenschreibens genau die Unterstützung gegeben die ich brauchte. Ich freue mich sehr darauf, mit die nächste Etappe meiner Lebensreise mit dir gemeinsam anzutreten, wo auch immer sie hingehen mag.

Bibliography

- Abay, G. (1999) Nachfrageabschätzung Swissmetro: Eine Stated-Preference Analyse, *Research Report*, Nationales Forschungsprogramm 41: Verkehr und Umwelt, **F1**, Swiss National Science Foundation, Planungsbüro Abay & Meier, Berne.
- Abbe, E., M. Bierlaire and T. Toledo (2007) Normalization and correlation of cross-nested logit models, *Transportation Research Part B: Methodological*, **41** (7) 795–808.
- Auld, J., C. Williams, A. K. Mohammadian and P. Nelson (2009) An automated GPS-based prompted recall survey with learning algorithms, *Transportation Letters*, **1** (1) 59–79.
- Azevedo, J. A., M. E. O. Santos Costa, J. J. E. Silvestre Madeira and E. Q. Vieira Martins (1993) An algorithm for the ranking of shortest paths, *European Journal of Operational Research*, **69** (1) 97–106.
- Balmer, M., A. Horni, K. Meister, F. Ciari, D. Charypar and K. W. Axhausen (2009) Wirkungen der Westumfahrung Zürich: Eine Analyse mit einer Agenten-basierten Mikrosimulation, *Final Report*, Baudirektion Kanton Zurich, IVT, ETH Zurich, Zurich, February 2009.
- Batsell, R. R. (1982) A multiattribute extension of the luce model which simultaneously scales utility and substitutability, in J. Huber (ed.) *The Effect of Item Similarity on Choice Probabilities: A Collection of Working Papers from a Conference at Quail Roost*, 4–23, Duke University, Durham.
- Bekhor, S. (2007) Estimation of the cross-nested logit model for a large number of alternatives, paper presented at the *86th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2007.
- Bekhor, S., M. E. Ben-Akiva and M. S. Ramming (2002) Adaptation of logit kernel to route choice situation, *Transportation Research Record*, **1805**, 78–85.

- Bekhor, S., M. E. Ben-Akiva and M. S. Ramming (2006) Evaluation of choice set generation algorithms for route choice models, *Annals of Operations Research*, **144** (1) 235–247.
- Bekhor, S. and J. N. Prashker (2008) GEV-based destination choice models that account for unobserved similarities among alternatives, *Transportation Research Part B: Methodological*, **42** (3) 243–262.
- Bekhor, S., L. Reznikova and T. Toledo (2007) Application of cross-nested logit route choice model in stochastic user equilibrium traffic assignment, *Transportation Research Record*, **2003**, 41–49.
- Bell, D. E. (1982) Regret in decision making under uncertainty, *Operations Research*, **30** (5) 961–981.
- Ben-Akiva, M. E. (1973) Structure of passenger travel demand models, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge.
- Ben-Akiva, M. E. (1974a) Note on the specification of a logit model with utility functions that include attributes of competing alternatives, *Working Paper*, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge.
- Ben-Akiva, M. E. (1974b) Structure of passenger travel demand models, *Transportation Research Record*, **526**, 26–42.
- Ben-Akiva, M. E., M. J. Bergman, A. J. Daly and R. Ramaswamy (1984) Modelling inter-urban route choice behaviour, in J. Volmuller and R. Hamerslag (eds.) *Proceedings of the Ninth International Symposium on Transportation and Traffic Theory*, chap. 15, 299–330, VNU Science Press, Utrecht.
- Ben-Akiva, M. E. and M. Bierlaire (1999) Discrete choice methods and their applications to short-term travel decisions, in R. Hall (ed.) *Handbook of Transportation Science*, chap. 2, 5–34, Kluwer, Dordrecht.
- Ben-Akiva, M. E. and D. Bolduc (1996) Multinomial probit with a logit kernel and a general parametric specification of the covariance structure, paper presented at the *3rd Invitational Choice Symposium*.
- Ben-Akiva, M. E. and S. R. Lerman (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge.
- Bernardin, V. L., F. S. Koppelman and D. Boyce (2009) Enhanced destination choice models incorporating agglomeration related to trip

- chaining while controlling for spatial competition, *Transportation Research Record*, **2132**, 143–151.
- Bhat, C. R. and J. Y. Guo (2004) A mixed spatially correlated logit model: Formulation and application to residential choice modeling, *Transportation Research Part B: Methodological*, **38** (2) 147–168.
- Biding, T. and G. Lind (2002) Intelligent Stöd för Anpassning av hastighet (ISA), Resultat av storskalig försöksverksamhet i Borlänge, Lidköping, lund och Umea under perioden 1999–2002, *Research Report*, Vaegverket, Borlaenge.
- Bierlaire, M. (2006) A theoretical analysis of the cross-nested logit model, *Annals of Operations Research*, **144** (1) 287–300.
- Bierlaire, M. (2009) *An introduction to BIOGEME (Version 1.8)*, URL <http://transp-or2.epfl.ch/biogeme/doc/tutorialv18.pdf>.
- Bierlaire, M., K. W. Axhausen and G. Abay (2001) The acceptance of modal innovation : The case of Swissmetro, paper presented at the *1st Swiss Transport Research Conference*, Ascona, March 2001.
- Bierlaire, M., E. Frejinger and J. Stojanovic (2006) A latent route choice model in Switzerland, paper presented at the *European Transport Conference*, Strasbourg, September 2006.
- Bierlaire, M., R. Hurtubia and G. Flötteröd (2009a) A comparative analysis of implicit and explicit methods to model choice set generation, paper presented at the *9th Swiss Transport Research Conference*, Ascona, September 2009.
- Bierlaire, M., J. Newman and J. Chen (2009b) A method of probabilistic map distribution of path likelihood, paper presented at the *9th Swiss Transport Research Conference*, Ascona, September 2009.
- Bliemer, M. C. J. and P. H. L. Bovy (2008) Impact of route choice set on route choice probabilities, *Transportation Research Record*, **2076**, 10–19.
- Bliemer, M. C. J., P. H. L. Bovy and H. Li (2007) Some properties and implications of stochastically generated route choice sets, paper presented at the *6th Triennial Symposium on Transportation Analysis (TRISTAN)*, Phuket Island, June 2007.

- Bohte, W. and K. Maat (2008) Deriving and validating trip destinations and modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands, paper presented at the *8th International Conference on Survey Methods in Transport*, Annecy, May 2008.
- Borgers, A. W. J. and H. J. P. Timmermans (1987) Choice model specification, substitution and spatial structure effects: A simulation experiment, *Regional Science and Urban Economics*, **17** (1) 29–47.
- Bovy, P. H. L. (2009) On modelling route choice sets in transportation networks: A synthesis, *Transport Reviews*, **29** (1) 43–68.
- Bovy, P. H. L., S. Bekhor and C. G. Prato (2008) The factor of revisited path size: Alternative derivation, *Transportation Research Record*, **2076**, 132–140.
- Bovy, P. H. L., S. Bekhor and C. G. Prato (2009) Route sampling correction for stochastic route choice set generation, paper presented at the *88th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2009.
- Bovy, P. H. L. and S. Fiorenzo-Catalano (2007) Stochastic route choice set generation: Behavioral and probabilistic foundations, *Transportmetrica*, **3** (3) 173–189.
- Bricka, S. and C. R. Bhat (2006) A comparative analysis of GPS-based and travel survey-based data, *Transportation Research Record*, **1972**, 9–20.
- Buliung, R. N., M. J. Roorda and T. K. Remmel (2008) Exploring spatial variety in patterns of activity-travel behaviour: Initial results from the Toronto Travel-Activity Panel Survey (TTAPS), *Transportation*, **35** (6) 697–722.
- Casas, J. and C. Arce (1999) Trip reporting in household travel diaries: A comparison to GPS-collected data, paper presented at the *78th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 1999.
- Cascetta, E., A. Nuzzola, F. Russo and A. Vitetta (1996) A modified logit route choice model overcoming path overlapping problems: Specification and some calibration results for interurban networks, in J. B. Lesort (ed.) *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, 697–711, Pergamon, Oxford.

- Cascetta, E., F. Pagliara and K. W. Axhausen (2007) The use of dominance variables in choice set generation, paper presented at the *11th World Conference on Transportation Research*, Berkeley, June 2007.
- Cascetta, E. and A. Papola (2001) Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand, *Transportation Research Part C: Emerging Technologies*, **9** (4) 249–263.
- Cascetta, E. and A. Papola (2009) Dominance among alternatives in random utility models, *Transportation Research Part A: Policy and Practice*, **43** (2) 170–179.
- Cascetta, E., F. Russo, F. A. Viola and A. Vitetta (2002) A model of route perception in urban road networks, *Transportation Research Part B: Methodological*, **36** (7) 577–592.
- Charypar, D. and K. Nagel (2005) Generating complete all-day activity plans with genetic algorithms, *Transportation*, **32** (4) 369–397.
- Chorus, C. G., T. A. Arentze and H. J. P. Timmermans (2008) A random regret-minimization model of travel choice, *Transportation Research Part B: Methodological*, **42** (1) 1–18.
- Chorus, C. G. and H. J. P. Timmermans (2009) An empirical study into the influence of travel behavior on stated and revealed mental maps, paper presented at the *88th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2009.
- Chu, C. (1989) A paired combinatorial logit model for travel demand analysis, paper presented at the *5th World Conference on Transportation Research*, Yokohama, June 1989.
- Chung, E.-H. and A. Shalaby (2005) A trip bases reconstruction tool for GPS-based personal travel surveys, *Transportation Planning and Technology*, **28** (5) 381–401.
- Daganzo, C. F. (1979) *Multinomial Probit: The Theory and Its Application to Demand Forecasting*, Academic Press, New York.
- Daganzo, C. F. and Y. Sheffi (1977) On stochastic models of traffic assignment, *Transportation Science*, **11** (3) 253–274.

- Daly, A. J. and M. Bierlaire (2006) A general and operational representation of Generalised Extreme Value models, *Transportation Research Part B: Methodological*, **40** (4) 285–305.
- de Jong, R. and W. Mensonides (2003) Wearable GPS device as a data collection method for travel research, *Working Paper, ITS-WP-03-02*, University of Sydney, Institute of Transport Studies, Sydney.
- de la Barra, T., B. Pérez and J. Añez (1993) Multidimensional path search and assignment, paper presented at the *21st Planning and Transport, Research and Computation (PTRC) Summer Meeting*, Manchester, September 1993.
- Debreu, G. (1960) Review of R.D. Luce individual choice behavior, *American Economic Review*, **50** (1) 186–188.
- Dijkstra, E. W. (1959) A note on two problems in connexion with graphs, *Numerische Mathematik*, **1**, 269–271.
- Doherty, S. T., C. Noel, M. E. H. Lee-Gosselin, C. Sirois, M. Ueno and F. Theberge (2001) Moving beyond observed outcomes: Integrating Global Positioning Systems and interactive computer-based travel behaviour surveys, *Transportation Research E-Circular*, **C026**, 449–466.
- Draijer, G., N. Kalfs and J. Perdok (2000) Global Positioning System as data collection method for travel research, *Transportation Research Record*, **1719**, 147–153.
- Du, J. and L. Aultman-Hall (2007) Increasing the accuracy of trip rate information from passive multi-day GPS travel datasets: Automatic trip end identification issues, *Transportation Research Part A: Policy and Practice*, **41** (3) 220–232.
- Dugge, B. (2006) Ein simultanes Erzeugungs-, Verteilungs-, Aufteilungs- und Routenwahlmodell, Ph.D. Thesis, Technical University Dresden, Dresden.
- Dugundji, E. R. and J. L. Walker (2005) Discrete choice with social and spatial network interdependencies: An empirical example using mixed GEV models with field and “panel” effects, *Transportation Research Record*, **1921**, 70–78.
- Feil, M., M. Balmer and K. W. Axhausen (2009) Generating comprehensive all-day schedules: Expanding activity-based travel demand

- modelling, paper presented at the *European Transport Conference*, Leeuwenhorst, October 2009.
- Fishburn, P. C. (1982) Non-transitive measurable utility, *Journal of Mathematical Psychology*, **26** (1) 31–67.
- Flamm, M. and V. Kaufmann (2007) Combining person based GPS tracking and prompted recall interviews for a comprehensive investigation of travel behaviour adaptation processes during life course transitions, paper presented at the *11th World Conference on Transportation Research*, Berkeley, June 2007.
- Fotheringham, A. S. (1983) A new set of spatial interaction models: The theory of competing destinations, *Environment and Planning A*, **15** (1) 15–36.
- Frejinger, E. and M. Bierlaire (2007) Capturing correlation with sub-networks in route choice models, *Transportation Research Part B: Methodological*, **41** (3) 363–378.
- Frejinger, E., M. Bierlaire and M. E. Ben-Akiva (2009) Sampling of alternatives for route choice modeling, *Transportation Research Part B: Methodological*, **43** (10) 984–994.
- Friedrich, M., I. Hofsäss and S. Wekeck (2001) Timetable-based transit assignment using branch and bound techniques, *Transportation Research Record*, **1752**, 100–107.
- Greenfeld, J. S. (2002) Matching GPS observations to locations on a digital map, paper presented at the *81st Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2002.
- Guo, J. Y. and C. R. Bhat (2007) Operationalizing the concept of neighborhood: Application to residential location choice analysis, *Journal of Transport Geography*, **15** (1) 31–45.
- Hannes, E., D. Janssens and G. Wets (2008) Does space matter? travel mode scripts in daily activity travel, *Transportation Research Record*, **2054**, 20–27.
- Hess, S. (2005) Advanced discrete choice models with application to transport demand, Ph.D. Thesis, Imperial College London, London.
- Hess, S., M. Bierlaire and J. W. Polak (2005) Capturing taste heterogeneity and correlation structures with Mixed GEV models, paper

- presented at the *84th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2005.
- Hess, S. and J. W. Polak (2006) Exploring the potential for Cross-Nesting structures in airport-choice analysis: A case-study of the Greater London area, *Transportation Research Part E: Logistics and Transportation Review*, **42** (2) 63–81.
- Hettinger, T. (2008) Konsistenz von Entscheidungsmodellen, *Semester Project*, **221**, IVT, ETH Zurich, Zurich, March 2008.
- Hägerstrand, T. (1970) What about people in regional science?, *Papers of the Regional Science Association*, **24** (1) 7–21.
- Hong, S. P., J.-h. Kim, H. Jang and S. Lee (2006) The roles of categorization, affective image and constraints on destination choice: An application of the NMNL model, *Tourism Management*, **27** (5) 750–761.
- Hoogendoorn-Lanser, S. and P. H. L. Bovy (2007) Modeling overlap in multi-modal route choice by inclusion of trip part specific path size factors, *Transportation Research Record*, **2003**, 74–83.
- Hoogendoorn-Lanser, S., R. van Nes and P. H. L. Bovy (2005a) Path size and overlap in multimodal transport networks, in H. S. Mahmassani (ed.) *Flow, Dynamics and Human Interaction - Proceedings of the 16th International Symposium on Transportation and Traffic Theory*, 63–83, Elsevier, Oxford.
- Hoogendoorn-Lanser, S., R. van Nes and P. H. L. Bovy (2005b) Path size modeling in multimodal route choice analysis, *Transportation Research Record*, **1921**, 27–34.
- Hoogendoorn-Lanser, S., R. van Nes and P. H. L. Bovy (2006) A rule-based approach to multi-modal choice set generation, paper presented at the *11th International Conference on Travel Behaviour Research (IATBR)*, Kyoto, August 2006.
- Horni, A., D. M. Scott, M. Balmer and K. W. Axhausen (2009) Location choice modeling for shopping and leisure activities with MATSim: Combining micro-simulation and time geography, *Transportation Research Record*, **2135**, 87–95.
- Horni, A., D. M. Scott, M. Balmer and K. W. Axhausen (forthcoming) Location choice modeling for shopping and leisure activities with

- MATSim: Combining micro-simulation and time geography, *Transportation Research Record*.
- Huybers, T. (2003) Destination choice modelling - to label or not to label?, paper presented at the *Tourism Modelling and Competitiveness Conference*, Paphos, Cyprus, November 2003.
- Joh, C.-H., T. A. Arentze and H. J. P. Timmermans (2001) A position-sensitive sequence alignment method illustrated for space-time activity-diary data, *Environment and Planning A*, **33** (2) 313–338.
- Jun, J., R. Guensler and J. Ogle (2007) Smoothing methods to minimize impact of Global Positioning System random error on travel distance, speed, and acceleration profile estimates, *Transportation Research Record*, **1972**, 141–150.
- Kahneman, D. and A. Tversky (1979) Prospect theory: An analysis of decisions under risk, *Econometrica*, **47** (2) 263–291.
- Kawasaki, T. and K. W. Axhausen (2009) Choice set generation from GPS data set for grocery shopping location choice modelling in canton Zurich: The comparison to Swiss Microcensus 2005, *Working Paper*, **595**, IVT, ETH Zurich, Zurich.
- Kitamura, R. (1984) Incorporating trip chaining into analysis of destination choice, *Transportation Research Part B: Methodological*, **18** (1) 67–81.
- Koppelman, F. S. and C.-H. Wen (2000) The paired combinatorial logit model: Properties, estimation and application, *Transportation Research Part B: Methodological*, **34** (2) 75–89.
- Kuby, M., X. Zhongyi and X. Xiaodong (1997) A minimax method for finding the k best 'differentiated' paths, *Geographical Analysis*, **29** (4) 298–313.
- Lawler, E. L. (1976) *Combinatorial Optimization: Networks and Matroids*, Holt, Rinehart & Winston, New York.
- Lefebvre, N. and M. Balmer (2007) Fast shortest path computation in time-dependent traffic networks, paper presented at the *7th Swiss Transport Research Conference*, Ascona, September 2007.

- Liao, L., D. Fox and H. Kautz (2007) Extracting places and activities from GPS traces using hierarchical conditional random fields, *International Journal of Robotics Research*, **26** (1) 119–134.
- Loomes, G. and R. Sugden (1982) Regret theory: An alternative of rational choice under uncertainty, *Economic Journal*, **92** (368) 805–842.
- Manski, C. F. (1977) The structure of random utility models, *Theory and Decision*, **8** (3) 229–254.
- Marchal, F., J. K. Hackney and K. W. Axhausen (2005) Efficient map matching of large Global Positioning System data sets: Tests on speed-monitoring experiment in Zürich, *Transportation Research Record*, **1935**, 93–100.
- Martínez, F., F. Aguila and R. Hurtubia (2009) The constrained multinomial logit: A semi-compensatory choice model, *Transportation Research Part B: Methodological*, **43** (3) 365–377.
- Martínez, F., E. Cascetta, F. Pagliara, M. Bierlaire and K. W. Axhausen (2008) An application of the Constrained Multinomial Logit (CMNL) for modelling dominated choice alternatives, paper presented at the *8th Swiss Transport Research Conference*, Ascona, October 2008.
- MATSim-T (2008) Multi Agent Transportation Simulation Toolkit, webpage, URL <http://www.matsim.org>.
- McFadden, D. (1974) Conditional logit analysis of qualitative choice-behaviour, in P. Zarembka (ed.) *Frontiers in Econometrics*, 105–142, Academic Press, New York.
- McFadden, D. (1975) On Independence, Structure, and Simultaneity in Transportation Demand Analysis, *Working Paper*, **7511**, Institute of Transportation Studies, University of California, Berkeley.
- McFadden, D. (1978) Modeling the choice of residential location, in A. Karlqvist (ed.) *Spatial Interaction Theory and Residential Location*, 75–96, North-Holland, Amsterdam.
- McFadden, D., A. Talvitie, S. Cosslett, I. Hasan, M. Johnson, F. A. Reid and K. E. Train (1977) Demand model estimation and validation, *Research Report*, Urban Travel Demand Forecasting Project Phase 1, **5**, National Science Foundation (USA), Institute of Transportation Studies, University of California, Berkeley and Irvine.

- McFadden, D. and K. E. Train (2000) Mixed MNL models for discrete response, *Journal of Applied Econometrics*, **15** (5) 447–470.
- Menghini, G., N. Carrasco, N. Schüssler and K. W. Axhausen (2009) Route choice of cyclists: Discrete choice modelling based on GPS-data, *Working Paper*, **544**, IVT, ETH Zurich, Zurich.
- Meyer, R. J. and T. C. Eagle (1982) Context-induced parameter instability in a disaggregate-stochastic model of store choice, *Journal of Marketing Research*, **19** (1) 62–71.
- Mohammadian, A. K., M. Haider and P. S. Kanaroglou (2005) Incorporating spatial dependencies in random parameter discrete choice models, paper presented at the *84th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2005.
- Moiseeva, A., J. Jessurun and H. J. P. Timmermans (2010) Semi-automatic imputation of activity travel diaries using GPS-traces, prompted recall and context-sensitive learning algorithms, paper presented at the *89th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2010.
- Mondschein, A., E. Blumenberg and B. D. Taylor (2008) Accessibility and cognition: Effect of transportation mode on spatial knowledge, paper presented at the *87th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2008.
- Morikawa, T. (1996) A hybrid probabilistic choice set model with compensatory and noncompensatory choice rules, in D. A. Hensher, J. King and T. Oum (eds.) *World Transport Research: Proceedings of the 7th World Conference on Transport Research*, vol. 1, 317–325, Pergamon, Oxford.
- Nielsen, O. A., C. Würtz and R. M. Jørgensen (2004) Improved map-matching algorithms for GPS-data - methodology and test on data from the AKTA roadpricing experiment in Copenhagen, paper presented at the *19th European Conference for ESRI Users*, Copenhagen, November 2004.
- Ochieng, W. Y., M. A. Quddus and R. B. Noland (2004) Map matching in complex urban road networks, *Brazilian Journal of Cartography*, **55** (2) 1–18.

- Ogle, J., R. Guensler, W. Bachman, M. Koutsak and J. Wolf (2002) Accuracy of Global Positioning System for determining driver performance parameters, *Transportation Research Record*, **1818**, 12–24.
- Papola, A. (2004) Some development on the cross-nested logit model, *Transportation Research Part B: Methodological*, **38** (9) 833–851.
- Pasquier, M., U. Hofman, F. H. Mende, M. May, D. Hecker and C. Körner (2008) Modelling and prospects of the audience measurement for outdoor advertising based on data collection using GPS devices (electronic passive measurement system), paper presented at the *8th International Conference on Survey Methods in Transport*, Nancy, May 2008.
- Pearson, D. (2001) Global Positioning System (GPS) and travel surveys: Results from the 1997 Austin household survey, paper presented at the *8th Conference on the Application of Transportation Planning Methods*, Corpus Christi, April 2001.
- Prato, C. G. (2009) Route choice modeling: Past, present and future research directions, *Journal of Choice Modelling*, **2** (1) 65–100.
- Prato, C. G. and S. Bekhor (2006) Applying branch-and-bound technique to route choice set generation, *Transportation Research Record*, **1985**, 19–28.
- Prato, C. G. and S. Bekhor (2007a) Modeling route choice behavior: How relevant is the choice set composition?, paper presented at the *86th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2007.
- Prato, C. G. and S. Bekhor (2007b) Modeling route choice behavior: How relevant is the composition of choice set?, *Transportation Research Record*, **2003**, 64–73.
- Pyo, J.-S., D.-H. Shin and T.-K. Sung (2001) Development of a map matching method using the multiple hypothesis technique, paper presented at the *Intelligent Transportation Systems Conference (ITSC)*, Oakland, August 2001.
- Quddus, M. A., W. Y. Ochieng and R. B. Noland (2007) Current map matching algorithms for transport applications: State-of-the-art and future research directions, *Transportation Research Part C: Emerging Technologies*, **15** (5) 312–328.

- Quddus, M. A., W. Y. Ochieng, L. Zhao and R. B. Noland (2003) A general map matching algorithm for transport telematics applications, *GPS Solutions*, **7** (3) 157–167.
- Quddus, M. A., W. Y. Ochieng, L. Zhao and R. B. Noland (2006) A high accuracy fuzzy logic-based map matching algorithm for road transport, *Journal of Intelligent Transportation Systems*, **10** (3) 103–115.
- Ramming, M. S. (2002) Network knowledge and route choice, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge.
- Rich, J., S. L. Mabit and O. A. Nielsen (2007) Route choice model for Copenhagen: A data-driven choice set generation approach based on GPS data, paper presented at the *6th Triennial Symposium on Transportation Analysis (TRISTAN)*, Phuket Island, June 2007.
- Rieser, M. and K. Nagel (2009) Combined agent-based simulation of private car traffic and transit, paper presented at the *12th International Conference on Travel Behaviour Research (IATBR)*, Jaipur, December 2009.
- Sazonov, E. S., P. Klinkhachorn, H. V. GangaRao and U. B. Halabe (2002) Fuzzy logic expert system for automated damage detection from changes in strain energy mode shapes, *Nondestructive Testing and Evaluation*, **18** (1) 1–17.
- Schlich, R. (2001) Analysing intrapersonal variability of travel behaviour using the Sequence Alignment Method, paper presented at the *European Transport Conference*, Cambridge.
- Schönfelder, S., H. Li, R. Guensler, J. Ogle and K. W. Axhausen (2006) Analysis of commute Atlanta instrumented vehicle GPS data: Destination choice behavior and activity spaces, paper presented at the *85th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2006.
- Schönfelder, S. and U. Samaga (2003) Where do you want to go today? - More observations on daily mobility, paper presented at the *3rd Swiss Transport Research Conference*, Ascona, March 2003.
- Scott, D. M. (2006) Constrained destination choice set generation: A comparison of GIS-based approaches, paper presented at the *85th Annual Meeting of the Transportation Research Board*, Washington, D.C., January 2006.

- Sivakumar, A. and C. R. Bhat (2007) Comprehensive, unified framework for analyzing spatial location choice, *Transportation Research Record*, **2003**, 103–111.
- Small, K. A. (1987) A discrete choice model for ordered alternatives, *Econometrica*, **55** (2) 409–424.
- Stopher, P. R. (2009) Collecting and processing data from mobile technologies, in P. Bonnel, J. Zmud, M. E. H. Lee-Gosselin and J.-L. Madre (eds.) *Transport Survey Methods: Keeping Up with a Changing World*, chap. 21, 361–391, Emerald, Bingley.
- Stopher, P. R., Q. Jiang and C. FitzGerald (2005) Processing GPS data from travel surveys, paper presented at the *2nd International Colloquium on the Behavioural Foundations of Integrated Land-use and Transportation Models: Frameworks, Models and Applications*, Toronto, June 2005.
- Stopher, P. R., Q. Jiang and C. FitzGerald (2007) Deducing mode and purpose from GPS data, paper presented at the *11th TRB National Transportation Planning Applications Conference*, Daytona Beach, May 2007.
- Stouffer, S. A. (1940) Intervening opportunities: A theory relating mobility and distance, *American Sociological Review*, **5** (6) 845–867.
- Swait, J. (2001) Choice set generation within the generalized extreme value family of discrete choice models, *Transportation Research Part B: Methodological*, **35** (7) 643–666.
- Swait, J. and M. E. Ben-Akiva (1987) Incorporating random constraints in discrete models of choice set generation, *Transportation Research Part B: Methodological*, **21** (2) 91–102.
- Swiss Federal Statistical Office (2000) Eidgenössische Volkszählung 2000, URL http://www.bfs.admin.ch/bfs/portal/de/index/infothek/erhebungen__quellen/blank/blank/vz/uebersicht.html.
- Swiss Federal Statistical Office (2006) *Ergebnisse des Mikrozensus 2005 zum Verkehrsverhalten*, Swiss Federal Statistical Office, Neuchatel.
- Train, K. E. (1986) *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand*, MIT Press, Cambridge.

- Tsui, S. Y. A. and A. Shalaby (2006) An enhanced system for link and mode identification for GPS-based personal travel surveys, *Transportation Research Record*, **1972**, 38–45.
- Upadhyay, D., N. Schüssler, K. W. Axhausen, M. Flamm and V. Kaufmann (2008) Optimal parameter values for mode detection in GPS post-processing: An experiment, *Working Paper*, **506**, IVT, ETH Zurich, Zurich.
- van der Zijpp, N. J. and S. Fiorenzo-Catalano (2005) Path enumeration by finding the constrained k-shortest paths, *Transportation Research Part B: Methodological*, **39** (6) 545–563.
- van Eggermond, M. (2007) Consumer choice behavior and strategies of air transportation service providers, Master Thesis, IVT, ETH Zurich and Institute for Transport Planning, Technical University Delft, Zurich and Delft.
- van Eggermond, M., N. Schüssler and K. W. Axhausen (2007) Accounting for similarities in air transport route choice, *Working Paper*, **496**, IVT, ETH Zurich, Zurich.
- Velaga, N. R., M. A. Quddus and A. L. Bristow (2009) Developing an enhanced weight-based topological map-matching algorithm for intelligent transport systems, *Transportation Research Part C: Emerging Technologies*, **17** (6) 672–683.
- Vovsha, P. (1997) Cross-Nested Logit model: An application to mode choice in the Tel-Aviv metropolitan area, *Transportation Research Record*, **1607**, 6–15.
- Vovsha, P. and S. Bekhor (1998) The link-nested logit model of route choice: Overcoming the route overlapping problem, *Transportation Research Record*, **1645**, 133–142.
- Vovsha, P., E. Petersen and R. Donnelly (2002) Microsimulation in travel demand modeling: Lessons learned from the New York best practice model, *Transportation Research Record*, **1805**, 68–77.
- Vrtic, M. (2003) Simultanes Routen- und Verkehrsmittelwahlmodell, Ph.D. Thesis, Technical University Dresden, Dresden.
- Vrtic, M., P. Fröhlich, N. Schüssler, S. Dasen, S. Erne, B. Singer, K. W. Axhausen and D. Lohse (2005) Erzeugung neuer Quell-/Zielmatrizen

- im Personenverkehr, *Research Report*, Swiss Federal Department for Environment, Transport, Energy and Communication, Swiss Federal Office for Spatial Development, Swiss Federal Roads Authority and Swiss Federal Office of Transport, IVT, ETH Zurich, Emch und Berger, Institute for Transportation Planning and Traffic, Technical University Dresden, Zurich.
- Wagner, D. P. (1997) Lexington area travel data collection test: GPS for personal travel surveys, *Final Report*, Office of Highway Policy Information and Office of Technology Applications, Federal Highway Administration, Battelle Transport Division, Columbus, September 1997.
- Walker, J. L. (2002) The mixed logit (or logit kernel) model: Dispelling misconceptions of identification, *Transportation Research Record*, **1805**, 86–98.
- Walker, J. L., M. E. Ben-Akiva and D. Bolduc (forthcoming) Identification of parameters in normal error component logit-mixture (NECLM) models, *Journal of Applied Econometrics*.
- Weis, C. (2006) Routenwahl im ÖV, Master Thesis, IVT, ETH Zurich, Zurich.
- Wen, C.-H. and F. S. Koppelman (2001) The generalized nested logit model, *Transportation Research Part B: Methodological*, **35** (7) 627–641.
- White, C. E., D. Bernstein and A. L. Kornhauser (2000) Some map matching algorithms for personal navigation assistants, *Transportation Research Part C: Emerging Technologies*, **8** (1–6) 91–108.
- Wilson, C. (1998) Analysis of travel behavior using sequence alignment methods, *Transportation Research Record*, **1645**, 52–59.
- Wilson, C. (2008) Activity patterns in space and time: Calculating representative Hagerstrand trajectories, *Transportation*, **35** (4) 485–499.
- Wilson, C., A. S. Harvey and J. Thompson (2005) ClustalG: Software for analysis of activities and sequential events, paper presented at the *Workshop on Sequence Alignment Methods*, Halifax, October 2005.
- Wolf, J. (2000) Using GPS data loggers to replace travel diaries in the collection of travel data, Ph.D. Thesis, Georgia Institute of Technology, Atlanta.

- Wolf, J. (2006) Applications of new technologies in travel surveys, in P. R. Stopher and C. C. Stecher (eds.) *Travel Survey Methods - Quality and Future Directions*, 531–544, Elsevier, Oxford.
- Wolf, J., R. Guensler and W. Bachman (2001) Elimination of the travel diary - experiment to derive trip purpose from Global Positioning System travel data, *Transportation Research Record*, **1768**, 125–134.
- Wolf, J., S. Hallmark, M. Oliveira, R. Guensler and W. Sarasua (1999) Accuracy issues with route choice data collection by using Global Positioning System, *Transportation Research Record*, **1660**, 66–74.
- Wolf, J., M. Oliveira and M. Thompson (2003) Impact of underreporting on mileage and travel time estimates - results from Global Positioning System-enhanced household travel survey, *Transportation Research Record*, **1854**, 189–198.
- Wolf, J., S. Schönfelder, U. Samaga, M. Oliveira and K. W. Axhausen (2004) Eighty weeks of Global Positioning System traces, *Transportation Research Record*, **1870**, 46–54.
- Yai, T., S. Iwakura and S. Morichi (1997) Multinomial probit with structured covariance for route choice behaviour, *Transportation Research Part B: Methodological*, **31** (3) 195–208.
- Yalamanchili, L., R. M. Pendyala, N. Prabakaran and P. Chakravarty (1999) Analysis of Global Positioning System-based data collection methods for capturing multistop trip-chaining behavior, *Transportation Research Record*, **1660**, 58–65.
- Zheng, Y., L. Liu, L. Wang and X. Xie (2008) Learning transportation mode from raw GPS data for geographic applications on the web, paper presented at the *17th World Wide Web Conference*, Beijing, April 2008.

