DISS. ETH NO. 26114

TRANSPORT DEMAND MODELS: A SPATIAL PERSPECTIVE

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES
of
ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

GEORGIOS SARLAS
M.Sc. Transport Systems, KTH, Sweden
Dipl. Engineer, NTUA, Greece

born on 12th March, 1986
citizen of Athens, Greece

accepted on the recommendation of

Prof. Dr. Kay W. Axhausen, examiner
Prof. Dr. Antonio Páez, co-examiner

2019

DOI: 10.3929/ethz-b-000370441

Et in Arcadia ego

ABSTRACT

Travel demand models have increased their data demands massively both in scope and scale as they have become more complex over the course of years. Against that trend, the current dissertation pursues the development of a direct demand modelling approach tailored for speed and volume prediction purposes. In this regard, the main objective of this dissertation is twofold. First, to investigate how the predictive accuracy of a direct demand model can be enhanced if viewed through the lens of a spatial perspective, and second to identify and resolve the statistical shortcomings that arise due to the spatial nature of data. Methodologically, the family of spatial regression models is exploited while the issue of endogeneity governing the relationship between the two modelled transport phenomena is taken into consideration.

On the spatial interaction side, a new framework that revisits the definition of the distance decay function is introduced. Subsequently, this revision is translated into a series of modified accessibility measures. Furthermore, a new indicator that combines the concepts of centrality and gravity-based accessibility in a unified measure is introduced. This indicator provides a richer picture of the ways a transport system generates connectivity and how accessibility is jointly generated by the network and the landscape of opportunities. In addition, the new centrality indicator is thoroughly tested for its ability to improve the predictive accuracy of direct demand models.

Finally, a comparison of the results of the developed modelling approach against the output of a traditional four-step model showcases that direct demand models can provide a trustworthy alternative to more advanced, but definitely more data demanding and computationally burdensome approaches. Especially in cases where the development of more advanced models is not possible, either due to data availability issues, or due to various limitations in place, a direct demand model can constitute a viable alternative.

ZUSAMMENFASSUNG

Die gegenwärtigen Verkehrsnachfragemodelle werden zunehmend komplexer und beruhen auf immer grösseren sowie detaillierteren Daten. Im Gegensatz zu diesem Trend konzentriert sich die vorliegende Arbeit auf die Entwicklung eines direkten Ansatzes zur Modellierung der Verkehrsnachfrage. Im Vordergrund steht dabei die Vorhersage von Verkehrsgeschwindigkeiten und -mengen. In diesem Zusammenhang ergeben sich zwei Hauptziele für diese Dissertation. Das erste besteht darin zu untersuchen, wie die Genauigkeit der Vorhersagen direkter Nachfragemodelle verbessert werden kann, wenn eine räumliche Perspektive eingenommen wird. Das zweite Ziel ist es, die statistischen Mängel zu identifizieren und beheben, die bei der Schätzung solcher Modelle aufgrund des räumlichen Bezugs der Daten entstehen. Zu diesem Zweck werden verschiedene räumliche Regressionsmodelle vorgestellt, getestet und genutzt, während gleichzeitig das Problem der Endogenität explizit angesprochen und berücksichtigt wird.

Für die Thematik der räumlichen Interaktion wird ein neuer Ansatz präsentiert, welcher verschiedene Erreichbarkeitsindikatoren überarbeitet und neu definiert, die auf distanzbasierten Widerstandsfunktionen basieren. Es wird ein neuer Indikator vorgestellt, der die zentralitäts- sowie gravitationsbasierte Erreichbarkeit als einheitliches Mass kombiniert. Dieser Indikator liefert ein besseres Bild davon, auf welche Art und Weise ein Verkehrssystem verbunden ist und wie die Erreichbarkeit durch das Netzwerk und die Aktivitäts-möglichkeiten gemeinsam erzeugt wird. Darüber hinaus wird der neue Zentralitätsindikator eingehend auf seine Fähigkeit getestet, die Vorhersagegenauigkeit von direkten Nachfragemodellen zu verbessern.

Als Abschluss zeigt ein Vergleich der Ergebnisse des neu entwickelten Modellierungsansatzes mit jenen eines traditionellen, vier-stufigen Modells, dass direkte Nachfragemodelle eine vertrauenswürdige Alternative zu komplexeren, aber weitaus daten- und rechenintensiveren Ansätzen darstellen können. Insbesondere in Fällen, in denen die Entwicklung von komplexen Modellen nicht möglich ist, sei es aus Gründen der Datenverfügbarkeit oder aufgrund verschiedener anderer Einschränkungen, kann ein direktes Nachfragemodell eine brauchbare Alternative darstellen.

ACKNOWLEDGEMENTS

The journey of the past years has been characterized by peaks and valleys. A number of exceptional individuals have aided into formulating the ideas in this thesis. Initially, I would like to express my sincere gratitude and appreciation to Prof. Kay W. Axhausen for offering me the opportunity to conduct my doctoral thesis in his group and also for his constructive guidance throughout my PhD studies. Giving me the independence to explore subjects that I was curious of investigating was definitely one of the key aspects that contributed positively to the overall PhD journey.

I would like to thank Prof. Antonio Páez for not only accepting to be my external supervisor but also hosting me for a brief but inspiring time at McMaster University. His valuable comments on my research are greatly appreciated. I want to thank my colleagues at the IVT for creating a fruitful and vibrant working environment and especially my office mates (even though sometimes they were more talkative than I would have hoped for). I am deeply grateful for the financial support that was granted by SNSF (Swiss National Science Foundation) through the project *Models without (personal) data?*.

I would also like to take the opportunity and thank my Diploma and MSc theses supervisors, the late Prof. Petros Vythoulkas, Prof. Constantinos Antoniou, and Prof. Haris Koutsopoulos, that have inspired me to pursue a PhD, each on his own way.

Moving to a new place is always a challenge. Nevertheless, I was fortunate enough to meet many new people and create strong friendships that helped me maintain a healthy work-life balance and enjoy my time in Zurich. So, special thanks go to Dimitris, Giannis, Lena, and the rest of the "gang".

A huge thank you goes to Eve for her support, patience and being by my side. Last, but not least I would like to thank my family for their constant support in the last 33 years.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

INTRODUCTION

1.1 MOTIVATION

Travel demand models keep increasing their complexity levels since the shift to the activity based paradigm in the early 1970's (Jones et al., 1983; Axhausen, 1998). As a result, their data demands have also increased massively, both in scale and scope in order to support this enhancement. While the accrued gains in policy sensitivity and plausibility are obvious, there has been little evidence that they actually improve the quality of public decision making. The obvious reluctance of the practice to adopt such advanced models has raised concerns on potentially widening the gap between academia and practice (US National Research Council , 2007). On the one hand, the increased data collection abilities of the field, along with the expected wave of "big data" might allow the (academic) field to continue on its current trajectory, but on the other hand the use and abuse of "big data" raises the danger of a sudden change in the course of public policy and the sudden lack of high quality alternatives to the existing state-of-the-art (*i.e.* academic) models.

A closer look at the history of transport demand models reveals that the advances in the field have been mainly materialized due the evolution of computational systems. Initially, the demand models had adopted a low level of spatial and temporal detail with simplified behavior considerations in place (de Dios Ortuzar and Willumsen, 2011). The joint increase in computing power, the enlargement of the range of policy instruments to be tested, and the growing prominence of geographers and economists among the transport planners led to more complex and theoretically better informed models: the temporal and spatial detail was increased, and the range of behavioral responses was broadened as well. The activity-based approach (Jones et al., 1983) was essentially merged in practice with the more or less simultaneous development of discrete choice models (Domencich and McFadden, 1975). Both the conceptual framework and the continuous advancement of the statistical tools (*e.g.* see Train (2009) for a summary of the relevant progress) invited a growing appetite for spatial and temporal detail, and more specific information about the decision processes modelled.

Nevertheless, the underlying core idea of all transport demand models is to examine and quantify the interaction between demand and supply at different levels. To this end, the most prominent modelling approaches rely on the use of simulation models, comprising a set of sub-models to quantify the different aspects of the transport system, and then through an iterative process facilitate the interaction between demand and supply until an equilibrium has been reached. Two broad categories of such models exist, differentiated on how they simulate the interactions within the system, hence involving and allowing different considerations and sub-model formulations. The first category, the macroscopic approach, focuses on the system as a whole and models its different components and their interactions in an aggregate way. The second category, the microscopic approach, considers the individual components of the system and models their behavior and interactions in a disaggregate way, making use of advanced statistical models.

A further distinction of the transport simulation models can be made based on how the demand aspect is modelled. On the one hand, models focusing on the operational side of the system consider fixed demand and focus on simulating individuals movements and interactions (*e.g.* traffic simulation models). On the other hand, models focusing on planning purposes, model the demand aspect of the system under the assumption that it is a derived need, and hence not fixed. Depending on how the generation of demand is formulated, a further distinction can be made between microscopic and macroscopic demand modelling approaches. A pivotal example of the former one is the traditional 4-step model, while in the case of the latter are the agent-based models (*e.g.* Horni et al. (2016)). Obviously, a microscopic demand model requires much more detailed data on a person level and the development of many statistical sub-models, increasing considerably the computational effort to reach the equilibrium point. While the increasing richness is obvious to the developers, it is also well documented that the practitioners do adopt these models only with very long delays, or if ever (US National Research Council , 2007).

However, when it comes to the appraisal of public transport projects, as Flyvbjerg et al. (2005) argue, the quality of the demand forecasts has not been improved over the years even though more complex and behaviorally sound models have been employed. In a similar line of thought, Dowling and Skabardonis (1993) highlight the fact that large scale planning models, which are only calibrated against volume estimates, typically fail to provide reasonable speed estimates. This aspect has been systematically neglected in the literature along with its implications. More specifically, the travel time and demand estimates constitute core elements when it comes to the system performance evaluation and the appraisal of new projects (*e.g.* estimation of travel time savings). Driven by these, it appears that the increased complexity of transport demand models has on the one hand contributed to the development of more elaborate and policy sensitive models, but on the other hand it has not improved the demand and travel time forecasts' quality, at least not in an analogous way to the required effort and data.

Based on the above, the question of potential modelling alternatives emerges as an important one. Nevertheless, any alternative modelling structure should be at least capable of making statements about the speed and the traffic volume on a link level, items that constitute the minimum requirements for the appraisal of transport projects. To this end, regression models have been employed for providing the required answers in recent years (*e.g.* Desyllas et al., 2003; Hackney et al., 2007), reviving essentially an old, but recently neglected tradition in transport planning of direct demand models (*e.g.* Quandt and Baumol, 1966; Talvitie, 1973; Crow et al., 1973; Gaudry and Wills, 1978; Oum, 1989).

1.2 RESEARCH OBJECTIVES AND OUTLINE

The formation of an alternative modelling framework in the spirit of direct demand models seems appealing for a number of reasons. First, it can offer a structural explanation of the modelled phenomena at any location on the network in a direct and straightforward manner. Second, the data and computational requirements are considerably lower than the prevailing simulation approaches. Third, it requires a substantially lower number of parameters to be estimated in order to provide predictions. Fourth, it can constitute a worthwhile alternative especially if its predictive performance is found to be within acceptable accuracy levels. Last, developing such a framework can provide valuable insights about the demand-supply interaction mechanism under study, which in turn can also supplement more elaborate approaches to improve their predictive capabilities and their overall performance.

Invoking the first law of geography stating that *"everything is related to everything else, but near things are more related than distant things"* (Tobler, 1970), the importance of distance as a fundamental concept of geography is emphasized. Besides the critical question of what is related to what (Miller, 2004), another dimension of distance is associated to the spatial interaction phenomenon.

Driven by all the above, the main objective of this dissertation is twofold. First, to investigate how the predictive accuracy of a direct demand model can be enhanced if viewed through the lens of a spatial perspective, and second to identify and treat for the implications that arise due to the spatial nature of data. Nevertheless, both of these aspects have in common the space dimension and as such the main hypothesis that this dissertation puts into testing is that space matters when it comes to the formulation of (direct) transport demand models.

More specifically, this dissertation pursues the formulation of a direct demand model capable of producing localized mean speed and volume predictions. Methodologically, the family of spatial regression models is employed for that purpose while the issue of endogeneity governing the relationship between these two transport phenomena is explicitly taken into consideration. On the spatial interaction side, a new framework that revisits the definition of distance decay function is introduced. Subsequently, this revision is translated into a series of modified accessibility measures which when combined with network theory concepts can provide a neat way of capturing the underlying demand patterns, and in turn can also improve the predictive accuracy of the developed direct demand model.

The overall structure of this dissertation takes the form of nine chapters, including this introductory chapter. Chapter 2 begins by discussing in parallel the literature on accessibility measures and spatial interaction models. The 3rd chapter revisits conceptually and methodologically the principles of spatial interaction and proposes an alternative based on the survival analysis concept. In chapter 4 the aforementioned revision is evaluated in terms of resulted interaction rates and values, while its application is also exemplified through the construction of different gravity-based accessibility measures. Chapter 5 introduces a new centrality indicator that associates accessibility to the network structure while a case study is designed to demonstrate and discuss its utility. In chapter 6 different spatial regression models for speed prediction purposes are estimated while the issue of the optimal construction of the spatial weight matrix is addressed. Chapter 7 focuses on the issue of traffic volume prediction and the output of both spatial and aspatial models are assessed in order to draw relevant conclusions. The 8th chapter proposes a speed model formulation that accounts both for spatial effects and endogeneity issues and when paired up with the model of the previous chapter can form a coherent direct demand modelling approach. The final chapter draws upon the entire dissertation, tying up the various theoretical and empirical contributions, and identifies areas for further research.

2

ACCESSIBILITY

2.1 INTRODUCTION

"Throughout the evolution of human settlements, there is only one factor which defines their extent: the distance man wants to go or can go in the course of his daily life. The shortest of the two distances defines the extent of the real human settlement, through definition of a daily urban system" (Doxiadis, 1970). Admittedly, transportation constitutes a core element of everyday life, having the ability to shape not only urban systems but also the regional form and function (Wachs and Kumagai, 1973).

Different ways have been introduced over the years in an attempt to quantify the access to spatially distributed opportunities, commonly referred to as accessibility. Seeking ways to increase accessibility and provide people with increased opportunities for employment, social participation, *etc.*, has always been at the forefront of transport planning. As a result, accessibility makes up a key concept that has found wide application in planning practice (Straatemeier, 2008). However, it is due to its popularity that it is often mentioned as a term but in several instances remains undefined or poorly measured (Axhausen, 2008). As Miller (2018) states, *"the 'simple' notion of accessibility becomes surprisingly difficult to operationalize"*.

Since Hansen (1959) first formulated accessibility in mathematical terms, different variants have emerged. Two broad categories of accessibility measures can be identified in the literature, differing on whether they draw on utility, or spatial interaction concepts. In the case of the former, the focus lies on quantifying the benefits that people derive from access to certain kind of opportunities W, facilitated through discrete choice modelling approaches. For example, in (Ben-Akiva and Lerman, 1979) the accessibility Acc_i^W of a location i is defined as the log-sum term of a destination choice model.

On the contrary, spatial interaction accessibilities are potential opportunities measures in the sense that they quantify the reachable opportunities from any location/person i. Different indicators have been proposed over the years, however what all of them have in common is that they rely conceptually on spatial interaction principles. A general accessibility formulation and in analogy to the workhorse of spatial interaction modelling, the gravity model (Wilson, 1967), can be formulated in the following way:

$$Acc_i^W = \sum_j W_j f(d_{ij}) \tag{2.1}$$

with Acc_i^W being the accessibility to opportunities W of a location/person i, defined as the sum of opportunities W_j at locations j (with $j \in N$, and N being the number of locations), weighed by the interaction intensity function $f(d_{ij})$. The interaction function depends on an attribute d_{ij}, specific to any particular interaction pair of locations/individuals. In summary, such measures offer an integrated way of quantifying land use and transport interaction in a concise and easily comprehensible manner, while their popularity stems to a large extent from their simplicity, making them by far the most commonly used in practice (Geurs and van Wee, 2004). Depending on whether or not the self-potential aspect of locations is taken into account, a further restriction can be imposed accordingly (*i.e.* $j \neq i$). A discussion on that issue can be found in Frost and Spence (1995).

As it can be seen in formula (2.1), spatial interaction accessibility measures rely on an interaction intensity function coupled with the spatial distribution of opportunity points of interest (*e.g.* population, work places, *etc.*). Depending on the scope and the availability of data, as noted by Páez et al. (2012), *"these two components can be deployed in a number of different ways to produce location- or person-based indicators of accessibility"*.

In general terms, a further classification can be made with respect to the analysis level. In this respect, two categories have emerged in the literature, namely the aggregate and the disaggregate accessibilities. In the case of the former, the analysis level is that of a zone, or a social group, and accessibility is viewed as a zonal, or group characteristic accordingly. On the contrary, disaggregate accessibility indicators focus on a finer analysis level such as individuals, or specific elements of space (*e.g.* points).

Nevertheless, both categories utilize observed mobility patterns and travel behavior aspects in order to capture information about how far people are willing, or have to travel. In particular, the aggregate measures are conceptually and mathematically driven by the advances in the field of spatial interaction modelling, while their determination takes place based on population aggregates. On the other hand, disaggregate measures have also developed along the same broad conceptual lines, entailing though the formulation of individual- and/or location-specific interaction functions.

In addition, further classification ways have been proposed over the years. Most notably, in the review paper by Páez et al. (2012) a distinction on the basis of whether the accessibility measure pertains to a normative or a positive situation was introduced. In this regard, the authors provide the following definition: *"normative accessibility measures are defined in terms of how far people ought to travel or how far it is reasonable for people to travel whereas positive accessibility measures are defined in terms of how far people actually travel"* (Páez et al., 2012). Nevertheless, highlighting this dimension of accessibility measures is of importance and relevance, especially for policy making applications. A thorough discussion of the literature on the topic of accessibility can be found in Geurs and Ritsema van Eck (2003); Páez et al. (2012), while in another paper by Van Wee (2016) different challenges associated with the formation and the use of accessibilities are presented.

2.2 AGGREGATE ACCESSIBILITY MEASURES

Three categories of aggregate accessibility measures have emerged, namely the gravity-based, the competition effects, and the cumulative opportunity ones. Their central difference lies in their underlying assumptions of spatial interaction, and how these are incorporated into their formulations. *"Spatial interaction is the process whereby entities at different points in physical space make contacts, demand/supply decisions or locational choices"* (Roy and Thill, 2003).

The importance of spatial interaction as a phenomenon has been acknowledged in different disciplines, varying from transportation, to migration and trade flows. In this regard, different models have been introduced over the years in an attempt to explain such processes. Nevertheless, the vast majority of those models have developed with an aim to describe aggregate human spatial actions (Sheppard, 1984). A theoretical analysis of the aggregate nature of spatial interaction models is presented in Webber (1980), while in another paper by Ubøe (2004) the associated shortcomings are highlighted. A thorough overview on the topic is given by Fotheringham and O'Kelly (1989), and Roy and Thill (2003).

2.2.1 *Gravity-based accessibility measures*

In the case of the gravity-based accessibility measures, essentially the interaction intensity function serves the purpose of discounting the access to the opportunities. The accessibility formulation is in this case based on the production-/ attraction-/ doubly-constrained forms of the gravity model, with accessibility being the corresponding production and attraction balancing factors (Wilson, 1971). It is worthwhile putting into perspective the gravity model formulation to allow for a more in-depth discussion. The unconstrained gravity model for trip distribution is given by equation (2.2), in accordance with Wilson (1971).

$$T_{ij} = K\, W_i^{(1)} W_j^{(2)} f(d_{ij}) \tag{2.2}$$

with T_{ij} being the interaction between locations i and j (trips for the transportation case), $W_i^{(1)}$ and $W_j^{(2)}$ are the mass measure terms associated with locations i and j respectively, and K is a proportionality constant. Last, $f(d_{ij})$ is the interaction intensity function, commonly referred to as distance decay or deterrence function, with d_{ij} normally being a measure of the distance, or (generalised) transport cost of travel between zones i and j.

An important aspect of the gravity model is the definition of mass terms $W_i^{(1)}$ and $W_j^{(2)}$. Conceptually these terms should correspond to measures of trip productiveness and attractiveness, respectively. Thereupon, they can be of two kinds: either a measure of total trip production or attraction per location, or some relevant factors (Wilson, 1971). In the case of the former, the total number of trips originating from each zone $O_i = \sum_j T_{ij}$, or terminating at each zone $D_j = \sum_i T_{ij}$, can be either observed, or estimated on the basis of modelling techniques (*i.e.* $W_i^{(1)} = O_i$ and $W_j^{(2)} = D_j$). In the factors' case, the employment of proxy variables serves the purpose of quantifying trip productiveness and attractiveness. Typically, different opportunity variables are used for that reason, such as population and employment opportunities (*i.e.* $W_i^{(1)} = Pop_i$ and $W_j^{(2)} = Empl_j$).

As Wilson (1971) points out, the first mass terms' kind has the advantage of constituting a well-defined measure, while the second one is by conception difficult to associate with a dimension. However, lacking dimensionality bears no modelling implications when production and attraction constraints are imposed. For instance, upon knowing the total number of trips O_i or/and D_j, we can integrate that knowledge in the form of constraints into formula (2.2), resulting in the constrained counterparts of it. In particular, the production-constrained form of it then becomes:

$$T_{ij} = A_i O_i W_j^{(2)} f(d_{ij}) \tag{2.3}$$

while the attraction-constrained form is:

$$T_{ij} = B_j D_j W_i^{(1)} f(d_{ij}) \tag{2.4}$$

with A_i and B_j a set of balancing factors that replace the proportionality constant K, and hence ensure that the aforementioned constraints are satisfied. The set of factors A_i is given by:

$$A_i = \frac{1}{\sum_j W_j^{(2)} f(d_{ij})} \tag{2.5}$$

while the set of B_j is:

$$B_j = \frac{1}{\sum_i W_i^{(1)} f(d_{ij})} \tag{2.6}$$

Therefore and as it can be observed in formulas (2.5) and (2.6), the denominators of the factors A_i and B_j correspond in essence to the well-known gravity-based accessibility measures (formula (2.1)), having their mass terms replaced by proxy variables instead of total trips' numbers. For instance for the case of the population and employment positions proxy variables, the balancing factors are:

$$A_i = \frac{1}{\sum_j Empl_j f(d_{ij})} \tag{2.7}$$

and

$$B_j = \frac{1}{\sum_i Pop_i f(d_{ij})} \tag{2.8}$$

In addition, the doubly constrained version of the gravity model can be formulated in the following way:

$$T_{ij} = A_i B_j O_i D_j f(d_{ij}) \tag{2.9}$$

where

$$A_i = \frac{1}{\sum_j B_j D_j f(d_{ij})} \tag{2.10}$$

$$B_j = \frac{1}{\sum_j A_i O_i f(d_{ij})} \tag{2.11}$$

In a similar way as before, the mass terms could have been replaced by the proxy variables instead. For instance and for the same proxy variables as before, the doubly-constrained gravity model is:

$$T_{ij} = A_i B_j Pop_i Empl_j f(d_{ij}) \tag{2.12}$$

where

$$A_i = \frac{1}{\sum_j B_j Empl_j f(d_{ij})} \tag{2.13}$$

$$B_j = \frac{1}{\sum_j A_i Pop_i f(d_{ij})} \tag{2.14}$$

The estimation of the doubly-constrained model can take place either in an iterative way, or based on the entropy maximization framework as presented by Wilson (1970). In this case, the balancing factors A_i and B_j can also be perceived as accessibility measures, incorporating both competition and capacity constraints. However, the drawback of these values is that they are not easily interpretable and also they constitute the output of an iterative process, making them rather unpopular in practice (Geurs and Ritsema van Eck, 2003). A recent application involving such accessibility measures can be found in Allen and Farber (2019), where the developed measures were used for examining spatial inequalities with respect to access to employment.

2.2.2 *Cumulative opportunities accessibility measures*

An alternative conceptual framework for studying spatial interaction is due to Stouffer (1940) who proposed that there is no necessary direct relationship between distance and interaction. More specifically, he introduced the concept of intervening opportunities stating *"that the number of persons* *going a given distance is directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities"* (Stouffer, 1940). Therefore, the proposed model assumes that the interaction between two locations (number of trips for the trip distribution case) is given by the following formula:

$$T_{ij} = K \frac{O_i D_j}{\sum\limits_k D_k} \tag{2.15}$$

with k being all the locations within a distance d_{ij} from location i while the term $\sum_k D_k$ quantifies the total number of intervening opportunities. Mathematically, the model can be perceived as a special case of the production-constrained gravity model (formula (2.3)), where geographical separation has no effect on interaction (hence $f(d_{ij}) = 1$), and both trip production and attraction is known *a priori*. An extension of the model to account for competition followed up (Stouffer, 1960) while recently it has regained popularity through a series of modified forms, denoted as radiation models (*e.g.* Simini et al., 2012; Yang et al., 2014). Essentially, these models operationalize the idea of replacing the mass terms by proxy variables.

Nevertheless, in analogy to the accessibility derivation for the gravity model, the denominator of formula (2.15) can be seen as the accessibility of location i to opportunities W within a distance d_{ij}, thus forming a cumulative opportunities measure, subject to a specific value of distance d_{ij} though. Therefore, the cumulative opportunities accessibility measure is a special case of the general spatial interaction accessibility formulation (2.1), having the following interaction intensity function in place:

$$f(d_{ij}) = \begin{cases} 0, & if \;\; d_{ij} > c \\ 1, & if \;\; d_{ij} \leq c \end{cases} \tag{2.16}$$

where c is the cut-off value that after that point no interaction is taken into consideration. Consequently, the function yields binary values of zero and one, hence classifying location pairs as interacting or non-interacting ones.

2.2.3 *Competition effects accessibility measures*

A major limitation of the gravity spatial interaction models is associated with lacking the ability to account for competition effects. As highlighted by Fotheringham (1983), this can be a source of serious misspecification, resulting in turn in biased interaction function parameter estimates. To overcome that shortcoming, Fotheringham (1983) proposed a new set of spatial interaction models that build on the concept of competing destinations. To that end, the proposed model takes into account the interlinked nature between spatial interaction and accessibility by incorporating that aspect into a modified origin-specific gravity model formulation. More specifically, the proposed model for the production-constrained case with known trip production takes on the following form:

$$T_{ij} = A_i O_i W_j^{(2)} \left(Acc_{ij}^{\mathrm{W}^{(2)}} \right)^{\delta_i} d_{ij}^{\alpha_i} \tag{2.17}$$

where

$$A_i = \frac{1}{\sum_j W_j^{(2)} \left(Acc_{ij}^{\mathrm{W}^{(2)}} \right)^{\delta_i} d_{ij}^{\alpha_i}} \tag{2.18}$$

The term $Acc_{ij}^{\mathrm{W}^{(2)}}$ quantifies the accessibility of destination j to all other destinations, defined as:

$$Acc_{ij}^{\mathrm{W}^{(2)}} = \sum_{k,k \neq i} W_k^{(2)} d_{jk}^{\sigma_i} \tag{2.19}$$

The parameter σ_i quantifies "*the importance of distance in determining the perception of accessibility*" (Fotheringham, 1983). The model includes origin-specific parameters (δ_i, α_i, and σ_i) which are obtained iteratively. Moreover, the expected sign of the parameter δ_i is negative, thus formula (2.18) can be rewritten in the following way:

$$A_i = \frac{1}{\sum_j \frac{W_j^{(2)} d_{ij}^{\alpha_i}}{\left(Acc_{ij}^{\mathrm{W}^{(2)}} \right)^{\delta_i}}} \tag{2.20}$$

Therefore, in this particular spatial interaction model formulation, the accessibility of origin i corresponds to the denominator of the aforementioned formula (2.20). Driven by such insights, a category of accessibility formulations accounting for competition effects has appeared in the literature.

For instance, in Joseph and Bantock (1982) an accessibility measure that accounts for competition effects at the destination level was introduced. Along the same line of thought, in Shen (1998) it is highlighted that a major limitation of traditional accessibility measures is that they incorporate only the supply side in their formulation, whereas the competition for available opportunities is not taken into account (demand side). Interestingly, in the same paper it is noted that the traditional accessibility measures remain valid only if at least one of the two following conditions is satisfied: a) the spatial distribution of demand is uniform, and b) there is no capacity constraint on the opportunities side. Failing to meet these conditions can result in misleading measures. Nonetheless, the roots of this argumentation are closely related to the limitations of the constrained forms of the gravity model. Furthermore, in the same paper an accessibility formulation accounting for the competition at destinations was proposed, aligned with the one presented in Joseph and Bantock (1982). In a similar vein, in a series of papers (Cheng and Bertolini, 2013; Geurs and Ritsema van Eck, 2003; Van Wee et al., 2001) various accessibility measures for analyzing job accessibility in the Netherlands were calculated, also demonstrating the importance of accounting for competition. A general form of the aforementioned category of accessibility measures is presented in formula (2.21).

$$cAcc_i^{\mathrm{W}} = \sum_j \frac{W_j}{Acc_j^{\mathrm{E}}} f(d_{ij}) \tag{2.21}$$

with Acc_j^{E} being the accessible potential opportunities E from location j, calculated based on formula (2.1). Essentially, this term can be perceived as a correction for the attractiveness of the opportunities at the destination level.

For the job competition case, W would correspond to the number of employment positions, while E would be the population. The outcome of formula (2.21) then can be interpreted as the number of potential regional employment opportunities per resident at location i. Thereupon, values of one indicate perfect regional balance whereas values higher than one reflect an excess of opportunities per resident. On the contrary, values lower than one indicate an imbalance of potential opportunities. An alternative way of accounting for spatial competition was introduced in Loder et al. (2017) where the accessibility is specified as a nominal variable based on combinations of ordinal values of cumulative accessibility variants. Earlier attempts for the cumulative opportunities case also exist in the literature (*e.g.* Stouffer, 1960).

2.2.4 *Interaction intensity function parameter(s)*

As was pointed out in a previous section, all three categories of aggregate accessibility measures rely on an interaction intensity function $f(d_{ij})$ to quantify the potential for interaction. The importance of this function is high since accessibility measures are found to be strongly influenced by the chosen parametric form along with the corresponding parameter(s) (Geurs and Ritsema van Eck, 2003; Reggiani et al., 2011; Vale and Pereira, 2017). Furthermore, it should be noted that transferability of the estimated functions among different regions and studies is not advisable since it can give rise to misleading and erroneous results (*e.g.* Pirie, 1979; Sarlas et al., 2015).

Nevertheless and as it was mentioned before, the interaction function depends on an attribute d_{ij}, specific to any particular pair of locations. On its most widely applied form, interaction intensity is viewed as a function of spatial segregation solely (*e.g.* Euclidean distance, network distance, or travel time/cost). Essentially, the function then takes on the form of a distance decay, or travel impedance function. Common forms are power, (power) exponential, logistic, (modified) Gaussian, and uniform.

Several approaches of obtaining the parameters of those functions have been proposed. In summary, they can be obtained either based on a gravity model (*e.g.* Fotheringham, 1983; Levinson, 1998; Simma and Axhausen., 2003; de Dios Ortuzar and Willumsen, 2011; Reggiani et al., 2011), or by fitting an empirical trip length distribution function (*e.g.* Zhao et al., 2003; Halás et al., 2014; Sarlas et al., 2015; Scott and Horner, 2008), or by specifying cut-off values (*e.g.* Gutierrez and Gomez, 1999; O'Kelly and Horner, 2003; Ribeiro et al., 2010).

In the first case, a gravity model formulation serves as the basic underlying aggregate relationship to model the interaction intensity while origin-destination (O-D) demand matrices are required as input (*e.g.* Levinson, 1998; Simma and Axhausen., 2003; Reggiani et al., 2011) along with census and count data. Moreover, the interaction function takes either a power form d_{ij}^{α}, or an exponential one $e^{\alpha d_{ij}}$, with α being an estimated parameter measuring the diminishing impact of distance/cost on interaction. Due to the nature of this relationship, α is expected to have negative values whereas its estimation is facilitated by means of either regression, or entropy maximization techniques. To this end, a commonly faced problem is the issue of observations with zero value (*e.g.* Burger et al., 2009). In many cases, gravity models are also stratified by mode, or trip purpose (*e.g.* commuting, leisure, *etc.*), hence resulting in mode- and trip purpose-specific parameter estimates.

The second approach is data-driven where empirical trip data is used. More specifically, the approach aims at fitting a function to the probability distribution function (pdf) of the observed trip lengths, or of the interaction rate, and can be stratified by mode, trip purpose, age, income, type of employment, *etc.* An important aspect is that the function can be fitted both in terms of linear and non-linear approaches. Due to this, more flexible parametric forms than the ones mentioned before can be accommodated. Normally, the data are aggregated in different intervals of equal size and the estimation is facilitated by taking into account one observation per interval (*e.g.* Zhao et al., 2003; Scott and Horner, 2008; Sarlas et al., 2015). The first bin should by definition be the one with the highest frequency value in order to ensure a monotonically decreasing function. In most of the cases though, trip length frequency distributions are found to be increasing up to a certain point and then decreasing. Therefore, the first interval has to be specified in a way that ensures that it has the maximum frequency, determining hence the intervals' size and subsequently the number of fitting points. In addition, a normalization of the interval frequencies is commonly applied, normally by the maximum frequency in order to obtain values ranging from zero to one. As a result, the estimated parameters are sensitive to the intervals' definition and the imposed normalization.

The last approach requires the definition of a cut-off value, or isochrones, after which it is assumed to see no possible interaction. Moreover, within that threshold all opportunities are weighed the same. However, as mentioned in Geurs and van Wee (2004), the arbitrary selection of isochrones along with the lack of differentiation between opportunities make such measures simplistic and problematic. Various attempts have been made to resolve those issues either by assigning weights within the threshold based on a decay function (*e.g.* Black and Conroy, 1977), or by assuming that people are indifferent to distances up to a certain point and after that the interaction intensity is determined based on a distance decay function, resulting in a modified Gaussian function (*e.g.* Vale and Pereira, 2017).

2.2.5 *Interaction intensity function parameter(s) and spatial structure*

The conceptual aspects of the distance decay parameter have attracted a lot of attention in the spatial interaction literature, initiated due to the variance of the reported parameters and the occasional existence of positive values (*e.g.* Griffith and Jones, 1980; Fotheringham, 1981; Sheppard, 1984; Ewing, 1986). In particular, the distance decay parameter should capture the effect that distance has on interaction, *ceteris paribus*. Driven by this, a purely behavioral relationship between these two has been posited for a long time (Roy and Thill, 2003).

However, the existence of a relationship between non-generic but location-specific parameters and spatial structure questions and contradicts fundamentally such statements (*e.g.* Fotheringham, 1981). Consequently and as Fotheringham (1981) notes: "*...it has been proposed that distance-decay parameter estimates are a function of spatial structure as well as a function of interaction behavior, although such theories have aroused much controversy*". This relationship was the subject of an intense debate within the relevant scientific community, an overview of which is given by Sheppard (1984). In the same paper, Sheppard discusses the potential sources of the underlying conceptual bias, attributed to mis-specification reasons. In addition, Fotheringham (1981) also discusses in detail the various, by that time, existing theories relating spatial structure and distance decay parameters.

In this regard, Curry et al. (1975) proposed capturing the spatial structure effect through a spatial autocorrelation function. Subsequently, in following up studies the issue of spatial autocorrelation was further investigated (*e.g.* LeSage and Pace, 2008; Schatzmann et al., 2019-01). Fotheringham and Webber (1980) suggested the use of a simultaneous equations model to properly account for the effect of spatial structure. Later on, Fotheringham (1983) proposed a new set of spatial interaction models that incorporate spatial structure effects into their formulation, in addition to origin-specific distance parameters (see section 2.2.3). Essentially, in this way both ends of the trip are taken into account.

Interestingly, an alternative approach, namely the expansion method, is due to Casetti (1972). More specifically, he suggested the spatial contextualization of the parameters to account for the underlying spatial structure. In essence, his approach involves the estimation of the parameter of interest as a function of spatial structure variables. As outlined by Roy and Thill (2003) though, this practical way has received very limited attention with the exception of few studies on spatial interaction modelling (*e.g.* Zhang and Kristensen, 1995; Roy, 2004).

In a study by Halás et al. (2014), the variance of empirical distance decay functions for the daily travel-to-work trips to various regional centers of Czechia was investigated. Moreover, the authors attempt the construction of a universal distance decay function which is in turn estimated as a function of distance, population and employment of the associated regional centers. However, the proposed universal function accounts only partially for the spatial structure since it focuses only on the destination end of the trips.

2.3 DISAGGREGATE ACCESSIBILITIES

A major limitation of the aggregate accessibilities is that they ascribe the same levels of accessibility on all individuals in each zone (Pirie, 1979). As a result, they lack the ability to capture personal differences on accessibility (Kwan, 1998). However and as pointed out in Weber (2003), households' and individuals' characteristics are found to be more important than locations' characteristics. Driven by such insights, a strand of literature has emerged aiming at developing and evaluating disaggregate accessibilities.

At the outset, the transition from a zonal to a point level of analysis was exploited to investigate spatial variations of the accessibility levels (*e.g.* Hanson and Schwab, 1987). Essentially, that transition was motivated by the need for analyzing urban settings. In Handy (1993), the distinction between regional and local accessibility measures is discussed.

A substantial number of papers have been based on the concept of the space-time prism as put forward by Hägerstrand (1970), resulting in various space-time accessibility implementations (Miller, 1999). More specifically, the central idea of those formulations is that individual-specific interaction intensity functions can be specified by taking into account various aspects such as individuals' spatio-temporal constraints and activity sequencing. For instance, in Kwan (1998) different individual accessibilities were developed by restricting and enumerating the feasible opportunity set based on individuals' space-time prisms. In a wider sense, such measures can be perceived as being theoretically aligned with the concept of cumulative opportunities accessibilities, with the major difference that they have individual-specific cut-off values in place.

The space-time prism approach was extended later on to account for non-uniform interaction intensity values, hence being more aligned with the gravity-based accessibility formulation. For instance, in Horner and Downs (2014) the space-time prism approach was refined by employing a network-based probability density function *"which indicates the relative likelihood that an object was at a particular location on the network"* as the interaction intensity function. The idea of employing a density function for determining the interaction intensity values was also exploited by Li et al. (2011). In that study, dynamic location-based accessibility measures were proposed that account for the temporal variations of speed, and thus of network connectivity and conditions. In another study of similar scope by Wang et al. (2018), a location-based accessibility formulation based on the concept of space-time utility perspective was introduced. Nonetheless, a major limitation of such space-time measures is associated with their data demanding nature. As Pirie (1979) states "*as they depend so heavily on large amounts of information about completed activities and trips they are probably best applied retrospectively.*"

Cascetta et al. (2016) introduced a new class of accessibility formulations. More specifically, the proposed formulation includes a behavioural interaction intensity function that takes into account various individual spatio-temporal constraints and perception of alternatives. A different dimension of distance decay functions was studied by Martínez and Viegas (2013). More specifically, the authors propose the construction of aggregate distance decay functions on the basis of survey-based individual psychological perceptions of distance.

Last, Páez et al. (2010) utilized the expansion method of Casetti (1972) in order to construct individual accessibilities of a cumulative form. More specifically, individual-specific cut-off values are predicted on the basis of a trip length regression model that accounts for the impact of both sociodemographic and neighborhood attributes on the trip length. A thorough discussion of individual accessibilities can be found in Kwan and Weber (2003); Neutens et al. (2007).

2.4 LIMITATIONS

Based on the previous overview of the accessibility literature, three major limitations can be identified. The first limitation concerns the aggregate gravity-based accessibility measures and in particular the calibration of their interaction intensity parameter. On one hand, this can entail the estimation of a gravity model which is conditional though on the availability of a (full) O-D trip matrix. However, such matrices are rarely directly available and they need to be generated through various modelling techniques thereon. Consequently, potential errors propagate to the subsequent models in a systematic manner, thus introducing bias in the parameter estimates. In addition, in many cases (*e.g.* developing countries) the lack of census and count data is restrictive. On the other hand and for the case that requires fitting a function to an empirical pdf, the data availability becomes less of an issue since a (representative) sample can suffice for that reason (*e.g.* household surveys). Nevertheless, the estimation process of the respective parameter(s) is not robust both against the intervals definition and the commonly applied normalizations.

The second limitation is associated with the interdependence between the interaction function and the spatial structure. In spite of constituting a well acknowledged issue in the literature of spatial interaction models, to date aggregate accessibility measures have not incorporated this dimension into their formulations.

Similarly but for the case of disaggregate accessibility indicators, individual characteristics, such as age and income, along with location-specific characteristics are undoubtedly important determinants of travel behavior. Presumingly, this can be viewed as a main source of variance on the individual accessibility levels. The literature on disaggregate accessibilities has evolved along the lines of specifying individual-specific interaction intensity functions in order to evaluate such variations. However, a major limitation lies in the fact that this specification is conditional on the provision of individual-specific constraints (*e.g.* travel time budget, spatio-temporal constraints, *etc.*), failing to explicitly account for the various determinants of spatial interaction.

Driven by the above, the concept of the interaction intensity function is revisited in the next chapter. More specifically, an alternative way of specification is introduced, capable of addressing the aforementioned limitations. In addition, the theoretical implications of this revision are discussed while a case study is designed to exemplify its application.

3

REVISITING SPATIAL INTERACTION

3.1 ALTERNATIVE SPECIFICATION

As highlighted in the previous chapter, the advances in the field of accessibility measures are mainly driven by those in spatial interaction modelling. Contradicting that trend, an alternative way of defining a spatial interaction approach can originate from the accessibility concept itself. To this end and by reiterating accessibility's definition, accessibility quantifies the potential to reach different opportunities. Therefore, spatial interaction can be seen as a two-step process. In the first step, the reachable opportunities from any location are taken into consideration, while in the second step the specific interaction probabilities for any given pair of locations are determined.

A key aspect of gravity-based accessibility measures is that they measure the potential for interaction between origins and destinations. In this regard, the interaction intensity function component can be viewed as a function of the number of people that travel up to a certain point in space. For instance, knowing that 80% of the trips have a length equal, or higher than a specific value, it becomes meaningful to assign a weight of 0.80 on the opportunities at that point since they constitute accessible opportunities for at least 80% of the trip makers. Nevertheless, the fact that they continue further away from that point can be attributed to spatial matching issues.

In particular, if we denote trip length as L, then the probability that an individual makes a trip with a length equal, or smaller, than some value l is $F(l) = Pr[L \leq l]$, which corresponds to the cumulative distribution function (cdf) of trip length. Since our focus lies on the rate of trip makers that continue after a certain value l, we are therefore interested in the survival function of trip length, which is defined as $S(l) = 1 - F(l)$. Consequently, the function $f(d_{ij})$ of the general gravity-based accessibility formulation (formula (2.1)) can now be replaced by $S(l_{ij})$. In a similar manner, the gravity model (formula (2.2)) can also be adjusted accordingly, resulting in a spatial interaction framework that relies on survival analysis and accessibility concepts. More specifically, the production-constrained case then can be formulated in the following way:

$$T_{ij} = A_i Pop_i Empl_j S(l_{ij}) \tag{3.1}$$

with

$$A_i = \frac{1}{\sum_j Empl_j S(l_{ij})} \tag{3.2}$$

while the attraction-constrained case is:

$$T_{ij} = B_j Empl_j Pop_i S(l_{ij}) \tag{3.3}$$

with

$$B_j = \frac{1}{\sum_i Pop_i S(l_{ij})} \tag{3.4}$$

The choice of defining the interaction intensity function as a survival function is appealing for a number or reasons. First, it is theoretically aligned with the perception of space and how people evaluate opportunities. Second, the function is by definition bounded between zero and one. Third, upon the availability of empirical trip data, the function can be obtained in terms of either a data-driven, or a model-based way.

In the case of the former, fitting a parametric function to the empirical survival one has the apparent advantage of yielding parameter estimates that are overall robust to the interval definition, while no normalization is required. Furthermore, a model-based way can be employed as well where the various determinants of trip length, such as spatial structure, location and individual characteristics, can be taken into account (*i.e.* $S(l_{ij}|x_k)$). To this end, survival analysis models can be utilized for that purpose. In particular, this choice facilitates a spatial and individual contextualization to take place, along the lines of the expansion method as put forward by Casetti (1972). To the best of author's knowledge, no prior studies have tackled this problem for the case of spatial interaction models for transportation. The same holds true for the case of the gravity-based accessibility measures as well. A previous study by Páez et al. (2010) has addressed methodologically the issue but only for the cumulative opportunities case.

As for the data requirements, the proposed framework requires only a (representative) sample of trips, which is normally available through travel surveys, unlike the gravity model that requires an O-D trip matrix. In summary, the proposed specification of the interaction intensity function can overcome the identified limitations of the accessibility measures, as discussed in section 2.4. It should be noted that the conceptual framework described above was first introduced by Sarlas and Axhausen (2019).

An important aspect of implementing accessibility measures with the proposed interaction intensity functions in place concerns the nature of the outcome measures. In this regard, the distinction between normative and positive implementations becomes of relevance to clarify the differences that arise due to the specification. It is worthwhile to put into perspective once again the definition of this distinction: *"normative accessibility measures are defined in terms of how far people ought to travel or how far it is reasonable for people to travel whereas positive accessibility measures are defined in terms of how far people actually travel"* (Páez et al., 2012). In a wider sense, this definition draws on the fundamental causes of the distance decay parameter debate, as briefly discussed in a previous section (section 2.2.5), and translates them into the accessibility literature.

In particular, a generic interaction intensity function, irrespective of the way its parameter is obtained, quantifies the effect that distance exerts on interaction on average. However, the actual effect normally varies in space, as demonstrated in Fotheringham (1981). Therefore, accessibility implementations with a generic specification in place can be perceived as normative accessibility measures, under the assumption that average interaction aligns with the expectation of what is considered to be reasonable in terms of distance. On the contrary, a spatially varying function captures more accurately how far people actually travel, hence leading to the implementation of positive accessibility measures.

Based on the above, it can be concluded that implementations with the data-driven specification in place result in normative accessibility measures, while the model-based in positive ones. Furthermore, the combination of both measures in a single measure can be of merit as well. For instance, forming a ratio allows the construction of relative accessibility measures, such as the ones presented in Páez et al. (2012). Such measures can supplement the traditional accessibility analysis and also allow policy relevant conclusions to be drawn, with respect to spatial and social aspects of accessibility.

3.2 SURVIVAL ANALYSIS MODELS

When it comes to the problem of duration data modelling, the use of traditional linear regression models can be problematic. In particular, this can be attributed to the underlying distribution of the dependent variable which can result in non-normally distributed error terms. Consequently, this violation of the normality assumption can give rise to different shortcomings, such as inaccurate statistical tests. In this respect, survival analysis models can be exploited for modelling travel distances (*e.g.* Anastasopoulos et al., 2012), being more flexible in the sense that they can accommodate a larger number of distributions. Survival analysis models have been applied to a plethora of problems in different domains, varying from medicine (*e.g.* Hosmer et al., 2008), to pedestrian aided wayfinding (*e.g.* Giannopoulos et al., 2017). An overview of transport related applications can be found in Bhat and Pinjari (2007).

Survival analysis models aim to model the hazard function of a process, with *"the term hazard being used to describe the risk of "failure" in an interval after time l, conditional on the subject having survived to time l"* (Hosmer et al., 2008). The mathematical definition of hazard function is $h(l) = f(l)/S(l)$. The hazard function can take both parametric and non-parametric forms, however the choice of a parametric function has the advantage that it can fully describe the basic underlying function of survival time, and thus of the error term. Furthermore, it quantifies the effect that different variables exert on it, allowing predictions and forecasts to be made. An analytical presentation of the various survival analysis models is given by Hosmer et al. (2008).

The class of fully parametric accelerated failure time (AFT) models are of interest for the particular modelling task at hand. More specifically, the underlying assumption of those models is that the variables have an acceleration effect on the survival function. Subsequently, the formulation of the survival function depends on the assumptions about the duration distribution. The most common distributions are the exponential, the Weibull, and the log-logistic ones. For instance, the model formulation for the latter case is:

$$ln(l) = \beta_0 + \beta_k x_k + \sigma\epsilon \tag{3.5}$$

where the error term ϵ follows the standard logistic distribution, σ is the scale parameter, x_k are the independent variables, and β are the estimated parameters. In that case, the survival function is given according to the formula:

$$S(l, x_k, \beta, \sigma) = [1 + e^{z}]^{-1} \tag{3.6}$$

with z being the standardized log-time outcome defined as:

$$z = \frac{y - \beta_0 - \beta_k x_k}{\sigma} \tag{3.7}$$

with $y = ln(l)$. If we replace z in formula (3.6) with its equivalent from formula (3.7), then we get:

$$\begin{aligned} S(l, x_k, \beta, \sigma) = \quad & \left[1 + e^{\frac{y - \beta_0 - \beta_k x_k}{\sigma}}\right]^{-1} = \left[1 + e^{y/\sigma}\, e^{-\beta_0/\sigma}\, e^{-\beta_k x_k/\sigma}\right]^{-1} = \\ & \left[1 + l^{\rho}\, e^{-\beta_0 \rho}\, e^{-\beta_k x_k \rho}\right]^{-1} \end{aligned} \tag{3.8}$$

with $\rho = 1/\sigma$ and $e^{y} = l$. As it can be seen, the included variables have an accelerating effect on the survival function. Furthermore, the equation for the median survival time is:

$$l_{50}(x_k) = exp(\beta_0 + \beta_k x_k) \tag{3.9}$$

3.3 CASE STUDY

In the following sections, the application of the proposed spatial interaction specification is exemplified with a case study. More specifically, by utilizing commuting data, three variants of mode-specific interaction intensity functions are constructed. The first variant employs a data-driven specification way while the other two a model-based one, differentiated by the type of variables that they take into consideration.

More specifically, the second variant employs only variables associated with locations' characteristics while the third one extents the previous one by accounting also for individuals' differing characteristics. In this way, the construction of both aggregate and disaggregate forms of interaction functions is enabled.

3.3.1 *Commuting data*

The data for this study come from a detailed household travel survey (micro-census) for the year 2010. The survey is carried out every five years by the Federal Office for Spatial Development, in cooperation with the Swiss Federal Statistical Office. It involved a number of approximately 60'000 households and 63'000 individuals were asked to report the trips they made on a pre-assigned date. For each household and individual, demographic data were collected. Given the focus of the study on commuting, only the individuals with a reported trip to their workplace are included in our sample. A further reduction of the sample size is due in order to exclude observations with ambiguous mode choices. In addition, the focus is centered only on those having a working place in a different zone/municipality than their residence.

In addition, utilizing the reported travel times and trip lengths is considered problematic because in many cases individuals make multi-chain trips on their way to/from the workplace. Therefore, identification of the actual duration of the trip to work is not possible. As a remedy, the travel time matrices of the calibrated nationwide four-step model[1] are utilized. The advantage of that choice is that generalized cost measures can be retrieved. Especially, for the case of public transport, the generalized cost can be perceived as being more reflective of the actual cost since it incorporates multiple cost components (*e.g.* in-vehicle time, waiting time, out-of-vehicle time, *etc.*).

In brief, the final sample size consists of 9'509 individuals, of which 70% commute by car with an average generalized cost of 12.5 minutes, while the remaining 30% commute by public transport with an average generalized cost of 64 minutes. The huge difference can be attributed to two main reasons. First, the generalized cost function of public transport takes into account more travel components, and thus resulting in overall higher generalized costs. Second, this finding can be explained partially by mode choice considerations. More specifically, people prefer public transport for commuting when the distances are relatively long due to the comfort factor. A note should be made here that in Switzerland the public transport system is very extensive, well-maintained, and efficient, viewed by many people as a desirable mode for commuting. The work and residence locations are spatially matched to the zonal system of the aforementioned transport model, consisting of almost 3'000 zones.

1 ARE; National Transport Model, 2010

3.4 DATA-DRIVEN INTERACTION FUNCTIONS

The first part of the analysis deals with the issue of fitting a generic interaction intensity function to the observed generalized cost frequencies. Given the large identified differences, the sample is stratified by mode in order to obtain mode-specific functions. Subsequently, the survival rates of the observed trip costs are utilized in order to fit the relevant interaction intensity functions in a data-driven way. More specifically, a generic function of the negative exponential family with two parameters is fitted to the actual values. The form of the function is given in formula (3.10), whereas similar functions have also been fitted in the recent literature (*e.g.* Halás et al., 2014).

$$f(gc) = e^{\alpha_1 gc^{\alpha_2}} \tag{3.10}$$

The nonlinear least-squares estimates of the parameters are calculated with the Gauss-Newton algorithm. For the car commuters case, the estimated parameters are $\alpha_1 = -0.031$ and $\alpha_2 = 1.362$, while for the public transport ones are $\alpha_1 = -0.000025$ and $\alpha_2 = 2.507$, respectively. The fitted functions are presented in figure 3.1, where as it can be observed the car intensity function is steeper while the public transport one naturally extends to substantially larger generalized cost values.

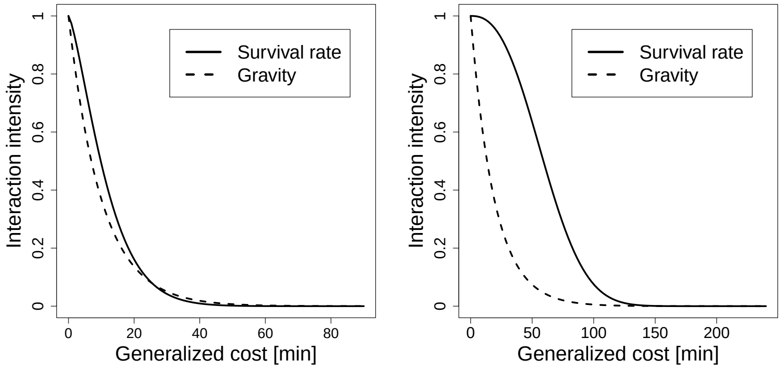
Furthermore, the interaction functions, as implied by the estimates of the corresponding gravity models, are also plotted for comparison purposes. The estimation of the gravity models takes place by utilizing the daily O-D matrices of a nationwide four-step model[2] (observations with zero value are excluded), while the parameter estimates are obtained by means of an ordinary least squares (OLS) estimator, having a logarithmic dependent variable in place. The estimation results are presented in table 3.1. It is worthwhile highlighting that the gravity functions are found to be steeper than their survival counterparts, especially for the public transport case.

3.4.1 *Accessibility measures*

The next step deals with the calculation of different accessibility measures. More specifically, two types are calculated, one in accordance with formula (2.1) measuring the absolute number of accessible potential opportunities, and one with formula (2.21) involving a normalization to account

2 ARE; National Transport Model, 2010

FIGURE 3.1: Interaction intensity functions



(a) Car (b) Public transport

TABLE 3.1: Gravity model estimates by mode

	Dependent variable: *Log(Trips)*	
Regressor	Car	PuT
Constant	-5.5879^{***}	-11.0667^{***}
	(0.0086)	(0.0106)
Generalized cost [min.]	-0.0997^{***}	-0.0520^{***}
	(0.0001)	(0.0001)
Log(Population)	0.5862^{***}	0.8009^{***}
	(0.0010)	(0.0011)
Log(Employment)	0.4750^{***}	0.6291^{***}
	(0.0008)	(0.0009)
Observations	1,609,126	1,167,410
Adjusted R^2	0.554	0.496
df	1,609,122	1,167,406
( ) Std. Errors, + p<0.1; * p<0.05; ** p<0.01; *** p<0.001		

for the competition at the destination level. It should be noted that the self-potential aspect is not taken into account in the accessibility calculations since the case study's focus lies on commuting trips outside the zone of residence.

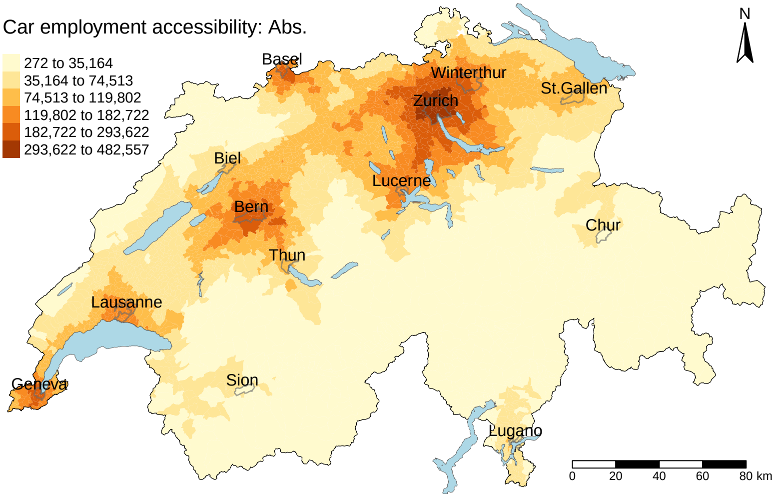
In total, eight gravity-based accessibility measures are calculated; four per mode, two per accessibility formulation, and two per opportunity kind (employment positions and population). Admittedly, accessibility measures with respect to both kinds of opportunities constitute important determinants of the regional form and function, in consequence this is also hypothesized to be the case for the commuting distance as well. The calculated accessibility measures are plotted spatially in the maps that follow (figures 3.2 and 3.3). Only the indicators corresponding to the employment opportunities are presented while similar patterns are identified for the population ones as well.

Interestingly, the patterns seem to be different when comparing visually each accessibility against its normalized version. In the case of the traditional formulation (denoted as absolute), the main cities of Switzerland along with their neighboring zones stand out as the ones having high accessibility values. A finding which is to a large extent anticipated given the high concentration of employment in those areas.

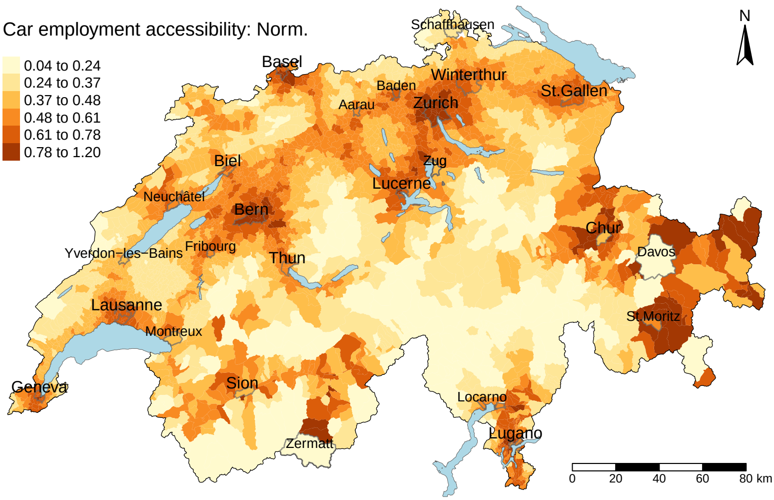
On the contrary and for the competition accessibility measures, substantially different patterns emerge. Apart from the main cities, also other locations surface as ones with high accessibility. More specifically, these locations correspond to smaller cities of high regional importance. Therefore, it can be concluded that by accounting for the spatial competition in the accessibility formulation, the regional role of a zone can be emphasized. It is worth pointing out that in the case of more remote areas (such as Zermatt and Davos), the zones themselves have lower accessibility values than their neighbors. This is due to the nature of the employed accessibility formulations, taking into account only the opportunities lying outside the zones (no self-potential). Furthermore, car accessibility values are found to be more spatially correlated. This can be justified by the spatial continuity of the road network, which is clearly not the case for public transport.

Last, the equivalent accessibility measures based on the gravity interaction functions are calculated as well (denoted as Acc_{G}^{W} and $cAcc_{G}^{W}$). A correlation analysis is conducted to quantify the similarity between them. The corresponding correlation matrices are given in tables 3.2 and 3.3, accordingly. As it can be observed, the survival accessibility measures are almost perfectly correlated with their gravity counterparts.

FIGURE 3.2: Car employment accessibility measures

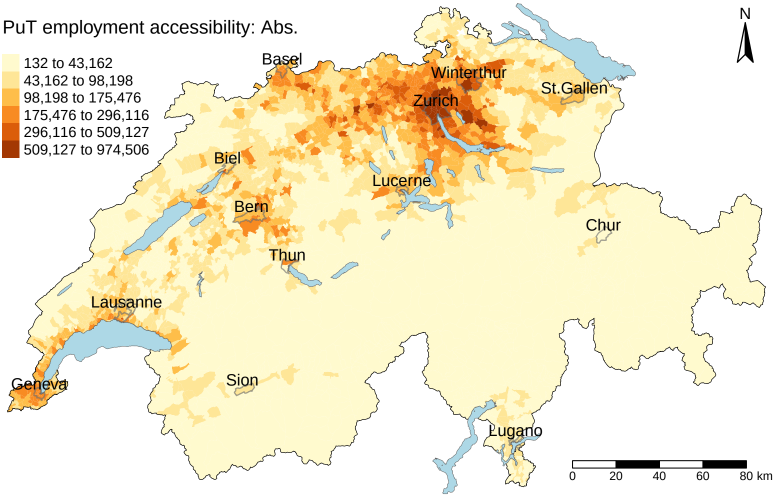


(a) Employment accessibility

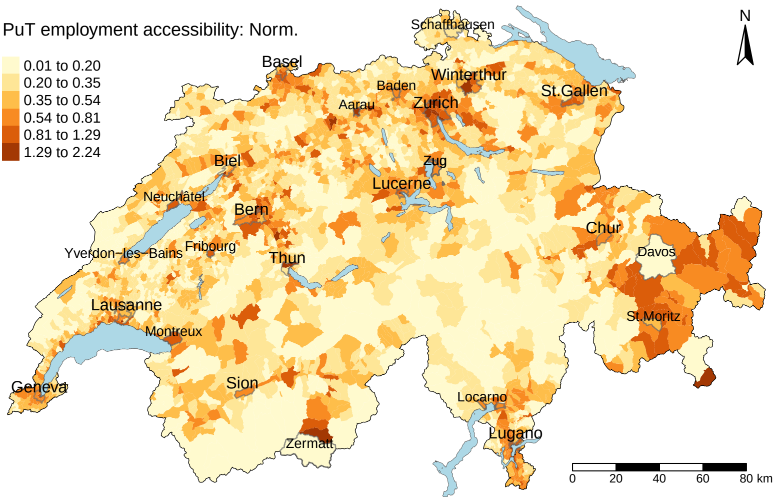


(b) Competition employment accessibility

Figure 3.3: Public transport employment accessibility measures



(a) Employment accessibility



(b) Competition employment accessibility

Table 3.2: Correlation matrix of car accessibility measures

	Acc^{P}	Acc^{E}	$cAcc^{E}$	$cAcc^{P}$	Acc^{E}_{G}	Acc^{P}_{G}	$cAcc^{E}_{G}$	$cAcc^{P}_{G}$
Acc^{P}	1	0.99	0.60	0.42	0.99	0.995	0.55	0.68
Acc^{E}	0.99	1	0.62	0.39	0.99	0.97	0.51	0.70
$cAcc^{E}$	0.60	0.62	1	0.86	0.59	0.57	0.88	0.98
$cAcc^{P}$	0.42	0.39	0.86	1	0.37	0.40	0.97	0.83
Acc^{E}_{G}	0.99	0.99	0.59	0.37	1	0.99	0.50	0.67
Acc^{P}_{G}	0.995	0.97	0.57	0.40	0.99	1	0.52	0.65
$cAcc^{E}_{G}$	0.55	0.51	0.88	0.97	0.50	0.52	1	0.89
$cAcc^{P}_{G}$	0.68	0.70	0.98	0.83	0.67	0.65	0.89	1

Table 3.3: Correlation matrix of public transport accessibility measures

	Acc^{P}	Acc^{E}	$cAcc^{E}$	$cAcc^{P}$	Acc^{E}_{G}	Acc^{P}_{G}	$cAcc^{E}_{G}$	$cAcc^{P}_{G}$
Acc^{P}	1	0.995	0.66	0.53	0.96	0.98	0.50	0.64
Acc^{E}	0.995	1	0.64	0.48	0.96	0.96	0.45	0.60
$cAcc^{E}$	0.66	0.64	1	0.95	0.65	0.68	0.90	0.97
$cAcc^{P}$	0.53	0.48	0.95	1	0.51	0.56	0.96	0.92
Acc^{E}_{G}	0.96	0.96	0.65	0.51	1	0.99	0.50	0.65
Acc^{P}_{G}	0.98	0.96	0.68	0.56	0.99	1	0.54	0.68
$cAcc^{E}_{G}$	0.50	0.45	0.90	0.96	0.50	0.54	1	0.93
$cAcc^{P}_{G}$	0.64	0.60	0.97	0.92	0.65	0.68	0.93	1

3.5 MODEL-BASED INTERACTION FUNCTIONS

The next two variants of spatial interaction functions are constructed in a model-based way by exploiting survival analysis techniques. In essence, trip length constitutes the observed outcome of the interaction between demand for mobility and supply, having received attention in the literature as a standalone topic. For instance, in Morency et al. (2011) a model for mean trip distance was estimated accounting for the impact of different individual socio-economic and demographic attributes, along with the spatial characteristics of the house location. Interestingly, the authors make the claim that traveled distance can be viewed as a proxy of activity spaces, a concept which has also found application in social exclusion research (*e.g.* Schönfelder and Axhausen, 2003). In a study of a similar scope, Mercado and Páez (2009) investigated the determinants of mean travel distance, focusing though on elderly people. Their results identify a negative relationship between age and travel distance, reaffirming the findings of previous studies.

In the same context but focusing on the commuting distance, a number of papers have attempted to quantify the impact of various socio-economic and spatial structure attributes on the commuting patterns (*e.g.* Gordon et al., 1989; Khattak et al., 2000; Shen, 2000; Manaugh et al., 2010; Sandow and Westin, 2010; Maoh and Tang, 2012; Axisa et al., 2012). For instance, Gordon et al. (1989) estimated aggregate mode-specific commuting distance models as a function of spatial structure variables, such as urbanized area, share of total population, job mixture and home ownership. In another study by Shen (2000), the spatial and social dimensions of commuting are descriptively analyzed and discussed.

In addition, a regression model was estimated to identify the main factors that explain variations in commuting times while urban structure was incorporated into the model through the inclusion of an employment accessibility variable at the home end. The importance of accounting for the home accessibility was also demonstrated by Manaugh et al. (2010). The issue of commuting duration and income was discussed in Sandow and Westin (2010), concluding that economic incentives are important determinants of long commuting times.

In a connection with the distance decay parameter debate (see section 2.2.5), Levinson (1998) estimated an OLS model to highlight the importance of accounting for location characteristics at both trip ends. In particular, this was facilitated by accounting for spatial structure effects through the formation of employment and population accessibility indicators at both trip ends. Levinson hypothesized the existence of strong effects of the corresponding spatial structure indicators, the direction of which was theoretically discussed and drawn based on the concept of spatial competition.

Later on, a similar analysis was conducted by Cui et al. (2019), differentiating between low- and high-income individuals, and employing cumulative opportunities measures. The authors identify the presence of stronger accessibility impacts for the former group of individuals. In the same spirit, in Anastasopoulos et al. (2012) a hazard-based model approach was employed to model urban travel times. The authors account for both individual and location characteristics but they include no accessibility measures in their model specification. Last, the same family of models was employed also in Ermagun et al. (2016) to analyze the tolerable walking distance for students while in another study by O'Sullivan and Morrall (1996) an empirical analysis of the walking distances to different kinds of public transport stops was conducted to identify varying stop catchment areas.

Based on the above, it can concluded that commuting distance can be perceived as a function of locations', individual's and household's characteristics. Therefore, a regression modelling approach can be employed to quantify that relationship, enabling us to make statements about the strength and the statistical significance of it, and at the same time to obtain statistically sound predictions.

At first, OLS models are estimated in order to employ them as the benchmark for the analysis. A special focus needs to be given to potential multicollinearity issues due to high correlation between the independent variables. Especially, the implications of multicollinearity if untreated are the difficulty of uncovering the partial effects of each variable on the response, and can inflate the error variance and consequently the standard errors of the estimated parameters (Wooldridge, 2012).

The estimation of the equivalent survival analysis models follows, thus allowing us to construct the model-based instances of the proposed interaction intensity functions. To that end, two variants are estimated, one accounting only for locations' characteristics (denoted as aggregate) and one that extents the previous one by accounting also for individuals' differing characteristics (denoted as disaggregate).

3.5.1 *Descriptive analysis and hypotheses*

As mentioned before, it is presumed that commuting distance can be perceived as a function of locations', individual's and household's characteristics. Driven by this, different variables need to be checked for their ability to constitute statistically significant determinants of commuting distance. A descriptive analysis precedes that step to gain insights about the sample and the variance of the different variables. The summary statistics are given in table 3.4.

TABLE 3.4: Summary statistics of employed variables

Variable	Car commuters		PuT commuters	
	Mean	Std. Dev.	Mean	Std. Dev.
Generalized Cost [min.]	12.50	11.07	63.78	30.26
Female [dummy]	0.40		0.53	
Household size	2.61	1.32	2.79	1.33
HH income: non-reported [dummy]	0.12		0.13	
HH income: <2 Kchf [dummy]	0.01		0.01	
HH income: 2-4 Kchf [dummy]	0.05		0.08	
HH income: 4-6 Kchf [dummy]	0.20		0.18	
HH income: 6-8 Kchf [dummy]	0.19		0.17	
HH income: 8-10 Kchf [dummy]	0.16		0.15	
HH income: 10-12 Kchf [dummy]	0.10		0.13	
HH income: >12 Kchf [dummy]	0.17		0.15	
Vehicles/licenses	0.91	0.35	0.51	0.44
Age: 18-30 [dummy]	0.16		0.27	
Age: 31-40 [dummy]	0.22		0.24	
Age: 41-50 [dummy]	0.30		0.24	
Age: 51-60 [dummy]	0.24		0.21	
Age: >60 [dummy]	0.08		0.04	
PhD degree [dummy]	0.02		0.04	
MSc degree [dummy]	0.08		0.15	
Self-employed [dummy]	0.05		0.03	
Part-time [dummy]	0.23		0.29	
Part-time (>1 jobs)[dummy]	0.01		0.01	
Working (Real estate)[dummy]	0.007		0.005	
Working (IT)[dummy]	0.03		0.04	

Table 3.4: Summary statistics of employed variables (continued)

Variable	Car commuters		PuT commuters	
	Mean	Std. Dev.	Mean	Std. Dev.
Working (hotel, restaurants)[dummy]	0.03		0.04	
Working (construction) [dummy]	0.07		0.02	
Working (health)[dummy]	0.12		0.15	
Pop. density: W [pop/hectare]	12.55	20.69	27.34	12.55
Empl. density: H [empl/hectare]	8.48	19.69	23.27	35.79
Population: home	9362.11	9919.54	16722.54	13671.01
Active pop. share: H	0.62	0.03	0.63	0.04
Empl./active pop.: W	1.12	0.53	1.34	0.51
Empl. pos./active pop.: H	0.79	0.46	1.05	0.53
Employment positions: W/H	15.54	54.46	12.3	50.45
Population: W/H	4.81	10.87	3.64	8.72
3^{rd} sector share: W/H	1.21	0.6	1.19	0.46
Car empl. access. norm.: H	0.60	0.18	0.76	0.21
Car pop. access. norm.: H	1.79	0.35	1.94	0.34
Car empl. access. norm.: W	0.55	0.15	0.64	0.19
Car pop. access. norm.: W	1.69	0.36	1.78	0.34
PuT empl. access. norm.: H	1.44	0.79	1.87	1.01
PuT pop. access. norm.: W	1.91	1.14	2.66	1.46
Log (car empl. access.): W	11.61	0.79	12.11	0.75
Log (car pop. access.): H	11.99	0.66	12.31	0.60
PuT/Car empl. access: H	1.21	0.71	1.42	0.81
Car/PuT pop. access.: W	1.03	0.74	0.85	0.54

Note: W=work location, H=home location

On the individual's characteristics front, the public transport users' sample includes a higher share of female users and also of individuals with lower vehicle access. In addition, the share of public transport users having a higher education is substantially higher than of the car users' group. Notably, a comparison of the employment and population density values between the two groups shows that public transport users are placed at more urbanized locations with higher accessibility levels. The ratio of employment positions between the workplace and the home location in both cases shows that people choose less dense areas for living, having to bear the cost of commuting to the regional centers. For the income, no clear difference between the two groups can be identified.

On the causality front, a negative relationship between age and commuting distance is anticipated on the basis that as people get older they are becoming less willing to travel long distances. As the household size increases, also commuting distance is expected to decrease due to increased responsibilities at home. The dummy variables associated with the different sectors are expected to capture that some types of jobs are more likely to be matched by the local population, given that they d not require highly skilled individuals. Moreover, highly skilled individuals are expected to travel longer distances in order to have a job that matches their skills. This can be captured by the dummy variables of having an MSc, or a PhD degree respectively. The same argument can also be made for high income individuals.

On the location characteristics side, the presence of different effects is expected. Reiterating the hypotheses formulated by Levinson (1998), it is presumed that individuals living in areas with high accessibility, in terms of population, will have to compete with a larger number of individuals for a job, and thus this would lead to increased commuting distances. On the contrary, accessibility to employment positions at the home location is hypothesized that it should have a negative relationship with commuting distance since more opportunities are regionally available. In the case of the work location, employment accessibility should have a positive impact on commuting distances since it captures the influence of competing workers. A different pattern is anticipated for the population accessibility since a negative relationship is hypothesized due to the larger number of competing individuals, and hence of a matching with shorter transportation cost. Following the same logic as above, population density at workplace and employment density at home end are also expected to exert a negative effect on distance. Differences on the signs between the variables of the two modes are not expected to exist as the direction of the causality should, at least in theory, be the same.

3.5.2 *Estimation of model-based interaction functions*

In this subsection the results of the model estimation are presented and discussed. Two categories of models are estimated, an aggregate and a disaggregate one in terms of analysis level. In the beginning, two OLS models per category are estimated in order to serve as benchmarks and also allow for a thorough investigation of potential multicollinearity issues. More specifically, the models employ a log-level functional form where the βs can be interpreted as semi-elasticity values. In the case of a log transformed independent variable although, the β estimate can be interpreted as an elasticity value.

The OLS estimates are presented in tables 3.5 and 3.6, accordingly. It should be noted that both the data preparation and the model estimation was conducted in R (R Core Team, 2018), making use of the 'survival' package (Therneau and Lumley, 2014) for the estimation of the AFT models. Last, the 'tmap' package (Tennekes, 2018) was used for the maps' production throughout the dissertation.

TABLE 3.5: Trip duration OLS estimates: aggregate

Dependent variable: Log(GC)	Car	PuT
Constant	1.21*** (0.23)	4.57*** (0.26)
Pop. density: W	−0.004*** (0.0004)	−0.002*** (0.0002)
Empl. density: H	−0.004*** (0.001)	−0.002*** (0.0003)
Log(population): H	0.08*** (0.01)	
Active pop. share: H	1.21*** (0.33)	0.59* (0.26)
Empl./active pop.: W	0.11*** (0.02)	
Population: W/H	0.01*** (0.002)	
Empl. pos./active pop.: H		0.05** (0.02)
Employment positions: W/H		0.0004+ (0.0002)
3^{rd} sector share: W/H	0.05** (0.02)	0.02 (0.02)
Car empl. access. norm.: H	−2.35*** (0.16)	
Car pop. access. norm.: H	0.39*** (0.06)	
Car empl. access. norm.: W	2.23*** (0.12)	
Car pop. access. norm.: W	−0.73*** (0.06)	
PuT empl. access. norm.: H		−0.31*** (0.04)
PuT pop. access. norm.: W		−0.07*** (0.01)
Log (car empl. access.): W		0.20*** (0.02)
Log (car pop. access.): H		−0.24*** (0.03)
PuT/Car empl. access: H		−0.02 (0.01)
Car/PuT pop. access.: W		0.10*** (0.02)
Observations	6,695	2,814
Adjusted R^2	0.13	0.31
Adjusted R^2	0.13	0.31
df	6683	2801
( ) Heterosc. robust std. errors,	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001	

Table 3.6: Trip duration OLS estimates: disaggregate

Dependent variable: Log(GC)	Car	PuT
Constant	1.31*** (0.23)	4.73*** (0.26)
HH income: non-reported	−0.15*** (0.04)	−0.02 (0.03)
HH income: <2 Kchf	−0.27* (0.14)	−0.07 (0.09)
HH income: 2-4 Kchf	−0.28*** (0.05)	−0.04 (0.03)
HH income: 4-6 Kchf	−0.19*** (0.03)	−0.07** (0.03)
HH income: 6-8 Kchf	−0.18*** (0.03)	−0.05* (0.03)
HH income: 8-10 Kchf	−0.05 (0.03)	−0.02 (0.03)
HH income: 10-12 Kchf	−0.05 (0.04)	0.01 (0.03)
HH income: >12 Kchf	*Ref.*	
Age: 18-30	0.15*** (0.04)	0.08* (0.04)
Age: 31-40	0.14*** (0.04)	0.09* (0.04)
Age: 41-50	0.11** (0.04)	0.05 (0.04)
Age: 51-60	0.02 (0.04)	0.04 (0.04)
Age: >60	*Ref.*	
Female: HH size	−0.03*** (0.01)	−0.02*** (0.005)
Vehicles/licenses	0.13*** (0.03)	−0.03+ (0.02)
Self-employed	−0.26*** (0.05)	−0.13** (0.04)
Part-time	−0.10*** (0.03)	
Part-time (>1 jobs)	−0.21* (0.10)	
Working (real estate)	−0.22* (0.11)	
Working (IT)	0.15* (0.06)	
Working (hotel, restaurants)	−0.16** (0.06)	−0.08* (0.04)
Working (construction)	−0.12** (0.04)	−0.06 (0.05)
Working (health)	0.06* (0.03)	−0.03 (0.02)
MSc degree		0.08*** (0.02)
PhD degree		0.20*** (0.05)
Pop. density: W	−0.004*** (0.0004)	−0.002*** (0.0002)
Empl. density: H	−0.003*** (0.001)	−0.002*** (0.0003)
Log(population): H	0.08*** (0.01)	
Active pop. share: H	1.05** (0.33)	0.50* (0.26)
Empl./active pop.: W	0.10*** (0.02)	
Population: W/H	0.01*** (0.001)	
Empl. pos./active pop.: H		0.04* (0.02)
Employment positions: W/H		0.0004+ (0.0002)

TABLE 3.6: Trip duration OLS estimates: disaggregate (continued)

Dependent variable: Log(GC)	Car	PuT
3^{rd} sector share: W/H	0.05** (0.02)	0.03 (0.02)
Car empl. access. norm.: H	−2.27*** (0.16)	
Car pop. access. norm.: H	0.37*** (0.06)	
Car empl. access. norm.: W	2.13*** (0.12)	
Car pop. access. norm.: W	−0.71*** (0.06)	
PuT empl. access. norm.: H		−0.31*** (0.04)
PuT pop. access. norm.: W		−0.07*** (0.01)
Log (car empl. access.): W		0.19*** (0.02)
Log (car pop. access.): H		−0.24*** (0.03)
PuT/Car empl. access: H		−0.03* (0.01)
Car/PuT pop. access.: W		0.10*** (0.02)
Observations	6,695	2,814
Adjusted R^2	0.16	0.33
df	6662	2782
( ) Heterosc. robust std. errors,	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001	

In summary, both models' estimates confirm the prior hypotheses about the causalities in place. Especially, the estimated parameters of the aggregate accessibility variables are all found to be of high statistical significance, emerging as important determinants of commuting distance. In the car model's case, the inclusion of all accessibility measures in a normalized form yields better results than the inclusion of their absolute versions. In addition, that choice results in no multicollinearity issues, which is not found to be the case for their other forms. It should be noted that multicollinearity is checked by the means of variance inflation factors (VIF), assuming that values higher than 5 are indicative of multicollinearity.

In the public transport case, the simultaneous inclusion of all accessibility variables in either form gives rise to extreme multicollinearity issues, resulting in statistically insignificant estimates. It is worth highlighting the fact that in a study of a similar scope by Levinson (1998), similar issues with respect to the significance of the public transport accessibility estimates were identified but remained untreated. As a remedy to that, a set of proxy variables is employed instead. More specifically, two of them are replaced with the log of their corresponding car accessibility measures along with two "balancing" ratios of mode accessibilities. Due to this replacement, it becomes possible to take into account all accessibility indicators at the same time while obtaining statistically significant parameter estimates that are in good agreement with the prior hypotheses.

Naturally, a comparison between the aggregate and disaggregate results shows that small differences exist between the corresponding parameters estimates. The high adjusted R^2 values of the aggregate models though reveal that the most important determinants of commuting distance are associated with locations' characteristics. Nevertheless, the aggregate models suffer from omitted variables bias. However and given the extremely low correlation among the omitted variables and the location-specific ones, this omission has negligible impacts on the estimation process. The results validate this to a great extent since only minor differences between the parameter estimates can be observed.

Following the estimation of the OLS models, the next step is the estimation of the survival analysis models with a maximum likelihood estimator. Different formulations of survival models are tested before concluding on the choice of a log-logistic AFT model based on an information criterion (AIC) goodness of fit measure (Akaike, 1974). Moreover, AFT models accounting for the heterogeneity of individuals (*e.g.* Weibull with Gamma heterogeneity as presented in Anastasopoulos et al. (2012)) could have been employed instead. However, the application of such models for prediction purposes is less straightforward. To mitigate, at least partially, the impact of heterogeneity, we are making use of a robust "sandwich" estimator.

Finally, it should be noted that we expect small differences in the estimates, in comparison to the corresponding OLS models, since the chosen AFT models formulation is actually also a log-level model (and thus its parameters can be interpreted in a similar manner as before). Nevertheless, the AFT models can be considered to be statistically more sound since they account properly for the error term distribution. In particular, the OLS error terms are tested on their compliance with the normality assumptions,in terms of kurtosis and skewness, which is not found to be the case (Pena and Slate, 2006). The estimated parameters are presented in tables 3.7 and 3.8, for the aggregate and disaggregate cases, accordingly.

TABLE 3.7: Trip duration AFT log-logistic estimates: aggregate

Dependent variable: Log(GC)	Car	PuT
Constant	1.09*** (0.22)	4.77*** (0.22)
Pop. density: W	−0.004*** (0.0004)	−0.002*** (0.0002)
Empl. density: H	−0.003*** (0.001)	−0.002*** (0.0003)
Log(population): H	0.08*** (0.01)	
Active pop. share: H	1.41*** (0.32)	0.64** (0.23)
Empl./active pop.: W	0.13*** (0.02)	
Population: W/H	0.01*** (0.001)	
Empl. pos./active pop.: H		0.05** (0.02)
Employment positions: W/H		0.0003* (0.0002)
3^{rd} sector share: W/H	0.05* (0.02)	0.03 (0.02)
Car empl. access. norm.: H	−2.87*** (0.14)	
Car pop. access. norm.: H	0.57*** (0.06)	
Car empl. access. norm.: W	2.58*** (0.11)	
Car pop. access. norm.: W	−0.84*** (0.05)	
PuT empl. access. norm.: H		−0.34*** (0.03)
PuT pop. access. norm.: W		−0.08*** (0.01)
Log (car empl. access.): W		0.24*** (0.02)
Log (car pop. access.): H		−0.30*** (0.02)
PuT/Car empl. access: H		−0.02+ (0.01)
Car/PuT pop. access.: W		0.10*** (0.02)
Scale σ	0.435	0.207
Observations	6,695	2,814
Nagelkerke R^2	0.15	0.36
df	6682	2800
( ) Heterosc. robust std. errors,	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001	

Table 3.8: Trip duration AFT log-logistic estimates: disaggregate

Dependent variable: Log(GC)	Car	PuT
Constant	1.21*** (0.22)	4.91*** (0.23)
HH income: non-reported	−0.14*** (0.04)	−0.03 (0.03)
HH income: <2 Kchf	−0.23+ (0.12)	−0.08 (0.07)
HH income: 2-4 Kchf	−0.27*** (0.05)	−0.05+ (0.03)
HH income: 4-6 Kchf	−0.18*** (0.03)	−0.08** (0.03)
HH income: 6-8 Kchf	−0.17*** (0.03)	−0.05* (0.02)
HH income: 8-10 Kchf	−0.04 (0.03)	−0.02 (0.03)
HH income: 10-12 Kchf	−0.04 (0.04)	0.01 (0.03)
HH income: >12 Kchf	*Ref.*	
Age: 18-30	0.15*** (0.04)	0.07+ (0.04)
Age: 31-40	0.15*** (0.04)	0.06+ (0.04)
Age: 41-50	0.10** (0.04)	0.04 (0.04)
Age: 51-60	0.02 (0.04)	0.02 (0.04)
Age: >60	*Ref.*	
Female: HH size	−0.02** (0.01)	−0.02*** (0.01)
Vehicles/licenses	0.12*** (0.03)	−0.03+ (0.02)
Self-employed	−0.26*** (0.04)	−0.11* (0.04)
Part-time	−0.10*** (0.03)	
Part-time (>1 jobs)	−0.19* (0.09)	
Working (real estate)	−0.22* (0.11)	
Working (IT)	0.15* (0.06)	
Working (hotel, restaurants)	−0.14* (0.05)	−0.08* (0.04)
Working (construction)	−0.10** (0.04)	−0.03 (0.05)
Working (health)	0.06* (0.03)	−0.02 (0.02)
MSc degree		0.06** (0.02)
PhD degree		0.17*** (0.04)
Pop. density: W	−0.004*** (0.0004)	−0.002*** (0.0002)
Empl. density: H	−0.003*** (0.001)	−0.002*** (0.0003)
Log(population): H	0.08*** (0.01)	
Active pop. share: H	1.22*** (0.31)	0.59* (0.23)
Empl./active pop.: W	0.12*** (0.02)	
Population: W/H	0.01*** (0.001)	
Empl. pos./active pop.: H		0.04* (0.02)
Employment positions: W/H		0.0003* (0.0001)

Table 3.8: Trip duration AFT log-logistic estimates: disaggregate (continued)

Dependent variable: Log(GC)	Car	PuT
3^{rd} sector share: W/H	0.05^{**} (0.02)	0.03^{+} (0.02)
Car empl. access. norm.: H	-2.74^{***} (0.13)	
Car pop. access. norm.: H	0.53^{***} (0.05)	
Car empl. access. norm.: W	2.45^{***} (0.11)	
Car pop. access. norm.: W	-0.80^{***} (0.05)	
PuT empl. access. norm.: H		-0.34^{***} (0.03)
PuT pop. access. norm.: W		-0.08^{***} (0.01)
Log (car empl. access.): W		0.23^{***} (0.02)
Log (car pop. access.): H		-0.29^{***} (0.02)
PuT/Car empl. access: H		-0.03^{*} (0.01)
Car/PuT pop. access.: W		0.10^{***} (0.02)
Scale σ	0.427	0.203
Observations	6,695	2,814
Nagelkerke R^2	0.18	0.38
df	6661	2781
( ) Heterosc. robust std. errors,	+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001	

Overall, small but not negligible differences can be observed between the two types of models. More particularly, the difference on the individual-specific attributes parameter estimates is generally small. However, it is interesting that in the case of the location variables, the OLS parameters are found to be on average 10% lower in absolute values and in both cases, than their AFT model counterparts. Therefore, it appears that the OLS models underestimate the impact of the location variables.

3.6 CONCLUSIONS

In this chapter, the concept of spatial interaction was revisited and formulated as a survival analysis one. The proposed methodology has the advantage of being able to address both aggregate and disaggregate cases, bridging to some extent the methodological gap between those two. Moreover, it allows overcoming certain limitations associated with other estimation approaches.

In the case study, the different model estimates reaffirmed the central role of accessibilities on the determination of commuting times, a finding which was facilitated mainly due to the proper treatment of collinearity issues. Moreover, the importance of employing accessibility indicators that account for competition was demonstrated. It total, three variants of interaction intensity functions were estimated; one generic, one location-specific, and one location- and individual-specific. In the following chapter, the aforementioned variants are evaluated within three instances; in terms of spatial interaction intensity values, rates, and resulted accessibility measures.

4

EVALUATION OF ALTERNATIVE SPECIFICATION

4.1 INTRODUCTION

In this chapter, the evaluation of the previously described spatial interaction specification way takes place. In this respect, the three estimated variants of interaction intensity functions are utilized for modifying the traditional gravity model along with the corresponding accessibility measures. Depending on the perspective, this modification is evaluated in two instances. On one hand, the output of the gravity model corresponds to predictions of spatial interaction values (*i.e.* flows). On the other hand, the interaction function defines essentially the interaction space in terms of rates and spatial extent, which can then be utilized for the construction of gravity-based accessibility measures.

More specifically, the aggregate variants are evaluated with respect to their ability to explain daily commuting flows by mode. To this end, a comparative analysis, including the results of two prevailing spatial interaction models, is conducted. In addition, a visual comparison of the predicted interaction spaces of all variants is conducted to exhibit the capabilities of each specification.

On the accessibility front, all three variants are employed for the implementation of aggregate and disaggregate gravity-based measures. Given their positive and normative nature, the construction of relative indicators is needed in order to investigate different aspects of accessibility. Last, the aforementioned results are accompanied by a thorough discussion in order to draw conclusions on the value of employing the proposed spatial interaction framework.

4.2 AGGREGATE SPATIAL INTERACTION CASE

4.2.1 *Predicted interaction values*

At the outset, the first two estimated aggregate variants of interaction functions, namely the generic and the location-specific model-based ones, are employed for obtaining spatial interaction values based on formulas (3.1) and (3.3). In addition, the resulted values are multiplied by a corresponding correction factor each. In the production-constrained case, the correction factor is defined as the share of the economically active population over the total population at the national level, while in the other case the factor quantifies the employment positions per economically active resident nationally. As active population is defined the number of individuals between the age of 18 to 65 years old. In total, we obtain six prediction matrices of values $\hat{T}_{ij}$ for each variant; three per mode, two per gravity formulation (production-constrained and attraction-constrained), and two formulated as the average of the corresponding two gravity outputs.

Moreover, the results are compared against the ones coming from a gravity and a radiation model. In the case of the former, the results of both their unconstrained and the doubly-constrained forms are reported. In the case of the latter, the reported results are calculated based on the formulation (4.1), as presented in (Simini et al., 2012) with the number of commuters per location being replaced by the population per location, multiplied by the active population correction factor (denoted as Cor_{act}).

$$\hat{T}_{ij} = Cor_{\text{act}}\, Pop_i \frac{Pop_i Pop_j}{(Pop_i + S_{ij})(Pop_i + Pop_j + S_{ij})} \tag{4.1}$$

with S_{ij} being the total population within a radius d_{ij} from location i, excluding the population at locations i and j. As d_{ij}, the previously employed generalized cost values are used.

Furthermore, the daily O-D matrices per mode of a nationwide four-step model[1] are utilized as ground truth data to draw conclusions with respect to the predictive accuracy of each approach. However, it should be noted that these matrices are estimates for daily traffic demand, estimated by means of gravity models and subsequently calibrated against count data. Nevertheless, details with respect to the exact estimation approach are not available to the author. As a result, it is anticipated that the gravity model results (as shown in 3.1) will outperform all other approaches since the ground truth data are the output of gravity models to a great extent. Last, the aforementioned O-D matrices are symmetric, with dimensions $n = 2944$. As a result and in order to make quantitatively comparable the various outputs, the prediction matrices are turned into symmetric ones by applying the following formula:

1 ARE; National Transport Model (2010): A 4-step model, implemented in VISUM

$$\hat{T}_{ij} = \hat{T}_{ji} = \frac{\hat{T}_{ij} + \hat{T}_{ji}}{2} \tag{4.2}$$

The evaluation of the predictive accuracy is conducted with three metrics, namely the root mean square error (RMSE), the common part of commuters (CPC), and their correlation. More specifically, the RMSE metric is calculated based on the following formula:

$$RMSE = \frac{\sqrt{\sum_{i}^{n} \sum_{j,j\neq i}^{n} (\hat{T}_{ij}^{2} - T_{ij}^{2})}}{n(n-1)} \tag{4.3}$$

The CPC metric has been widely employed for quantifying the goodness of flow estimation (*e.g.* Yang et al., 2014; Lenormand et al., 2016; Barbosa et al., 2018). It builds upon the Sørensen index (Sørensen, 1948), and quantifies the share of correctly identified flows (formula (4.4)).

$$CPC = \frac{2\sum_{i}^{n} \sum_{j,j\neq i}^{n} \min(T_{ij}, \hat{T}_{ij})}{\sum_{i}^{n} \sum_{j,j\neq i}^{n} T_{ij} + \sum_{i}^{n} \sum_{j,j\neq i}^{n} \hat{T}_{ij}} \tag{4.4}$$

The different evaluation metric results are reported in tables 4.1 and 4.2.

Table 4.1: Evaluation of different trip distribution models for car commuting

Alternative	RMSE	CPC	Corr.
Generic: production	18.277	0.657	0.786
Generic: attraction	16.901	0.652	0.778
Generic: average	17.405	0.658	0.785
Model-based: production	21.074	0.439	0.674
Model-based: attraction	13.712	0.518	0.789
Model-based: average	15.717	0.490	0.742
Radiation model	16.175	0.466	0.704
Gravity model: unconstrained	13.464	0.596	0.796
Gravity model: doubly-constrained	10.075	0.712	0.910

Table 4.2: Evaluation of different trip distribution models for public transport commuting

Alternative	RMSE	CPC	Corr.
Generic: production	21.848	0.360	0.657
Generic: attraction	19.459	0.411	0.671
Generic: average	20.485	0.385	0.668
Model-based: production	19.277	0.330	0.658
Model-based: attraction	11.494	0.377	0.641
Model-based: average	14.250	0.363	0.677
Radiation model	16.155	0.402	0.541
Gravity model: unconstrained	7.025	0.380	0.654
Gravity model: doubly-constrained	5.439	0.587	0.812

As it can be observed, in both cases the gravity models produce predictions closer to the ground truth commuting values, as it was anticipated. Attempting a closer look though at the results of the other models, it appears that the different metrics paint a different picture with respect to identifying the model with the highest predictive accuracy. More specifically, for the car case (table 4.1), the attraction constrained gravity model with the model-based interaction function has the lowest RMSE values along with the highest correlation with the "true" commuting values. Its counter form, involving a generic function, scores high as well with a substantial higher value of CPC than the model-based cases. Furthermore, the radiation model has the second lowest CPC value while also in terms of correlation and RMSE it is outperformed by most of the models.

For the public transport case (table 4.2), similar patterns as before can be noticed with the traditional gravity models exhibiting the highest predictive accuracy. Among the rest, most interestingly the attraction constrained version of the model-based function outperforms by far the other models in terms of RMSE. In general, the RMSE metric is more sensitive to the existence of large outliers than the other metrics. The radiation model ranks among the ones with the highest values of both CPC and RMSE. However, in terms of correlation is the worse performing trip distribution alternative. Nevertheless, the results of both modes highlight that the introduced spatial interaction models bear the ability of capturing relatively well the actual commuting patterns. A further exploration, involving another ground truth data set in place, would be required though to draw solid conclusions.

Nevertheless, one critical aspect that the various estimated spatial interaction approaches fail to address is mode choice. Typically and for the traditional transport demand models case, this happens sequentially once the total demand has been generated and distributed over space. To examine the extent to which a simplistic approach can capture mode choice considerations, the proposed spatial interaction framework is further modified. More specifically, the mode-specific interaction rate for each pair of locations is defined as:

$$S_{ij}^{\text{mode}} = \max(S_{ij}^{\text{car}}, S_{ij}^{\text{PuT}}) \frac{S_{ij}^{\text{mode}}}{S_{ij}^{\text{car}} + S_{ij}^{\text{PuT}}} \tag{4.5}$$

Subsequently, the evaluation happens in two ways; one that focuses on the mode-specific flow predictions, and one on the total demand irrespective of modes. The results corresponding to the average case are reported in table 4.3. As it can be seen, small differences can be identified among the different versions of model while the total demand predictions have the highest correlation and CPC values. A comparison of the mode-specific values with the previous results (tables 4.1 and 4.2) reveals different patterns. For instance, the mode-choice public transport generic version of the model outperforms its previous mode-specific form, in terms of RMSE and CPC. The same can be observed for the case of car but only according to the RMSE metric. On the other hand, the model-based variants of the mode-choice formulation seem to be giving rise to slightly worse predictions than before. Nonetheless, and as mentioned before, a further exploration with another data set with true commuting flows would be required to draw definite conclusions[2].

4.2.2 *Predicted interaction spaces*

The next part of the analysis deals with the interaction space aspect of the proposed spatial interaction functions. To this end, the predictions of the two estimated aggregate variants are employed for the determination of the interaction space. More specifically, the focus in this particular case lies

2 First informal tests on a data set with mobile-based generated O-D matrices show interesting results.

TABLE 4.3: Evaluation of different trip distribution models with mode-choice considerations

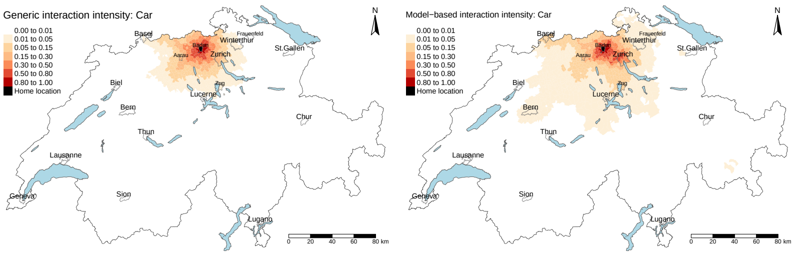
	RMSE	CPC	Corr
Generic: average (PuT)	17.037	0.531	0.663
Generic: average (car)	15.275	0.535	0.760
Generic: average (total)	19.039	0.619	0.764
Model-based: average (PuT)	16.986	0.451	0.668
Model-based: average (car)	17.044	0.350	0.703
Model-based: average (total)	19.757	0.476	0.750
Previous predictions			
Generic: average (PuT)	20.485	0.385	0.668
Generic: average (car)	17.405	0.658	0.785
Model-based: average (PuT)	14.250	0.363	0.677
Model-based: average (car)	15.717	0.490	0.742

on both the spatial extent along with the interaction rates, as implied by the predictions. In order to examine the differences between the two specification ways of the interaction function, the mode-specific predictions for a given location are produced. In particular, the chosen location is Baden, which is a medium-sized city between Zurich and Basel. In figures 4.1 and 4.2, maps of absolute interaction intensity rates per variant, along with their absolute differences, are presented.

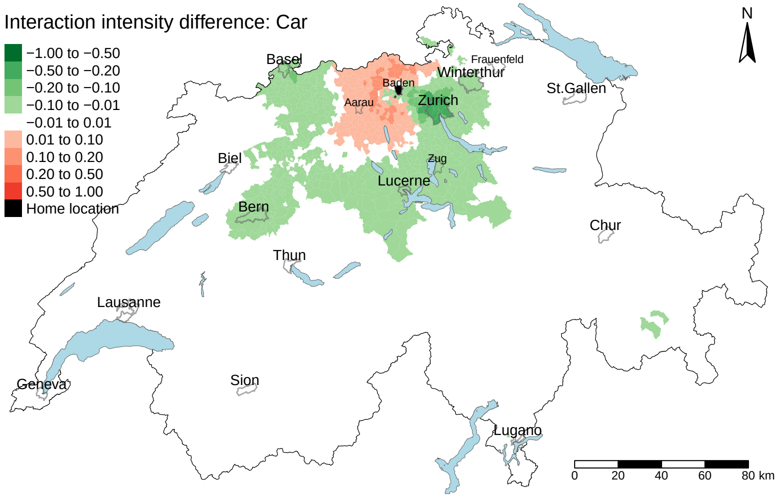
As it can be seen in these figures, substantial differences exist between the interaction intensity predictions of the two variants. In the case of both modes, the generic predictions highlight a smaller space of potential interaction with a much steeper reduction of rates. Interestingly, in both cases and especially in the public transport case, the model-based interaction intensity variant assigns (high) values to more distant zones.

More specifically, if we focus on the influence that major cities with high concentration of economic activity, such as Zurich, Basel, Bern and Lucerne, exert on the predicted interaction intensity rates, it appears that the generic rates are substantially lower than the model-based ones. For instance, the city of Bern is not classified as interacting one according to the generic variants while that is not found to be the case for the model-based ones. In a similar spirit, the results for the city of Zurich surface

FIGURE 4.1: Car interaction rates predictions for a specific location

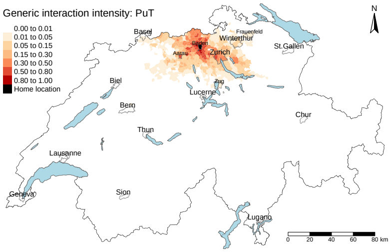


(a) Generic predictions (b) Model-based predictions

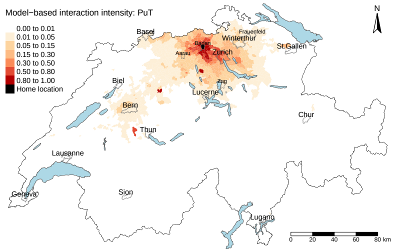


(c) Absolute differences (generic - model-based rate predictions)

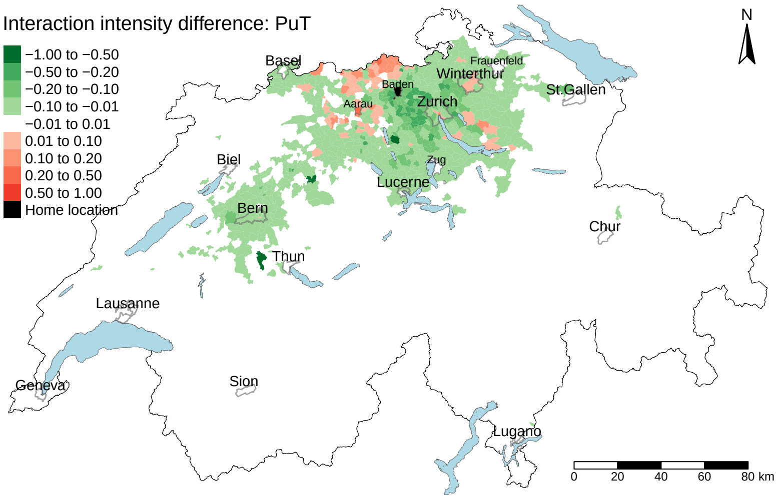
Figure 4.2: Public transport interaction rates predictions for a specific location



(a) Generic predictions



(b) Model-based predictions



(c) Absolute differences (generic - model-based rate predictions)

the differences between the two variants. In conclusion, it appears that the model-based specification yields more realistic and meaningful results than the generic one.

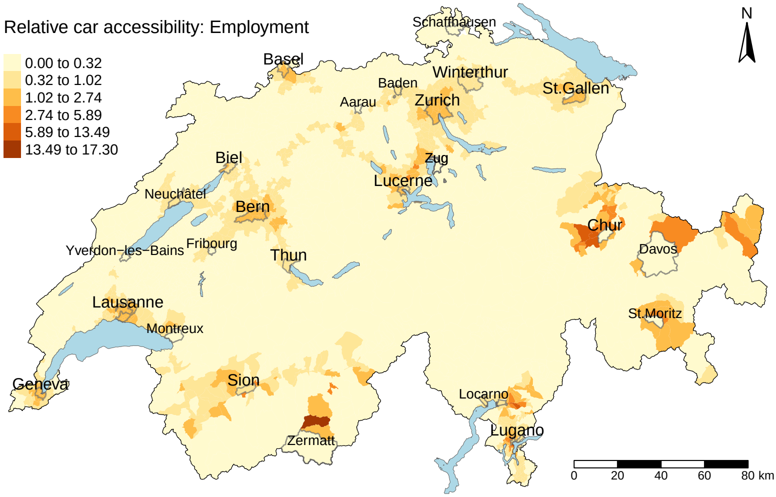
4.2.3 *Relative accessibility measures*

As mentioned in the previous chapter (section 3.1), an important aspect of accessibility measures concerns their positive or normative character. Please note that it is presumed that implementations with a generic interaction function in place are perceived as normative, while the ones with a model-based function as positive. Driven by this, the two types of measures can be combined in a single relative accessibility measure. A discussion on the construction and interpretation of relative accessibility measures can be found in (Páez et al., 2010) where the use of such indicators was demonstrated for examining social exclusion issues.

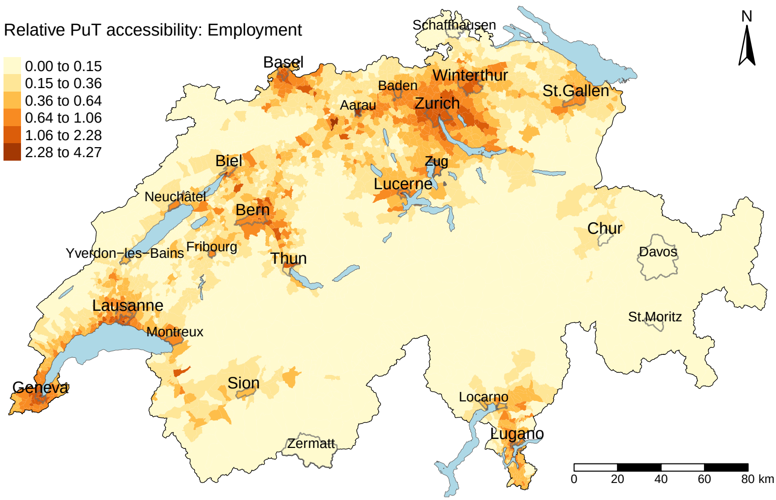
Driven by this, a relative accessibility indicator can be defined as the ratio between the normative and the positive implementations of the gravity-based accessibility measure. Essentially, the normative measure quantifies the average access to opportunities. On the other hand, the positive measure quantifies the actual access to opportunities that people have due to the spatial structure and spatial matching issues. Therefore, values lower than one reflect a location with overall limited access to opportunities, where its residents have to travel further away than on average to gain the required level of access to opportunities. On the contrary, values higher than one indicate a location that has very good access to opportunities and therefore its residents have to travel relatively small distances to obtain the needed accessibility level. Furthermore, the formation of such indicators can also assist in the study of the phenomenon of excess commuting (*e.g.* Ma and Banister, 2006; Schürmann, 2015). The relative accessibility values to employment opportunities are plotted spatially in figure 4.3.

For the car case, the relative accessibility values reveal that mainly the urban areas, along with their neighbors, are the ones with the highest values. On the other hand, locations further away from the main cities have the lowest level of relative access to employment opportunities. This finding is to a large degree anticipated though since people residing in such places typically have to commute long distances to gain access to the job market. In the case of the public transport relative accessibility indicator, similar patterns emerge but only for the main cities and their surrounding

Figure 4.3: Relative accessibility measure values



(a) Car



(b) Public transport

areas while more remote areas have substantially lower values of relative access to employment opportunities.

4.3 DISAGGREGATE SPATIAL INTERACTION CASE

4.3.1 *Predicted interaction rates*

The disaggregate variant of estimated interaction intensity functions can be employed in a similar manner as before for the prediction of the interaction space of individuals. The key difference with the aggregate model-based variant lies on the fact that individual characteristics are now taken into account, allowing in sequence the specification of individual-specific functions. The application of the estimated interaction function for predicting the interaction rates for an individual with specific characteristics is demonstrated in figures 4.4 and 4.5, accordingly.

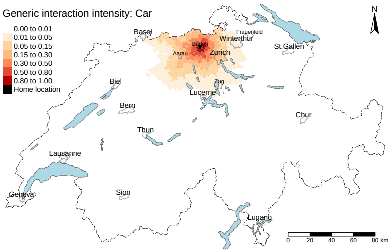
More specifically, the person is a female at the age of 25, being part of 3-persons household that includes a child, working part-time, having a bachelor degree, while it resides in Baden. As the corresponding figures show, similar patterns as before can be identified.

4.3.2 *Relative accessibility measures*

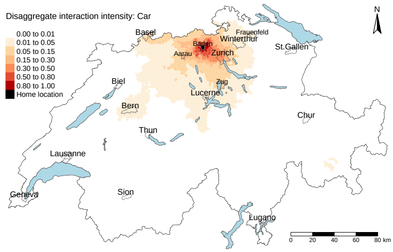
Finally, the application of the estimated individual-specific interaction intensity functions for the construction of disaggregate accessibility indicators is demonstrated. To that end, these indicators are employed for studying differences on the individual levels of regional accessibility. The employment accessibility values for a person with varying characteristics are calculated. More specifically, for the base case the person is a female at the age of 25, being part of 3-persons household that includes a child, working part-time, and having a bachelor degree. In the scenario case, the characteristics of the person remain the same with the only difference that she now has an MSc degree, works full-time and is 35 years old.

Based on the disaggregate model estimates (table 3.6), the age increase has a negative impact on the accessibility levels while a positive impact is implied for the acquisition of the MSc degree and the full-time employment status. The disaggregate accessibility values are calculated for all zones, both for the base case and the scenario.Subsequently, a relative accessibility indicator can be formulated as the ratio of the scenario accessibility values to the base ones. More specifically, in this case the indicator would capture the relative percent difference on the positive accessibility levels. The results are presented in figure 4.6.

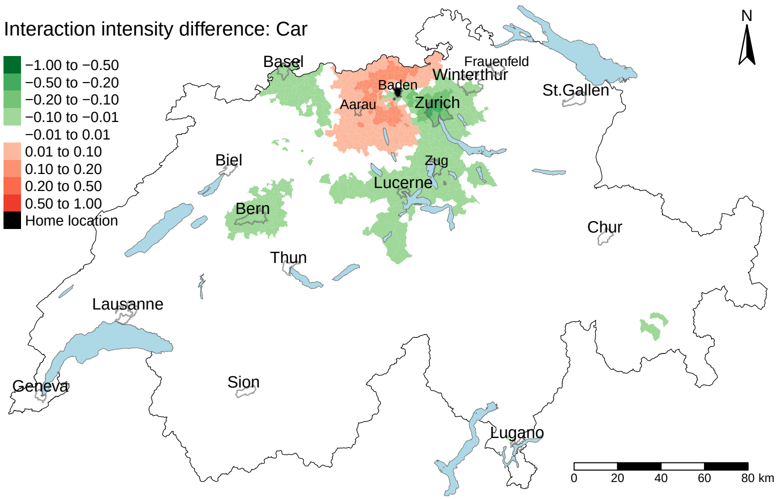
FIGURE 4.4: Car interaction rates predictions for a specific location and an individual with certain characteristics



(a) Generic predictions

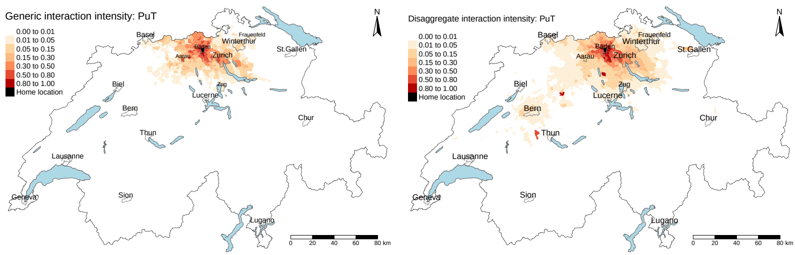


(b) Model-based predictions

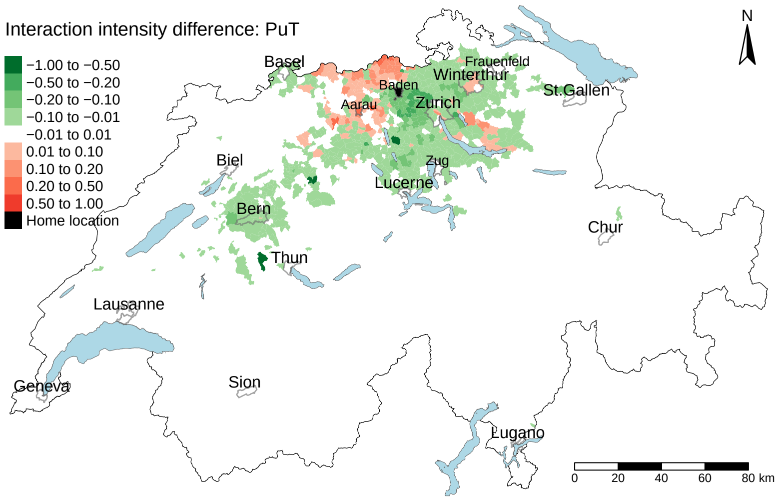


(c) Absolute differences (generic - model-based rate predictions)

Figure 4.5: Public transport interaction rates predictions for a specific location and an individual with certain characteristics

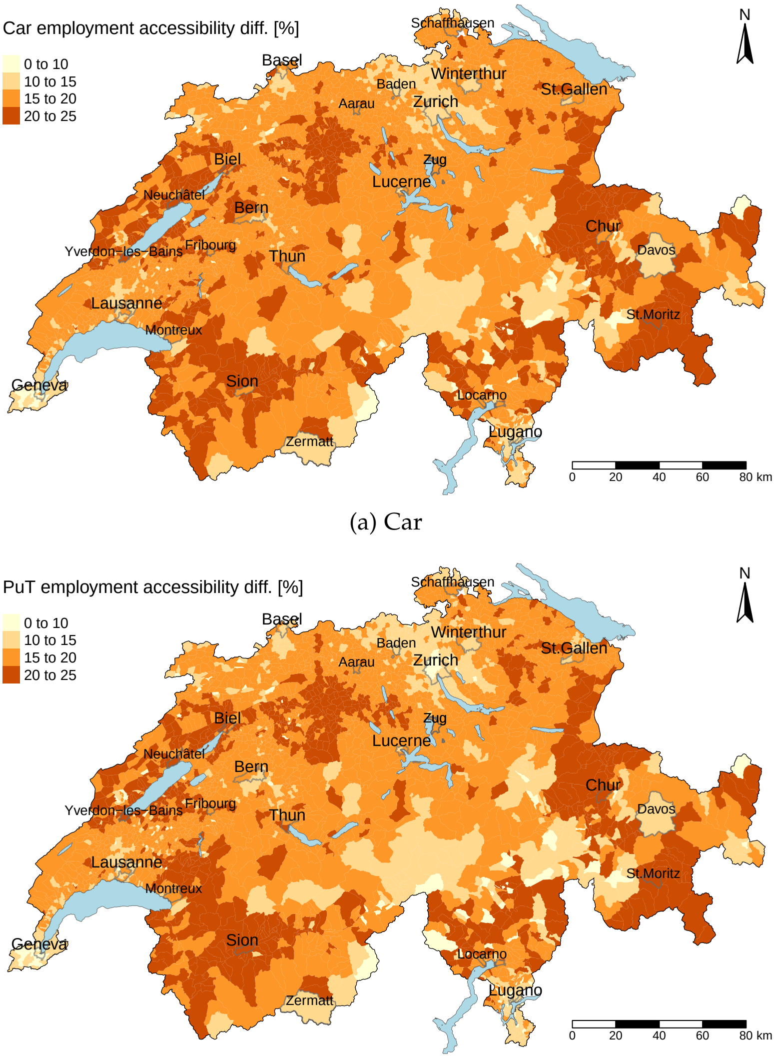


(a) Generic predictions (b) Model-based predictions



(c) Absolute differences (generic - model-based rate predictions)

Figure 4.6: Relative accessibility measure values



(a) Car

(b) Public transport

As it can be seen in figure 4.6, substantial differences on the individual regional accessibility levels exist. Interestingly, the differences are found to lie in the range of 1 to 25%, revealing strong spatial variability, while for both modes the patterns are similar. More specifically, smaller differences can be identified in the main cities while the differences are magnified in the cases of rural areas. Based on this, it can be concluded that the impact of the individual characteristics is much more important for the less urbanized areas, implying the existence of a large variance on the accessibility levels of individuals residing to such areas.

On the spatial interaction side, the estimated model can be applied as well for predicting interaction values in line with section 4.2.1. However, in this case knowledge of the specific characteristics of all individuals per zone would be required as input. To overcome this massive requirement in terms of data, the characteristics of a sample of individuals can be utilized instead, with a proper weighting of the observations though to account for the fact that they constitute a sample (*e.g.* synthetic population). Nevertheless, agent-based populations would allow a disaggregate application.

4.4 CONCLUSIONS

In this chapter the application of a new spatial interaction specification way was illustrated for different instances. Depending on the perspective of analysis, specifying the interaction intensity function as a survival one can tackle both aggregate and disaggregate cases, while it also paves the way for exploiting the construction of various relative accessibility measures. In this regard, the proposed specification way is flexible in the sense that it allows for the construction of generic, location-, and individual-specific interaction functions.

In conclusion, the different results demonstrated the capacity of the proposed specification to be employed for examining various aspects associated with spatial interaction phenomena such as commuting and access to different opportunities. Especially for the commuting case, the results highlighted that a gravity-based model with a survival function in place can produce reasonably good predictions by quantifying the potential for interaction between origins and destinations. Furthermore, the results attest to the ability of the model-based variants to produce reasonable and realistic predictions of the potential interaction space. To that end, the ability to account for location and/or individual characteristics within a gravity-based accessibility measure constitutes an aspect which was not addressed in the literature to date. Last, the application of the proposed specification for studying differences on the individual levels of regional accessibility, revealed interesting spatial insights regarding the impact of the individual characteristics.

5

ACCESSIBILITY AND CENTRALITY

5.1 INTRODUCTION

As discussed in the previous chapter, accessibility is a central concept in transportation research. Gravity-based measures are by far the most commonly used in practice, due to their ease of implementation, interpretation, and communicability (Geurs and van Wee, 2004). As presented in formula (2.1), it can be seen that such measures are the summation over j of $W_j f(d_{ij})$ for a given i. This represents the potential of interaction of i with the rest of the system, in terms of opportunities W. In summary, a key aspect of gravity-based accessibility measures is that they measure the potential for interaction between origins and destinations.

An alternative approach to understand spatial interaction is due to Stouffer (1940), who talked about intervening opportunities (also see Black and Conroy, 1977; Cheung and Black, 2008), as discussed in section 2.2.2. Stouffer in his work was interested, in addition to the origins and destinations of trips, in the pass-through flows, that is, what opportunities were available for trips beween origin i and destination j. This aspect of the system is important not only for moderating trip length (something that is captured in gravity-based accessibility measures by a properly calibrated interaction function $f(d_{ij})$), but also as a measure of exposure of travelers to opportunities. This is a perspective that seems to be missing in current accessibility research and the work presented in this chapter proposes a way of addressing.

This idea finds an analog in the literature on graph theory and social network analysis. In these streams of literature, the relevance of an element in a network is variously measured by means of centrality and betweenness indicators (*e.g.* Freeman et al., 1979; Borgatti, 2005). To that end, among the most popular ones is the degree centrality that quantifies the number of links incident upon a node (Freeman et al., 1979). In that sense, degree centrality is similar to accessibility in that it counts the potential for interaction in a network, albeit in a more limited way (*e.g.* by considering only the first order neighbors in a graph). A cumulative opportunities accessibility indicator, on the other hand, considers all neighbors that satisfy some cost threshold c.

Betweenness measures, in contrast, are concerned with flow-through traffic. One example is the stress-centrality (Shimbel, 1953) and the betweenness-centrality (Freeman et al., 1979) indicators. These indicators are based on the number of paths that pass through an element of the network, say i, when interaction between (all) elements j and k is established. See from this perspective, gravity-based accessibility measures are similar to centrality indicators used in network analysis in the sense that they count the potential of interaction between locations of the system. Apart from their conceptual similarity though, a critical difference concerns their character with the former ones assuming a location-based point of view while the latter ones a network-based one. Nevertheless, an implicit connection between these well-established concepts exists which has not been acknowledged in the literature.

Driven by the above, the objective of this chapter is to propose and discuss a new measure of centrality called betweenness-accessibility. This measure is inspired by the measure of betweenness-centrality used in social network analysis (see Wasserman et al., 1994; Degenne and Fors, 1999), which in turn is closely related to the concept of stress-centrality (also see Shimbel, 1953; Brandes, 2001). Betweenness-accessibility, it will be seen, is a useful tool to investigate the incidental impacts of accessibility on a network. This measure, when used in conjunction with accessibility analysis, provides a richer picture of the ways a transportation system operates to generate connectivity in the system and how the landscape of opportunities in turn impacts the network. It should be noted that the formulation of the betweenness-centrality centrality was first introduced in Sarlas and Axhausen (2015b).

In the following sections, a brief overview of the concept of centrality precedes the introduction of the betweenness-accessibility measure. A real-life case study is designed to demonstrate the application of the introduced measure along with discussing its construction, what exactly it represents, and last its potential utility in terms of applications.

5.2 NETWORK AND BETWEENNESS-ACCESSIBILITY CENTRALITY

5.2.1 *Centrality: stress and betweenness*

Suppose that there is a graph representation of a network composed of nodes and edges that connect some or all of those nodes. Under this setup, the betweenness of a node is a measure based on the frequency with which a node p_i is located on the shortest path that connects the pair of nodes p_j and p_k (Freeman, 1977). This index can be conceptualized as a measure of network flow control (Wasserman et al., 1994), since it is calculated based on the number of paths $\sigma(p_j, p_k \mid p_i)$ where p_i acts as a bridge between p_j and p_k. Freeman (1977) defines betweenness as follows:

$$C_B(p_i) = \sum_{j,k \in V, j \neq k \neq i} \frac{\sigma(p_j, p_k \mid p_i)}{\sigma(p_j, p_k)} \tag{5.1}$$

where $C_B(p_i)$ corresponds to the betweenness of node p_i, $\sigma(p_j, p_k \mid p_i)$ is the number of shortest paths between nodes p_j and p_k that pass through node p_i, $\sigma(p_j, p_k)$ is the total number of shortest paths between p_j and p_k, N is the total number of nodes in the network, and $j, k = \{1, 2, ..N\}$. Subsequently, a normalization term $\frac{1}{(N-1)(N-2)}$ can be included in order to turn the measure into the system-wide proportion of shortest paths that pass through node p_i.

A similar measure but for the case of edges l_i was proposed by Girvan and Newman (2002), whereby betweenness now represents the number of paths between nodes that pass through a specific link.

$$C_B(l_i) = \sum_{j,k,j \neq k} \frac{\sigma(p_j, p_k \mid l_i)}{\sigma(p_j, p_k)} \tag{5.2}$$

Normalization can be implemented in this case by multiplying the measure by the term $\frac{1}{N(N-1)}$ to obtain the system-wide proportion of paths that traverse l_i.

Along the same line of thinking, another popular centrality indicator is the stress-centrality which was proposed by Shimbel (1953), denoted as $C_S(l_i)$ for the case of links. The calculation of that indicator takes place according to formula (5.3) that follows.

$$C_S(l_i) = \sum_{j,k,j \neq k} \sigma(p_j, p_k \mid l_i) \tag{5.3}$$

As it can observed, the difference between the two measures lies on the fact that C_B takes into account the fraction of the shortest paths for every pair of nodes passing through a specific node/link, whereas C_S takes into account the corresponding absolute number of shortest paths. Nevertheless, in cases where only one shortest path exists per pair of nodes, both measures yield identical results. Both indicators conceptualize though that all sets of interacting nodes contribute the same towards the centrality value of each link ($\frac{1}{N(N-1)}$).

Apart from the aforementioned betweenness centrality measures, variants of the original formulation have emerged over the years. Brandes (2008) provides a coherent review of those variants, discussing also their algorithmic implementations. One prominent example of such measures is the so-called *distance-scaled betweenness* where the involved shortest paths are weighed inversely proportional to their length (Borgatti and Everett, 2006). Essentially, this weighting operationalizes the idea that longer paths should count less towards the calculation of betweenness. A further modification on the aforementioned betweenness measure was introduced by Geisberger et al. (2008) where the weights are defined based on the relative position of a node p_i in the relevant shortest paths.

Along the same line of thinking but for the transportation networks' case, Lowry (2014) modified the stress-centrality indicator by applying a weighting on the basis of trip production and attraction rates at both shortest paths' ends.

Nevertheless and based on the above, a general weighted formulation can be introduced, including a weighting component that allows to account for the non-uniform impact of each pair on the centrality indicators (formula (5.4)).

$$C_{\mathrm{B_W}}(l_i) = \sum_{j,k,j\neq k} \frac{\sigma(p_j, p_k \mid l_i)}{\sigma(p_j, p_k)} w_{ij} \tag{5.4}$$

Subsequently, a normalized indicator can be obtained by dividing the measure with the sum of the employed weights.

$$C^{\mathrm{N}}_{\mathrm{B_W}}(l_i) = \frac{1}{\sum\limits_{j,k,j\neq k} w_{ij}} \sum_{j,k,j\neq k} \frac{\sigma(p_j, p_k \mid l_i)}{\sigma(p_j, p_k)} w_{ij} \tag{5.5}$$

5.2.2 *Betweenness-accessibility centrality*

For the case of transportation networks, the utilization of centrality indicators can be perceived as problematic mainly due to two main conjoint issues, as identified by Sarlas and Axhausen (2015b). The first one is associated with the aspect of travel demand and especially with the fact that trip production and attraction is not spatially uniform. The second issue relates to the trip distribution aspect. Therefore, taking into account all interactions between nodes along with imposing an equal weighting principle that assumes that all interactions account the same towards the calculation of the centrality indicators, can be viewed as problematic and potentially misleading.

A way to relax these assumptions is through the formulation of a weighted betweenness centrality indicator, as presented in formula (5.5). More specifically, the conceptualized idea is that a definition of the weights on the basis of the accessibility formulation can provide insights on how network's connectivity can generate accessibility and also investigate the incidental impacts of accessibility on the network elements.

In practical terms, a restriction with respect to the number of interacting nodes can be imposed by defining a subset of interacting nodes such as $N_S \subset N$. In its most general form, betweenness would account for interaction between all pairs of nodes. As an alternative, only the shortest paths $\{p_j, p_k\}, \forall p_j, p_k \in N_S$ can be considered instead, resulting in a betweenness indicator denoted as $C_{\mathrm{B_{OD}}}^{\mathrm{N}}$. Furthermore, similar to the gravity model, it is assumed that the mass of the nodes W_j and W_k is a useful proxy for trip production and attraction levels. Commonly used measures of mass in the gravity model framework are population (pop) and employment ($empl$).

An additional issue that needs to be resolved relates to the specification of weights to facilitate the formulation of the weighted betweenness indicator(s). Drawing on the geographical concept of distance decay firmly embedded in the gravity model, a connection between the centrality and the accessibility concepts can be established. More specifically, the interaction between nodes p_j and p_k is assumed to be determined based on the connection cost as mediated by a distance decay function $f(d_{jk})$. Thereupon, the weights can be formulated by accounting for the spatial distribution of masses (i.e., pop and $empl$). Based on this, two types of betweenness-accessibility indicators can be specified.

The first type of indicators quantifies the potential level of demand in terms of trip production (population) and attraction (employment), respectively. In particular, the formulation of the indicators is presented below, in line with formula (5.5). Both the unnormalized and the normalized versions are of potential interest since the first one yields absolute values, whereas the second relative values of potential demand.

$$C^{\mathrm{N}}_{\mathrm{pop}}(l_i) = \frac{1}{TotalPop} \sum_{j,k,j \neq k} \frac{\sigma(p_j, p_k \mid l_i)}{\sigma(p_j, p_k)} pop_j \frac{Empl_k f(d_{jk})}{Acc^{\mathrm{empl}}_j} = [\%\mathrm{TotalPop}] \tag{5.6}$$

$$C^{\mathrm{N}}_{\mathrm{empl}}(l_i) = \frac{1}{TotalEmpl} \sum_{j,k,j \neq k} \frac{\sigma(p_j, p_k \mid l_i)}{\sigma(p_j, p_k)} empl_j \frac{Pop_k f(d_{jk})}{Acc^{\mathrm{pop}}_j} = [\%\mathrm{TotalEmpl}] \tag{5.7}$$

with

$$Acc^{\mathrm{pop}}_i = \sum_{j,j \neq i} Pop_j f(d_{ij})$$

$$Acc^{\mathrm{empl}}_i = \sum_{j,j \neq i} Empl_j f(d_{ij})$$

and $TotalPop = \sum_j Pop_j$, $TotalEmpl = \sum_j Empl_j$.

Given the normalization of the weights, the interpretation of $C^{\mathrm{N}}_{\mathrm{pop}}(l_i)$ is as the estimated proportion of the total population that is allocated to link l_i when the potential for interaction across the system is considered. Conversely, the interpretation of $C^{\mathrm{N}}_{\mathrm{empl}}(l_i)$ is as the estimated proportion of jobs that are serviced by link l_i when the potential for interaction across the system is considered.

The second type of indicator is formulated in a way that quantifies the importance of the elements of the network with respect to their ability to generate accessibility. Thereupon, the focus of this type lies on the potential opportunities that can be reached through each network element, and hence it quantifies the incidental impact of the accessibility on the network. The relevant formulations are presented below.

$$C^{\mathrm{N}}_{A_{\mathrm{pop}}}(l_i) = \frac{1}{TotalAcc^{pop}} \sum_{j,k,j \neq k} \frac{\sigma(p_j, p_k \mid l_i)}{\sigma(p_j, p_k)} Pop_k f(d_{jk}) = [\%\mathrm{TotalAcc}^{\mathrm{pop}}] \tag{5.8}$$

$$C^{\mathrm{N}}_{A_{\mathrm{empl}}}(l_i) = \frac{1}{TotalAcc^{empl}} \sum_{j,k,j \neq k} \frac{\sigma(p_j, p_k \mid l_i)}{\sigma(p_j, p_k)} Empl_k f(d_{jk}) = [\%\mathrm{TotalAcc}^{\mathrm{empl}}] \tag{5.9}$$

with

$$TotalAcc^{pop} = \sum_j Acc_j^{pop}$$

and

$$TotalAcc^{empl} = \sum_j Acc_j^{empl}$$

The interpretation of $C^{\mathrm{N}}_{A_{\mathrm{pop}}}(l_i)$ is as the proportion of the system-wide accessibility to population that is supported by link l_i, whereas $C^{\mathrm{N}}_{A_{\mathrm{empl}}}(l_i)$ is the proportion of system-wide accessibility that is supported by link l_i.

5.3 CASE STUDY

In this section, the application of betweenness-accessibility in a real-life urban network is demonstrated. This allows to examine different aspects of the proposed measures, along with their potential utility in practice.

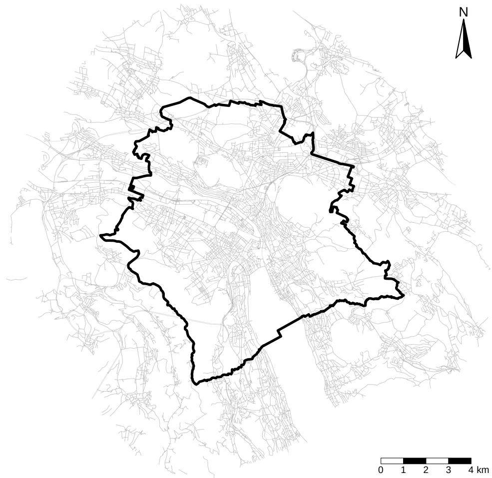
Zurich is employed as the case study. Zurich is the largest city of Switzerland with a population of over 400 thousands while its agglomeration area has a population of almost 1.4 million inhabitants[1]. In addition to its role as a major city in Switzerland, Zurich constitutes one of the main economic hubs in central Europe. The study network is of navigational-quality, which is commercially available from Tom-Tom, including all the links and the nodes within the boundaries of the city. The network includes all links and nodes within the boundaries of the city. In addition, to minimize the impact of boundary effects, the case study includes a buffer area of two kilometers. The final network consists of over 48'000 links and 23'000 nodes (see figure 5.1).

5.3.1 *Centrality indicators*

In summary, the introduced indicators build upon the concepts of centrality and accessibility, combining them in a unified measure. Therefore, apart from the network data population and employment data need to be utilized for the involved accessibility calculations. One major issue that arises though, is the selection of the accessibility analysis level along with the computational and conceptual implications that this brings in. Accessibility can take either a zonal, or a point level of analysis, depending on the

1 Statistical Data on Switzerland 2018, Federal Statistical Office

FIGURE 5.1: Network overview and city boundaries



objectives and the availability of data. Especially, the latter has emerged due to the need to analyze urban settings (*e.g.* Hanson and Schwab, 1987).

In the case of the point level analysis, the opportunities (*e.g.* population, employment positions) would need either to be reported in a spatially disaggregated way, or to make assumptions about their spatial distribution in order to assign their values on the nodes lying in their vicinity. However, in that case the construction of the betweenness-accessibility indicators would become computationally burdensome since that would require the identification and processing of the shortest paths between all pairs of network nodes (or at least of the ones with allocated opportunities), meaning $N(N-1)$ shortest paths for a directed network with N nodes.

In the case of the zonal analysis level, the definition of the zones normally coincides with different administrative or physical/built boundaries (*e.g.* neighborhoods, blocks, etc.), whereas the socio-demographic variables are typically available in an aggregated way for those zones (census data). Nevertheless, conditionally on a sufficient number of zones, zonal accessibility analysis can be useful for urban settings. In that case and in analogy to the issues faced in the traditional four-step model, the opportunities can be assigned either to few chosen nodes (*e.g.* centroid(s)), or to all nodes uniformly, or based on some assumptions about their spatial distribution. Therefore, in that case a total of Z^2 shortest paths, or $Z(Z-1)$ if self-potential is not accounted for, need to be processed, with Z being the number of zones.

For the given case study, the latter configuration is pursued, including a directed network and without taking into account the self-potential. More specifically, the zoning system of the Swiss national transport model[2] is employed as the zonal analysis level. The city of Zurich is divided into 308 zones with reported population and employment positions for the year 2015. Therefore, the identification and processing of $308 * 307 = 94556$ shortest paths is required for the construction of the introduced indicators. In order to account for the spatial distribution of opportunities within each zone, for each pair of zones a randomly chosen node within each zone is assigned as the starting and ending node respectively. Subsequently, the shortest paths are identified in terms of free-flow travel time, and thus the construction of the origin-destination travel time matrix can take place, which is a prerequisite for the accessibility calculations.

Another issue that arises is the choice of the interaction intensity function of the accessibility formulation. Given the absence of a calibrated function, the methodology for estimating the parameters in a model-based way based on the survival analysis notion is adopted (see chapter 3). More specifically, data from a household survey[3] are utilized, making use only of the observations of individuals residing in Zurich and commuting by car. The chosen interaction intensity function is of the negative exponential family, involving two parameters (formula (3.10)). The calibrated function is presented below:

$$f(gc) = e^{-6.164*10^{-4}\, tt^{3.5875}}$$

with tt the free-flow travel time in minutes.

Having defined the zonal analysis level along with the accessibility components, and having identified the required shortest paths, the next step concerns the construction of the various betweenness indicators. For comparison purposes, the results of the betweenness centrality indicator C_B are calculated and presented as well.

The data and network analysis described throughout this chapter is performed using R (R Core Team, 2018), and utilizing the *igraph* package (Csárdi and Nepusz, 2006). In addition, a further customization of those functions necessitated to assist with the construction of the various indicators. Furthermore, the presented analysis takes a link-level analysis point

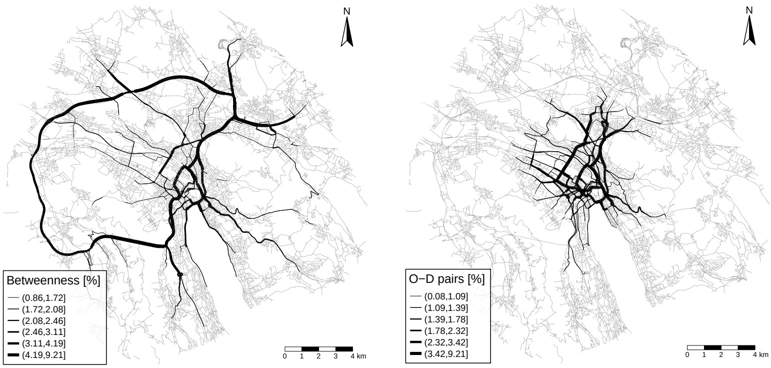
2 ARE; National transport model, 2015
3 Swiss Micro-census 2015, ARE

of view, however the same analysis could be easily conducted for a node-level point as well. The various indicators are calculated based on the normalized formulas as presented before.

5.3.2 *Analysis and results*

Having calculated the different indicators, we can now proceed to the evaluation of the different results in terms of both a visual and a correlation analysis. The various specified indicators are presented in figures 5.2 and 5.3.

FIGURE 5.2: Comparison of the C_B^N and $C_{B_{OD}}^N$ indicators

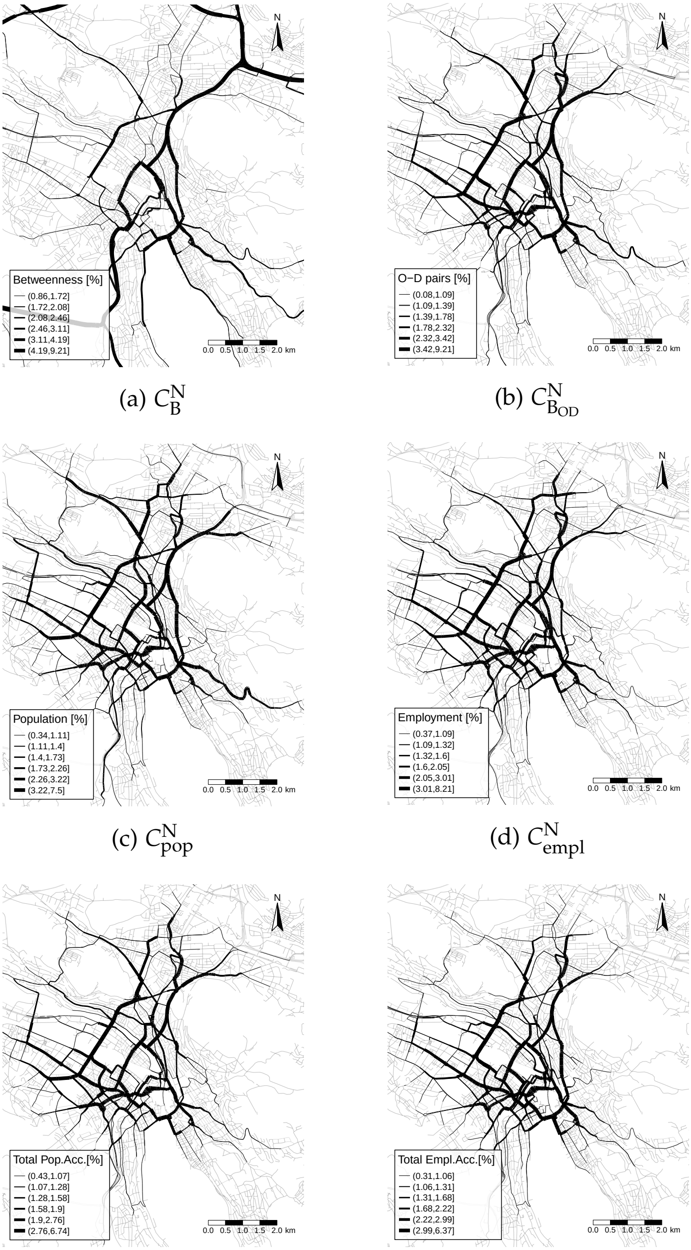


(a) C_B^N (b) $C_{B_{OD}}^N$

Interestingly, the different indicators paint a different picture of the importance of the links. A comparison between the C_B^N and the $C_{B_{OD}}^N$ (figure 5.2) for the whole network, shows that in the case of former, the links on the periphery along with some in the central area of Zurich, emerge as of high importance. In the case of the $C_{B_{OD}}^N$, the links of high importance are fully concentrated in the central area of the city. The range of the values is also of apparent interest, revealing that the 1% of the links with the highest values per case, have relative shares of $4.19 - 9.21\%$ and $3.42 - 9.21\%$ of the total shortest paths that pass through them.

The differences among the $C_{B_{OD}}^N$ and the betweenness-accessibility indicators are not visible for the whole network case due to the scale of the maps. Therefore, the focus is centered on the central area of Zurich for

Figure 5.3: The different centrality indicators



(a) $C_{\mathrm{B}}^{\mathrm{N}}$ (b) $C_{\mathrm{B_{OD}}}^{\mathrm{N}}$

(c) $C_{\mathrm{pop}}^{\mathrm{N}}$ (d) $C_{\mathrm{empl}}^{\mathrm{N}}$

(e) $C_{A_{\mathrm{pop}}}^{\mathrm{N}}$ (f) $C_{A_{\mathrm{empl}}}^{\mathrm{N}}$

highlighting visually the differences among the different indicators (Fig. 5.3). At first sight, all indicators appear to be yielding similar results. However, differences can be identified especially with respect to the corresponding shares.

In the case of $C^{\mathrm{N}}_{\mathrm{pop}}$, the highest 1% of the links have a a potential number of users in the range of $3.22 - 7.5\%$ of the total population of the city. In the case of $C^{\mathrm{N}}_{\mathrm{empl}}$, the corresponding relative share is $3.01 - 8.21\%$, revealing a rather high share of the total employment positions. Of similar magnitude are also the results of the indicators $C^{\mathrm{N}}_{A_{\mathrm{pop}}}$ and $C^{\mathrm{N}}_{A_{\mathrm{empl}}}$ but slightly lower. Interestingly, it appears that the connectivity generated through specific links serves up to almost 6.8% and 6.4% respectively of the total accessibility per case.

Last, the previous results are supplemented with a correlation analysis in order to identify similarities and also the extent to which the various indicators are related to each other. The correlation analysis is conducted by means of the Pearson correlation coefficient and is presented in table 5.1. As it can be seen, the normal betweenness indicator has the lowest degree of correlation with the other indicators. This is to a large extent anticipated due to the substantial differences in their formulations. Among the remaining indicators, in general they exhibit high values of correlation which can be attributed to the fact that they all utilize the same shortest paths, though having different weighting schemes in place.

TABLE 5.1: Correlation matrix of betweenness accessibility measures

	$C^{\mathrm{N}}_{\mathrm{B}}$	$C^{\mathrm{N}}_{\mathrm{B_{OD}}}$	$C^{\mathrm{N}}_{\mathrm{pop}}$	$C^{\mathrm{N}}_{\mathrm{empl}}$	$C^{\mathrm{N}}_{\mathrm{A_{empl}}}$	$C^{\mathrm{N}}_{\mathrm{A_{pop}}}$
$C^{\mathrm{N}}_{\mathrm{B}}$	1	0.63	0.55	0.55	0.54	0.57
$C^{\mathrm{N}}_{\mathrm{B_{OD}}}$	0.63	1	0.89	0.91	0.92	0.96
$C^{\mathrm{N}}_{\mathrm{pop}}$	0.55	0.89	1	0.69	0.93	0.78
$C^{\mathrm{N}}_{\mathrm{empl}}$	0.55	0.91	0.69	1	0.83	0.97
$C^{\mathrm{N}}_{\mathrm{A_{empl}}}$	0.54	0.92	0.93	0.83	1	0.86
$C^{\mathrm{N}}_{\mathrm{A_{pop}}}$	0.57	0.96	0.78	0.97	0.86	1

5.4 APPLICATION TO VULNERABILITY ANALYSIS

In the last part, the focus is turned on the utility and applicability of the introduced betweenness-accessibility indicators. More specifically, their ability to address the criticality issue of the network elements is tested. In this regard, a plethora of different approaches exist in the literature to evaluate the impact on a network of loss of functionality commonly referred as vulnerability analysis. As defined by Berdica (2002), *"Vulnerability in the road transportation system is a susceptibility to incidents that can result in considerable reductions in road network serviceability"*.

In general, the different vulnerability approaches have developed along two main lines of thinking (Mattsson and Jenelius, 2015). The first and most popular one identifies the most critical links by focusing on the demand and supply interaction mechanism (*e.g.* Scott et al., 2006; de Oliveira et al., 2014), normally by utilizing the output of a demand model (*e.g.* simulation). The second approach emerged mainly due to the need to analyze urban networks and draws on their topological properties, such as betweenness centrality, to identify the most critical elements (*e.g.* Duan and Lu, 2015; Demšar et al., 2008; Akbarzadeh et al., 2017; Sarlas and Kouvelas, 2019).

For instance, in a recent paper by López et al. (2017), the issue of the interaction between network topology and flow autocorrelation was investigated to draw conclusions with respect to the vulnerability of nodes. Nevertheless, the utilization of the introduced betweenness-accessibility indicators for vulnerability analysis purposes can be seen as a way to bridge the gap between the two approaches, mitigating the associated shortcomings of each, at least to some extent. A thorough overview of the literature on the topic can be found in Mattsson and Jenelius (2015) and Reggiani et al. (2015).

The most widely applied medium for studying the vulnerability aspect of a network is the interdiction of its elements, and measuring its performance in terms of various indicators such as network connectivity and travel time (*e.g.* Jenelius et al., 2006; Scott et al., 2006). Different strategies have been proposed over the years with respect to both the network performance and the interdiction aspects. For instance, in the paper by Holme et al. (2002) four different attack strategies based on topological indicators are presented, whereas the network performance is measured also in terms of graph theory concepts (average inverse geodesic length and network's largest connected subgraph).

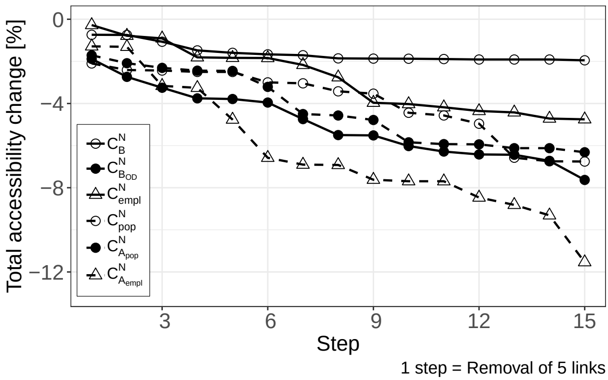
Furthermore, a strand of literature (*e.g.* Taylor et al., 2006; Chen et al., 2007; Sohn, 2006; Taylor, 2017) acknowledges that a limitation of the vulnerability analysis approaches is lacking considerations of the various associated socio-economic impacts of network's degradation. To this regard, different indicators based on the accessibility concept have been proposed to overcome this limitation.

In order to test the ability of the introduced indicators to quantify the criticality of the links, the performance of the network is studied subject to attacks on its links (edges). As highlighted by Holme et al. (2002), this evaluation approach (so-called "attack vulnerability") has originated from the field of computer networks and quantifies the decrease of network performance due to the removal of specific elements of the network (Barabási and Albert, 1999). For this particular case, the identification of the under attack links is based on their ranking in terms of the six presented centrality indicators. In addition, the removal process takes place in 5 links increments where the new shortest paths are identified, and hence a new O-D travel time matrix has to be calculated.

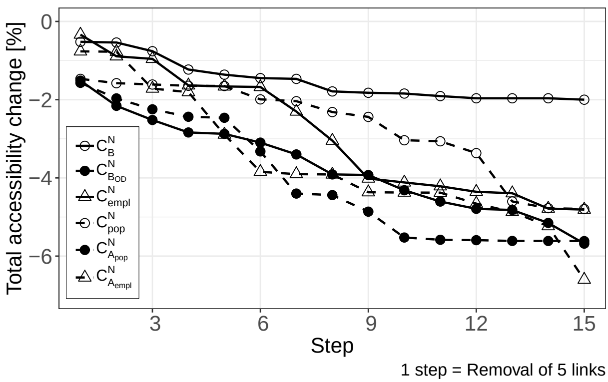
In total, 15 removal steps per indicator are evaluated whereas the robustness and the performance of the network is evaluated in two ways. The first one involves an accessibility-based approach where the impact on the total accessibility of the system serves as the chosen network performance metric. The second one utilizes the total travel time change as the performance indicator.

In particular, the problem is formulated as a transportation problem (Hitchcock, 1941) where the number of residents per zone is scaled up proportionally to match the total number of employment positions within the city of Zurich. Subsequently, an optimization process is called upon minimizing the system-wide total travel time, and hence calculating the corresponding demand matrix. In order to account for the dimensionality of demand, the optimization problem is formulated as two distinct problems. The first one (denoted as population- employment scenario), the zonal population serves as the trip production and the employment positions as the trip attraction, and vice versa (denoted as employment- population scenario). The results of the different evaluation criteria for the six centrality indicators are presented in figures 5.4 and 5.5.

FIGURE 5.4: Accessibility-based performance indicator



(a) Employment accessibility

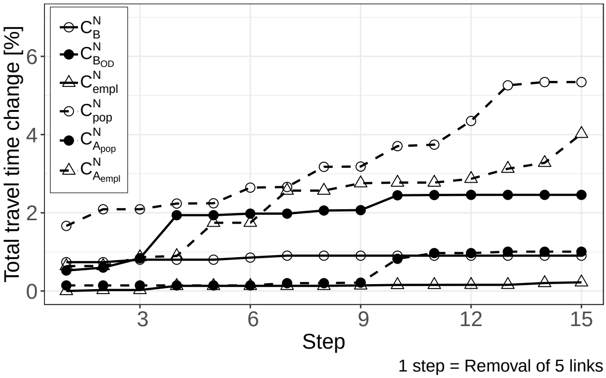


(b) Population accessibility

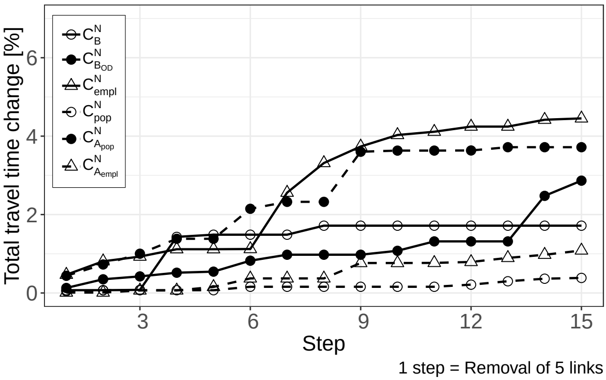
5.4.1 *Results*

For the case of the accessibility-based performance measure (figure 5.4), the network deteriorates much faster and to a greater extent when the attacks take place based on the betweenness-accessibility indicators. In particular, removing links based on the formulation with the embedded employment accessibility weights ($C^N_{A_{\text{empl}}}$) has the highest impact in terms of total accessibility reduction. Especially, for the employment accessibility case (figure 5.4a), the reduction on the accessibility levels is substantial

FIGURE 5.5: Total travel-time performance indicator



(a) Population-employment scenario



(b) Employment-population scenario

and of almost 12% with a difference of about 4% from the second best performing criterion.

In addition, it is worth highlighting that the removal of the 10 highest links results in a reduction of more than 2% of serviceability, revealing a network that is highly susceptible to the interdiction of a few links. Regarding the other ranking indicators, the $C^N_{B_{OD}}$ indicator outperforms most of the centrality indicators in both cases. Furthermore, it is noteworthy that the simple betweenness indicator (C^N_B) yields by far the worst results in terms of identifying critical links.

For the case of the total travel time performance measures (figure 5.5), the ranking based on the indicators that capture the dimensionality of the demand are the most effective ones. For instance, in the first variation (Fig. 5.5a), the transportation problem has been set up in a way to minimize the total travel time from the residence zones towards the employment zones. Thereupon, it is meaningful that the rankings based on the employment accessibility concept ($C^{\mathrm{N}}_{A_{\mathrm{empl}}}$ and $C^{\mathrm{N}}_{\mathrm{pop}}$) are the ones with the highest impact. Between those two, the addition of the trip production variable in the weighting formulation of the betweenness indicator ($C^{\mathrm{N}}_{\mathrm{pop}}$) enhances the ability to identify more critical links. The same patterns can be identified for the other variation as well (figure 5.5b) for the formulations with the population accessibility weights in place. Interestingly, the worst ranking results are identified for the formulations involving weights based on the accessibility in terms of the corresponding trip production variables.

5.5 CONCLUSIONS-DISCUSSION

In this chapter, a new indicator that combines the concepts of centrality and gravity-based accessibility in a unified measure was introduced. In particular, the indicator bears the ability of addressing inherent shortcomings associated with other centrality indicators. Furthermore, it allows to investigate the incidental impacts of accessibility on a network, facilitating a network-based presentation of the potential interactions that arise in a system. Overall, the betweenness-accessibility indicator provides a richer picture of the ways a transportation system operates to generate connectivity. This dimension of accessibility and especially how it is jointly generated by the transportation system and the landscape of opportunities, constitutes an important aspect which was neither acknowledged, nor addressed in the literature to date, at least to the best of author's knowledge.

In the case study different variants of the indicator were calculated for a real-life urban network in order to surface different aspects of its function and performance. Its utility was demonstrated through a vulnerability analysis where the most critical links were removed accordingly. In summary, the undertaken network performance evaluation highlighted that the introduced indicator is more reflective than the traditional betweenness centrality indicator of a transportation networks' function. In following chapters of the dissertation, the betweenness-accessibility indicator is thoroughly tested on its ability to improve flow and speed estimation.

In conclusion, it can be said that the value of the newly introduced indicator, especially on its general weighted version, can potentially extend beyond the scope of transportation research as it can pave the way for examining different aspects of various kinds of networks (*e.g.* social) where interaction among network elements happens in a disproportional way.

6

SPEED ESTIMATION

6.1 INTRODUCTION

As highlighted in the introduction chapter, travel demand models have increased their data demands massively both in scope and scale while overall they have become more computationally burdensome over the course of years. Contradicting that trend, the current dissertation pursues the formation of a demand modelling approach which can provide speed and volume predictions in a direct manner. To that end, the deployment of linear regression models constitutes an alternative aligned with the main objective of the thesis. In essence, such models allow to quantify the impact of different variables directly on the outcomes of interest, and hence provide a coherent framework for obtaining localized predictions.

Turning to the speed prediction issue, a closer look at the nature of the task points out that speed is essentially the outcome of the interaction between supply and demand. Driven by this realization, different ways of capturing the demand aspect within the model formulation have been proposed in the literature. More specifically, a strand of literature has resorted to the use of proxy variables for traffic volume, typically operationalized in the form of spatial density values of various socio-demographic variables (*e.g.* population, employment positions) along with land-use characteristics (*e.g.* Hackney et al., 2007; Sarlas and Axhausen, 2015a). Yet and as identified in Sarlas and Axhausen (2015b), such variables fail to capture the directionality and the complexities associated with the interregional demand, and thus can suffice only for small area cases. Thereupon, another way to tackle methodologically the problem is by accounting for the endogenous nature of demand through the use of appropriate modelling techniques, an approach though which has received relatively low attention for such problems (*e.g.* Sarlas and Axhausen, 2017).

Besides the demand aspect of the speed modelling problem, another issue is related with the spatial nature of the speed observations. In particular, the main implication of modelling such data is the existence of spatial dependence, thereby pointing to non-independent observations. For instance, the correlation of speed observations was demonstrated by Bernard et al. (2006), pointing out the necessity of accounting for spatial dependency when it comes to the estimation of speed. Moreover, Hackney et al. (2007) demonstrated the plausibility of accounting for the spatial dependence in the estimation of speed where three types of spatial regression models were estimated. Cheng et al. (2011) examined the spatio-temporal dependence structure of road networks, arguing on the need of incorporating a dynamic spatial weight matrix when it comes to forecasting on real-time data. Along a similar line of thinking, Jenelius and Koutsopoulos (2013) introduced a statistical network model for travel time estimation, allowing for correlation between travel times on different links based on a spatial moving average structure. In a different spirit, the issue of dependence was also acknowledged for the case of operating speed modelling (Park and Saccomanno, 2006). A large number of studies has focused on this issue while an overview can be found in (Highway Capacity Manual, 2010). However, their scope differs substantially than the one of the current study, since their purpose is to evaluate solely the impact of design characteristics on speed values.

In summary, a number of applications of spatial modelling techniques can be found in the urban analysis area. A comprehensive review of such applications is presented by Páez and Scott (2004). However, the presence of spatial effects constitutes a dimension which normally is neglected in the existing transport modelling approaches. Black (1992) introduced and described the existence of autocorrelation among the variables in the context of networks, stating that *"spatial autocorrelation usually concerns itself with variable values at given locations being influenced by variable values at nearby or (contiguous) locations in a spatial context. Network autocorrelation concerns the dependence of variable values on given links to such values on other links to which it is connected in a network context."*

Wang et al. (2012) review and assess the methodological issues that arise from the application of spatial models in transport. In a different context, Lopes et al. (2014) examined the spatial dependence effect on transportation demand models and specifically in the trip generation phase of the four step model. Furthermore, spatial regression models have also been applied for various transport related issues, such as traffic counts prediction (e.g. Selby and Kockelman, 2013; Zhao and Park, 2004; Sarlas and Axhausen, 2015b), and road crash predictions (e.g. Song et al., 2006; Aguero-Valverde and Jovanis, 2010). Nevertheless, there is a relatively limited number of applications employing spatial regression models for the explanation of how transport related phenomena, such as speed or flows, occur and evolve over the space.

Interestingly, the presence of endogeneity issues was demonstrated in a number of studies of a different scope though, involving treatment for the simultaneity between mean speed and speed deviations (*e.g.* Shankar and Mannering, 1998; Himes and Donnell, 2010). The same issue was also acknowledged for the case of accident models, accounting for the simultaneity between speed and accident rates (*e.g.* Cheng et al., 2013; Quddus, 2013). In another study by Porter and Wood (2013), a simultaneous equation modelling approach was proposed to explore the relationships between mean speed, standard deviation of speed and work zone design characteristics. Last, another strand of literature is concerned with modelling of mean speed values for emission models. A review of such applications is given by Boulter et al. (2007).

In conclusion, two main considerations should be made with regard to the deployment of linear regression techniques for speed prediction purposes. First, the model should account for spatial dependence issues. The second one relates to the endogenous character of volume in a speed model formulation. Nevertheless, both of the issues have the capacity of giving rise to various statistical shortcomings, such as invalid statistical testing, and inconsistent and biased parameter estimates, if remain untreated. In this chapter, the focus is centered on the first consideration. More specifically, building upon the work of Hackney et al. (2007), a larger network is employed in conjunction with a different source of speed data to enhance the understanding of the application of spatial regression models for speed predictions. The inclusion of various variables in the model specification is explored while a particular focus is given to the construction of the weighting matrices and the identification of the optimum number of neighbors. The analysis that described in this chapter is based on Sarlas and Axhausen (2015a). Last, it should be noted that the simultaneous treatment of both considerations takes place in a latter chapter of the thesis.

6.2 METHODOLOGY

6.2.1 *Spatial regression models*

Spatial econometrics was popularized by Anselin (1988), defined by the same author as *"the domain that deals with the peculiarities caused by space in the statistical analysis of regional science models"*. More specifically, these peculiarities are caused by the dependence and the heterogeneity of data in space. *"As spatial dependence, it can be considered to be the existence of a* *functional relationship between what happens at one point in space and what happens elsewhere. Spatial heterogeneity is considered to be the lack of structural stability of the various phenomena over space, and also the lack of homogeneity of the spatial units of the observations."* (Anselin, 1988).

Traditional statistical and econometric models have evolved over the years to account for the spatial effects. In this regard, spatial regression models are defined as the use of regression models that account for the impact of spatial effects in their specification and estimation. A prominent example of such models is the family of spatial simultaneous autoregressive models (SAR) which account for the spatial dependence by the inclusion of relevant spatial autoregressive components. A thorough overview and discussion of SAR models can be found in Anselin (1988); LeSage (1999); Elhorst (2014).

Spatial dependence issues arise due to the presence of spatial correlation on the dependent variable which fails to be fully explained by the different variables included in the model specification. As a result, the remaining correlation is "transmitted" to the residuals, leading to a violation of the independent and identically distributed (iid) assumption of OLS (autocorrelation). The presence of autocorrelated residuals gives rise to statistical problems such as unreliable statistical tests and biased and inconsistent parameter estimates.

SAR models constitute a modelling medium allowing to treat for this issue, assuming different underlying mechanisms that generate the spatial dependence. As suggested by Ord (1975), the estimation should be conducted by means of maximum likelihood since the ordinary least square (OLS) estimation produces inconsistent estimates. In brief, the assumption of these models is that the response variable at each location is a combination of the explanatory variables at that location but also of the response of neighboring locations.

Three main types of SAR models can be found in the literature, each one having different characteristics based on their underlying assumptions about where the autoregressive process occurs (Kissling and Carl, 2007; LeSage and Pace, 2004). At first, the spatial error autoregressive model (SAR error) assumes that the spatial dependence is in the error term of the model, and thus the spatial autoregressive process is applied to it. As Elhorst (2014) states, *"Interaction effects among the error terms are consistent with a situation where determinants of the dependent variable omitted from the model are spatially autocorrelated, or with a situation where unobserved shocks follow a spatial pattern"*. The formulation of the model is:

$$Y = \beta X + u \tag{6.1}$$

with $u = \lambda W u + \epsilon$, where Y is a vector with N values of the dependent variable, β is a vector with the regression coefficients, X is a matrix with the independent variables, u the residual term, λ the spatial autoregressive coefficient, W a matrix with the contiguity structure having dimensions N^2, and ϵ a vector of independent and identically distributed (iid) error terms.

The spatial lag autoregressive model (SAR lag) assumes that the spatial dependence exists in the response variable (endogenous interaction effects), and applies the spatial autoregressive process to the response variable, treating it as a lagged variable. The formulation of the model is:

$$Y = \rho W Y + \beta X + \epsilon \tag{6.2}$$

where ρ is the spatial autocorrelation parameter, and WY is the term corresponding to the lagged dependent variable.

The third type of autoregressive model, namely X the spatial lag of x (SLX), assumes that the spatial dependence exists in the independent variables (exogenous interaction effects), and thereupon applies the spatial autoregressive process to them. The formulation of the model is:

$$Y = \beta X + W X \gamma + \epsilon \tag{6.3}$$

with γ a vector of parameters.

Apart from the three main SAR models, where each one assumes different interaction effects, SAR models assuming more than one interaction effects can be formulated. More specifically, the spatial mixed autoregressive model (SARmix, also denoted as spatial Durbin model in some applications; *e.g.* in LeSage and Pace (2004)) assumes that the spatial dependence exists in both the response and the independent variables. The formulation of the model then becomes:

$$Y = \rho W Y + \beta x + W X \gamma + \epsilon \tag{6.4}$$

The spatial autocorrelation model (SAC) assumes that the spatial dependence exists both in the response variable and the error term. The formulation of the model in that case is:

$$Y = \rho W Y + \beta x + u \tag{6.5}$$

On the front of spatial heterogeneity, geographically weighted regression (GWR) constitutes a technique which allows different relationships to exist in space, instead of a global relationship (Brunsdon et al., 1996). The formulation of the model corresponds to an extended simple linear model with spatially-varying parameters and is given in formula (6.6).

$$Y_k = \beta_k x + \epsilon \tag{6.6}$$

with k the location index. The estimation of GWR models typically takes place using a weighted least squares approach where the weights are defined by means of a weighting function of distance (*e.g.* negative exponential).

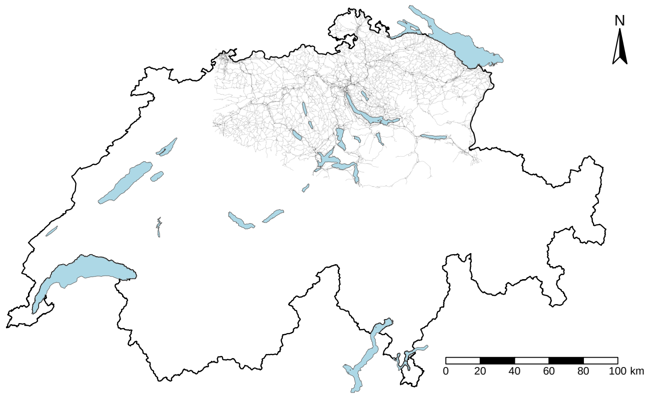
6.2.2 *Spatial weight matrices*

A key aspect of the spatial regression models is to account for the spatial structure of the data. This is facilitated by the inclusion of a spatial weight matrix, denoted as W, in the model specification. In summary, the spatial weight matrix W serves a two-fold purpose. First, it specifies the neighborhood of each location, and second it assigns weights on the neighboring locations on the basis of different schemes (*e.g.* binary, inverse distance weighted *etc*).

In the transport network case, it specifies the expected direction and mechanism of influence. An overview of the standard approaches concerning the construction of the W can be found in Harris et al. (2011) while the identification of the "true" spatial matrix issue constitutes a relatively understudied topic, lacking formal guidance (Anselin, 2002). Nevertheless, the majority of the identification approaches have developed along the lines of minimizing the Akaike criterion (AIC) (e.g. Seya et al., 2013; Herrera et al., 2012).

It should be noted that even though spatial econometrics have found wide application in applied research (*e.g.* regional studies), their application has not been without theoretical objections and skepticism. For instance, in Gibbons and Overman (2012) it is argued that identification problems are inherent to the estimation of such models for reasons associated with the formation of the W (*e.g.* unknown true weights, endogeneity, *etc.*). In light of the criticism, an interesting brief overview of the main points raised against the use of spatial econometrics is presented by Partridge et al. (2012).

Figure 6.1: Case study network



6.3 case study

In order to assess the plausibility of applying SAR models for localized speed prediction purposes, a large-scale case study is conducted. A part of the national network of Switzerland is selected, including the canton of Zurich and the neighboring cantons. In particular, the full road network of the North-East Switzerland is included in the chosen network. A navigational network is used, commercially available by Tom-Tom, including average daily speed values, estimated based on GPS measurements. In detail, the study network includes approximately 220'000 links (having excluded the secondary, or less important links) while the remaining links are classified based on five available types. In addition to the estimated speeds, the set speed limit is available. A map of the study network can be seen in figure 6.1.

The average daily speed of a typical weekday is the dependent variable of interest for the regression. The regression yields two speed components; first, the average road speed which is a function of the speed limit, the link type, and the length, and constitutes a non-spatial quantity. Spatial variation is added to the link speed estimates in the second component through the spatially resolved explanatory variables. Spatially resolved road and public transport network densities represent the effect of road supply on speed while spatial data on population and employment densities are taken to be indicative of the intensity of local activities, hence reflecting travel demand locally (Hackney et al., 2007).

6.3.1 *Spatially resolved variables*

Apart from the network data that presented above, the spatially resolved variables constitute an important component of the regression model since they introduce variation in the estimated average values, as these result from the non-spatial component of the model. At first, the road and public transport densities are of apparent interest since they represent the effect of road supply and also the spatial competition between the private and public modes, especially in urban areas. The road density is estimated as the total length of links within a given area and it is calculated for different radii.

The full navigational network is used for the density calculation. Besides the network's density, the density of ramp links is calculated as well as it is expected to have local impact on speed by providing access to the motorway network. In the case of accounting for the impact of the public transport network, it is less straightforward the way that a pertinent variable can be constructed. As an approximation, the density of public transport stops within a given area, is considered to serve as a good starting point.

Another source of spatially resolved variables corresponds to the demand aspects. More specifically, the socio-demographic data of interest are the population and the employment locations, aggregated per hectare available from the BfS[1]. The population data were collected in the year 2011, while the employment data in the year of 2008. Given the disaggregate level of these data (hectare based), they are also taken into account as densities over different radii. In addition to the normal densities, Gaussian kernel densities are calculated as well to account for the diminishing impact of the socio-demographic data over the space. The choice of interacting the various demand densities with the link types allows to diffuse the demand into the network in a disproportionate way.

At last, the spatially resolved variables need to be associated to the links of the network. Thereupon, each link of the network is associated with the hectare (cell) values of each spatial variable, closest to the upstream endpoint of the link.

6.4 MODEL ESTIMATION

In this section the different regression models estimates are presented and compared to assess the impacts of accounting for the spatial dependence of

1 Swiss Federal Statistical Office

speeds. More specifically, a standard linear regression model is estimated in terms of OLS, while four SAR models are estimated subsequently. A comparison of the estimated models is conducted in order to shed some light on the plausibility of the SAR models to predict traffic related variables, as speed, and also to what extent they can accomplish that.

At first, an OLS model is estimated to serve as the basis for testing the necessity of accounting appropriately for spatial dependence issues (autocorrelation). It is expected that the OLS model is going to give rise to biased and inconsistent estimates, and thus the resulted adjusted coefficient of determination will be inconsistent and not true.

In addition, OLS predicted values are going to be used for testing if spatial association exists in the residuals by estimating Moran's I measure. Depending on these results, a justified explanation of whether or not the need to account for spatial dependence arises. The independent variables that are included in the model are determined based on their predictive power and in accordance to the appropriate statistical tests, avoiding to give rise to multi-collinearity issues (none of the correlations is higher than 0.41). The summary statistics of the included independent variables are presented in table 6.1, while the specification of the model along with the estimated coefficients are presented in table 6.2.

As it can be seen in table 6.2, the adjusted R^2 is extremely high while the estimated parameters are all statistically significant, a fact which can be attributed to the large sample size. Employment and population densities are not used at the same time due to their high correlation. Notably, a differentiation of the employed densities radius for different links' types is found to be more appropriate and thus chosen, instead of a fixed radius density for all links' types. This finding exhibits that depending on the type of the link, the impact of spatial resolved variables on speed is not homogeneous, indicating a rather localized impact in the case of lower link types. Ramps' density variables have a negative impact on speed which can be explained by the fact that the higher the density of ramps, more vehicles are diffused in the adjacent local road network, leading to higher traffic loads.

The negative sign of the line density is not according to our expectations since it would be more reasonable to assume that the higher the local supply of roads, the more alternative routes exists, and thus lower congestion occurs. However, the sign of line density exhibits the opposite which reveals that locally, the higher number of roads corresponds to more intersections where the lower classified links often have to yield priority. The

Table 6.1: Summary statistics of employed variables

Variable	Mean	St.Dev.
Average daily speed [km/h]	48.82	15.63
Speed-limit [km/h]	52.93	13.28
Highways [dummy]	0.027	-
Trunk roads [dummy]	0.006	-
Collector roads [dummy]	0.008	-
Distributor roads [dummy]	0.392	-
Urban roads [dummy]	0.567	-
Road curveness [degrees]	0.048	0.186
Distributor: Public transp. stops density, r=0.5km [$stops/km^2$]	3.242	3.566
Urban: Public transp. stops density, r=0.2km [$stops/km^2$]	5.17	7.37
Highways: Population density, r=5km [pop/km^2]	683.394	820.818
Trunk roads: Population density, r=2km [pop/km^2]	955.504	1365.736
Collector roads: Employment positions density, r=2km*[$empl/km^2$]	726.682	1798.524
Distributor roads: Employment positions density, r=1km*[$empl/km^2$]	927.006	2408.971
Urban roads: Employment positions density, r=0.5km*[$empl/km^2$]	1114.429	3003.314
Urban roads: ramps' density, r=1km [m/km^2]	0.128	0.316
Distributor roads: road density, r=500 [m/km^2]	16.769	7.67
Urban roads: road density, r=100 m [m/km^2]	28.841	11.215
Highways with length less than 0.1 km [dummy]	0.441	-
Trunk roads with length less than 0.1 km [dummy]	0.641	-
Collector roads with length less than 0.1 km [dummy]	0.762	-
Distributor roads with length less than 0.1 km [dummy]	0.739	-
Urban roads with length less than 0.1 km [dummy]	0.705	-
Highways with length between 0.1 km and 0.2 km [dummy]	0.169	-
Trunk roads with length between 0.1 km and 0.2 km [dummy]	0.103	-
Collector roads with length between 0.1 km and 0.2 km [dummy]	0.064	-
Distributor roads with length between 0.1 km and 0.2 km [dummy]	0.067	-
Urban roads with length between 0.1 km and 0.2 km [dummy]	0.082	-

Note: *=kernel weighted

Table 6.2: Estimated OLS coefficients

Dependent variable: Average daily speed	Coef.	Std. Error
Speed-limit	0.472^{***}	(0.003)
Highways: Constant	73.801^{***}	(0.945)
Trunk roads: Constant	52.559^{***}	(1.181)
Collector roads: Constant	54.919^{***}	(1.002)
Distributor roads: Constant	45.655^{***}	(0.235)
Urban roads: Constant	34.975^{***}	(0.199)
Road curveness	-10.420^{***}	(0.109)
Distributor: PuT stops density, r=0.5km	-0.339^{***}	(0.011)
Urban: PuT stops density, r=0.2km	-0.149^{***}	(0.004)
Highways: ln(population), r=5km	-3.795^{***}	(0.134)
Trunk roads: ln(population), r=2km	-2.939^{***}	(0.162)
Collector roads: ln(employment), r=2km,kernel weighted	-3.529^{***}	(0.127)
Distributor roads: ln(employment), r=1km, kernel weighted	-1.822^{***}	(0.03)
Urban roads: ln(employment), r=0.5km, kernel weighted	-0.937^{***}	(0.017)
Urban roads: Ramps' density, r=1km	-0.666^{***}	(0.119)
Distributor roads: Road density, r=500 m	-0.263^{***}	(0.006)
Urban roads: Road density, r=100 m	-0.152^{***}	(0.003)
Highways: Length < 0.1 km, Dummy	-3.897^{***}	(0.282)
Trunk roads: Length < 0.1 km, Dummy	-10.902^{***}	(0.733)
Collector roads: Length < 0.1 km, Dummy	-11.611^{***}	(0.798)
Distributor roads: Length < 0.1 km, Dummy	-6.439^{***}	(0.114)
Urban roads: Length < 0.1 km, Dummy	-4.648^{***}	(0.09)
Highways:0.2 km > Length > 0.1 km, Dummy	-2.812^{***}	(0.339)
Trunk roads: 0.2 km > Length > 0.1 km, Dummy	-4.222^{***}	(0.868)
Collector roads: 0.2 km > Length > 0.1 km, Dummy	-6.447^{***}	(0.955)
Distributor roads: 0.2 km > Length > 0.1 km, Dummy	-0.263^{***}	(0.006)
Urban roads: 0.2 km > Length > 0.1 km, Dummy	-1.989^{***}	(0.103)
Adjusted R^2		0.964
Log-likelihood		-807852
AIC		1615760
Observations		220599
df		220572

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

impact of short length variables differs among the link types and shows that the short length has an impact on the speed, possibly because of the close proximity of intersections.

6.4.1 *Estimation of spatial regression models*

The first key aspect before proceeding to the estimation of the spatial regression models is to examine the existence of spatial autocorrelation, and thus justify if the need for the estimation of spatial regression models arises. Spatial autocorrelation is normally measured in terms of the Moran's I index which quantifies the degree of autocorrelation on the residuals of a model (0 value indicates no autocorrelation, while 1 or -1 perfect autocorrelation) (Anselin, 2001). However, in order to facilitate this calculation, the spatial structure of data should be defined beforehand, in the form of an adjacency matrix.

The inclusion of the adjacency in the model specification incorporates information in the model about the extent of the neighborhood, the type of the adjacency, and the relative weight that should be assigned on the neighboring locations. In order to identify the optimum spatial matrix for the problem at hand, the impact of different adjacency matrices type is assessed. Consequently, three different adjacency matrices schemes are constructed and tested thoroughly; one that identifies all the k-nearest neighbors based on the Euclidean distance, one that identifies the k-th order nearest neighbors in terms of network distance, and last, one where only the k-th order straight movements are included in the adjacency matrix.

For all three schemes, two variations of the adjacency matrix are examined to conclude if the assigned weight should be uniform for all observations (denoted as normal), or calculated based on a weighting function aiming in capturing the diminishing dependence of links over the distance or order (distance and order decay). In the first case, the weight is defined based on the inverse distance of each link from the midpoint of the base link. In the case of the other two weighting schemes, the weight is assigned as the inverse of the order of connectivity (k). Moreover, in all tested cases, each row of the weighing matrices is standardized to one.

Subsequently, the spatial autocorrelation of the OLS residuals is estimated for the different developed spatial weighting schemes, and it is found to lie within the range of 0.5 to 0.7, being statistically highly significant. This finding reveals the presence of strong spatial autocorrelation in the residuals of the OLS models, that should be treated through the estimation of SAR models.

In light of this, four SAR models are estimated for different number of neighbors and weighting schemes, namely the SARerror, the SARlag, the SAC, and the Durbin model. The optimum number of neighbors for each model is identified on the basis of minimizing the Akaike Criterion. The estimation of the SAR models and the construction of the weighting matrices was conducted in R (R Core Team, 2018), making use of the package "spdep" (Bivand et al., 2011). It should be noted that to facilitate computationally the estimation of the SAR models, the lower upper (LU) method for the decomposition of sparse matrix is used (LeSage, 1997). The analytical results are presented in table 6.3.

At first, the optimum number of neighbors for each scheme and variation is found to be the same for all the estimated models. More specifically, for the case of the k-nearest adjacency matrix, the optimum number of neighbors is equal to six for the normal weighting scheme, while for the distance decay weighting scheme it is eight. In the case of the other two weighting schemes, the optimum number of k-th order neighbors is found to be one. An interesting finding is that the application of an order decay function produces better results compared to the normal weighting function, a trend which is not present in the case of the k-nearest weighting scheme. Nevertheless, it can be concluded that the third weighting scheme, including the 1^{st} order straight links, gives the best results and hence is the one employed for the comparison and the evaluation of the SAR models that follows.

6.5 RESULTS - DISCUSSION

In tables 6.4 and 6.5, the estimated coefficients, along with the relevant goodness of fit measurements, can be seen. In summary, the coefficients of the OLS model differ significantly from the corresponding ones of the SAR models, reflecting that omitting to take into account the spatial dependence, the estimated coefficients are inconsistent and biased since more (or less) explanatory power is attributed to them. SAR models are significantly better than the OLS one, all of them having smaller values (in absolute terms) of both the AIC and the Log-likelihood measure.

It should be noted that the formulation of the model remains the same in the different model estimations in purpose, in order to allow a comparison

TABLE 6.3: Measures of quality of fit for SAR models for different weight matrices

AIC		SARerror		SARlag		SAC		Durbin	
Weighting scheme	k-nearest	normal	decay	normal	decay	normal	decay	normal	decay
k-nearest	3	1474896	1498912	1507270	1525942	1474050	1492775	1460516	1484429
	4	1459532	1479500	1497931	1512246	1459090	1473969	1446664	1465823
	5	1458598	1473010	1498341	1507921	1458480	1469365	1446866	1460196
	6	**1454827**	1467068	**1496637**	1504080	**1454806**	1464987	**1444357**	1455146
	7	1457784	1465396	1499480	1503302	1457630	1464304	1448472	1454361
	8	1459152	**1463702**	1501433	**1502640**	1458607	**1463252**	1450904	**1453379**
	9	1463834	1463964	1505937	1503471	1463135	1463809	1456580	1454368
	10	1466634	1463821	1508881	1503907	1465394	1463811	1460209	1454815
	k-th order	normal	decay	normal	decay	normal	decay	normal	decay
network	1	**1426432**	**1426432**	**1459870**	**1459870**	**1425438**	**1425438**	**1414796**	**1414796**
	2	1463463	1443308	1498131	1481536	1459081	1437002	1456881	1436600
	3	1489760	1456989	1523922	1497499	1485152	1448503	1485180	1452220
	4	1508012	1466917	1541482	1509124	1503793	1457342	1504598	1463234
	5	1521512	1474590	1554187	1518153	1517834	1464451	1518795	1471537
	k-th order	normal	decay	normal	decay	normal	decay	normal	decay
straight network	1	**1410453**	**1410453**	**1466226**	**1466226**	**1388647**	**1388647**	**1390509**	**1390509**
	2	1454948	1434346	1489788	1480813	1424690	1404782	1433716	1415466
	3	1481920	1447590	1505595	1490135	1449457	1415911	1459457	1429576
	4	1499128	1455559	1516544	1496516	1466190	1423346	1475865	1438366
	5	1510986	1460843	1524538	1501180	1478142	1428652	1487132	1444354

Table 6.4: Estimated coefficients for the SAR error and SAR lag model

Dependent variable: Average daily speed	SAR error coef.	SAR lag coef.
Speed-limit	0.254***	0.272***
Highways: Constant	96.456***	38.421***
Trunk roads: Constant	56.704***	26.84***
Collector roads: Constant	54.042***	30.047***
Distributor roads: Constant	38.941***	24.363***
Urban roads: Constant	30.332***	20.189***
Road curveness	-3.592***	-4.248***
Distributor: PuT stops density, r=0.5km	-0.083***	-0.186***
Urban: PuT stops density, r=0.2km	-0.095***	-0.073***
Highways: ln(population), r=5km	-7.978***	-2.073***
Trunk roads: ln(population), r=2km	-3.602***	-1.497***
Collector roads: ln(employment), r=2km,kernel weighted	-3.429***	-2.04***
Distributor roads: ln(employment), r=1km, kernel weighted	-1.081***	-0.881***
Urban roads: ln(employment), r=0.5km, kernel weighted	-0.501***	-0.404***
Urban roads: Ramps' density, r=1km	0.346*	-0.054
Distributor roads: Road density, r=500 m	-0.271***	-0.133***
Urban roads: Road density, r=100 m	-0.112***	-0.093***
Highways: Length < 0.1 km, Dummy	-0.713***	-1.723***
Trunk roads: Length < 0.1 km, Dummy	-2.064***	-4.967***
Collector roads: Length < 0.1 km, Dummy	-3.109***	-5.915***
Distributor roads: Length < 0.1 km, Dummy	-2.645***	-4.147***
Urban roads: Length < 0.1 km, Dummy	-3.622***	-4.127***
Highways:0.2 km > Length > 0.1 km, Dummy	-0.725***	-0.797***
Trunk roads: 0.2 km > Length > 0.1 km, Dummy	-1.632***	-2.64**
Collector roads: 0.2 km > Length > 0.1 km, Dummy	-3.047***	-3.148***
Distributor roads: 0.2 km > Length > 0.1 km, Dummy	-1.931***	-2.285***
Urban roads: 0.2 km > Length > 0.1 km, Dummy	-2.474***	-2.258***
λ	0.928***	-
ρ	-	0.459***
Log-likelihood	-705197	-733084
AIC	1410453	1466226
Residuals spatial autocorrelation	0.013***	0.342***
Observations		220599
df	220571	220571

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Table 6.5: Estimated coefficients for the SAC and Durbin model

Dependent variable: Average daily speed	SAC	Durbin	
	coef.	coef.	lag. coef.
Speed-limit	0.260***	0.267***	-0.161***
Highways: Constant	83.897***	93.021***	-76.444***
Trunk roads: Constant	51.514***	53.107***	-40.41***
Collector roads: Constant	51.287***	52.499***	-39.803***
Distributor roads: Constant	38.95***	36.618***	-25.288***
Urban roads: Constant	30.428***	29.003***	-19.623***
Road curveness	-3.597***	-4.147***	1.477***
Distributor: PuT stops density, r=0.5km	-0.143***	-0.079***	-0.007***
Urban: PuT stops density, r=0.2km	-0.094***	-0.087***	0.051***
Highways: ln(population), r=5km	-5.962***	-7.776***	7.026***
Trunk roads: ln(population), r=2km	-3.15***	-3.21***	2.58***
Collector roads: ln(employment), r=2km,kernel weighted	-3.452***	-3.25***	2.625***
Distributor roads: ln(employment), r=1km, kernel weighted	-1.244***	-1.009***	0.635***
Urban roads: ln(employment), r=0.5km, kernel weighted	-0.554***	-0.477***	0.302***
Urban roads: Ramps' density, r=1km	-0.049	0.543***	-0.51***
Distributor roads: Road density, r=500 m	-0.256***	-0.225***	0.165***
Urban roads: Road density, r=100 m	-0.115***	-0.117***	0.058***
Highways: Length < 0.1 km, Dummy	-0.859***	-1.315***	-0.23*
Trunk roads: Length < 0.1 km, Dummy	-2.368***	-3.554***	-0.177*
Collector roads: Length < 0.1 km, Dummy	-3.218***	-4.912***	0.336**
Distributor roads: Length < 0.1 km, Dummy	-2.786***	-3.573***	0.913***
Urban roads: Length < 0.1 km, Dummy	-3.823***	-3.994***	2.293***
Highways:0.2 km > Length > 0.1 km, Dummy	-0.769***	-0.843***	0.298***
Trunk roads: 0.2 km > Length > 0.1 km, Dummy	-1.835***	-2.168***	0.618*
Collector roads: 0.2 km > Length > 0.1 km, Dummy	-3.027***	-3.039***	2.377***
Distributor roads: 0.2 km > Length > 0.1 km, Dummy	-2.009***	-2.108***	1.287***
Urban roads: 0.2 km > Length > 0.1 km, Dummy	-2.56***	-2.477***	1.884***
λ	0.742***	-	
ρ	0.215***	0.722***	
Log-likelihood	-694294	-695199	
AIC	1388647	1390509	
Residuals spatial autocorrelation	-0.034	0.101***	
Observations	220599		
df	220570	220554	

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001

of all models in terms of identifying the impacts that the four different SAR models have both on the estimated coefficients and on the results.

In addition, the in-sample predictive power of each model is calculated, in order to facilitate their comparison and draw some conclusions also with respect to their ability to make accurate predictions. The predictive accuracy, in terms of predicted values that are within different specified ranges, is presented in table 6.6. In summary, OLS model performs relatively bad since less than 40% of the predictions fall within a range of 10%. On the other hand, the predictive accuracy of the SAR models is much better and it is clearly reflected that accounting for the spatial dependence, can lead to significantly improved predictions. The summary statistics of the percent error term also provide support to this argument.

Between the first two SAR models, clearly the SARerror model is better than the SARlag model in terms of AIC, indicating that accounting for the spatial dependence in the error terms of the model is more important than accounting for the spatial dependence in the response variable. Nevertheless, the SAC model gives the best results and improves further the results of SARerror model which is logical since both of the models account for the spatial dependence in the error terms, while the slight improvement in terms of AIC and predictive power can be attributed to the additional accounted spatial interaction between the dependent variables.

On the other hand, Durbin model gives slightly worse results in terms of AIC and similar predictive results. Attempting a closer look at the estimated coefficients of the Durbin model, the majority of the lagged variables' coefficients have opposite sign, compared to the variables' coefficients at the response location, which matches our expectations due to the formulation of the model. However, the magnitude of the spatial autocorrelation parameter (ρ) indicates that alarmingly high weight is attributed to the dependent variable of the neighbor(s), which as mentioned earlier is endogenous interaction and thus can raise concerns, while the impact of the independent variables at the response location is outweighed significantly. In addition, the nature of the included variables in the model specification, especially of the socio-demographic ones, can give rise to multicollinearity issues since they are not truly independent in space. Moreover, a substantial and statistically significant spatial autocorrelation remains in the residuals of the SARlag and the Durbin model, that can be taken as indicative of biased coefficients' estimation. On the other hand, the SARerror model has relatively low remaining spatial autocorrelation (0.01), while the

SAC model has statistically insignificant remaining spatial autocorrelation which shows that the residuals have spatial randomness.

Table 6.6: Predictive accuracy of estimated models in terms of predicted speeds within specified range of actual speeds, and summary statistics of their errors

Model	2%	5%	10%	15%	20%	30%	SDE	ME
OLS	8.01%	20.35%	39.86%	57.07%	70.36%	84.69%	27.25%	-5.13%
SARerror	21.25%	47.20%	69.89%	81.07%	87.21%	93.68%	16.81%	-2.05%
SARlag	14.57%	35.27%	61.09%	75.31%	82.82%	90.88%	19.33%	-2.58%
Durbin	20.63%	46.19%	70.04%	81.18%	87.39%	93.95%	16.81%	-2.05%
SAC	21.09%	47.26%	70.99%	81.92%	87.84%	94.05%	17.04%	-1.92%

Note: SDE=standard deviation of error; ME=mean error

6.6 CONCLUSIONS

In the present chapter, a methodology for predicting localized speed estimates for a large scale network was presented. The alternative of SAR models as direct demand model was examined and evaluated. In summary, the presented results of the SAR models highlighted the impact of accounting for spatial dependence issues when modelling transport related data. Furthermore, the SAC model was found to be the most appropriate among the SAR models for the given prediction task. In general, all SAR models outperformed substantially the simple OLS model, both in terms of goodness of fit measures, and in terms of predictive ability. Furthermore, the inconsistency of the estimated parameters of the OLS model surfaced, providing solid evidence in favor of accounting for spatial dependence issues when modelling speed values.

A particular focus was given to the identification of the optimum W. In this regard, three formulations were checked, conceptually driven by different underlying hypotheses. In conclusion, it was found that the most appropriate one for the given situation was the one based on network connectivity. Moreover, the process of identifying the optimum adjacency matrix based on a goodness-of-fit measure criterion was presented analytically.

In a following chapter (chapter 8), the speed estimation task is revisited to account for the endogenous nature of demand. In addition, the implications of this revision are investigated, both in terms of model specification, estimation, and optimum spatial weight matrix.

7

AADT ESTIMATION

7.1 INTRODUCTION

In this chapter, the second part of the developed direct demand modelling approach, concerning the issue of volume predictions, is presented. Many studies in the field of transport modelling have dealt with the issue of annual average daily traffic (AADT) prediction, developing different methodologies to tackle the problem. In general, two main streams of literature can be found. One that exploits different modelling techniques aiming at resolving the issues of spatial dependence and heterogeneity, while in the second stream the construction of variables describing the demand patterns is investigated. In general terms, the employed methodologies vary from aspatial regression to statistical techniques accounting for the presence of spatial effects. In particular, the latter encompass two different approaches. The first one utilizes a data-driven approach of spatial statistics called kriging, while the second one the geographically weighted regression of the class of spatial econometric models. In summary, the vast majority of the studies have proposed methodologies tailored for small, or medium, scale level of analysis in terms of network size. In addition, the focus of those studies has been mainly on the interpolation of AADT from known to unmeasured locations.

More specifically, Xia et al. (1999) developed a multiple regression model for estimating AADT on non-state roads of Florida and found that the most important contributing predictors are the roadway characteristics along with the area type, while surprisingly socioeconomic variables were found to have an insignificant impact on AADT. Similarly, Mohamad et al. (1998) developed a multiple regression model for AADT prediction for county roads in Indiana, incorporating various demographic variables which in this case were found to be statistically significant. In a similar context, Desyllas et al. (2003) developed a multiple regression analysis model for pedestrian flows.

The plausibility of applying GWR for estimating AADT was demonstrated in another study by Zhao and Park (2004) where it was shown that it can lead to the enhancement of the prediction accuracy, compared to the aspatial ordinary linear regression. Eom et al. (2006) exploited ordinary kriging for interpolating AADT for non-freeway facilities in Wake County, North Carolina, concluding that such models outperform the ordinary regression models in terms of predictive capability. In accordance with the same line of thought, Wang and Kockelman (2009) applied kriging-based methods for AADT prediction at unmeasured locations, making use of Texas highway count data, and highlighted further the utility of applying kriging for prediction purposes on a statewide network. Similarly, Selby and Kockelman (2013) explored the application of two spatial methods for prediction of AADT on the same statewide network (universal kriging and GWR), and they concluded that both methods reduce predictions errors over aspatial regression techniques whereas the predictive capabilities of kriging exceed those of GWR. Interestingly, employing network distances instead of Euclidean ones for the kriging models showed no enhanced performance.

Furthermore, Pulugurtha and Kusam (2012) developed a generalized estimating equation model to estimate AADT using integrated spatial data based on multiple network buffer bandwidths. In particular, spatial data included off-network characteristics, such as demographic, socio-economic and land use characteristics, captured over multiple network buffer bandwidths around a link and integrated by the employment of distance decreasing weights. The methodology was applied on a city level (Charlotte, North Carolina). As a continuation of the previous study, Duddu and Pulugurtha (2013) exploited the application of the principle of demographic gravitation to estimate AADT based on land-use characteristics on the same network. A negative binomial model was estimated along with neural network models. Interestingly, the results obtained showed that the developed models yielded substantially lower errors in comparison to outputs from the traditional four-step method.

In a study by Lowry (2014), a new method for interpolating AADT was presented, tailored for communities where attributes such as roadway characteristics, land-use etc., are uniform over space, and thus their inclusion in the model bears no explanatory power. The new method used novel explanatory variables that are derived through a modified form of stress centrality, a network analysis metric that quantifies the topological importance of a link in a network. The presented case study showed high quality results while later on the same methodology found application for estimating directional bicycle volumes as well (McDaniel et al., 2014). In a recent study by Jayasinghe et al. (2019), a coherent modelling approach for AADT prediction purposes was proposed that relies solely on different network analysis metrics, thus being suitable for cases where data scarcity is the norm. Likewise, the application of such metrics was also demonstrated for the case of public transport relative ridership (Luo et al., 2019).

7.1.1 *Framework*

This chapter focuses on developing a direct demand modelling approach for prediction of AADT on a nationwide network, an issue which has not been sufficiently addressed in the existing literature. The particularity of the nationwide network level case stems from the incapability of the spatial densities of different socioeconomic data to capture adequately the demand patterns that occur on the links, since they fail to bear explanatory power with respect to high volume of interregional through traffic. Naturally, the construction of a variable that can account for interregional flows necessitates, taking into consideration the direction of potential interactions, allowing in turn to capture the demand capacity interaction at the core of transport modelling. More specifically, the previously introduced betweenness-accessibility variable (chapter 5) is exploited for the particular problem at hand.

In addition to the already tested models in the literature, the family of spatial simultaneous autoregressive (SAR) models is exploited with their capability to be applied for AADT prediction purposes. The advantage of such models is that they can resolve spatial dependence issues, offering a structural explanation of the AADT and since their estimated coefficients are unbiased and consistent, they can fulfill both interpolation and forecasting purposes which is important for policy evaluation and project appraisal purposes. In summary, a set of different models is estimated and evaluated in order to draw sound conclusions on the newly constructed variable and also on models' capabilities to be employed for AADT prediction purposes and thus highlight in a quantifiable way their strengths and weaknesses. At last, a comparison of models predictive accuracy to the output of a traditional four-step model is conducted to show to what extent such models can constitute a trustworthy alternative to more advanced, but definitely more data demanding and computationally burdensome, models. It should be noted that this chapter is based on a working paper of the author with a revisited analysis in place (Sarlas and Axhausen, 2015b).

7.2 METHODOLOGY

As mentioned earlier on, the quantification of travel demand constitutes an important and critical dimension with respect to the performance of the direct demand modelling approaches. To that end, one potential alternative is exploiting network analysis indicators. To that end, centrality is an index that aims to identify the most influential persons in the context of a social network. Different centrality indices have been introduced over the years, aiming at the identification and the quantification of the importance of a particular person in a social network. In general, centrality indices take into account the number of shortest paths that pass through a given link/node. In the case where a capacity constraint exists in the form of a particular weight/cost associated with each link/node, then this weight should be considered in the routing algorithm for the identification of the shortest paths.

Departing from the social sciences questions, centrality indices are meaningful for all networks' analysis. From this viewpoint, centrality indices are meaningful for the analysis of transport networks as well and can provide a quantifiable measure of the importance of links, calculated based on the network structure and the cost of traversing each link (distance or time) . In the case of transportation, networks correspond to directed networks, given the allowed and prohibited turning movements on its vertices (nodes), and are typically modelled as higher level networks in order to account for them.

By definition, higher hierarchical links have high centrality values, while that might be the case as well for lower hierarchical links given the network structure. In the case of transport networks, the hierarchy is given by the functional class of the roads where their importance is normally matched by the number of trips using the given link. Naturally, two issues with respect to the application of the stress centrality index for transport networks come to the surface. First, the issue of travel demand since not all nodes are attracting, or producing the same number of trips. Second, interaction between nodes tends to diminish and becomes very small as the distance between them increases. Based on the above, utilizing the betweenness-accessibility variable advances as an apparent choice. It should be noted that the constructed variable mirrors to a great extent the first two steps of the traditional four-step model, however this is inevitable due to the nature of the relationships that this variable attempts capturing.

7.2.1 *Modelling Approaches*

In order to test the predictive accuracy of models for AADT prediction, the application of different models is examined. In particular, the classical OLS model constitutes the starting point due to its simplicity, where the dependent variable Y is described by a linear function of independent variables X with the parameters β being the least squares estimates, and ϵ the corresponding residual terms (7.1). One of the main assumptions of the model requires that the error should be spherical, meaning that they should be homoscedastic and not auto-correlated.

$$Y = \beta X + \epsilon \tag{7.1}$$

where Y is a vector with N values of the dependent variable, β is a vector with the regression coefficients, X is a matrix with the independent variables, while ϵ is a vector with the residuals.

However, and as mentioned before the application of the OLS estimator for the statistical analysis of spatial data can give rise to residuals that are not independent, but spatially correlated, thus leading to the violation of the assumptions of the OLS estimator. In such cases, SAR and GWR models become of high relevance in order to properly treat for spatial autocorrelation. Apart from those models, another popular modelling approach is the kriging which is a geostatistical technique used for interpolation purposes. In the case of ordinary kriging, the assumption is that the unobserved value is decomposed into two terms, the local trend βX, and the error terms which are spatially correlated and their variance is assumed to follow a semivariogram relation $\gamma(d_{ij})$, as a function of the distance d between the points. Two of the most popular types of semivariogram functions are namely the Gaussian and the spherical ones which are presented in formulas (7.2 and 7.3). More information regarding kriging and the involved parameters can be found in Oliver and Webster (1990).

$$\gamma(d_{ij}; c_0, c_e, \alpha_s) = c_0 + c_e \left(1.5\frac{d_{ij}}{\alpha_s} - 0.5\left(\frac{d_{ij}}{\alpha_s}\right)^3\right) \tag{7.2}$$

$$\gamma(d_{ij}; c_0, c_e, \alpha_s) = c_0 + c_e \left(1 - e^{-\frac{d_{ij}}{\alpha_s}}\right) \tag{7.3}$$

Last, the nature of AADT data governs the choice of another modelling approach. More specifically, the particularity of using count data, such as AADT, as the dependent variable in the context of linear regression models, stems from their non-negative character which can lead to a number of shortcomings (Winkelmann, 2015). To that end, negative binomial regression is one of the most widely used models for such data. More information with respect to those models can be found in Winkelmann (2008).

The involved data processing and model estimation is undertaken with the statistical programming language R (R Core Team, 2018), making use of the additional packages "igraph" (Csárdi and Nepusz, 2006), "spdep" (Bivand et al., 2011), "spgwr" (Bivand et al., 2017), and "gstat" (Pebesma and Heuvelink, 2016).

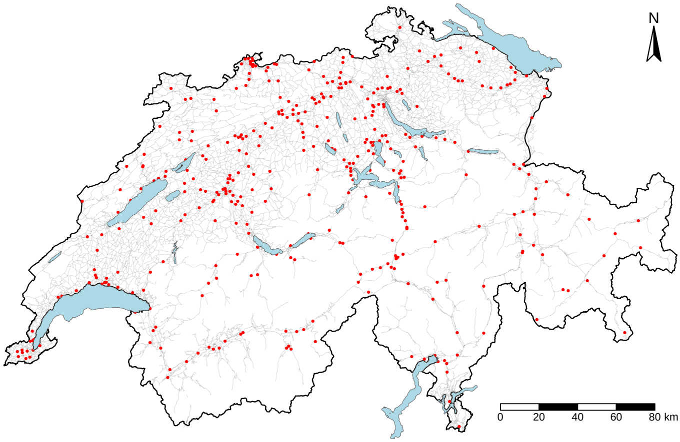
7.3 CASE STUDY

In order to assess the plausibility of applying a direct demand modelling approach for prediction of AADT on a nationwide network, along with evaluating the capability of the accessibility-weighted centrality measure to enhance the predictive accuracy of such models, a case study is designed and conducted. More specifically, a transport planning network of Switzerland consisting of approximately 40'000 directed links is employed as the study network[1], while the Federal Roads Office collects count data at various locations of the network and calculates AADT values. Moreover, the count data are supplemented by additional AADT data, freely available from various cantonal offices. Subsequently, the count locations are matched to the employed network.

As the basis year, the year 2010 is chosen in order to be comparable with the output of a calibrated four-step model[2]. In particular, AADT data on 397 links exist which are used for the model estimation as dependent values. It should be mentioned that for each count location with bidirectional traffic, only one of the two directions is randomly chosen and included in the sample. This choice is made because the available AADT data are reported per location, and not per link. Given the absence of specific information regarding the shares per direction, the obtained AADT values are divided in half and assigned equally on both directions. A map of the study network along with the spatial distribution of the count locations can be seen in figure 7.1.

1 ARE; National Transport Model (2010)
2 ARE; National Transport Model (2010): A 4-step model, implemented in VISUM

FIGURE 7.1: Case study network with the count locations



7.3.1 *Betweenness-accessibility centrality*

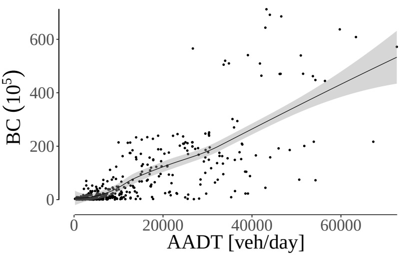
The first step is to proceed to the construction of the betweenness- accessibility centrality measure for the study network. A prerequisite for that is the determination of the origin and the destination nodes of the network that their shortest paths need to be identified. Given the interregional character of the trips, a convenient choice is to employ a zonal level according to the administrative level of municipalities. Driven by this, the zonal level of the national transport model is employed as the chosen one, including 2'944 zones in total. Subsequently, a node close to the centroid of each zone is assigned as the origin and the destination node, respectively, for the trips of each zone, associating on it the population and the employment positions of each zone. The advantage of that choice is the availability of socioeconomic data aggregated on this level while the methodology can be easily applied if more disaggregate data (*e.g.* on a hectare level) exist along with the identification of different population and employment clusters, which can then replace the employed zonal analysis level. In total, $2944 * 2943$ (due to excluding self-potential) shortest paths are identified in terms of free-flow travel time.

Finally, in the next step the calculation of the betweenness-accessibility variables of interest take place. In particular, they are calculated based on the unnormalized formulations of potential demand and in accordance with formulas (5.6) and (5.7).

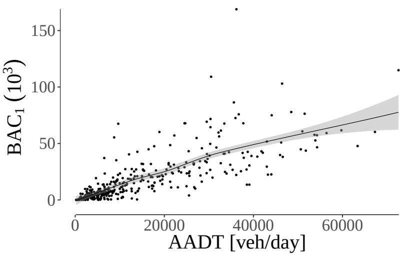
A closer look at the involved terms (*i.e.* $pop_j \frac{Empl_k f(d_{jk})}{Acc_j^{\text{empl}}}$ and $empl_j \frac{Pop_k f(d_{jk})}{Acc_j^{\text{pop}}}$) points out that they correspond to the output of formulas (3.3) and (3.1), accordingly. Thereupon, the previously predicted interaction values based on the newly introduced spatial interaction framework (section 4.2.1) can be utilized for that purpose. More specifically, two betweenness-accessibility variables are calculated; one including the generic average interaction values by car (denoted as BAC_1), and one that includes the generic average values irrespective of mode (denoted as BAC_2), and they correspond to the values evaluated in table 4.3. Both cases have mode considerations in place, and hence allow to incorporate at least partially such aspects in the model specification.

Besides the aforementioned centrality values, also the normal link betweenness centrality values are calculated as well, in line with formula (5.2) and denoted as BC. In figure 7.2, the scatter plots of AADT values against the BAC_1 and BC values are presented. Furthermore, a smoothing line is added along with the relevant 95% confidence bands (grey area) to assist with the visual interpretation of the potential existing relationships between the variables. To this end, it appears that a strong linear relationship underlines both cases, while interestingly in the BAC_1 case the confidence bands are much narrower than the BC case, revealing a smaller variance of values.

FIGURE 7.2: Scatter plots of AADT



(a) BC



(b) BAC_1

7.3.2 *Independent variables*

In essence, the regression yields two components; one that captures the impact of supply on AADT, and one that captures the impact of demand allowing to model their interaction. On the supply side, variables describing the road capacity are put to use. More specifically, the functional class of the road along with the number of lanes are the chosen explanatory variables. On the demand side, the constructed accessibility-weighted centrality measures are introduced for incorporating information about the magnitude and the direction of the spatial interactions, serving as an approximation of spatial flows. Additional spatial variation is added on the demand side by the inclusion of the population density in the vicinity of each road (within 5 km radius), as indicative of the intensity of local demand.

Last, a descriptive analysis precedes the model estimation. In this regards, the summary statistics of the different employed variables are given in table 7.1, while in table 7.2 the corresponding correlation matrix of the continuous variables is presented. In the case of the former, the range of BAC_1 and AADT values appear to be in good agreement with each other, something which obviously cannot be the case for BC. This can be seen as a first indication of the suitability of BAC_1 to serve as a proxy variable of the demand. The correlation matrix results provide further support to this argument since the correlation between the two is found to be equal to 0.75, while BAC_2 has even higher correlation (0.85). The correlation between the different centrality variables points out that no alarmingly high value exist, besides the two versions of betweenness-accessibility centrality variables. To a large extent, this was anticipated given the fact that they have a high similarity in terms of utilized interaction rates. Nevertheless, in order to avoid giving rise to multicollinearity issues due to their simultaneous inclusion in the model formulation, a ratio indicator is introduced capable of treating for mode choice related aspects.

7.3.3 *Predictive accuracy*

The evaluation of the predictive accuracy of the developed models takes place by utilizing five different accuracy measures. Mean percentage error (MPE) and mean absolute percentage error (MAPE) are easily interpretable measures, having the main disadvantage though that they can be heavily influenced by outliers. On the other hand, symmetric mean absolute percentage error (SMAPE) is a similar measure which has the advantage that it corrects for outlier's influence. In a similar manner, median absolute percentage error (MdAPE) also has the advantage that it is not influenced by outliers and can provide an overview of the distribution of the errors in

Table 7.1: Summary statistics of employed variables

Statistic	Mean	St. Dev.
AADT [veh/day]	14,431.55	14,155.99
Freeway-Highway [dummy]	0.46	-
Major road [dummy]	0.31	-
Rural major road [dummy]	0.20	-
Urban arterial road [dummy]	0.03	-
2-lane road [dummy]	0.42	0.49
Population density: 10 km* [res/km^2]	575.92	634.52
BC	8,411,925.00	13,277,549.00
BAC_1 [generic:average (car)[1]]	9,211.05	11,021.13
BAC_2 [generic:average (total)[1]]	24,724.15	31,466.01
BAC_1/BAC_2	0.45	0.16

Note: *=kernel weighted
1: with mode choice considerations, table 4.3

Table 7.2: Correlation matrix of employed variables

	AADT	Pop. dens.	BC	BAC_1	BAC_2	BAC_1/BAC_2
AADT	1	0.52	0.76	0.75	0.85	-0.35
Pop. dens.	0.52	1	0.19	0.53	0.50	-0.08
BC	0.76	0.19	1	0.59	0.75	-0.35
BAC_1	0.75	0.53	0.59	1	0.87	-0.10
BAC_2	0.85	0.50	0.75	0.87	1	-0.37
BAC_1/BAC_2	-0.35	-0.08	-0.35	-0.10	-0.37	1

conjunction with MPE. On the contrary, mean squared error (MSE) because of the quadratic term is influenced heavily by the outliers.

An overview of the employed accuracy measures is given by Makridakis and Hibon (1995), where they conclude that for forecasting purposes, MSE and SMAPE are found to be the more preferable measures. The formulas of the accuracy measures are given below with Y_i being the true values, while $\hat{Y}_i$ are the predicted ones.

$$MPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\hat{Y}_i - Y_i}{Y_i}100 \tag{7.4}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{Y}_i - Y_i}{Y_i}\right|100 \tag{7.5}$$

$$SMAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{Y}_i - Y_i}{\frac{\hat{Y}_i + Y_i}{2}}\right|100 \tag{7.6}$$

$$MdAPE = median\left(\left|\frac{\hat{Y}_i - Y_i}{Y_i}\right|100\right) \tag{7.7}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2 \tag{7.8}$$

7.4 MODEL ESTIMATION

7.4.1 *OLS models*

In this section, a set of different models is estimated and evaluated in order to draw safe conclusions on both the newly constructed variable and also on models' capabilities. In addition to the already tested models in the literature, the family of spatial simultaneous autoregressive (SAR) models is tested as well. An assessment of models' predictive accuracy and comparison to the output of a traditional four-step model is conducted to show to what extent such models can be employed for such purposes.

Initially, three OLS variants are estimated and reported in table 7.3. The models employ a log-level functional form, hence βs can be interpreted as semi-elasticity values for the case of untransformed independent variables, and as elasticity values for the log transformed ones. An important but normally neglected aspect governing the estimation of regression models is the identification of extreme outliers. In this regard, two observations are omitted from the data set since they are found to exert high leverage (higher than 6%) on the parameter estimates, identified on the basis of the Cook's distance diagnostic (Cook and Weisberg, 1982).

The first two are almost identical, albeit differentiated by means of employed betweenness centrality variable. A closer look at the results shows that the specification with the BAC_1 variable has higher adjusted R^2 value than the model with the BC variable in place. Among the three models, the third one, having also the most elaborate specification, is found to be the best one in terms of goodness-of-fit.

As expected, all model formulations yield statistically significant and positive parameter estimates for the demand relevant variables. However, the negative relationship between the AADT and the betweenness-centrality ratio can be attributed to the following reasoning. Essentially, that ratio quantifies the average weighted car share of the involved interactions. Therefore, cases where the ratio is found to be close to 1 they reveal interaction cases with car prevailing as the chosen mode, while the other way around for values lower than 0.5.

Two potential reasons for the negative sign can be posited. First, because of the simplistic incorporation of mode-choice considerations in the variable, cases with high car shares are potentially underestimating the impact of public transport, and *vice versa*. Second, another interpretation is that cases with low car shares capture interactions from/towards urban centers, or in general places with good public transport connectivity. Nevertheless, such places typically generate high intra-zonal and inter-zonal (from neighboring zones) car traffic which the variable fails to capture. In conclusion, the ratio variable can be perceived as a correction factor for over-/under-estimating the potential for interaction by car.

The results of the remaining parameter estimates are aligned with theoretical hypotheses, having the expected order of magnitude and sign. More specifically, the functional class parameters have the expected order of magnitude since they quantify the impact of the functional class, relative to the reference group. For instance, links with a lower hierarchical level than of a highway, typically have lower capacity and also less utility in terms of serving interregional demand. Therefore, the identified negative relationship is justified. Based on a similar reasoning, the positive impact of the number of lanes is also found to be in line with expectations.

Last, the OLS models are also evaluated with respect to their in-sample predictive accuracy (table 7.4). It should be noted that given the log transformation of the dependent variable, when back transforming to the original scale the fact that the model predicts the geometric mean instead of the arithmetic one is accounted through a proper correction, as suggested by Wooldridge (2012). As shown in table 7.4, Model (3) yields the lowest values of all measures while it produces predictions with an almost 3.5% lower mean absolute error. A comparison between model (1) and (2) shows that no clear conclusions can be drawn since the employed accuracy measures paint a different picture. Nevertheless, model (3) outperforms the other two models in a substantial manner, therefore it constitutes the model specification that is further tested in terms of other modelling techniques.

7.4.2 *Spatial models*

On the spatial dependence front, the existence of spatial autocorrelation in the residuals is examined in order to justify if the need for the estimation of spatial regression models arises. The spatial autocorrelation is calculated in terms of the Moran's I measure while three spatial weight matrices W are formulated and tested. The three matrices are constructed based on Euclidean distance, and network distance in terms of shortest path length and free-flow travel time. The identification of the spatial extent of the neighborhood takes place through a trial-and-error procedure that utilizes the goodness-of-fit measures along with Moran's I measure to reveal the optimum distance per case.

In particular, for both the Euclidean and the network distance, the Moran's I measure exhibits that the autocorrelation exists up to a radius of 10 kilometers. In the case of network time, the autocorrelation remains significant up to a radius of 5 minutes of free-flow travel time. The last part of the construction of the spatial weight matrices is to determine the weight that should be assigned to each neighboring location. Making use again of the Moran's I measure, we conclude that the inverse distance metric along with a normalization of the sum of the weights of the neighboring locations to one, is the more appropriate to capture the spatial structure. The Moran's I values are reported at the lower part of table 7.3 where as it is shown the network matrices W yield higher autocorrelation values. Nevertheless, all W formulations show the existence of statistically significant autocorrelation, varying between 0.16 and 0.23. The implication of this, as

Table 7.3: OLS estimates

	Dependent variable:Log(AADT)		
	OLS		
Regressor	(1)	(2)	(3)
Constant	4.92^{***}	5.86^{***}	5.78^{***}
	(0.30)	(0.17)	(0.29)
Freeway-Highway		*Ref*	
Major road	-0.51^{***}	-0.57^{***}	-0.49^{***}
	(0.09)	(0.08)	(0.08)
Rural major road	-0.68^{***}	-0.75^{***}	-0.63^{***}
	(0.10)	(0.09)	(0.09)
Urban arterial road	-0.31^{*}	-0.51^{***}	-0.38^{**}
	(0.15)	(0.13)	(0.13)
2-lane road	0.47^{***}	0.51^{***}	0.41^{***}
	(0.09)	(0.08)	(0.08)
Log(Population)[10 km]	0.36^{***}	0.22^{***}	0.26^{***}
	(0.02)	(0.02)	(0.02)
Log(BC)	0.15^{***}		0.05^{**}
	(0.02)		(0.02)
Log(BAC_1)		0.23^{***}	
		(0.02)	
BAC_1/BAC_2			-0.86^{***}
			(0.13)
Observations			395
Adjusted R^2	0.86	0.87	0.88
AIC	472.77	431.27	395.31
df			386
Moran's I Eucl. W [10km]	0.18^{***}	0.16^{***}	0.16^{***}
Moran's I Ntw. W [10km]	0.20^{***}	0.19^{***}	0.20^{***}
Moran's I Fftt W [5min.]	0.20^{***}	0.22^{***}	0.23^{***}

( ) Heterosc. corrected std. errors, + p<0.1; * p<0.05; ** p<0.01; *** p<0.001

TABLE 7.4: OLS models' predictive accuracy

	MdAPE	MPE	MAPE	MSE	SMAPE
OLS 1	26.81	17.15	39.36	40,225,804.00	0.080
OLS 2	25.62	17.24	38.17	43,732,672.00	0.079
OLS 3	24.96	14.50	34.79	35,558,953.00	0.073

mentioned before, is that the estimates are biased and inconsistent and more (or less) explanatory power is attributed to them than it should.

The type of SAR model is chosen on the basis of Lagrange multiplier tests (Anselin et al., 1996) which point towards the application of spatial error models. Furthermore, the validity of this outcome is also tested by estimating all SAR models, and then concluding on the most appropriate SAR formulation based on the significance of the corresponding spatial parameters along with the goodness-of-fit measures. The results of the different models validate the previous choice of the spatial error model as the most appropriate to address the underlying spatial autocorrelation issues. The results of the spatial error models with the three spatial weight matrices in place are presented in table 7.5 where similar patterns as before can be observed in their estimated coefficients. In all cases, the spatial autoregressive coefficient λ is found to be statistically significant at the 0.1% significance level. In terms of goodness-of-fit measures, the AIC measure shows that the spatial error model is the best one among the three.

The next estimated model is the GWR, which aims to resolve spatial heterogeneity issues and it is calculated by taking into account an adaptive radius bandwidth, identified in terms of AIC. The results are reported in table 7.6 where the summary statistics of the parameter estimates are reported.

The parameter estimates of the negative binomial regression are reported in table 7.7 where similar relationships as before are identified.

7.5 PREDICTIVE ACCURACY

Finally, an evaluation of the predictive accuracy of the developed models follows in order to draw conclusions with respect to the ability of the different models to be employed for AADT predictions. In addition to the previously estimated models, the accuracy results of two kriging models

Table 7.5: Spatial error models' estimates

	Dependent variable:Log(AADT)		
	Spatial Error Models		
Regressor	(Eucl. W)	(Fftt W)	(Ntw. W)
Constant	5.68^{***}	5.70^{***}	5.67^{***}
	(0.29)	(0.29)	(0.29)
Freeway-Highway		*Ref*	
Major road	-0.47^{***}	-0.46^{***}	-0.47^{***}
	(0.08)	(0.08)	(0.08)
Rural major road	-0.64^{***}	-0.63^{***}	-0.62^{***}
	(0.09)	(0.09)	(0.09)
Urban arterial road	-0.32^{*}	-0.32^{*}	-0.31^{*}
	(0.13)	(0.13)	(0.13)
2-lane road	0.39^{***}	0.37^{***}	0.38^{***}
	(0.08)	(0.08)	(0.08)
Log(Population)[10 km]	0.25^{***}	0.25^{***}	0.25^{***}
	(0.02)	(0.02)	(0.02)
Log(BC)	0.07^{***}	0.07^{**}	0.07^{***}
	(0.02)	(0.02)	(0.02)
Log(BAC_1)	0.17^{***}	0.18^{***}	0.18^{***}
	(0.02)	(0.02)	(0.02)
BAC_1/BAC_2	-0.84^{***}	-0.92^{***}	-0.87^{***}
	(0.14)	(0.14)	(0.14)
λ	0.24^{***}	0.31^{***}	0.29^{***}
Observations			395
AIC	381.43	367.74	373.66
df			385
Moran's I	-0.01	0.003	-0.001

( ) Std. errors, + p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Table 7.6: GWR estimates

Statistic	Median	Min	1^{st} Quart.	3^{rd} Quart.	Max
Constant	4.877	1.747	3.956	5.797	7.540
Major road	−0.468	−1.303	−0.656	−0.237	0.174
Rural major road	−0.586	−1.641	−0.852	−0.265	0.105
Urban arterial road	−0.054	−0.707	−0.337	0.179	0.987
2-lane road	0.279	−0.297	0.132	0.430	0.741
Log(Population)[10 km]	0.284	−0.059	0.225	0.341	0.540
Log(BC)	0.124	−0.063	0.061	0.182	0.341
Log(BAC_1)	0.147	−0.101	0.108	0.195	0.400
BAC_1/BAC_2	−0.709	−1.622	−1.024	−0.665	2.883
Adaptive radius=5.06% of the observations					

Table 7.7: Negative binomial regression estimates

Regressor	Estimate	Std. Error
Constant	6.12***	(0.28)
Major road	−0.50***	(0.08)
Rural major road	−0.68***	(0.09)
Urban arterial road	−0.43***	(0.13)
2-lane road	0.42***	(0.07)
Log(Population)[10 km]	0.27***	(0.02)
Log(BC)	0.04*	(0.02)
Log(BAC_1)	0.16***	(0.02)
BAC_1/BAC_2	−0.87***	(0.12)
Observations		395
Log Likelihood		−3,751.05
θ		7.18*** (0.50)
AIC		7,520.11
df		386
+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001		

are also presented. The evaluation takes place both for in-sample and out-of-sample prediction by calculating the various set accuracy measures in line with formulas (7.4) - (7.8). The in-sample results are presented in table 7.8.

Subsequently and for the out-of-sample case, the evaluation takes place based on the result of a leave-one-out analysis. More specifically, the different models are estimated for a sample size that equals to the original size, minus one observation. Subsequently, the fitted models are applied for predicting the value of the left out observation. A number of 1000 replications is performed and the corresponding accuracy measures are reported in table 7.9.

TABLE 7.8: In-sample predictive accuracy results

Model	MdAPE	MPE	MAPE	MSE	SMAPE
OLS	24.96	14.50	34.79	35,558,953	0.073
Neg. Binomial	24.62	16.54	35.83	36,428,562	0.074
SAR error: eucl. W	21.84	14.74	33.75	32,540,692	0.071
SAR error: fftt W	20.63	14.67	32.70	30,606,564	0.069
SAR error: ntw. W	21.86	15.19	33.44	31,051,310	0.070
GWR	17.51	11.09	25.51	20,662,449	0.056
National model (4-step)	4.72	3.85	12.64	3,313,040	0.028

TABLE 7.9: Out-of-sample predictive accuracy results

Model	MdAPE	MPE	MAPE	MSE	SMAPE
OLS	35.11	14.45	35.11	33,817,370.00	0.074
Neg. Binomial	35.84	16.40	35.84	34,359,908.00	0.074
SAR error: eucl. W	36.13	16.50	36.13	32,762,172.00	0.075
SAR error: fftt W	34.31	14.84	34.31	32,549,484.00	0.072
SAR error: ntw. W	34.31	14.86	34.31	32,549,238.00	0.072
Kriging: spherical	33.47	8.40	33.47	35,037,767.00	0.074
Kriging: Gaussian	33.30	8.02	33.30	35,682,116.00	0.074
GWR	32.83	14.16	32.83	27,884,970.00	0.070
National model (4-step)	13.65	5.30	13.65	3,947,758.00	0.029

A comparison of the accuracy measures results reveals similar patterns for both in-sample and out-of-sample. In particular, among the variations of spatial error models, the one that employs a spatial matrix based on the free-flow time distance gives slightly better results. Interestingly, the negative binomial model yields the results with the lower predictive accuracy for the in-sample case, while it also performs poorly for the out-of-sample case.

Among the various estimated models, GWR has the highest in-sample and out-of sample accuracy. However, in terms of MPE kriging models have the lowest values while both kriging models yield similar results. The spatial models are all performing much better than the aspatial models in the in-sample case they, while in the out-of-sample case the differences are much smaller but still evident. Furthermore, the spatial error models with the network-based constructed W outperform the one with the Euclidean-based W, providing solid evidence in favor of the former way of accounting for spatial dependence. In terms of SMAPE, small differences can be identified for the out-of-sample case which paint though the same picture as the other measures.

In general, it appears that even though GWR exhibits the highest predictive accuracy, the differences with the spatial error models are only marginal for the out-of-sample case. Moreover, counting in the fact that GWR and kriging models are aimed for interpolation purposes and not for forecasting ones, while spatial error models bear the ability to be applied for forecasting purposes as well since their parameters are unbiased and consistent, the utility of spatial error models is highlighted.

Attempting a comparison with the results of a study of a similar scale by Selby and Kockelman (2013) where kriging models were estimated and the MAPE was calculated and found to be close to 60%, the difference in the magnitude of the accuracy can be attributed to a great extent to the inclusion of the betweenness - accessibility centrality measure. In the case of the study conducted by Lowry (2014) and for a community network, the reported in-sample MdAPE values of 28%, are slightly larger but still of similar overall magnitude with the current results.

Last, a comparison with the Swiss national model's[3] accuracy, which corresponds to the state-of-practice four-step model used for AADT estimation, reveals that the national model outperforms all the estimated models. In general, it has higher accuracy than the other models but at the same it has to be pointed out that the model has been calibrated against

3 ARE; National Transport Model (2010)

the employed count data while it also requires much more data and effort to set up at the first place. In addition, a potential source of introduced bias might have resulted from not accounting for international commuters which can lead to underestimation of AADT close to the borders, an aspect that the Swiss national model takes into account.

7.6 CONCLUSIONS

In this chapter a direct demand modelling approach for AADT prediction on a nationwide network was presented. It was exhibited that the construction of a variable that can account for interregional flows, such as the betweenness-accessibility centrality measure, can lead to a significant enhancement on the accuracy of the models. In addition to the already tested models in the literature, the spatial error model was estimated and it was shown that GWR and kriging models are more appropriate for interpolation purposes while spatial error and OLS models have the potential to be applied for forecasting purposes as well since their estimated parameters are unbiased and consistent.

Under this consideration, spatial error models can be utilized within a two-step statistical procedure to make statements about speed and AADT values on a link level. In particular, this procedure can account for their interdependent nature on top of other consideration such as spatial autocorrelation and is the focus of the next chapter of this dissertation.

Last, the developed methodology can be easily applied to different scales of network (*e.g.* urban cases), where a finer zonal analysis level can be employed. Moreover, it has the apparent advantage that it requires only publicly available socioeconomic data and can be easily tuned in to different networks (*e.g.* Open street map). Especially this dimension needs to be highlighted since in many cases, either due to data scarcity, or budgeting reasons, the development of four-step or agent-based models is not possible, while flow predictions are still needed to support various operational and policy related applications.

8

SPEED ESTIMATION WITH ENDOGENOUS DEMAND

8.1 INTRODUCTION

In a previous chapter of the dissertation the issue of speed modelling was discussed and the results of different linear regression models were evaluated. However and as mentioned earlier, the deployment of traditional linear regression models for tackling the issue of speed modelling can be perceived as problematic mainly due to two issues, namely of the spatial dependence and endogeneity issues which if present can lead to different undesired statistical shortcomings.

To this end, in chapter 6 a particular focus was given to the treatment of the former issue while endogeneity was treated indirectly by employing assumed exogenous variables as proxy variables of localized demand levels. Endogeneity arises due to the demand and supply interaction mechanism. Therefore, the utilization of actual traffic volume data instead of proxy variables within a regression model should be made in a way that ensures that the endogenous nature of traffic volume is taken into account. Driven by this, this chapter extends the previously presented approach by proposing an alternative that can address both spatial effects and endogeneity issues. An empirical case study is designed to demonstrate its application while in a similar manner as before the prediction results are compared against the output of a four-step model that corresponds to a state-of-the-practice regional transport planning model.

On that note, it is of interest to stress out that typically such models are only calibrated against volume estimates, and as a consequence they fail drastically to provide reasonable speed estimates. A discussion on that issue can be found in Dowling and Skabardonis (1993) while even though a considerate time has passed since then, there has been little evidence in the literature that actions have been taken to tackle this.

Essentially, this chapter constitutes the second part of the direct demand modelling approach that this dissertation develops. Therefore, when paired up with the previously presented volume regression model it can form a coherent direct demand modelling approach, suitable both for prediction and forecasting applications. Last, this chapter is based on Sarlas and Axhausen (2017) with a revised analysis, results and discussion in place.

8.2 METHODOLOGY

8.2.1 *Estimators*

As mentioned in the prior section, there are two considerations associated with the choice of a mean speed model. The first one relates to the endogenous character of volume in the speed model. More specifically, estimation by means of ordinary least-squares constitute the standard for linear regression models. However, in the case of endogenous variable, the main OLS assumption of uncorrelated error terms with the independent variables is violated (Wooldridge, 2012). As a consequence, this violation turns OLS to an inconsistent and biased estimator, and should be thoroughly tested and treated if present. Methodologically, this issue is dealt with by accounting for the endogeneity via utilizing instrumental variable(s) (IV), normally within a two-stage least-squares (2SLS), or a control function (CF) approach. Under specific conditions these two approaches are essentially the same. In both cases though, a strong prerequisite is that the chosen instruments must be uncorrelated with the error term but substantially correlated with the endogenous variable. This estimation approach allows to obtain consistent and unbiased parameter estimates which can then be applied for forecasting purposes.

Secondly, the main implication of modelling data of a spatial nature is the existence of spatial effects, thereby pointing to non-independent observations. The implications of that were already discussed in a previous chapter. In brief, the existence of spatial dependence bears the ability of leading to a violation of the independent and identically distributed (iid) assumption of OLS. Spatial simultaneous autoregressive (SAR) models constitute a modelling medium allowing to treat for this issue in two main ways, assuming different underlying mechanisms that generate the spatial dependence. On the one hand, when a spatial variable has been omitted from the model specification, the error terms tend to be spatially autocorrelated, creating a need of an error term that inherently considers this (spatial error model). On the other hand, when neighboring locations' response variable has an indirect effect on the response at the location, then the inclusion of a spatially lagged dependent variable can mitigate the spatial dependence issues, hence facilitate the estimation of explanatory variables' direct effects on the response variable (spatial lag model). A combined treatment of both aforementioned spatial dependencies is also possible within a model formulation (spatial Durbin model). The exact formulation of those models can be found in section 6.2.1.

The prevailing estimation approach of SAR models is by means of maximum likelihood. However, this entails a number of drawbacks, such as being computationally infeasible for large samples, and most importantly lacking the ability to account for the presence of endogenous regressor(s) and heteroscedastic disturbances. Kelejian and Prucha (1999) suggested a generalized method of moments (GMM) estimator, capable of addressing the former issue, and also paved the way for addressing the latter shortcoming as well. Admittedly, this estimator can be seen as a major breakthrough in the field of spatial econometrics.

More specifically, in a follow-up paper the same authors developed a methodology for accounting for unknown forms of heteroscedasticity in conjunction with an IV estimator for the parameters (Kelejian and Prucha, 2010). Later on, Drukker et al. (2012) extended their work by developing a two-step generalized method of moments and instrumental variable estimator (2IV/GMM), capable of treating for endogeneity issues and heteroscedastic innovations, in addition to a spatially lagged variable.

In summary, their estimator involves four steps. Initially, a two stage least-squares (2SLS) approach is applied to obtain the starting values of the parameters of the model (betas), similar to the traditional IV estimation approach. In the next step, a GMM estimator is applied to obtain the value of the autoregressive coefficient λ. The moment conditions are defined on the basis of conforming to the orthogonality assumption, imposing the independence of the residuals with their first and second order neighbors' counterparts. In the third step, a generalized spatial two-stage least-squares estimator is applied on a Cochrane-Orcutt transformed model to obtain the new values of betas along with the residuals. In the final step, the residuals from the previous step are utilized within a GMM estimator to obtain the true value of λ, imposing the same moment conditions as before.

Besides spatial dependence, spatial heterogeneity is another potential cause for spatial autocorrelation. To that end, GWR has been extensively applied for exploring the existence of spatially varying relationships. However, to date no study has jointly addressed the issue of spatial heterogeneity and endogeneity, at least to the best of author's knowledge. From a methodological point of view, GWR estimation takes place using a weighted least squares approach which is similar in principle to the simple OLS one.

Thereupon, the problem can be formulated as a control function approach where in the first stage the endogenous variable is regressed on the set of all exogenous variables, including the instruments, based on an OLS estimator. This formulation of the endogenous variable is referred to as reduced form. Subsequently, the reduced form residuals are plugged into the main regression equation, including the endogenous variable, and the estimation then takes place using the normal weighted least squares approach. In essence, the approach entails the inclusion of the endogenous part of the independent variable as a standalone variable in the regression in order to ensure that the endogenous variable is truly independent with the residuals. Of course, spatial heterogeneity can be addressed also within the reduced form regression by replacing the OLS estimator with a weighted least squares one. More information on the control function can be found in Wooldridge (2011), while the previously described GWR estimator is henceforth denoted as CF-GWR.

8.2.2 *Model formulation*

Previous attempts to model average speed (e.g. Sarlas et al., 2015; Hackney et al., 2007) have resorted to the use of proxy variables for the traffic volume, operationalized in the form of spatial density values of various socio-demographic variables (*e.g.* population, employment). Yet, and as discussed in a previous chapter (chapter 7), such variables fail to capture the directionality and the complexities of the interregional demand, and thus can suffice only for small area cases. Based on this, traffic volume is instrumented on a set of variables capable of capturing the interregional demand aspects. To that end, the instrumentation can take place by utilizing the identified set of independent variables used for obtaining AADT predictions, given that the instruments comply with the strict exogeneity assumption.

In contrast to prior studies on the topic, a distinct approach is followed concerning the dependent variable. Specifically, and in line with the BPR functions' formulation, the dependent variable is specified as the mean travel time difference, defined as the difference between the free-flow travel time and the mean travel time. The apparent advantage of employing such a formulation is that it can better capture the relation between volume and speed. Furthermore, this way the problem can be easily transformed into a linear model by applying a logarithmic transformation, hence overcoming the non-linearity issues present. Subsequently, this choice also allows to exploit the 2IV/GMM and CF-GWR estimators, in order to account and treat both for endogeneity and spatial effects. More specifically, the model formulation is presented below:

$$tt = tt_0 + tt_0^{\alpha} \frac{AADT^{\beta}}{Capacity^{\gamma}} \tag{8.1}$$

$$log(tt - tt_0) = \alpha log(tt_0) + \beta log(AADT) - \gamma log(Capacity) \tag{8.2}$$

From a conceptual viewpoint , the various forms of BPR functions attempt to model the interaction between demand and supply on a link level by incorporating various congestion functions. The majority of those quantify the congestion as a function of the volume to capacity ratio. Capacity values are normally calculated on the basis of standardized values coming from the Highway Capacity Manual (2010). A viable alternative is to employ a set of typical explanatory variables as proxy ones for the capacity, say number of link types, legal speed limit, *etc.* Consequently, the need of a priori estimating capacity values diminishes. Especially, if we factor in the fact that for the case of national planning models, which involve some inherent degree of abstraction, a link might actually correspond to a set of links in reality, therefore determining a single capacity value for a set of links with varying attributes can be challenging and potentially troublesome.

8.3 CASE STUDY

An empirical case study is designed and conducted in order to model mean speed values in a set up that resembles a regional planning model's configuration. Particularly, the network of Switzerland is exploited as the study network in the form that is present in a state-of-the-practice transport model[1] (NTM). The network consists of approximately 40'000 directed links, while the links are classified into four hierarchical types, namely highway, major road, rural main road, and urban arterial road. In brief, the case study is identical to the one described in chapter 7. Two independent sources of volume and speed data are utilized to facilitate the construction of the observations' database. At first, the previously described traffic

1 ARE; National Transport Model (2010): A 4-step model, implemented in VISUM

volume observations in the form of AADT values are employed for that purpose (section 7.3).

On the speed front, a commercial floating car data source that has emerged in the last years is utilized. Tom-Tom provides historical travel time databases for the entire network of Switzerland, including daily mean speed estimates (see section 6.3 for a description). The acquired data correspond also to the basis year, however they are reported on a navigational network, which is considerably more detailed than the study network (1.5 million links). Naturally, the need for matching the two network arises in order to be able to integrate the historical travel time observations to the study network.

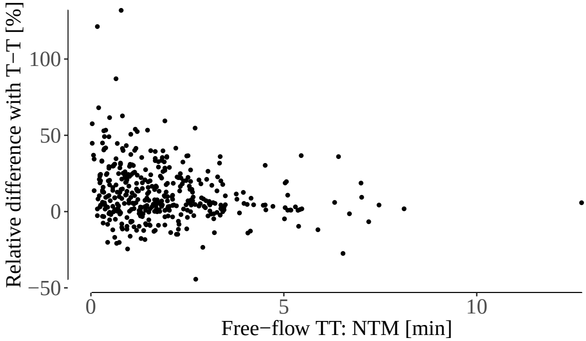
8.3.1 *Network Matching*

The network matching is facilitated by developing an automated procedure that incorporates an adaptive radius search, operationalizing an edge matching approach. In summary, the matching is established based on two assumptions. First, the nodes of the study network should correspond, or at least be nearby to the actual nodes. Second, the links in the study network can be perceived as paths, meaning that their reported length should correspond to the total length of the links composing the path.

At the first step, a circle with a 50 meters radius is drawn around each node of the study network in order to identify the Tom-Tom nodes that lie in the encompassed area (matched nodes). In the second step, for each pair of starting and ending nodes of the study network's links, the shortest path between all pairs of matched nodes are identified. Subsequently, the path with the lowest absolute deviation from the study network's link length is chosen as the most probable one. However, if the deviation is higher than a threshold value (set to 2%), or a path is not identified at all, then the radius increases and the procedure starts anew. The radius increase happens in increments of 50 meters and it continues up to a maximum distance of 300 meters, allowing for a maximum number of 6 iterations.

Nonetheless, the incompatibility of the two networks surfaces in various instances (*e.g.* presence of a node where no actual intersection exists), giving rise to erroneous matching. As a remedy, the matched paths are checked visually to conclude on whether or not a correct identification is achieved. For the given problem at hand involving 790 links, the accuracy of the developed matching routine is found to be close to 70%. For the

FIGURE 8.1: Relative free-flow travel time difference between Tom-Tom and NTM networks



remaining cases, the matching is conducted manually to ensure that no systematic error is introduced due to mismatching reasons.

Interestingly, a comparison of the common attributes between the two networks reveals that in many cases the values of attributes such as free flow speed, speed limit, number of lanes, *etc.* ,are not found to be aligned. This is explained by the fact that in less detailed networks a single link can essentially correspond to a number of links with varying attributes.

As a remedy and in order to mitigate the impact of a wrong specification on the attributes side, the length weighted attributes of the links forming each path are adopted as the corresponding attributes of the study network's links. An exception is made on the length attribute which is assumed to be correctly specified in the study network. In the case of free-flow travel time, calculated based on the posted speed limits per case, the mean relative difference between the two is found to be 13.30%, with a standard deviation of 19.70%, supporting the argument that less detailed networks have a higher degree of abstraction on their attributes specification, which can be a potential source of error. The relative difference between the corresponding free-flow travel times between the Tom-Tom and the NTM is given in figure 8.1.

8.3.2 *Explanatory Variables*

Having as an objective to model mean speed values on a nationwide network, a set of explanatory variables need to be included in the model specification, either directly if they comply with the exogeneity assumption, or indirectly as instruments of the traffic volume. Before proceeding further, a closer look at the phenomenon we are aiming to model can provide valuable insights. As mentioned before, mean speed is the outcome of the interaction between the two interrelated mechanisms of supply and demand.

On the supply side, link characteristics associated with the design and the operation aspects are the main determinants of the link's capacity. Therefore, variables such as the link's hierarchy type (*e.g.* highway, *etc.*), the free-flow travel time and speed, determine to a large extent the capacity. Variables such as curveness, and link type (tunnel or not) are expected to have a two-way impact on the capacity. At first, indirectly through affecting the actual free flow speed values, and directly either by affecting the driving behavior (*e.g.* more alert drivers), or by the existence of driving restrictions (*e.g.* prohibited overtaking).

On the demand side, variables such as the population density and the potential number of persons passing through (betweenness-accessibility centrality) can clearly be identified as the main determinants of the travel demand. Variables associated with the network design (*e.g.* betweenness centrality) are also expected to exert some influence on the demand. Variables such as the population density can be considered that they bear the ability of capturing the character of the surrounding area, and thus of different demand aspects. Thereupon, the instrumentation can take place by utilizing the identified set of independent variables used for obtaining AADT predictions in a previous chapter, under the assumption which needs to be thoroughly tested that the instruments comply with the strict exogeneity assumption.

Clearly, supply and demand are interrelated. Nevertheless, in the case of a regional planning network we can assume that the interaction between demand and supply on a link level, is not affecting the demand. The summary statistics of the different employed variables are presented in table 8.1. Last, it should be noted that the estimation of the models takes place by making use of the R packages "AER" (Kleiber and Zeileis, 2008), "sphet" (Piras et al., 2010), and "spgwr" (Bivand et al., 2017).

TABLE 8.1: Summary statistics of employed variables

Statistic	Mean	St. Dev.
Travel time difference [sec]	5.93	8.83
Free flow travel time [sec]	122.65	99.36
Speed limit [km/h]	79.46	24.22
Curvedness [degrees]	0.04	0.07
Link's tunnel share	0.15	0.26
AADT [veh/day]	14,431.55	14,155.99
Freeway-Highway [dummy]	0.46	-
Major road [dummy]	0.31	-
Rural major road [dummy]	0.20	-
Urban arterial road [dummy]	0.03	-
2-lane road [dummy]	0.42	0.49
Population density: 10 km* [res/km^2]	575.92	634.52
BC	8,411,925.00	13,277,549.00
BAC_1 [generic:average (car)[1]]	9,211.05	11,021.13
BAC_1/BAC_2	0.45	0.16

Note: *=kernel weighted
1: with mode choice considerations, table 4.3

8.4 MODEL ESTIMATION

8.4.1 *OLS model*

Having identified the set of potential variables, the next step concerns the model estimation. In total, our sample consists of 395 links. It should be mentioned that for each count location with bidirectional traffic, only one of the two directions is randomly chosen and included in the sample. This choice is made because the available AADT data are reported per location, and not per link (see section 7.3).

At first, an OLS model is estimated in order to highlight the bias of the parameter estimates when failing to account for the endogeneity and the spatial dependence issues. In a similar manner as before, variance inflation factors (VIF) are utilized for ensuring that no multicollinearity issues are present. However, due to the lack of a constant term in the regression model, variance inflation factors are not yielding correct values. Nevertheless and in order to test for multicollinearity, an equivalent model but with a constant term in place (which is found to be statistically insignificant, hence its exclusion from the specification) is estimated and tested for multicollinearity. In particular, the corresponding model estimates indicate that no such issues exist (VIF smaller than 5). To mitigate the impact of outliers (which can alter substantially the estimates), 8 observations with high leverage (higher than 6% according to Cook's distance diagnostic (Cook and Weisberg, 1982)) are excluded. The OLS parameter estimates are reported in table 8.2.

8.4.2 *Instrumental variables model*

The next step concerns the treatment of the endogeneity. At the outset, an IV model is estimated by means of a 2SLS estimator to account for the endogeneity of AADT. The 2SLS model serves as the benchmark model for checking for the presence of spatial autocorrelation, hence drawing conclusions on the need to utilize the aforementioned 2IV/GMM and GWR estimators.

The variables presented in the second part of table 8.1 are chosen as instruments, whereas a bigger set of variables was thoroughly tested as well on their ability to serve as instruments. A number of statistical tests is performed in order to conclude on the presence of endogeneity, and on the ability of the instruments to comply with the prerequisites. At first, a weak instruments test is performed through the formation of an F-test on the instruments. More specifically, the null hypothesis of weak instruments is rejected with a lower than 0.1% p-value.

The presence of endogeneity is checked with the Wu-Hausman test (DeMin, 1973; Hausman, 1978) and the null hypothesis of no endogeneity is rejected at the 5% level. Last, the validity of the instruments is tested with the Sargan test (Sargan, 1958). The null hypothesis of the instruments validity (exogenous) fails to be rejected at any of the examined levels. In summary, the performed statistical endogeneity tests demonstrate clearly that AADT is indeed endogenous, while the chosen instruments are found to be statistically valid. The existence of multicollinearity issues is checked with variance inflation factors, and found not be the case. Last, the presence of simultaneity bias between speed and AADT is tested by formulating an AADT model and instrumenting the speed. The results validate our hypothesis that for the current setting there is only a one-way endogeneity issue. Nonetheless, if that was not the case the estimation of a structural equation model with spatial considerations would constitute the appropriate way to tackle the problem.

Last, a major problem associated with the employment of IV approaches lies on the fact that the standard error estimates are biased due to finite sampling issues. A discussion on the specifics of the problem is given in Stock et al. (2002); Camponovo and Otsu (2015). To correct for that, the standard errors are estimated based on a bootstrapping approach that involves 500 iterations with random re-sampling with replacement. Essentially, this approach serves the purpose of approximating the distribution of the parameter estimates, and hence of the standard errors and the statistical significance levels. The followed bootstrapping procedure is described in Efron (1979). The 2SLS model is reported in table 8.2.

As it can be seen in table 8.2, the estimated parameters are in line with our expectations about how the different variables affect the travel time difference. All the variables are found to be statistically significant at different levels. A comparison with the OLS estimates highlights the existence of substantial differences. Interestingly, the impact of the endogenous variable (AADTT) is much lower when accounting for endogeneity.

8.4.3 *Spatial instrumental variables models*

In order to check for the presence of spatial autocorrelation on the residuals of the 2SLS model, and justify the choice to proceed to the estimation of the 2IV/GMM model, Moran's I measure is utilized. More specifically, different spatial matrices variants are constructed and tested, based on Euclidean and network distances. In the case of the latter, two distance metrics are employed, the free-flow travel time and the network traveled distance. In brief, the variant with a neighborhood defined on the basis of network free-flow travel time of up to 5 minutes is found to be the optimum one based on two set of criteria (goodness-of-fit measures and Moran's I measure).

Finally, the last part of the construction of the spatial weight matrices concerns the determination of the weights that should be assigned to the neighboring locations. Based on two set of criteria (goodness-of-fit measures and Moran's I measure), an inverse distance metric is concluded to be the most appropriate to capture the spatial structure. Moreover, in order to avoid having misspecification issues as those highlighted in Kelejian and Prucha (2010), a so-called min-max normalization of the weights

Table 8.2: Speed models' estimates

	Dependent variable: Log(tt-tt_0)		
Regressor	(OLS)	(2SLS)	(2IV/GMM)
Log(Speed limit)	-1.44^{***}	-1.21^{***}	-1.21^{***}
	(0.12)	(0.13)	(0.13)
Log(AADT)	0.31^{***}	0.21^{***}	0.20^{***}
	(0.05)	(0.06)	(0.05)
Urban arterial road	0.50^{*}	0.60^{**}	0.61^{**}
	(0.21)	(0.22)	(0.20)
Rural major road	0.36^{***}	0.27^{**}	0.24^{*}
	(0.09)	(0.09)	(0.09)
Major road	0.26^{***}	0.25^{***}	0.24^{***}
	(0.07)	(0.07)	(0.07)
Log(fftt)	1.03^{***}	1.01^{***}	1.03^{***}
	(0.04)	(0.04)	(0.04)
Curvedness	-1.32^{*}	-1.34^{**}	-1.22^{**}
	(0.56)	(0.51)	(0.56)
Link's tunnel share	-0.35^{+}	-0.36^{+}	-0.31
	(0.20)	(0.19)	(0.21)
λ			0.55^{**}
Observations			387
Adjusted R^2	0.87	0.87	-
Moran's *I*	0.13^{***}	0.14^{***}	-0.03
df	379	379	378
Weak instruments (*df1*,*df2*)	-	(5,375) 107.18^{***}	(5,375) 107.18^{***}
Wu-Hausman (*df1*,*df2*)	-	(1,378) 9.81^{**}	(1,377) 8.278^{**}
Sargan (*df*)	-	(4) 1.54	(4) 1.50

+ $p<0.1$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

OLS: ( ) Heterosc. corrected std. errors

2SLS: ( ) Bootstrapped std. errors

2IV/GMM : ( ) Heterosc. corrected & bootstrapped std. errors is applied. Among the tested spatial weighting schemes, the one based on the free-flow travel time is concluded to be the most pertinent one with a neighborhood extent of 5 minutes. The calculated Moran's I measure for this spatial matrix indicates that spatial autocorrelation is statistically significant with a value of 0.14 (table 8.2). Therefore, the presence of spatial effects has to be treated by using the 2IV/GMM estimator.

Initially, a model with the spatial Durbin formulation is estimated. However, the spatial autocorrelation ρ parameter is found to be statistically insignificant while this is not the case for the spatial autoregressive parameter. This finding indicates that a spatial error formulation should be adopted, dropping the spatially lagged dependent variable. In addition, it points towards the case of omitted spatial variable(s) as the underlying source of dependence. The new parameter estimates differ slightly in comparison to the previous ones. The results of the spatial error model with endogenous AADT, and heteroscedasticity corrected standard errors, are also presented in table 8.2. On the endogeneity diagnostics front, the Wu-Hausman and the Sargan tests have to be modified accordingly to take into consideration the errors of the spatial error model. The results of the new versions of the tests validate our prior results on the endogeneity presence and the exogeneity of the chosen instruments.

In a different manner, a second treatment of the spatial effects takes place by utilizing the CF-GWR estimator. More specifically, the model estimates are reported in table 8.3 while the model is estimated with an adaptive radius bandwidth in place that is identified on the basis of minimizing the AIC of the model, and hence ensuring maximum goodness-of-fit values. In this regard, 89% of the total observations are taken into account for the estimation of the localized paramater estimates for each location. Nevertheless, the low variance of the reported parameter estimates reveals that spatial heterogeneity is not present, or at least not to a large extent.

8.5 PREDICTIVE ACCURACY

Finally, the predictive performance of the estimated models is evaluated in order to draw solid conclusion with respect to their capability to be applied for speed prediction purposes. Furthermore, the results are also compared against the calibrated output of a four-step model[2] (NTM). Thereof, the predictive performance is evaluated in terms of the different accuracy measures, as defined in a previous section (formulas (7.4-7.8)).

2 ARE; National Transport Model (2010): A 4-step model, implemented in VISUM

TABLE 8.3: GWR estimates

Statistic	Min	1^{st} Quart.	Median	3^{rd} Quart.	Max	Global
Log(Speed limit)	−1.23	−1.22	−1.22	−1.20	−1.18	−1.21
Log(AADT)	0.20	0.21	0.22	0.22	0.22	0.21
Urban arterial road	0.50	0.53	0.58	0.61	0.64	0.60
Rural major road	0.23	0.23	0.25	0.26	0.29	0.27
Major road	0.22	0.23	0.23	0.24	0.26	0.25
Log(fftt)	1.00	1.01	1.01	1.01	1.02	1.01
Curvedness	−1.38	−1.36	−1.34	−1.33	−1.31	−1.34
Link's tunnel share	−0.44	−0.43	−0.41	−0.38	−0.32	−0.36
RF residuals	0.26	0.27	0.29	0.30	0.31	0.28
Adaptive radius=89.15% of the observations						

More specifically, besides measuring the accuracy of predicting actual travel time differences, the predictions of absolute mean speed is of apparent interest. It should be noted that given the log transformation of the dependent variable, when back transforming to the original scale we account for the fact that the model predicts the geometric mean instead of the arithmetic one, in the way suggested by Wooldridge (2012). Omitting this correction will give rise to systematic underestimation problems.

In a similar manner as before, the evaluation takes place both for in-sample and out-of-sample predictions. More specifically, the out-of-sample evaluation employs a leave-one-out strategy, including a number of 1000 replications. Furthermore, an important aspect of the evaluation concerns the endogenous variable itself. It is reminded that the overall objective of this dissertation is to form a coherent direct demand modelling framework capable of providing both speed and volume localized predictions. In this regard, the evaluation happens under the assumption that only predicted values of AADT exist, since this constitutes a choice that resembles the envisaged application of the developed methodology. Therefore, the AADT predictions based on a spatial error model, as presented in section 7.4, are exploited for that purpose. The calculated predictive accuracy measures are presented in tables 8.4 and 8.5, for the in- and out-of-sample cases accordingly.

The results clearly demonstrate that the estimated statistical models outperform the 4-step model, while among the estimated models the spatial ones exhibit the highest predictive accuracy for both cases and samples. CF-

TABLE 8.4: In-sample predictive accuracy

Model	MdAPE	MAPE	MPE	MSE	SMAPE
Travel time difference					
OLS	39.83	72.14	47.56	14.36	11.67
2SLS	41.45	72.66	49.02	14.12	11.70
2IV/GMM	39.49	70.57	45.50	13.91	11.61
CF-GWR	38.71	68.22	42.05	14.04	11.51
NTM	98.63	143.30	50.98	184.92	29.21
Mean speed					
OLS	1.64	2.22	-0.09	3.94	0.55
2SLS	1.69	2.21	-0.09	3.97	0.55
2IV/GMM	1.62	2.19	0.03	3.97	0.54
CF-GWR	1.57	2.16	0.13	3.91	0.54
NTM	12.26	17.00	10.18	189.52	3.91

TABLE 8.5: Out-of-sample predictive accuracy

Model	MdAPE	MAPE	MPE	MSE	SMAPE
Travel time difference					
OLS	40.92	74.10	48.24	15.69	11.95
2SLS	41.72	74.60	49.57	15.44	11.98
2IV/GMM	40.93	73.14	47.09	15.29	11.93
CF-GWR	40.27	70.81	43.19	15.48	11.85
NTM	98.03	143.15	51.28	185.61	29.14
Mean speed					
OLS	1.66	2.30	0.08	4.24	0.57
2SLS	1.69	2.29	-0.08	4.28	0.57
2IV/GMM	1.65	2.28	0.00	4.26	0.57
CF-GWR	1.60	2.26	0.13	4.26	0.56
NTM	12.23	17.02	10.15	189.95	3.92

GWR yields slightly better results in all cases while it is clearly highlighted that the sole treatment of endogeneity without accounting for spatial dependence issues lacks the ability of improving the prediction accuracy, in comparison to the OLS model. In the case of the symmetric mean absolute error, a measure which is less influenced by the presence of outliers, the magnitude of the measure is substantially lower for all cases, but still the statistical models outperform the 4-step model by more than a factor of two. This extremely high difference can be attributed to a large extent to the wrongly specified free-flow speed values of the latter, but nevertheless the results demonstrate that a simpler in conception model can provide much more reasonable predictions. In addition, one has to remember that the 4-step model is typically calibrated against volumes and not against speeds.

8.6 CONCLUSIONS

In the present chapter a methodology to estimate mean speed values on a large scale network was presented, treating both the endogeneity and spatial effects aspects. These two aspects are quite often acknowledged in empirical studies but the simultaneous treatment of both is rather rare in the literature, especially in the transport relevant one. Methodologically, the well-known GWR estimator was extended in a simple way by building upon the control function approach to account for the presence of endogenous variable, hence providing a way to address both spatial heterogeneity and endogeneity aspects.

On the instrumentation front, a particular focus was given on the selection process of the instruments to allow obtaining consistent and unbiased parameter estimates, thus making the model capable both for prediction and forecasting applications. To this end, the modification of the relevant statistical tests necessitated in order to make them suitable for the case of the 2IV/GMM estimator. In conclusion, the findings suggest that a correctly specified statistical model has the ability to provide accurate estimates, outperforming a much more complex and data demanding transport planning model, even though the superiority of such models is taken for granted in many cases.

Nonetheless, the low predictive accuracy of the 4-step model is alarming and it raises some well-founded concerns regarding the reasonableness of the speed predictions of such models, an aspect which is normally neglected in the calibration processes. Taking into account that such models normally constitute the medium for the evaluation of different policies and projects on a national level, the implications of unreasonable speed predictions can be rather huge.

In addition, the developed modelling approach when coupled with an AADT regression model, as the one proposed in chapter 7, it forms a coherent direct demand modelling approach which makes use only of aggregated data. A direct demand modelling approach has the apparent advantage that it can be set up within a short time frame with very low associated computational, maintenance, and monetary costs, while it can still provide the required answers for a number of transport planning problems.

9

DISCUSSION - CONCLUSIONS

9.1 OUTLOOK

The main objective of this dissertation was the formulation of a direct demand modelling approach capable of making statements about the mean speed and volume at each location of a network. In this regard, a spatial perspective was adopted in order to enhance the performance of such models. More specifically, the main hypothesis that this dissertation put into testing was that space matters when it comes to the formation of transport (direct) demand models. To this end, this dissertation evolved along two main directions. The first direction concerned the quantification of demand based on spatial interaction and graph theory concepts. The second direction focused on the existence of spatial effects, accounting appropriately for them in the estimation process in order to avoid giving rise to statistical shortcomings.

9.1.1 *Demand aspects*

The first direction focused on the demand side of the problem and in particular on how a quantification of the potential demand levels at various locations of a network can take shape. Even though previous studies had resorted to the use of various spatial density variables as proxy variables of demand, this choice entails a number of limitations. For instance, the demand on locations further away from residential zones but still along corridors of high inter-regional demand was found to be systematically underestimated. This realization led to the need for constructing variables capable of capturing the spatial interaction aspects in a more elaborate but still direct way. To this end, the accessibility measures were exploited since they provide a convenient way of measuring the potential for interaction between locations.

The overview of the accessibility literature revealed three major limitations with respect to the formulation of gravity-based measures. In brief, these are related to the estimation of the involved interaction intensity function parameter(s), which transpired as a crucial issue. As a remedy, an alternative way of specifying the interaction intensity function was proposed. More specifically, its definition was revisited and formulated as a survival analysis concept which can be specified both in a data-driven and model-based manner, depending on the analysis objective.

From a different point of view, accessibility measures can be perceived as a by-product of spatial interaction models. However, even though the interlinked nature of those two measures has been acknowledged in the literature for a long time (*e.g.* see distance decay debate 2.2.5), a few approaches have considered this aspect, either directly (*e.g.* competition effects accessibility measures), or indirectly (*e.g.* commuting duration modelling). Nonetheless, the proposed model-based specification of the interaction intensity function sets the stage for treating these two sides a the same problem. More specifically, the results of the case study attested to the ability of the proposed interaction function variants to produce reasonable and realistic predictions of the potential interaction space. To that end, the ability to account for location and/or individual characteristics within a gravity-based accessibility measure constitutes an aspect which was not addressed in the literature to date. Furthermore, this allows to examine different dimensions of accessibility, as demonstrated with the construction of relative indicators.

In addition, a new spatial interaction framework was introduced that originated from the accessibility concept itself. A nationwide case study was designed to illustrate its application while the results highlighted that it has the competence to be employed for examining various aspects associated with spatial interaction phenomena such as commuting and in general access to different opportunities. Nevertheless, the framework still requires to undergo further testing and development before drawing definite conclusions for its predictive accuracy and value.

Finally, a new indicator, called betweenness-accessibility, that combines the concepts of centrality and gravity-based accessibility in a unified measure was introduced. Subsequently, this indicator was utilized for enhancing the predictive accuracy of the developed model, both for volume and speed purposes. The hypotheses that such a variable can capture the directionality and magnitude of demand were validated. Moreover, the indicator allows to investigate the incidental impacts of accessibility on a network and provides a richer picture of the ways a transportation system operates to generate connectivity. This dimension of accessibility and especially how it is jointly generated by the transportation system and the landscape of opportunities, constitutes an important aspect which was not acknowledged in the literature to date. In conclusion, the value of the newly introduced indicator, especially on its general weighted version, can potentially extend beyond the scope of transportation research as it can pave the way for examining different aspects of various kinds of networks (*e.g.* social) where interaction among network elements happens in a disproportional way.

9.1.2 *Spatial effects*

The second dimension of this dissertation dealt with the question of what is related with what in space and what implications does this relationship entail from an econometric point of view. In brief, the framework was developed along the lines of accounting for both the spatial autocorrelation and for the interdependence issues that arise due to the nature of the modelled phenomena, while a number of methodological (*e.g.* suitability of estimators) and computational challenges (*e.g.* convergence) arose during the process.

On the spatial dependence front, the family of spatial autoregressive models was exploited where the focus was centered on three key aspects associated with their application. At first the issue of identifying the "true" spatial weight matrix W was investigated. In particular, this was facilitated based on the development of data-driven routines that utilize various goodness-of-fit and spatial autocorrelation metrics to uncover the optimum W per case. The relevant results highlighted that the analysis scale along with the spatial sparsity of observations are main determinants of that. In addition, the different case studies clearly point out that in the case of networked systems, autocorrelation is best described in terms of network distances.

The second aspect concerned the identification of the underlying cause of spatial autocorrelation and particularly of what type of spatial process gives rise to such issues. This issue was tackled by forming and testing hypotheses about the statistical significance of the relevant autoregressive parameters by utilizing different diagnostics (*e.g* Lagrange multiplier tests) and examining the goodness-of-fit and the predictive accuracy of the different estimated models. In conclusion, the issue of spatial sparsity also emerged as an important determinant while the results pointed towards the existence of spatially omitted variables (spatial error models) for the case of the developed direct demand modelling framework. This finding came as no surprise since, by definition, direct demand models are simplistic in nature. Even though different ways of quantifying the potential demand levels were introduced, still they cannot capture the full magnitude and complexities of demand. Therefore, it is presumed that this is one of the main underlying causes of spatial autocorrelation for this specific problem.

The third aspect revolved around the issue of the endogeneity. In particular, maximum likelihood estimators are found to be lacking the ability to handle such cases. Thereupon, the application of more elaborate estimators necessitated to account for the endogenous character of demand with respect to congestion, and subsequently of speed values.

On the spatial heterogeneity front, the application of GWR was demonstrated both for AADT and speed predictions purposes. Moreover, the GWR estimator was modified along the lines of a control function to account for endogeneity issues. In general, the GWR models yielded the best results among the spatial regression models in terms of predictive accuracy. However, it should be noted that GWR has limited abilities when it comes to the task of forecasting since the parameter estimates are heavily dependent on the identified adaptive radius bandwidth. Therefore, GWR can be summarized as a powerful spatial analysis tool that can be utilized for exhibiting trends and revealing spatial patterns that can point to the omission of certain explanatory variables, and its value was found to be high for interpolation purposes. In this respect, GWR was applied for testing whether or not the presumed causality aspects vary over space. Interestingly, in the case of AADT model the spatial variance of the parameters was found to be substantially higher than for the case of the speed model. This finding can also be interpreted in a similar manner as the choice of the SAR model formulation. More specifically, it is presumed that demand estimation has higher inherent uncertainty, which as a result appeared in the form of spatial heterogeneity on the relevant parameter estimates.

9.2 CONCLUSIONS

A comparison of the results of the developed direct demand modelling approach against the output of a traditional four-step model showcased that direct demand models can constitute a trustworthy alternative to more advanced and highly data demanding approaches. More specifically, its out-of-sample AADT predictive accuracy was found to be below 15% in terms of mean percent error, while in terms of speed values the predictions outperformed by far the output of a calibrated four-step model. Conceptually, it is arguable that a simplified approach cannot exhibit the predictive accuracy and the sensitivity of the prevailing approaches (four-step or agent-based models). However and as it was shown, accounting for spatial dependence and endogeneity issues, along with constructing variables that can quantify the peculiarities of transport demand, allows for the formulation of a direct demand model with the capacity to provide the required answers. In conclusion, a model as such can constitute a viable alternative in cases where either due to data availability issues, or due to various limitations (*e.g.* financial), the development of more advanced models is not possible.

Nevertheless, it should be stressed out that the effectiveness of a number of policies might not be able to be addressed in the context of a direct demand model (*e.g.* congestion pricing), at least not in a direct way but perhaps through the incorporation of spatially resolved (endogenous) variables that constitute the causal outcome of such policies (*e.g.* reduction of car ownership). Besides the above, the value of the developed approach is not only as a competing alternative to existing methods but it can also serve to point directions regarding the importance of accounting properly for the existence of spatial effects on various transport related modelled processes. In this regard, it can provide insights on how the existing approaches can be benefited by adopting a spatial perspective.

In line with this, the potential to contribute to the further improvement of the existing models needs to be investigated and evaluated (*e.g.* facilitate a quicker convergence by setting the initial values in iterative processes). Last, the simplicity of direct demand models allows them to be incorporated within a land use transport interaction models framework and it needs to be examined how the developed direct demand model can contribute to the advance of such models. An initial idea of the potential advantages of such integration is given by Zeiler et al. (2014).

9.3 FUTURE WORK

The results presented in this dissertation highlight the importance of space when it comes to the formulation of a direct demand model for volume and speed purposes. To this end, a number of interesting research questions and ideas about potential future applications have emerged that were outside the scope of this dissertation. At first, it would be of high interest to test the performance of the developed approach in the context of a cost-benefit analysis. More specifically, upon the availability of a given set of actual public projects, conducting a cost-benefit analysis based on the output of different modelling approaches would allow for a quantification in terms of resulted cost and benefits, and hence of project rankings. Subsequently and also taking into account the involved costs of setting up each model, recommendations can be made with respect to the strengths and limitations of each approach.

The application of the developed direct demand modelling approach for the case of urban networks would also be worthwhile to be investigated. Furthermore, this thesis focused only on the prediction of mean daily values. Future work could also explore the temporal dimension of this problem and in particular how this translates into the construction of the optimum spatial weight matrix, which should be heavily affected by the evolution of the demand patterns throughout the day. On this note, another important question that arises is the issue of the endogenous spatial weight matrices, which if it is indeed the case can have serious statistical implications (e.g. Kelejian and Piras, 2014).

Furthermore, and as mentioned before, the introduced spatial interaction framework based on the survival analysis concept requires further testing and development before drawing strong conclusions with respect to its predictive accuracy and value. In the same spirit, the specification way of the interaction intensity functions needs to be refined in order to address potential spatial dependence issues. To date though, survival analysis models cannot accommodate such issues. Last, the formation of the betweenness-accessibility centrality indicator sets the stage for exploring the function of different network structures where interaction among network elements happens in a disproportional way. This may as well constitute the object of future directions.

BIBLIOGRAPHY

Aguero-Valverde, J. and P. P. Jovanis (2010) Spatial correlation in multilevel crash frequency models, *Transportation Research Record: Journal of the Transportation Research Board*, **2165**, 21–32.

Akaike, H. (1974) A new look at the statistical model identification, *Selected Papers of Hirotugu Akaike*, 215–222.

Akbarzadeh, M., S. Memarmontazerin, S. Derrible and S. F. S. Reihani (2017) The role of travel demand and network centrality on the connectivity and resilience of an urban street system, *Transportation*, 1–15.

Allen, J. and S. Farber (2019) A measure of competitive access to destinations for comparing across multiple study regions, *Geographical Analysis*, 1–18.

Anastasopoulos, P. C., M. B. Islam, D. Perperidou and M. G. Karlaftis (2012) Hazard-based analysis of travel distance in urban environments: longitudinal data approach, *Journal of Urban Planning and Development*, **138** (1) 53–61.

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*, Kluwer, Dordrecht.

Anselin, L. (2001) A companion to theoretical econometrics, *Spatial econometrics*, 310–330.

Anselin, L. (2002) Under the hood Issues in the specification and interpretation of spatial regression models, *Agricultural Economics*, **27** (3) 247–267.

Anselin, L., A. Bera, R. Florax and M. Yoon (1996) Simple diagnostic tests for spatial dependence, *Regional science and urban economics*, **26**, 77–104.

Axhausen, K. (2008) Accessibility: long-term perspectives, *Journal of Transport and Land Use*, **1** (2) 5–22.

Axhausen, K. W. (1998) Can we ever obtain the data we would like to have, *Theoretical foundations of travel choice modeling*, 305–323.

Axisa, J. J., D. M. Scott and K. B. Newbold (2012) Factors influencing commute distance: a case study of Toronto's commuter shed, *Journal of Transport Geography*, **24**, 123–129.

Barabási, A.-L. and R. Albert (1999) Emergence of scaling in random networks, *science*, **286** (5439) 509–512.

Barbosa, H., M. Barthelemy, G. Ghoshal, C. R. James, M. Lenormand, T. Louail, R. Menezes, J. J. Ramasco, F. Simini and M. Tomasini (2018) Human mobility: models and applications, *Physics Reports*, **734**, 1–74.

Ben-Akiva, M. E. and S. Lerman (1979) Disaggregate travel and mobility choice models and measures of accessibility, in D. A. Hensher and P. Stopher (eds.) *Behavioural travel modelling*, Croom Helm.

Berdica, K. (2002) An introduction to road vulnerability: What has been done, is done and should be done, *Transport Policy*, **9** (2) 117–127.

Bernard, M., J. Hackney and K. Axhausen (2006) Correlation of link travel speeds, paper presented at the *6th Swiss Transport Research Conference. Ascona, Switzerland*.

Bhat, C. R. and A. R. Pinjari (2007) Duration Modeling, in *Handbook of Transport Modelling*, 105–131, Emerald Group Publishing Limited.

Bivand, R., L. Anselin, O. Berke, A. Bernat, M. Carvalho, Y. Chun, C. Dormann, S. Dray, R. Halbersma, N. Lewin-Koh et al. (2011) spdep: Spatial dependence: weighting schemes, statistics and models.

Bivand, R., D. Yu, T. Nakaya, M.-A. Garcia-Lopez and M. R. Bivand (2017) Package 'spgwr', *R software package*.

Black, J. and M. Conroy (1977) Accessibility measures and the social evaluation of urban structure, *Environment and Planning A*, **9** (9) 1013–1031.

Black, W. R. (1992) Network autocorrelation in transport network and flow systems, *Geographical Analysis*, **24** (3) 207–222.

Borgatti, S. P. (2005) Centrality and network flow, *Social Networks*, **27** (1) 55–71.

Borgatti, S. P. and M. G. Everett (2006) A Graph-theoretic perspective on centrality, *Social Networks*, **28** (4) 466–484.

Boulter, P., I. McCrae and T. Barlow (2007) A review of instantaneous emission models for road vehicles, *Technical Report*.

Brandes, U. (2001) A faster algorithm for betweenness centrality, *Journal of Mathematical Sociology*, **25** (2) 163–177.

Brandes, U. (2008) On variants of shortest-path betweenness centrality and their generic computation, *Social Networks*, **30** (2) 136–145.

Brunsdon, C., A. S. Fotheringham and M. E. Charlton (1996) Geographically weighted regression: a method for exploring spatial nonstationarity, *Geographical analysis*, **28** (4) 281–298.

Burger, M., F. van Oort and G. J. Linders (2009) On the specification of the gravity model of trade: Zeros, excess zeros and zero-inflated estimation, *Spatial Economic Analysis*, **4** (2) 167–190.

Camponovo, L. and T. Otsu (2015) Robustness of bootstrap in instrumental variable regression, *Econometric Reviews*, **34** (3) 352–393.

Cascetta, E., A. Cartenì and M. Montanino (2016) A behavioral model of accessibility based on the number of available opportunities, *Journal of Transport Geography*, **51**, 45–58.

Casetti, E. (1972) Generating models by the expansion method: applications to geographical research, *Geographical analysis*, **4** (1) 81–91.

Chen, A., C. Yang, S. Kongsomsaksakul and M. Lee (2007) Network-based acessibility measures for vulnerability analysis of degradable transportation networks, *Networks and Spatial Economics*, **7** (3) 241–256.

Cheng, J. and L. Bertolini (2013) Measuring urban job accessibility with distance decay, competition and diversity, *Journal of Transport Geography*, **30**, 100–109.

Cheng, T., J. Haworth and J. Wang (2011) Spatio-temporal autocorrelation of road network data, *Journal of Geographical Systems*, **14** (4) 389–413.

Cheng, W., J.-H. Wang, G. Bryden, X. Ye and X. Jia (2013) An Examination of the Endogeneity of Speed Limits and Accident Counts in Crash Models, *Journal of Transportation Safety & Security*, **5** (4) 314–326.

Cheung, C. and J. Black (2008) A reappraisal of the intervening opportunities model of commuter behaviour, *Road & Transport Research: A Journal of Australian and New Zealand Research and Practice*, **17** (2) 3.

Cook, R. D. and S. Weisberg (1982) *Residuals and influence in regression*, New York: Chapman and Hall.

Crow, R. T., K. H. Young and T. Cooley (1973) Alternative demand functions for "abstract" transportation modes, *Transportation Research*, **7** (4) 335 – 354.

Csárdi, G. and T. Nepusz (2006) The igraph software package for complex network research, *InterJournal Complex Systems*, **1695**, 1695.

Cui, B., G. Boisjoly, A. El-Geneidy and D. Levinson (2019) Accessibility and the journey to work through the lens of equity, *Journal of Transport Geography*, **74**, 269–277.

Curry, L., D. A. Griffith and E. S. Sheppard (1975) Those gravity parameters again, *Regional Studies*, **9** (3) 289–296.

de Dios Ortuzar, J. and L. G. Willumsen (2011) *Modelling Transport, Fourth Edition*, John Wiley & Sons.

De-Min, W. (1973) Alternative tests of independence between stochastic regressors and disturbances, *Econometrica (pre-1986)*, **41** (4) 733.

de Oliveira, E. L., L. d. S. Portugal and W. P. Junior (2014) Determining critical links in a road network: Vulnerability and congestion indicators, *Procedia – Social and Behavioral Sciences*, **162**, 158–167.

Degenne, A. and M. Fors (1999) *Introducing Social Networks*, Introducing Statistical Methods, SAGE Publications, London.

Demšar, U., O. Špatenková and K. Virrantaus (2008) Identifying critical locations in a spatial network with graph theory, *Transactions in GIS*, **12** (1) 61–82.

Desyllas, J., E. Duxbury, J. Ward and A. Smith (2003) Pedestrian demand modelling of large cities: an applied example from london, *Centre for Advanced Spatial Analysis (UCL)*.

Domencich, T. A. and D. McFadden (1975) *Urban travel demand : a behavioral analysis*, Amsterdam : North-Holland Pub. Co. ; New York : American Elsevier.

Dowling, R. and A. Skabardonis (1993) Improving average travel speeds estimated by planning models, *Transportation Research Record*, 68–74.

Doxiadis, C. A. (1970) Ekistics, the science of human settlements, *Science*, **170** (3956) 393–404.

Drukker, D. M., P. Egger and I. R. Prucha (2012) On two-step estimation of a spatial autoregressive model with autoregressive disturbances and endogenous regressors, *Econometric Reviews*, **32** (5-6) 686–733.

Duan, Y. and F. Lu (2015) Robustness analysis of city road network at different granularities, in *Space-Time Integration in Geography and GIScience*, 127–143, Springer.

Duddu, V. and S. Pulugurtha (2013) Principle of demographic gravitation to estimate annual average daily traffic: Comparison of statistical and neural network models, *Journal of Transportation Engineering*, **139** (6) 585–595.

Efron, B. (1979) Bootstrap methods: Another look at the jackknife, *The Annals of Statistics*, **7** (1) 1–26.

Elhorst, J. P. (2014) *Spatial econometrics: from cross-sectional data to spatial panels*, Springer.

Eom, J., M. Park, T.-Y. Heo and L. Huntsinger (2006) Improving the prediction of annual average daily traffic for nonfreeway facilities by applying a spatial statistical method, *Transportation Research Record*, **1968** (1968) 20–29.

Ermagun, A., A. Samimi and T. H. Rashidi (2016) How Far Is Too Far?, *Transportation Research Record: Journal of the Transportation Research Board*, **2586**, 72–82.

Ewing, G. (1986) Spatial pattern in distance-deterrence parameters and Fotheringham's theory of competing destinations, *Environment and Planning A*, **18** (4) 547–551.

Flyvbjerg, B., M. K. Skamris Holm and S. L. Buhl (2005) How (in) accurate are demand forecasts in public works projects?: The case of transportation, *Journal of the American planning association*, **71** (2) 131–146.

Fotheringham, A. S. (1981) Spatial structure and distance-decay parameters, *Annals of the Association of American Geographers*, **71** (3) 425–436.

Fotheringham, A. S. (1983) A new set of spatial-interaction models: the theory of competing destinations, *Environment and Planning A*, **15** (1) 15–36.

Fotheringham, A. S. and M. E. O'Kelly (1989) *Spatial interaction models: formulations and applications*, Kluwer Academic Publishers Dordrecht.

Fotheringham, A. S. and M. J. Webber (1980) Spatial structure and the parameters of spatial interaction models, *Geographical Analysis*, **12** (1) 33–46.

Freeman, L. C. (1977) A set of measures of centrality based on betweenness, *Sociometry*, 35–41.

Freeman, L. C., D. Roeder and R. R. Mulholland (1979) Centrality in social networks: Ii. experimental results, *Social networks*, **2** (2) 119–141.

Frost, M. and N. Spence (1995) The rediscovery of accessibility and economic potential: the critical issue of self-potential, *Environment and Planning A*, **27** (11) 1833–1848.

Gaudry, M. J. and M. J. Wills (1978) Estimating the functional form of travel demand models, *Transportation Research*, **12** (4) 257 – 289.

Geisberger, R., P. Sanders and D. Schultes (2008) Better approximation of betweenness centrality, paper presented at the *Proceedings of the Meeting on Algorithm Engineering & Expermiments*, 90–100.

Geurs, K. T. and J. R. Ritsema van Eck (2003) Evaluation of accessibility impacts of land-use scenarios: The implications of job competition, land-use, and infrastructure developments for the Netherlands, *Environment and Planning B: Planning and Design*, **30** (1) 69–87.

Geurs, K. T. and B. van Wee (2004) Accessibility evaluation of land-use and transport strategies: review and research directions, *Journal of Transport Geography*, **12** (2) 127–140.

Giannopoulos, I., D. Jonietz, M. Raubal, G. Sarlas and L. Stähli (2017) Timing of pedestrian navigation instructions, paper presented at the *13th International Conference on Spatial Information Theory (COSIT 2017)*, vol. 86, 16.

Gibbons, S. and H. G. Overman (2012) Mostly pointless spatial econometrics?, *Journal of Regional Science*, **52** (2) 172–191.

Girvan, M. and M. E. J. Newman (2002) Community structure in social and biological networks, *Proceedings of the National Academy of Sciences*, **99** (12) 7821–7826.

Gordon, P., A. Kumar and H. W. Richardson (1989) The influence of metropolitan spatial structure on commuting time, *Journal of urban economics*, **26** (2) 138–151.

Griffith, D. A. and K. G. Jones (1980) Explorations into the relationship between spatial structure and spatial interaction, *Environment and Planning A*, **12** (2) 187–201.

Gutierrez, J. and G. Gomez (1999) The impact of orbital motorways on intra-metropolitan accessibility: the case of Madrid's M-40, *Journal of Transport Geography*, **7** (1) 1–15.

Hackney, J. K., M. Bernard, S. Bindra and K. W. Axhausen (2007) Predicting road system speeds using spatial structure variables and network characteristics, *Journal of Geographical Systems*, **9** (4) 397–417.

Hägerstrand, T. (1970) What about people in Regional Science?, *Papers of the Regional Science Association*, **24** (1) 6–21.

Halás, M., P. Klapka and P. Kladivo (2014) Distance-decay functions for daily travel-to-work flows, *Journal of Transport Geography*, **35**, 107–119.

Handy, S. (1993) Regional versus local accessibility: Implications for network travel, *Transportation Research Record*, **1400**, 58–66.

Hansen, W. G. (1959) How accessibility shapes land use, *Journal of the American Institute of Planners*, **25**, 73–76.

Hanson, S. and M. Schwab (1987) Accessibility and Intraurban Travel, *Environment and Planning A*, **19** (6) 735–748.

Harris, R., J. Moffat and V. Kravtsova (2011) In search of 'W', *Spatial Economic Analysis*, **6** (3) 249–270.

Hausman, J. A. (1978) Specification tests in econometrics, *Econometrica: Journal of the econometric society*, 1251–1271.

Herrera, M., J. Mur and M. Ruiz Marin (2012) Selecting the most adequate spatial weighting matrix: A study on criteria, *Technical Report*, University Library of Munich, Germany.

Highway Capacity Manual (2010) Transportation research board of the national academies, *Washington, DC*, 2010.

Himes, S. C. and E. T. Donnell (2010) Speed prediction models for multilane highways: simultaneous equations approach, *Journal of Transportation Engineering*, **136** (10) 855–862.

Hitchcock, F. L. (1941) The distribution of a product from several sources to numerous localities, *Journal of mathematics and physics*, **20** (1-4) 224–230.

Holme, P., B. J. Kim, C. N. Yoon and S. K. Han (2002) Attack vulnerability of complex networks, *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, **65** (5) 14.

Horner, M. W. and J. Downs (2014) Integrating people and place: A density-based measure for assessing accessibility to opportunities, *Journal of Transport and Land Use*, **7** (2) 23.

Horni, A., K. Nagel and K. W. Axhausen (2016) *The Multi-Agent Transport Simulation MATSim*, Ubiquity Press, London.

Hosmer, D. W., S. Lemeshow and S. May (2008) *Applied Survival Analysis*, 2 edn., Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, NJ, USA.

Jayasinghe, A., K. Sano, C. C. Abenayake and P. Mahanama (2019) A novel approach to model traffic on road segments of large-scale urban road networks, *MethodsX*, **6**, 1147–1163.

Jenelius, E. and H. N. Koutsopoulos (2013) Travel time estimation for urban road networks using low frequency probe vehicle data, *Transportation Research Part B: Methodological*, **53**, 64–81.

Jenelius, E., T. Petersen and L. Mattsson (2006) Importance and exposure in road network vulnerability analysis, *Transportation Research Part A: Policy and Practice*, **40** (7) 537–560.

Jones, P., M. Dix, C. M.I. and H. I.G. (1983) *Understanding Travel Behaviour*, Gower, Aldershot.

Joseph, A. E. and P. R. Bantock (1982) Measuring potential physician accessibility to general practitioners in rural areas: a method and case study, *Social Science & Medicine*, **16** (1) 85–90.

Kelejian, H. H. and G. Piras (2014) Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes, *Regional Science and Urban Economics*, **46** (1) 140–149.

Kelejian, H. H. and I. R. Prucha (1999) A generalized moments estimator for the autoregressive parameter in a spatial model, *International economic review*, **40** (2) 509–533.

Kelejian, H. H. and I. R. Prucha (2010) Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances, *Journal of Econometrics*, **157** (1) 53–67.

Khattak, A. J., V. J. Amerlynck and R. G. Quercia (2000) Are travel times and distances to work greater for residents of poor urban neighborhoods?, *Transportation Research Record*, **1718** (1) 73–82.

Kissling, W. D. and G. Carl (2007) Spatial autocorrelation and the selection of simultaneous autoregressive models, *Global Ecology and Biogeography*, **17** (1) 59–71.

Kleiber, C. and A. Zeileis (2008) *Applied Econometrics with R*, Springer-Verlag, New York.

Kwan, M.-P. (1998) Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework, *Geographical Analysis*, **30** (3) 191–216.

Kwan, M.-p. and J. Weber (2003) Individual accessibility revisited: Implications for geographical analysis in the twenty-first century, *Geographical Analysis*, **35** (4) 341–353.

Lenormand, M., A. Bassolas and J. J. Ramasco (2016) Systematic comparison of trip distribution laws and models, *Journal of Transport Geography*, **51**, 158–169.

LeSage, J. and R. Pace (2004) Spatial and spatiotemporal econometrics, *Advances in Econometrics*, **18** (4) 1–32.

LeSage, J. P. (1997) Regression analysis of spatial data, *Journal of Regional Analysis and Policy*, **27** (1100-2016-89650) 83.

LeSage, J. P. (1999) *Spatial econometrics*, Regional Research Institute, West Virginia University Morgantown, WV.

LeSage, J. P. and R. K. Pace (2008) Spatial econometric modeling of origin-destination flows, *Journal of Regional Science*, **48** (5) 941–967.

Levinson, D. M. (1998) Accessibility and the journey to work, *Journal of Transport Geography*, **6** (1) 11–21.

Li, Q., T. Zhang, H. Wang and Z. Zeng (2011) Dynamic accessibility mapping using floating car data: A network-constrained density estimation approach, *Journal of Transport Geography*, **19** (3) 379–393.

Loder, A., R. Tanner and K. W. Axhausen (2017) The impact of local work and residential balance on vehicle miles traveled: A new direct approach, *Journal of Transport Geography*, **64**, 139–149.

Lopes, S., N. Brondino and A. da Silva (2014) Gis-based analytical tools for transport planning: Spatial regression models for transportation demand forecast, *ISPRS International Journal of Geo-Information*, **3** (2) 565–583.

López, F. A., A. Páez, J. A. Carrasco and N. A. Ruminot (2017) Vulnerability of nodes under controlled network topology and flow autocorrelation conditions, *Journal of Transport Geography*, **59**, 77–87.

Lowry, M. (2014) Spatial interpolation of traffic counts based on origin–destination centrality, *Journal of Transport Geography*, **36**, 98–105.

Luo, D., O. Cats and H. van Lint (2019) Can passenger flow distribution be estimated solely based on network properties in public transport systems?, *Transportation*, 1–20.

Ma, K.-R. and D. Banister (2006) Excess commuting: a critical review, *Transport reviews*, **26** (6) 749–767.

Makridakis, S. and M. Hibon (1995) Evaluating accuracy (or error) measures, *Insead*, 1–41.

Manaugh, K., L. F. Miranda-Moreno and A. M. El-Geneidy (2010) The effect of neighbourhood characteristics, accessibility, home–work location, and demographics on commuting distances, *Transportation*, **37** (4) 627–646.

Maoh, H. and Z. Tang (2012) Determinants of normal and extreme commute distance in a sprawled midsize Canadian city: evidence from Windsor, Canada, *Journal of Transport Geography*, **25**, 50–57.

Martínez, L. M. and J. M. Viegas (2013) A new approach to modelling distance-decay functions for accessibility assessment in transport studies, *Journal of Transport Geography*, **26**, 87–96.

Mattsson, L. G. and E. Jenelius (2015) Vulnerability and resilience of transport systems - A discussion of recent research, *Transportation Research Part A: Policy and Practice*, **81**, 16–34.

McDaniel, S., M. B. Lowry and M. Dixon (2014) Using origin–destination centrality to estimate directional bicycle volumes, *Transportation Research Record*, **2430** (1) 12–19.

Mercado, R. and A. Páez (2009) Determinants of distance traveled with a focus on the elderly: a multilevel analysis in the Hamilton CMA, Canada, *Journal of Transport Geography*, **17** (1) 65–76.

Miller, E. J. (2018) Accessibility: measurement and application in transportation planning, *Transport Reviews*, **38** (5) 551–555.

Miller, H. J. (1999) Measuring space-time accessibility benefits within transportation networks: Basic theory and computational procedures, *Geographical Analysis*, **31** (2) 187–212.

Miller, H. J. (2004) Tobler's First Law and Spatial Analysis, *Annals of the Association of American Geographers*, **94** (2) 284–289.

Mohamad, D., K. C. Sinha, T. Kuczek and C. F. Scholer (1998) Annual average daily traffic prediction model for county roads, *Transportation research record*, **1617** (1) 69–77.

Morency, C., A. Paez, M. J. Roorda, R. Mercado and S. Farber (2011) Distance traveled in three Canadian cities: Spatial analysis from the perspective of vulnerable population segments, *Journal of Transport Geography*, **19** (1) 39–50.

Neutens, T., F. Witlox and P. Demaeyer (2007) Individual accessibility and travel possibilities : A literature review on time geography, *European Journal of Transport and Infraestructure Research*, **7** (1967) 335–352.

O'Kelly, M. E. and M. W. Horner (2003) Aggregate accessibility to population at the county level: Us 1940–2000, *Journal of Geographical Systems*, **5** (1) 5–23.

Oliver, M. a. and R. Webster (1990) Kriging: a method of interpolation for geographical information systems, *International journal of geographical information systems*, **4** (3) 313–332.

Ord, K. (1975) Estimation methods for models of spatial interaction, *Journal of the American Statistical Association*, **70** (349) 120–126.

O'Sullivan, S. and J. Morrall (1996) Walking distances to and from light-rail transit stations, *Transportation Research Record: Journal of the Transportation Research Board*, **1538**, 19–26.

Oum, T. H. (1989) Alternative demand models and their elasticity estimates, *Journal of Transport Economics and Policy*, **23** (2) 163–187.

Páez, A., R. G. Mercado, S. Farber, C. Morency and M. Roorda (2010) Relative accessibility deprivation indicators for urban settings: Definitions and application to food deserts in Montreal, *Urban Studies*, **47** (7) 1415–1438.

Páez, A. and D. M. Scott (2004) Spatial statistics for urban analysis: A review of techniques with examples, *GeoJournal*, **61** (1) 53–67.

Páez, A., D. M. Scott and C. Morency (2012) Measuring accessibility: positive and normative implementations of various accessibility indicators, *Journal of Transport Geography*, **25**, 141–153.

Park, Y. J. and F. F. Saccomanno (2006) Evaluating speed consistency between successive elements of a two-lane rural highway, *Transportation Research Part A: Policy and Practice*, **40** (5) 375–385.

Partridge, M. D., M. Boarnet, S. Brakman and G. Ottaviano (2012) Introduction: Whither spatial econometrics?, *Journal of Regional Science*, **52** (2) 167–171.

Pebesma, E. and G. Heuvelink (2016) Spatio-temporal interpolation using gstat, *RFID Journal*, **8** (1) 204–218.

Pena, E. A. and E. H. Slate (2006) Global validation of linear model assumptions, *Journal of the American Statistical Association*, **101**, 341–354.

Piras, G. et al. (2010) sphet: Spatial models with heteroskedastic innovations in R, *Journal of Statistical Software*, **35** (1) 1–21.

Pirie, G. H. (1979) Measuring accessibility: A review and proposal, *Environment and Planning A*, **11** (3) 299–312.

Porter, R. J. and J. S. Wood (2013) Exploring endogeneity of macroscopic speed parameters: Empirical study during low volume conditions in construction work zones, *Transportation Letters*, **5** (1) 27–37.

Pulugurtha, S. S. and P. R. Kusam (2012) Modeling annual average daily traffic with integrated spatial data from multiple network buffer bandwidths, *Transportation Research Record*, **2291** (1) 53–60.

Quandt, R. E. and W. J. Baumol (1966) The demand for abstract transport modes: theory and measurement, *Journal of Regional Science*, **6** (2) 13–26.

Quddus, M. (2013) Exploring the relationship between average speed, speed variation, and accident rates using spatial statistical models and gis, *Journal of Transportation Safety & Security*, **5** (1) 27–45.

R Core Team (2018) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, `https://www.R-project.org/`.

Reggiani, A., P. Bucci and G. Russo (2011) Accessibility and impedance forms: Empirical applications to the German commuting network, *International Regional Science Review*, **34** (2) 230–252.

Reggiani, A., P. Nijkamp and D. Lanzi (2015) Transport resilience and vulnerability: The role of connectivity, *Transportation Research Part A: Policy and Practice*, **81**, 4–15.

Ribeiro, A., A. P. Antunes and A. Páez (2010) Road accessibility and cohesion in lagging regions: Empirical evidence from Portugal based on spatial econometric models, *Journal of Transport Geography*, **18** (1) 125–132.

Roy, J. (2004) *Spatial Interaction Modelling: A Regional Science Context*, Springer.

Roy, J. R. and J. C. Thill (2003) Spatial interaction modelling, *Papers in Regional Science*, **83** (1) 339–361.

Sandow, E. and K. Westin (2010) The persevering commuter–duration of long-distance commuting, *Transportation Research Part A: Policy and Practice*, **44** (6) 433–445.

Sargan, J. D. (1958) The estimation of economic relationships using instrumental variables, *Econometrica*, **26** (3) 393.

Sarlas, G. and K. W. Axhausen (2015a) Localized speed prediction with the use of spatial simultaneous autoregressive models, paper presented at the *94th Annual Meeting of the Transportation Research Board, Washington, DC, USA*.

Sarlas, G. and K. W. Axhausen (2015b) Prediction of AADT on a nationwide network based on an accessibility-weighted centrality measure, *Arbeitsberichte Verkehrs-und Raumplanung, IVT, ETH Zurich*, **1094**.

Sarlas, G. and K. W. Axhausen (2017) Mean speed prediction with endogenous volume and spatial autocorrelation: a Swiss case study, *Arbeitsberichte Verkehrs-und Raumplanung, IVT, ETH Zurich*, **1275**.

Sarlas, G. and K. W. Axhausen (2019) Commuting distance and individual accessibility, paper presented at the *98th Annual Meeting of the Transportation Research Board, Washington, DC, USA*.

Sarlas, G., R. Fuhrer and K. W. Axhausen (2015) Quantifying the agglomeration effects of Swiss public transport between 2000 and 2010, paper presented at the *15th Swiss Transport Research Conference (STRC 2015), Ascona, Switzerland*.

Sarlas, G. and A. Kouvelas (2019) Analysis of urban traffic network vulnerability and classification of signalized intersections, paper presented at the *MT-ITS 2019, Krakow, Poland*.

Schatzmann, T., G. Sarlas and K. W. Axhausen (2019-01) Spatial modelling of origin-destination commuting flows in Switzerland, paper presented at the *98th Annual Meeting of the Transportation Research Board (TRB 2019), Washington, DC, USA*.

Schönfelder, S. and K. W. Axhausen (2003) Activity spaces: Measures of social exclusion?, *Transport Policy*, **10** (4) 273–286.

Schürmann, V. (2015) Übermässiges Pendeln in der Schweiz, Master Thesis, ETH Zürich, IVT.

Scott, D. and M. Horner (2008) Examining the role of urban form in shaping people's accessibility to opportunities: An exploratory spatial data analysis, *Journal of Transport and Land Use*, **1** (2) 89–119.

Scott, D., D. Novak, L. Aultman-Hall and F. Guo (2006) Network Robustness Index: A new method for identifying critical links and evaluating the performance of transportation networks, *Journal of Transport Geography*, **14** (3) 215–227.

Selby, B. and K. M. Kockelman (2013) Spatial prediction of traffic levels in unmeasured locations: applications of universal kriging and geographically weighted regression, *Journal of Transport Geography*, **29**, 24–32.

Seya, H., Y. Yamagata and M. Tsutsumi (2013) Automatic selection of a spatial weight matrix in spatial econometrics: Application to a spatial hedonic approach, *Regional Science and Urban Economics*, **43** (3) 429–444.

Shankar, V. and F. Mannering (1998) Modeling the endogeneity of lane-mean speeds and lane-speed deviations: a structural equations approach, *Transportation Research Part A: Policy and Practice*, **32** (5) 311–322.

Shen, Q. (1998) Location characteristics of inner-city neighborhoods and employment accessibility of low-wage workers, *Environment and Planning B: Planning and Design*, **25** (3) 345–365.

Shen, Q. (2000) Spatial and social dimensions of commuting, *Journal of the American Planning Association*, **66** (I) 68–83.

Sheppard, E. S. (1984) The distance-decay gravity model debate, *Spatial Statistics and Models*, 367–388.

Shimbel, A. (1953) Structural parameters of communication networks, *The Bulletin of Mathematical Biophysics*, **15** (4) 501–507.

Simini, F., M. C. González, A. Maritan and A.-L. Barabási (2012) A universal model for mobility and migration patterns, *Nature*, **484** (7392) 96.

Simma, A. and K. W. Axhausen. (2003) Interactions between travel behaviour, accessibility and personal characteristics: The case of the Upper Austria Region, *Arbeitsbericht Verkehrs-und Raumplanung, ETH Zürich, Zürich*, **182**.

Sohn, J. (2006) Evaluating the significance of highway network links under the flood damage: An accessibility approach, *Transportation Research Part A: Policy and Practice*, **40** (6) 491–506.

Song, J., M. Ghosh, S. Miaou and B. Mallick (2006) Bayesian multivariate spatial models for roadway traffic crash mapping, *Journal of Multivariate Analysis*, **97** (1) 246–273.

Sørensen, T. (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons, *I kommission hos E. Munksgaard*.

Stock, J. H., J. H. Wright and M. Yogo (2002) A survey of weak instruments and weak identification in generalized method of moments, *Journal of Business & Economic Statistics*, **20** (4) 518–529.

Stouffer, S. (1940) Intervening opportunities: a theory relating mobility and distance, *American sociological review*, **5** (6) 845–867.

Stouffer, S. A. (1960) Intervening opportunities and competing migrants, *Journal of Regional Science*, **2** (1) 1–26.

Straatemeier, T. (2008) How to plan for regional accessibility?, *Transport Policy*, **15** (2) 127–137.

Talvitie, A. (1973) A direct demand model for downtown work trips, *Transportation*, **2**, 121–152.

Taylor, M. A. (2017) *Methods for Vulnerability Analysis*, Elsevier, ISBN 9780128110102.

Taylor, M. A., S. V. Sekhar and G. M. D'Este (2006) Application of accessibility based methods for vulnerability analysis of strategic road networks, *Networks and Spatial Economics*, **6** (3-4) 267–291.

Tennekes, M. (2018) tmap: Thematic maps in r, *Journal of Statistical Software*, **84** (6) 1–39.

Therneau, T. M. and T. Lumley (2014) Package 'survival', *Survival analysis Published on CRAN*.

Tobler, W. R. (1970) A computer movie simulating urban growth in the detroit region, *Economic geography*, **46**, 234–240.

Train, K. E. (2009) *Discrete Choice Methods with Simulation*, 2 edn., Cambridge University Press.

Ubøe, J. (2004) Aggregation of gravity models for journeys to work, *Environment and Planning A*, **36** (4) 715–729.

US National Research Council (2007) Metropolitan Travel Forecasting: Current Practice and Future Direction, *Transportation Research Board*, **288**.

Vale, D. S. and M. Pereira (2017) The influence of the impedance function on gravity-based pedestrian accessibility measures: A comparative analysis, *Environment and Planning B: Urban Analytics and City Science*, **44** (4) 740–763.

Van Wee, B. (2016) Accessible accessibility research challenges, *Journal of transport geography*, **51**, 9–16.

Van Wee, B., M. Hagoort and J. A. Annema (2001) Accessibility measures with competition, *Journal of Transport Geography*, **9** (3) 199–208.

Wachs, M. and T. G. Kumagai (1973) Physical accessibility as a social indicator, *Socio-Economic Planning Sciences*, **7** (5) 437–456.

Wang, C., M. A. Quddus, T. Ryley, M. Enoch and L. Davison (2012) Spatial models in transport: a review and assessment of methodological issues, paper presented at the *91st Annual Meeting of the Transportation Research Board, Washington, DC, USA*.

Wang, X. and K. M. Kockelman (2009) Forecasting network data: Spatial interpolation of traffic counts from texas data, *Transportation Research Record*, **2105** (1) 100–108.

Wang, Y., B. Y. Chen, H. Yuan, D. Wang, W. H. Lam and Q. Li (2018) Measuring temporal variation of location-based accessibility using space-time utility perspective, *Journal of Transport Geography*, **73**, 13–24.

Wasserman, S., K. Faust and D. Iacobucci (1994) *Social Network Analysis: Methods and Applications*, Cambridge University Press, Cambridge.

Webber, M. J. (1980) A theoretical analysis of aggregation in spatial interaction models, *Geographical Analysis*, **12** (2) 129–141.

Weber, J. (2003) Individual accessibility and distance from major employment centers: An examination using space-time measures, *Journal of Geographical Systems*, **5** (1) 51–70.

Wilson, A. (1967) A statistical theory of spatial distribution models, *Transportation Research*, **1** (3) 253–269.

Wilson, A. (1970) The use of the concept of entropy in system modelling, *Operational Research Quarterly (1970-1977)*, **21** (2) 247–265.

Wilson, A. G. (1971) A Family of Spatial Interaction Models, and Associated Developments, *Environment and Planning A: Economy and Space*, **3** (1) 1–32.

Winkelmann, R. (2008) *Econometric analysis of count data*, Springer Science & Business Media.

Winkelmann, R. (2015) Counting on count data models, *IZA World of Labor*.

Wooldridge, J. (2011) Control function and related methods, *LABOUR lectures: slides*, 1–67.

Wooldridge, J. M. (2012) *Introductory econometrics: A modern approach*, Cengage Learning.

Xia, Q., F. Zhao, Z. Chen, L. D. Shen and D. Ospina (1999) Estimation of annual average daily traffic for nonstate roads in a florida county, *Transportation research record*, **1660** (1) 32–40.

Yang, Y., C. Herrera, N. Eagle and M. C. González (2014) Limits of predictability in commuting flows in the absence of data for calibration, *Scientific reports*, **4**, 5662.

Zeiler, A., G. Sarlas, K. M, B. BR, S. B, J. Bode, F. P and K. W. Axhausen (2014) FaLC Transport Simulation Module: How accurate can simplified travel time calculations be?, paper presented at the *14th Swiss Transport Research Conference, Ascona*.

Zhang, J. and G. Kristensen (1995) A gravity model with variable coefficients: The eec trade with third countries, *Geographical Analysis*, **27** (4) 307–320.

Zhao, F., L.-F. Chow, M.-T. Li, A. Gan and I. Ubaka (2003) Forecasting transit walk accessibility : A regression model alternative to the buffer method, *Transportation Research Record: Journal of the Transportation Research Board*, **1835** (03) 16.

Zhao, F. and N. Park (2004) Using geographically weighted regression models to estimate annual average daily traffic, *Transportation Research Record: Journal of the Transportation Research Board*, **1879**, 99–107.