Spatial disaggregation in transport modelling

Modelling Europe with more than 100,000 travel zones

Zur Erlangung des akademischen Grades eines
Doktors der Ingenieurwissenschaften
(Dr.-Ing.)

bei der Fakultät für Wirtschaftswissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte
Dissertation
von

Diplom-Informatiker Jan Ihrig

Karlsruhe 2018

Referent:	Prof. Dr. Kay Mitusch
Korreferentin:	Prof. Dr. Stéphanie Souche-Le Corvec
Prüfer:	PD Dr. Patrick Jochem
Vorsitzender:	Prof. Dr. Martin Klarmann

Tag der mündlichen Prüfung: 19.11.2018

Vorwort

Die vorliegende Doktorarbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Netzwerkökonomie des Instituts für Volkswirtschaftslehre des Karlsruher Instituts für Technologie. Im Rahmen der beiden Forschungsprojekte ETISplus und HIGH-TOOL beschäftigte ich mich intensiv mit der Überarbeitung von bestehenden Konzepten in der europäischen Personenverkehrsmodellierung und erkannte das Potenzial einer feineren räumlichen Auflösung der Verkehrszellen für die Einsatzmöglichkeiten solcher Modelle. Diese Erkenntnis mündete in die vorliegende Dissertation.

Bei der Erstellung dieser Arbeit wurde mir vielfältige Unterstützung zuteil. Mein besonderer Dank gilt Prof. Dr. Kay Mitusch für die wohlwollende Betreuung meiner Arbeit, Prof. Dr. Stéphanie Souche-Le Corvec für die Übernahme des Korreferats, PD Dr. Patrick Jochem, Prof. Dr. Werner Rothengatter und Dr. Benedikt Mandel für ihre wertvollen und kritischen Anregungen zu meiner Arbeit. Gleichfalls danke ich allen derzeitigen und ehemaligen Kollegen, insbesondere Markus Kraft, Tilman Matteis und Dr. Eckhard Szimba. Gerne erinnere ich mich an unsere gemeinsamen Radtouren und die Diskussionsrunden freitagabends zurück.

Mein innigster Dank gilt meiner Ehefrau Miluška, meinen Eltern, Schwestern und Freunden für ihr Verständnis und ihre vollumfängliche Unterstützung während der Entstehung dieser Arbeit. Diesem Rückhalt in Familie und Freundeskreis verdanke ich sehr viel.

Karlsruhe, im Dezember 2018
Jan Ihrig

Abstract

European transport policy pursues several targets, such as reducing CO_2 emissions and improving the efficiency of transport systems. In order to investigate the effectiveness as well as the welfare impacts of potential policy measures, transport models are applied in policy consulting. However, the travel zones used in recent transport models operating at European scale are often too large, mainly due to complexity and data availability. These models can provide only limited insights into regional traffic flows.

In this thesis, an innovative transport modelling approach called HIPAT is introduced. HIPAT is based on disaggregated and homogeneous travel zones, and thus facilitates the modelling of European traffic flows, including long-distance, regional and short-distance passenger trips with a single, consistent transport model. This enables European policy makers to assess also regional welfare impacts when prioritising, for instance, investments in the trans-European transport network. The quantum leap from 1,500 travel zones at NUTS-3 level to more than 100,000 at LAU-2 level, while simultaneously reducing the model runtime from several days to one hour, can be facilitated by solving the trip distribution problem very efficiently at different spatial levels. In comparison to recent European transport models, the consistency of the overall modelling approach was also improved, by integrating the trip distribution, the modal split and the network assignment models.

Within the last nine years most parts of this thesis were researched in the course of the author's involvement as a transport modelling and data specialist in the two European research projects ETISplus and HIGH-TOOL. In a first step, a modelling database was established covering regional indicators at NUTS-3 level and harmonised mobility indicators from travel surveys at country level. In a second step, the currently applied four-step transport modelling approach was intensively revised and the IPAT passenger transport model was developed. Its methodology has successfully been validated in the course of publishing the IPAT model with the HIGH-TOOL policy assessment model that was awarded the German Mobility Prize in 2017. In a third step, the introduced modelling database was disaggregated from NUTS-3 to LAU-2 level for an important long-distance road transport corridor between Paris and Budapest (the "Magistrale"). The compilation of this database was a key precondition in order to realise a prototype implementation of the HIPAT approach. The model operates on about 33,000 travel zones and has a runtime of two minutes. It was intensively tested at NUTS-2, NUTS-3 and LAU-2 levels and the results clearly demonstrate the advantages of smaller travel zones.

Solely LAU-2 level enables the modelling of regional and short-distance traffic flows. Hence, the implemented HIPAT model provides, for the first time, a sound basis for the assessment of regional welfare impacts of European transport policies. In a next step, the scope of this model should be increased to cover the whole of Europe, thus encompassing more than 100,000 travel zones.

Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

AADT ..... Annual Average Daily Traffic
A/E ..... Access/Egress (cost)
CLC ..... CORINE land cover (database)
cnst ..... Constant term (in algebraic expressions)
EC ..... European Commission
EMC ..... Expected Minimum Cost
ETISplus ..... Research project between 2009 and 2012
EU ..... European Union / European
Eurostat ..... Statistical bureau of the European Commission
G/A ..... Generation/Attraction (trip matrix)
GDP ..... Gross Domestic Product
GIS ..... Geographic Information System
GTT ..... Generalised Travel Times
Gpkm ..... 1000 million passenger-kilometres
GVA ..... Gross Value Added
HIPAT ..... Hierarchical integrated passenger transport (model)
HIGH-TOOL ..... Research project between 2013 and 2016
HSR ..... High-speed rail
IPAT ..... Integrated passenger transport model applied in HIGH-TOOL
IWW ..... Institute for Economic Policy Research (former KIT department)
KIT ..... Karlsruhe Institute of Technology
LoS ..... Level of Service
LAU ..... Local Administrative Unit
MAUP ..... Modifiable Areal Unit Problem
MNL ..... Multinomial Logit (model)
NMNL ..... Nested Multinomial Logit (model)
NUTS ..... Nomenclature of Territorial Units for Statistics
O/D ..... Origin/Destination (trip matrix)
PAT ..... Passenger transport model applied in ETISplus
pkm ..... Passenger-kilometres
TEN-T ..... Trans-European transport network
TLD ..... Trip Length Distribution
TPM ..... Transport policy measure
TRANS-TOOLS ... Reference model of the EC for transport policy assessment
VACLAV ..... EU transport model developed at KIT
vkm ..... Vehicle-kilometres
VoT ..... Value of Time

Chapter 1

Introduction

Dealing with congestion in road transport is an important problem for road passengers and road haulage operators as well as a challenge for transport policy makers. Notably, commuting trips contribute to capacity overloads on major roads. A study turned out by the European Commission estimates the cost of road congestion in Europe to be equivalent to 1% of GDP (Christidis and Ibañez-Rivas, 2012). For 2016, this would amount to a total cost of 139.6 billion Euro (EC, 2016).

Increasing the efficiency of the trans-European transport network is therefore an important objective of European transport policy. It can be achieved by implementing adequate infrastructure policy measures. In order to assess the effectiveness of potential measures and to prioritise potential infrastructure projects, transport models are being developed. For complexity and for data availability reasons, however, European models are often based on large travel zones. Hence, these models can only assess transport policies which concern long-distance trips but not regional and short-distance trips. They thus fail to capture the important issue of road congestion caused by commuting trips.

In the European reference transport model TRANS-TOOLS and in the European Transport Policy Information System (ETIS), for instance, Europe is modelled by approximately 1,500 NUTS-3 regions with an average diameter of over 50 km (EC, 2008).[1] The spatial disaggregation of travel zones provides an enormous potential to improve the accuracy of current European transport models, given that over 90% of the trips carried out with passenger cars are shorter than 50 km.[2] However, modelling Europe with more than 100,000 LAU-2 regions is not possible in either model due to the complexity of the so-called O/D trip matrix[3] that increases quadratically with the number of travel zones. A current way out of this dilemma to enable a partly inconsistent but integrated assessment of transport policies is to link different models focusing solely on long-distance, regional or short-distance transport. In the European transport model VACLAV, for instance, the trip matrix is provided at NUTS-3 level but is subsequently disaggregated during the network assignment step (Schoch, 2004).[4] The development of an innovative and consistent approach to model Europe with more than 100,000 travel zones is therefore an interesting research question, which is tackled in this thesis.

[1] The NUTS classification distinguishes four regional levels, NUTS-0 to NUTS-3, and two local levels, LAU-1 and LAU-2. For instance, NUTS-0 refers to countries, NUTS-3 to cities, and LAU-2 to communes.

[2] E. g. 90.8% in Germany (infas and DLR, 2010a), 93.3% in Denmark (Christiansen, Hjalmar, 2011), 92.0% in the Netherlands (MVW, 2010), 94.7% in the United Kingdom (DfT, 2008), and 98.1% in Spain (Pérez Lou et al., 2007).

[3] The origin/destination matrix provides information on the number of trips between each pair of travel zones.

[4] The network assignment step produces information on estimated traffic loads at the level of network links (e. g. road segments). It is computed based on the O/D trip matrix and shortest paths of these trips in the transport networks.

1.1 Structure and milestones

This thesis introduces an innovative hierarchical approach for European passenger transport modelling relying on more than 100,000 homogenised and disaggregated travel zones. The hierarchical approach, operating at different levels (e. g. NUTS-0 and LAU-2), is capable of handling long-distance, regional and also short-distance trips. The distinction between these three distance categories within this thesis was largely made by taking into account the average diameter of about 50 km of a NUTS-3 region. Short-distance trips are less than 50 km, regional trips between 50 and 200 km, and long-distance trips longer than 200 km. The modelling of urban trips is, however, beyond the scope of this thesis.

The implemented hierarchical HIPAT model is the result of a long research process of about nine years, in which three successive model versions following the standard four-step approach in transport modelling were developed. These models are:

- the PAT model during the ETISplus project (2009-2012);
- the IPAT model during the HIGH-TOOL project (2013-2016); and finally
- the HIPAT model.

For the development of a European transport model, a critical aspect is the number of travel zones or, to be more precise, the size of the O/D trip matrix and runtime of the model. Given the large runtime problems of existing European transport models built on only 1,500 travel zones, in the development of the three models particular attention was paid to the trip distribution model computing the trip matrix. The HIPAT model has been designed to process 150,000 travel zones, which is why the complexity of the trip matrix will increase by a factor of 10,000 compared to existing models.[5] It was therefore necessary to significantly revise the methodology of the standard trip distribution model.

Reading this thesis, it is important to bear in mind that the three developed transport models and the related theoretical background are discussed one after another. The thesis is therefore subdivided into four parts. Part I starts with a general introduction to transport modelling and continues with the ETISplus project in which a European data basis for transport modelling was established at NUTS-3 level. It then discusses the PAT model that was developed in order to generate passenger demand data for ETISplus, namely O/D trip matrices. Though the PAT model was developed in close accordance with existing transport modelling approaches, its validation revealed several methodological issues that are summarised at the end of chapter 2. These issues have been considered in the development of the IPAT and the HIPAT models.

Part II starts with a review of several important modelling concepts in chapter 3. The starting point of this review are the many difficulties and problems experienced during the setup of the PAT model. In particular, the revision of the multinomial logit model is relevant for this thesis as it enables the transferability of a European transport model between different distance categories (e. g. long-distance and short-distance trips) and different hierarchical levels (e. g. NUTS-0 and LAU-2).[6] In chapter 4, the methodology of the trip distribution model is investigated and revised in detail. Paying attention

[5] The NUTS-2013 classification for the EU-28 comprises 118,504 LAU-2 regions (Eurostat, 2015b). Further travel zones are required in order to model other countries, e. g. Switzerland and Norway.

[6] For the PAT model transferability was not given. The corresponding problems were circumvented by introducing additional calibration parameters.

to the important role of the deterrence function in this model, the concept of a more sophisticated "deterrence model" is developed in chapter 5. The methodological revisions introduced in Part II are implemented in the IPAT model, which is discussed and validated in chapter 6. As part of the European HIGH-TOOL transport policy assessment model that was developed on behalf of the European Commission, the IPAT model has been validated successfully by a broader audience. Moreover, the HIGH-TOOL model, including all datasets, is offered free of charge and the IPAT model can therefore be further investigated and developed by anyone.

Part III starts with a discussion of the spatial disaggregation of travel zones in chapter 7. In addition, this chapter provides deeper insights into the theory of georeferenced data (GIS data) which are required for generating artificial travel zones and for ensuring the availability of regional indicators at LAU-2 level.[7] Chapter 8 introduces the hierarchical modelling approach. It focuses on the development of a methodology and an efficient algorithm for computing trip distribution at different hierarchical levels. Computation of long-distance trips, for instance, at the NUTS-1 level and short-distance trips at the LAU-2 level enables the number of travel zones to be increased from 1,500 NUTS-3 to more than 100,000 LAU-2 regions. The result, that also incorporates several other methodological improvements compared to the IPAT model, is called the HIPAT modelling approach, and is discussed in chapter 9.

This thesis concludes with Part IV summarising the main results and giving an outlook for the continuation of this research.

Important milestones forming the groundwork for the development of the three transport models addressed within this thesis include:

- implementing a consistent and complete database for transport modelling at European scale, covering regional indicators at NUTS-3 level and mobility indicators at country-level;
- revising the methodology of the trip distribution and the modal split models[8], and integrating both models to improve the consistency of the overall transport modelling approach;
- developing an innovative algorithm for solving the trip distribution problem efficiently at different spatial levels;
- improving the homogeneity of travel zones by generating artificial travel zones;[9]
- ensuring the availability of regional indicators for over 100,000 LAU-2 regions by downscaling;
- developing thread-safe data structures and algorithms in order to benefit from parallel computing and to achieve a runtime of two minutes of the HIPAT prototype model; and
- implementing these models, e. g. the IPAT model by approximately 20,000 lines of Java code.

The next section introduces the core topics and challenges of this thesis. With this context, the interrelation between Parts II to III (i. e. between the models PAT, IPAT and HIPAT) is described in more detail in section 1.3.

[7] Only such datasets were used that were provided free of charge.

[8] Identified issues relate to the deterrence function and the logit model that will be explained shortly.

[9] According to the NUTS-2013 classification, for instance, the German capital Berlin, with about 3.5 million inhabitants, is only reflected by one single LAU-2 region. Thus, Berlin should be modelled by about 1,000 artificial travel zones in order to ensure compatibility with other travel zones at LAU-2 level.

1.2 Background and motivation

The basic idea to improve European transport models by systematically disaggregating and homogenising the travel zones as well as by systematically overcoming existing methodological inconsistencies stems from experience acquired during prior research projects with high-resolution image recording systems for industrial visual inspection at the Fraunhofer IOSB Research Institute.[10] Systems of this kind are needed for monitoring production processes to ensure a high, constant product quality by detecting error patterns in the recorded images. The systematic approach in the optimisation of visual inspection systems includes:

- increasing the spatial resolution of the recorded images;
- homogenisation the image acquisition parameters such as temperature and illumination;[11]
- preprocessing the image data to overcome known inconsistencies such as vignetting[12]; and
- training pattern recognition algorithms based on sample data.

This is very similar to the systematic procedure that has been followed in the development of the HIPAT modelling approach. It includes:

- disaggregating the travel zones;
- homogenising their size;
- overcoming existing methodological inconsistencies (e. g. by revising and integrating the methodologies of the trip distribution and the modal split models); and
- calibrating the overall transport model based on observed transport demand data.

The impact of the average zone size on the accuracy of the model output is illustrated by Figure 1.1, which compares two output indicators that were computed at different levels, i. e. NUTS-2 and LAU-2.[13] The two images on the left show the distribution of the indicator "density of labour force by region" (Fig. 1.1a). While the image produced at NUTS-2 level is dominated by two large and uniform areas belonging to the same density category, the image produced at LAU-2 level resembles a patchwork of heterogeneous areas encompassing all density categories. In particular, urban and non-urban areas can be distinguished at LAU-2 level. Given that the mobility demand is different in urban and non-urban areas, LAU-2 travel zones provide the better basis for transport demand modelling and for investigating the regional effects of transport.

The main benefit, however, of applying more disaggregated travel zones concerns the trip distribution model computing the O/D trip matrix that is the basis for the network assignment (Fig. 1.1b). If the transport model is applied at NUTS-2 level, the majority of the trips are intra-zonal. Accordingly, the

[10] Fraunhofer Institute of Optronics, System Technologies and Image Exploitation; URL: `www.iosb.fraunhofer.de` (Retrieved October 10, 2017).

[11] See Ihrig (2009).

[12] Vignetting refers to a reduction of the image brightness from the centre of the image towards its edges.

[13] The example refers to the two NUTS-2 regions Karlsruhe (DE12) and Stuttgart (DE11). It is produced with the HIPAT prototype model and will be discussed in more detail in section 9.4.2.

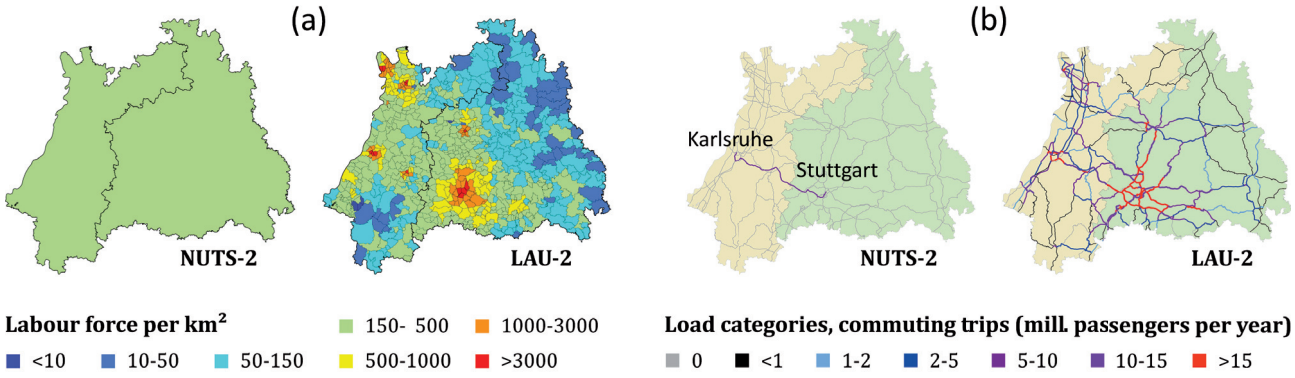

Figure 1.1: Comparison of transport demand indicators at different spatial levels

computed network assignment looks almost empty as only two inter-zonal trips are considered.[14] In the drafted example, traffic loads are computed only for those road links that are part of the shortest path between the two travel zones (i. e. the motorway between the cities of Karlsruhe and Stuttgart), while all the other existing road links are unloaded. In addition, the modelled traffic loads seem to be largely underestimated. This can be explained by the representative travel distance of over 70 km between the two illustrated NUTS-2 regions or, to be more precise, between the two cities of Karlsruhe and Stuttgart which are the two centroids of the travel zones. On the other hand, if the transport model is applied at LAU-2 level, 658 travel zones and 8,794 different routings are considered for the computation of the network assignment. It has to be emphasised that, in this case, almost any possible trip distance (1-240 km)[15] in and between the regions Karlsruhe and Stuttgart can be modelled, given the small size of the respective LAU-2 regions. Accordingly, traffic loads can be computed for many road network links, as shown on the right-hand map, and the network assignment looks more realistic.

Figure 1.2 shows a second example which demonstrates the limitation of current transport models operating at NUTS-3 level with regard to regional trips. The example refers to the existing Rhine bridge in Karlsruhe, which is a bottleneck of the regional transport system. The bridge is used by daily commuters, and increasing traffic volumes frequently lead to long traffic jams. The construction of a second Rhine bridge has therefore been under discussion by regional policy makers for several years. The construction of a second Rhine bridge near Karlsruhe could also be relevant for trans-European passenger and freight transport, which is why this infrastructure project should also be investigated from a European perspective, i. e. by a European transport model. As recent European transport models are operating at NUTS-3 level, regional trips can only be reproduced with limited accuracy, and congestion effects cannot be investigated.

In the drafted example, the left-hand image (Fig. 1.2a) shows the real routing of many commuting trips. The trips start in the western district of Karlsruhe, cross the Rhine and end in a large industrial area that is located close to the city of Wörth am Rhein.[16] The observed trip length is 7 km. If we try to reproduce these trips by a NUTS-3-based transport model (Fig. 1.2b), we have to identify a large

[14] Intra-zonal trips, starting and ending in same centroid of a travel zone, do not contribute to the assignment.
[15] E. g. Illingen/ Vaihingen an der Enz: 5 km; Alpirsbach/ Wertheim: 240 km.
[16] The industrial area comprises a large assembly plant for trucks, providing over 11,000 jobs. The locations of the cities Karlsruhe and Wörth are indicated in parts (b) and (c) of Figure 1.2.

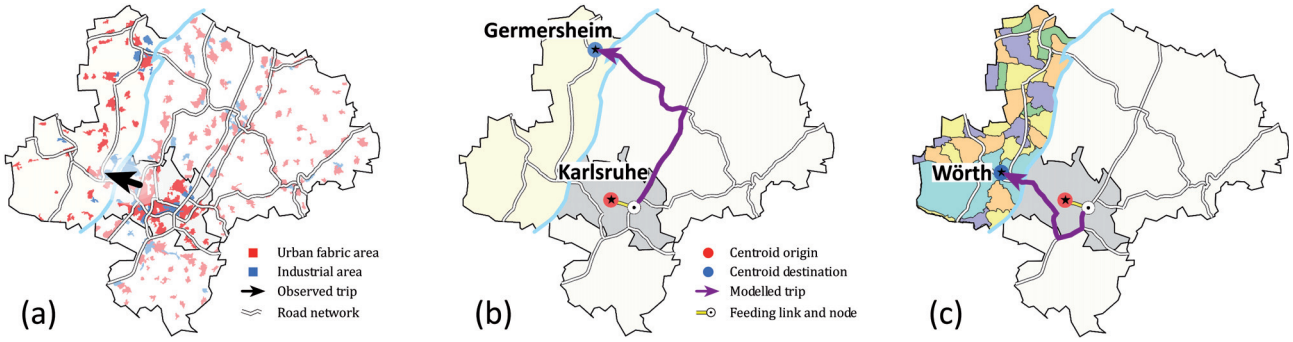

Figure 1.2: Example of modelled traffic volumes for the Rhine bridge in Karlsruhe

discrepancy between the observed and the modelled trips. In the transport model, the trips start in the centre of Karlsruhe, cross the Rhine in the north and end in the city of Germersheim. Due to this routing, the modelled trip length amounts to 40 km and does not depict the importance of the southern Rhine bridge. The discrepancy between the two routings could be improved by adjusting the centroid of Karlsruhe and connecting the travel zone to the motorway crossing the western district of Karlsruhe (not shown). In this case, the routing of the modelled trips would pass via the Karlsruhe Rhine Bridge. However, the discrepancy between the two trip lengths (7 km for the observed trips and 40 km for the modelled trips) could only be reduced if the centroid representing the NUTS-3 region Germersheim (DEB3E) were shifted from the capital to the city of Wörth. This would solve the issue for the Karlsruhe Rhine Bridge but would lead to several follow-up problems (e. g. modelling traffic flows for the Rhine bridge in Germersheim and from Germersheim to the north).

Given that particularly regional commuting road trips between the two drafted NUTS-3 regions contribute to the daily congestion patterns on the Karlsruhe Rhine Bridge, we have to conclude that the discussed infrastructure project of building a second Rhine bridge cannot be investigated by a current European transport model. The main reason for this is the large size of the NUTS-3-based travel zones. Thus, we can only expect better results if we reduce the size of the travel zones.

According to Figure 1.2c, the NUTS-3 region Germersheim is modelled by 31 travel zones at LAU-2 level. Even though the size of the LAU-2 region Wörth am Rhein is quite large, the industrial area in which the observed commuting trips end is well represented by the corresponding centroid. The main shortcoming of applying LAU-2-based travel zones is that the NUTS-3 region Karlsruhe (DE122) is not disaggregated. Given this limitation, the modelled trip length amounts to 20 km if we do not adjust the centroid representing Karlsruhe. For this reason, the NUTS-3 region Karlsruhe should be disaggregated by hand into several artificial city districts.[17] In contrast to recent European transport models operating at NUTS-3-level, the HIPAT model relying on LAU-2-based travel zones and artificial city districts (not shown) can therefore reproduce the observed commuting trips quite accurately. This enables the investigation of regional infrastructure projects like the second Karlsruhe Rhine bridge from a European perspective. Moreover, the HIPAT model operating at LAU-2 level can also support European decision makers prioritising investments in the development of the trans-European transport network.

[17] In the HIPAT model Karlsruhe is modelled by 16 artificially generated travel zones corresponding to LAU-2 level.

Increasing the number of travel zones in a European transport model from 1,500 NUTS-3 regions to more than 100,000 LAU-2 regions with a standard modelling approach is not possible due to complexity and runtime problems. Such problems already exist at NUTS-3 level. For instance, the runtime of the PAT model applied in ETISplus was several hours and the runtime of the TRANS-TOOLS v2.1.9 model exceeds two days (Ibañez-Rivas, 2010).[18]

A general bottleneck of transport models following a standard modelling approach is the complexity of the O/D trip matrix computed by the trip distribution model, given that the number of O/D relations increases quadratically with the number of travel zones. Hence, the more travel zones used, the more complex the trip distribution problem. In order to tackle complex problems, a fundamental principle in computer science is the so-called "divide and conquer strategy" (see, for example, Mehlhorn and Sanders (2008); Ottmann and Widmayer (2002)).[19] However, this principle cannot be applied in our case due to competition between travellers for destination opportunities. For instance, travellers from Antwerp, Berlin and Paris compete for accommodation facilities in Brussels in accordance with the first-come first-served principle. Hence, the respective trip distribution "sub-problems" for Belgium, Germany and France cannot be solved independently by following the divide and conquer strategy. Within this thesis, the competition of travellers for destination opportunities could finally be modelled by using a revised accessibility indicator that is introduced in section 3.3.2.

In order to accomplish the primary objective of this thesis, namely consideration of more than 100,000 LAU-2-based travel zones for transport modelling, a multitude of challenges had to be tackled. These challenges belong to different research fields, including data engineering, software engineering, transport modelling and transport economics. Many of these challenges could be solved during the author's involvement in the two European research projects ETISplus and HIGH-TOOL between 2009 and 2016, in which the two models PAT and IPAT were developed. The evaluation of the two models revealed a significant conceptual advance from PAT to IPAT, encouraging the development of the HIPAT modelling approach and the implementation of a prototype model between 2016 and 2018.

[18] In the TRANS-TOOLS model the majority of the runtime is related to the network assignment model following an excessively ambitious approach. Given the availability of very efficient routing algorithms solving the shortest path problem almost in real time (see, for example, Batz and Sanders (2012); Delling et al. (2013); KIT (2012)) the runtime of the assignment model could be significantly reduced.

[19] The divide and conquer strategy is based on the idea of breaking down a large problem into two or more sub-problems and solving these sub-problems independently, if possible by parallel computing.

1.3 Conceptual advances from PAT to IPAT and HIPAT

The primary objective of developing the PAT model was the computation of the O/D trip matrices for the base year 2010 for the ETISplus database, namely the passenger demand dataset (see sec. 2.3.4). The ETISplus database is the result of a joint European research project. As the demands of intra-zonal and air passenger trips were computed by other project partners, the scope of the PAT model was limited to the demand of inter-zonal rail and road passenger trips between the European NUTS-3 regions. The PAT model follows a tailored four-step approach in which the trip generation and trip distribution models were developed from scratch. The modal split model was developed based on the existing VACLAV model that was also used as a network assignment model.

Evaluation of the PAT model revealed several methodological issues, which are described and discussed in section 2.4 of Part I of this thesis. They include: (i) restriction to inter-zonal rail and road passenger trips, (ii) non-consideration of important explanatory factors in the trip distribution model, such as competition of travellers for destination opportunities, (iii) complexity of the deterrence function used in the trip distribution model, (iv) independent calibration of the trip distribution and modal split models leading to inconsistencies, (v) distance-dependent bias of the logit model used in the modal split model, (vi) application of travel zones that are far too large and too heterogeneous. The issues (i) to (v) are solved in the IPAT model, whereas issue (vi) is solved in the HIPAT model.

However, a milestone for this thesis was the elaboration of a consistent and complete input database for the base year 2010 for the PAT model. This database covers regional indicators for the European NUTS-3 regions as well as comprehensive and harmonised mobility indicators at national level for all European countries.[20] This database is also used for the IPAT and HIPAT models.[21]

In the development of the IPAT model, particular attention was paid to the revision of the methodology of the trip distribution and the modal split models, and to the integration of both models. Throughout this revision process, the distance-dependent bias of the logit mode choice model could be solved by varying the scaling parameter that is related to the error term of the logit model in relation to the travel distance (cf. sec. 3.4.2). Joint calibration of the trip distribution and modal split models was facilitated by using the expected minimum cost indicator computed by the modal split model in the trip distribution model (cf. sec. 3.4.4). In order to overcome the issue of complexity of the deterrence function, the concept of a composite deterrence function was developed (cf. sec. 4.3.3). In addition, the deterrence function was complemented by other explanatory factors, in particular by a scaling function (cf. sec. 5.1). The scaling function establishes the transferability of the deterrence function between different hierarchical levels (e. g. NUTS-2, NUTS-3 and LAU-2), which is a crucial prerequisite for the HIPAT model operating at different hierarchical levels. In order to model the competition of travellers in the trip distribution model, a revised accessibility indicator was developed (cf. sec. 3.3.2). The IPAT model is implemented at NUTS-2 level and calibrated to the base year 2010. As part of the HIGH-TOOL transport policy assessment model, the IPAT model computes the impact of transport policies on passenger transport demand and produces forecasts for the years 2015-2050 in terms of

[20] The collection of regional indicators is described in section 2.3.2, and the compilation of the input database in section 2.3.3.

[21] The two projects ETISplus and HIGH-TOOL, in which the PAT and IPAT models were developed, both referred to the base year 2010. Thus, the base year 2010 was also used to realise a prototype implementation of HIPAT.

O/D trip matrices covering rail, road, air and coach transport. The methodological advances and the IPAT model itself are described in Part II of this thesis.

In the HIPAT modelling approach, the challenge of increasing the number of travel zones from 1,500 NUTS-3 regions to over 100,000 LAU-2 regions and, at the same time, avoiding runtime and complexity problems, is solved by a hierarchical transport modelling approach operating at different spatial levels. Its implementation basically started as a copy of the IPAT programming code, in which standard data structures were replaced by hierarchical data structures. For instance, a hierarchical zoning system was introduced, comprising travel zones at seven different levels. The key innovation of the HIPAT modelling approach is a hierarchical trip distribution model facilitating the computation of long-distance, regional and short-distance trips at different levels which, at best, might reduce the number of O/D relations to be computed to less than 0.1%. In addition, a set of spatial algorithms was developed in order to generate artificial travel zones and to disaggregate the existing input database of the PAT model to LAU-2 level. Within this thesis, a prototype implementation of the HIPAT approach has been realised. The model operates on about 33,000 travel zones with a runtime of two minutes. It was intensively tested and the results clearly demonstrate the advantages of LAU-2-based travel zones for European transport modelling. The HIPAT modelling approach and its prototype implementation are extensively discussed in Part III of this thesis.

It should be noted that the travel zones could be disaggregated even further, e. g. to the level of homogeneous grid raster cells with an edge length of 1 km and in accordance with the spatial resolution of the European dataset on the population density (Bloch et al., 2012). This would certainly improve the accuracy of the transport model. However, it has to be stressed that the development of a European transport model covering urban trips as well is clearly beyond the scope of this thesis.

Part I

Transport modelling basis

Chapter 2

Introduction to transport modelling

This chapter provides a general introduction into transport modelling. It describes several fundamental modelling concepts in the first section and then discusses two existing European transport models. The ETISplus Passenger model (PAT), which was developed between 2009 and 2012 (Ihrig, 2012), is then discussed. With regard to the methodology of the PAT model, several concessions had to be made due to specific requirements of the ETISplus project. This resulted in several limitations of the PAT model. The last section summarises the identified issues of current European transport models and formulates the main research topics that must be further investigated in order to improve the methodology of the four-step transport model. The development of the second transport model in this thesis, the IPAT model, will therefore start in chapter 3 with a literature review and a revision of important economic and spatial concepts that are applied in the four-step model.

2.1 Standard modelling concepts

In this section, several fundamental methodological concepts in transport modelling as well as several existing transport models are discussed. The first section briefly explains the advantages and disadvantages of macroscopic, microscopic and mesoscopic approaches for transport modelling. The classic four-step approach in transport modelling is then briefly introduced. Subsequently, the specific meaning of the basic data items (e. g. travel zones and feeding links) underlying aggregated transport models is described. Furthermore, the challenges to the modelling of regional trips, arising from the definition of very large travel zones, are revealed.

2.1.1 Classification of modelling approaches

Transport models are developed in order to explain observed traffic flows and to compute forecasts. One of the main outputs of aggregated transport models relying on the concept of travel zones is the so-called origin/ destination (O/D) trip matrix. The matrix is commonly distinguished by transport mode and trip purpose and stores the number of trip estimates for all O/D relations that are computed by the transport model. Based on the information of the matrix, the most likely route for all trips can be computed and the traffic loads can be derived at the level of network links (e. g. the number of vehicles crossing a particular section of the road). The computation of this so-called network assignment enables, for instance, the modelling of congestion effects, the identification of infrastructure

bottlenecks, the computation of external costs (e. g. CO_2 emissions), the planning of infrastructure investments, and the assessment of transport policies. Depending on the purpose of the transport model, basically three alternative modelling approaches can be identified. These are the macroscopic, microscopic and mesoscopic modelling approach, each of which provides different advantages and disadvantages.

2.1.1.1 Macroscopic modelling approaches

Macroscopic models operate on a set of travel zones and compute trip volumes between these regions. Each travel zone is represented by its centroid that frequently refers to the centre of the largest city in this region. For this reason, the precise starting and ending point of the trips are unknown. This is of particular relevance for the modelling of intra-zonal trips, which originate and end in the same region, but also for the modelling of trips between neighbouring regions. Macroscopic approaches are particularly suited for European transport modelling. For instance, the travel zones of the two European transport models TRANS-TOOLS and VACLAV are defined in accordance with the European NUTS-3 regions.[1] However, the average diameter of these NUTS-3 regions is over 50 km and both models can therefore only explain the demand for inter-regional and long-distance trips.

2.1.1.2 Microscopic modelling approaches

In contrast to macroscopic models, microscopic approaches pay attention to individual components of traffic such as individual travellers and vehicles, individual decisions, as well as their interaction in the transport system, e. g. the competition between travellers for available destination opportunities. Microscopic models enable the in-depth investigation of traffic congestion, which occurs, for instance, due to zip merging on motorways. However, they are rather demanding with regard to the availability of input data, the complexity of the modelling approach, and, accordingly, the computing power needed. The application of microscopic models is therefore often restricted to territorially limited transport problems like the modelling of urban traffic flows and congestion within a city and its surroundings.

2.1.1.3 Mesoscopic modelling approaches

Mesoscopic models are a third group which can be categorised between macroscopic and microscopic models. Like the macroscopic models, they rely on the concept of travel zones. However, the travel zones chosen are much smaller, and particular attention is paid to the criterion of intra-region homogeneity. Within this thesis the most important criterion for measuring the intra-region homogeneity of a travel zone is the spatial distribution of the population and therefore the discrepancy between the location of the centroid and the exact starting and ending points of each trip. For instance, large travel zones, which cover several cities, cannot adequately be represented by a single centroid. Accordingly, trip estimates which are derived for neighbouring travel zones do not reflect the reality very well. Generally, the smaller the travel zones chosen, the better their intra-region homogeneity. Thus, even short-distance trips between two neighbouring regions can be modelled with sufficient accuracy by mesoscopic models.

[1] Source: NUTS-2006 zoning system (EC, 2008); geographical scope EU-28, Norway and Switzerland.

The size of the travel zones cannot be reduced arbitrarily, since the complexity of a transport model depends quadratically on the number of travel zones. For instance, if we divided an area of 100 square kilometres (a city) into quadratic districts with an edge length of 100 metres, we would generate 10,000 travel zones and accordingly 100 million O/D relations. Taking into account the fact that complexity and runtime problems can already be observed for the TRANS-TOOLS and VACLAV models[2], which only deal with approximately 1,500 regions and accordingly with 2.25 million O/D relations, it is quite obvious that an innovative transport modelling approach has to be developed in order to increase the number of travel zones up to 10,000 or even up to 100,000.

2.1.1.4 Hierarchical modelling approaches

The hierarchical transport modelling approach, which will be developed within this thesis, can be categorised to the group of mesoscopic models. The smallest travel zones underlying the hierarchical transport model correlate with the LAU-2 regions which are defined at the level of communes for the European member states.[3] At the same time, travel zones are also defined at the four NUTS levels (NUTS-0 to NUTS-3) and the LAU-1 level. The definition of travel zones at six different hierarchical levels facilitates the modelling, for instance, of long-distance trips at NUTS-1 level and short-distance trips at LAU-2 level. Following this approach, the number of travel zones can be increased from 1,500 to over 100,000, given that only a fraction of the ten thousand million resulting O/D relations is considered by the hierarchical approach as avoiding runtime problems. The development of the hierarchical transport model will be discussed in detail in Part III of this thesis.

2.1.2 The four-step model

The four-step approach is a standard concept in transport modelling. It explains the process of appearance of transport by four consecutively applied sub-models (Ortúzar and Willumsen, 2001). The results that are computed in each step feed into the next one. In the first step, the trip generation model computes the trip demand for all origin zones. In the second step, the trip distribution model computes trip estimates for all O/D relations. In the third step, the modal split model computes market shares of transport modes. In the last step, the network assignment model derives traffic loads for all network links. Each of the four sub-models can be developed and executed independently but it is also possible to integrate several of them. The integration of several sub-models increases the consistency of the overall modelling approach. At the same time, this process of determining modelling parameters can also be very difficult because all integrated sub-models have to be calibrated together. Nevertheless, it is also possible to follow a non-integrated approach in order to simplify the estimation of the modelling parameters. However, since each sub-model is then calibrated independently, the consistency between them is not ensured.

Depending on the purpose of a transport model, different approaches for integrating two or more sub-models may be appropriate. Known implementations of passenger transport models at European and national scales include the following models: the TRANS-TOOLS model integrating the first three steps (Petersen et al., 2009), the HIGH-TOOL model integrating the trip distribution and modal split

[2] E. g. TRANS-TOOLS v2.1.9 has runtime of 2.4 days (Ibañez-Rivas, 2010).

[3] This level of detail provides a sufficient basis for the modelling of regional trips, which cannot be modelled by NUTS-3-based transport models, such as TRANS-TOOLS and VACLAV.

models (Szimba et al., 2016), and the GENDIS model integrating the first two steps (Gaudry et al., 1994b). A modelling framework integrating all four sub-models has been developed recently for the identification of critical road infrastructure in the German state Baden-Württemberg (Schulz, 2012).

An important output of transport models is the so-called trip matrix. These matrices are produced within the first three steps (trip generation, trip distribution and modal split) and provide information on trip volumes for each pair of travel zones by mode of transport.[4] Trip matrices can be provided in two different formats. If the trip volumes stored for each O/D relation only refer to outward trips (i. e. trips leaving the travel zones) but not to returning trips (i. e. trips ending in the travel zones), it is commonly called a generation/attraction (G/A) matrix. Hence, G/A matrices only refer to half of the trip demand and it is necessary to add the returning trips when computing output indicators. If the stored trip volumes refer to outward and returning trips, it is called an origin/destination (O/D) matrix. O/D matrices (M^{OD}) can be derived from G/A matrices (M^{GA}) according to (2.1)

$$M_{ij}^{\mathrm{OD}} = M_{ij}^{\mathrm{GA}} + M_{ji}^{\mathrm{GA}} \tag{2.1}$$

where i refers to the origin and j to the destination of an O/D relation. Given that G/A matrices distinguish between outward and returning trips, they provide a better basis than O/D matrices for the modelling of traffic loads at the level of network links by the assignment model, in particular with regard to morning and afternoon traffic peaks caused by commuting trips.

It should be emphasised that the terms O/D and G/A trip matrices are not used consistently in the literature. Most often the term O/D trip matrix is used even though a matrix is provided in the G/A format and returning trips have to be added. Within this thesis, the trip matrices produced by the PAT, IPAT and HIPAT models refer to the G/A format. They are referred to as O/D matrices in some discussions in order to indicate the origin/destination character of the stored trip volumes.

2.1.2.1 Trip generation

The trip generation model is the first sub-model of the four-step approach. It computes the trip demand T for each travel zone i, which is generated by the population. T is frequently computed by a trip production function according to (2.2):

$$T_i = \sum_{a,g} R_i^{ag} \, \mathrm{pop}_i^{ag} \tag{2.2}$$

where R refers to a region-specific trip rate factor[5] and pop to the population. Both parameters are distinguished by age group a and gender g. The specific trip rate factors R rely on known average trip rate factors $\bar{R}$. Their values can be computed by a regression model, which is outlined by (2.3):

$$R_i^{ag} = \bar{R}^{ag} \left(\beta_1^{ag} z_{1i} + \beta_2^{ag} z_{2i} + \ldots + \mathrm{cnst}^{ag}\right) \tag{2.3}$$

$$z_{ki} = \frac{x_{ki} - \mu_k}{\sigma_k} \tag{2.4}$$

[4] Information on trip chains like home-work-shopping-home is often not included in trip matrices. For the modelling of trip chains, agent-based simulation frameworks like MATSim may be appropriate (Horni et al., 2016).

[5] Trip rate factors are frequently indicated in “annual trips per capita” (cf. Fig. 2.1).

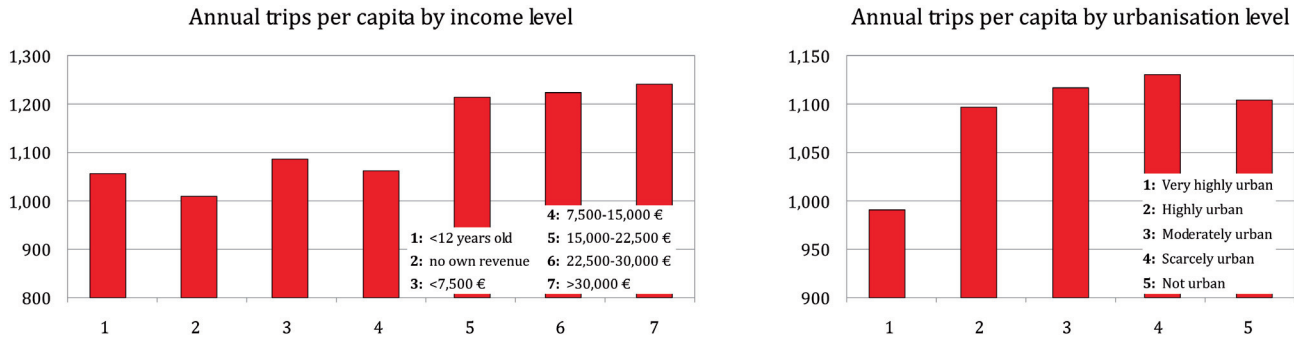

Figure 2.1: Non-linear dependency of trip rates on two explanatory variables[6]

where the β coefficients refer to weighting factors, x_{ki} to explanatory variables indexed by k for the origin i, and cnst identifies the constant term of the regression model. The explanatory variables x_{ki} are considered in the standardised form z_k, which is computed according to (2.4), where μ_k is their mean value and σ_k their standard deviation. They are also known as mobility drivers and refer to geographical, economic and other regional indicators of the origin i. Based on observed travel diary data, which is provided, for example, by travel surveys, the regression coefficients β can be estimated.

Figure 2.1 shows the dependency of trip rates on the explanatory variables of income level and degree of urbanisation at the point of origin.[7] Both variables are measured on an ordinal scale and therefore must be transformed in order to be taken into account in the linear regression model (2.3). If correlated variables are to be included in the regression model, e. g. average income per capita and unemployment level, statistical methods such as factor analysis must be applied (see, for example, Bortz (2005); Everitt and Hothorn (2011); Lohninger (2012); Washington et al. (2004)).

Besides the generation of trip demand, the trip generation model frequently also computes the attraction of each travel zone, in order to estimate the number of trip endings D_j in the destination zone j. The attraction indicators are commonly distinguished by purpose, e. g. workplaces are particularly related to commuting trips and accommodation opportunities to vacation trips. It is important to note that a general transferability of the trip generation model between countries is not possible due to differences in lifestyles and transport systems (Schulz, 2006).

2.1.2.2 Trip distribution

The trip distribution model is the second sub-model of the four-step approach. It computes trip estimates T_{ij} for each pair of travel zones based on information on trip generation T_i in origin i, estimated trip endings D_j in destination j, and generalised cost c_{ij} of travelling according to (2.5):

$$T_{ij} = f(T_i, D_j, c_{ij}) \tag{2.5}$$

[6] Data is from the Dutch survey: Mobiliteitsonderzoek Nederland 2007 (MVW, 2010).

[7] For the mobility driver "income per capita", the pattern is quite comprehensible. For instance, the higher trip rates for the categories 5-7 can be largely explained by the higher demand of commuting trips. For the explanatory variable "urbanisation level of the origin", the pattern is comprehensible at second glance. It can be largely explained by an overlapping of two effects. The first effect is related to the demand of leisure trips. It is significantly lower for very highly urban areas and increases with decreasing urbanisation level. The second effect is related to the demand of shopping trips. It is almost at same level for all urban areas but significantly lower for rural areas.

where $f(..)$ refers most often to a gravity model.[8] The computed trip estimates T_{ij} are stored in the O/D trip matrix that is provided in the G/A format, as returning trips from the destination to the origin are not considered by (2.5). The trip distribution model is a bottleneck for the computation of the four-step model, given that the number of O/D relations stored in the matrix depends quadratically on the number of travel zones. This is of particular relevance for European transport models which are currently limited to operate at the level of NUTS-3 regions.

2.1.2.3 Modal split

The modal split model is the third sub-model of the four-step approach. It computes the market shares of transport modes for each O/D relation on the basis of information on the mode-specific travel cost. In this step, the trip matrix, which was computed by the trip distribution model in the previous step, is disaggregated into several mode-specific trip matrices. In many cases, a logistic regression model is applied for computing the market share P_A for the transport mode A which is better known as the logit model (2.6):

$$P_A = \frac{e^{-\mu C_A}}{\sum\limits_{m \in M} e^{-\mu C_m}} \quad \text{with} \quad C_m = \sum_k \beta_{km}\, x_{km} \tag{2.6}$$

where M identifies the set of transport modes (e. g. $M = \{\text{rail, road, air, coach}\}$), C_m the travel cost[9] related to the transport mode $m \in M$, and $\mu > 0$ is a heterogeneity parameter. Travel cost C is frequently computed according to a linear cost function, where x refers to specific cost indicators k, e. g. travel time and fuel cost, which are weighted by the β coefficients. The concept of cost functions is further discussed in section 3.2.

The heterogeneity parameter μ determines the sensitivity of the logit model to cost differences $(C_\text{A} - C_\text{B})$. It is commonly set to 1 for estimating the logit model, or respectively the β coefficients which weight the cost indicators x (see Ben-Akiva and Lerman, 1985, p. 107). However, the assumption $\mu = 1$ is a crucial model restriction, which disables the transferability of the logit model between different cost levels. For this reason, the logit model (2.6) with μ set to 1 cannot be applied simultaneously for short, medium and long-distance trips but has to be extended.

The impact of the heterogeneity parameter μ on the model sensitivity can be illustrated when formulating the logit model for two transport modes (2.7) and when deriving its response curve $g_\mu(\Delta c)$ (2.8).

$$P_1 = \frac{e^{-\mu C_1}}{e^{-\mu C_1} + e^{-\mu C_2}} = \frac{1}{1 + e^{-\mu (C_2 - C_1)}} \tag{2.7}$$

$$g_\mu(\Delta c) = \frac{1}{1 + e^{-\mu\, \Delta c}} \quad \text{with} \quad \Delta c = (C_2 - C_1) \tag{2.8}$$

Figure 2.2 shows the response curve for three different values, $\mu = \{0.01, 1, 10\}$ where the y-axis shows the value of the function $g_\mu(\Delta c)$ and the x-axis Δc. Setting the heterogeneity parameter to 0.01, the logit model is almost insensitive to cost differences Δc, and it derives equal market shares for

[8] Generalised costs are discussed in detail in section 3.2, and the gravity model $f(..)$ in section 4.2.

[9] The travel cost C is computed by O/D relation and by purpose.

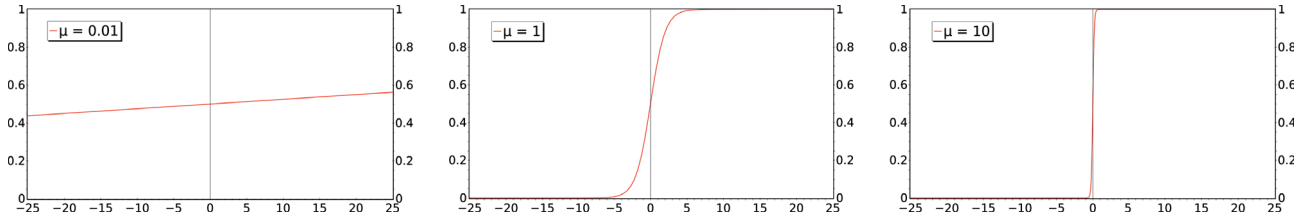

Figure 2.2: Response curve of the logit model for the heterogeneity parameters $\mu \in \{0.01, 1, 10\}$

both transport modes.[10] When we apply the default value $\mu = 1$ the logit model is sensitive to cost differences within the range $(-5, +5)$. When we apply $\mu = 10$ the logit model responds hypersensitively to any difference between the mode-specific travel costs and it computes a market share of 100 % for the cheaper transport mode. It is important to emphasise that these findings are also valid when the logit model is applied for more than two transport modes (see Ben-Akiva and Lerman, 1985, p. 106).

Given the need to apply the logit model at various cost levels in a European transport model, its theory and its mathematical derivation is investigated in detail in section 3.4. This section also discusses the derivation of the EMC (expected minimum cost) indicator, which enables the consistent integration of the trip distribution and the modal split model, and the derivation of the so-called Nested Multinomial Logit model. Besides the logit model, other models can be applied for the computation of the market shares such as the Kirchhoff model, which is sensitive to relative cost differences, and the Box-Cox model (Zumkeller, 2002). Since these models are not relevant for the development of the hierarchical trip distribution model, they are not further investigated within this thesis.

2.1.2.4 Network assignment

The network assignment model is the last sub-model of the four-step approach. For each transport mode, it derives the traffic volumes at the level of network links based on information on the routing of each trip. The most cost-effective routing for each O/D relation can be computed by shortest path algorithms, e. g. the Dijkstra's algorithm (Dijkstra, 1959). Information on the most cost-effective routing is also relevant for deriving the travel impedances for each O/D relation, e. g. travel time and travel distance. These travel impedance indicators are required to compute the generalised cost of travelling which is applied in the trip distribution and the modal split models. For this reason, the network assignment model is typically applied at least once prior to the actual application of the four-step transport model in order to compute the travel impedances for all O/D relations.

It is important to emphasise that in order to compute more accurate assignment results, in terms of traffic volumes at the level of network links, different routing options should be considered for each O/D relation. Given capacity constraints on the infrastructure, the most cost-effective routing varies depending on the current traffic loads and on the level of congestion. The European transport model VACLAV, for instance, follows an incremental approach where the network assignment is computed in several incremental steps (Szimba and Kraft, 2011). In the first step, the best routing is derived for all O/D relations, based on an unloaded network. In this step, a specific share of the trip volumes (e. g. 50%), which are reported by the trip matrix, is assigned to the level of network links. The cost

[10] The response curve in the left-hand image runs almost parallel to the x-axis ($g_\mu(\Delta c) \approx 0.5$) for $\Delta c \in (-25; +25)$.

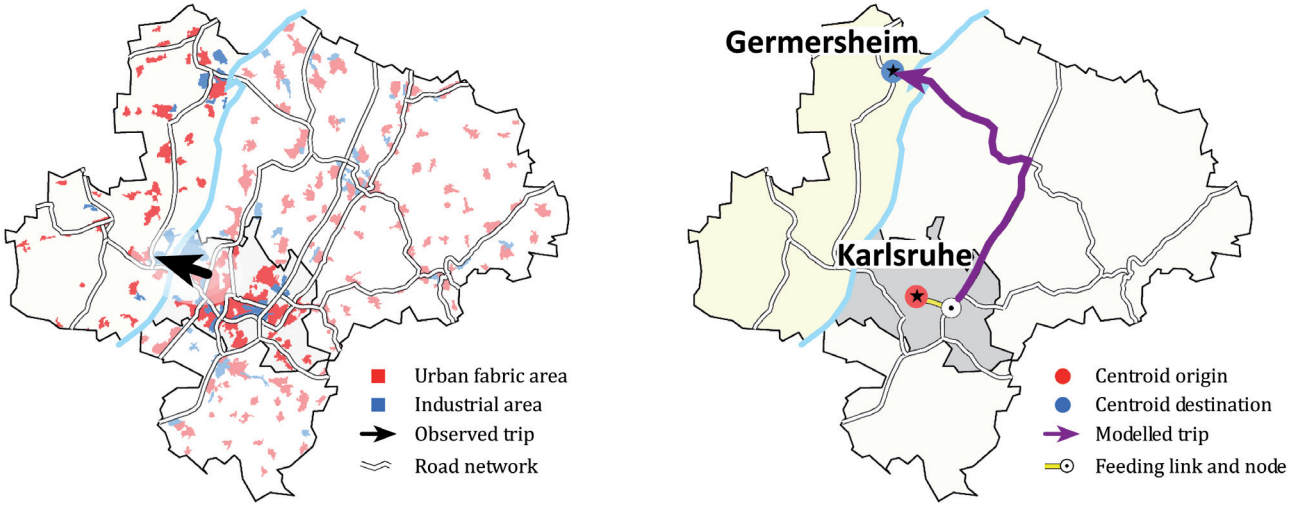

Figure 2.3: Lack of information due to the application of aggregated travel zones

information is then updated, based on the current network loads for the second step. Subsequently, the best routing is again computed for all O/D relations. In this step, another share of the reported trip volumes (e. g. 25%), which is smaller than the first one, is assigned to the level of network links. This procedure is repeated until the reported trip volumes are completely assigned.

Other approaches, which are not relevant for this thesis, include stochastic assignment models and route choice models. Stochastic models consider different perceptions that travellers have of the routing options by adding normally distributed random variables to specific travel impedance indicators. Following this approach, user-specific travel costs can be computed and accordingly different routing options can be considered within the assignment model. Route choice models can be formulated in a very similar way to mode choice models. Accordingly, the shares of the different routing options are frequently computed by a logit model.

2.1.3 Data and accuracy of aggregated transport models

Network-based transport models following macroscopic and mesoscopic approaches basically rely on the concepts of aggregated travel zones and network modelling graphs. They compute the number of trip estimates for each pair of travel zones and the vehicle flows for each network link, based on regional indicators for each travel zone and travel impedance indicators for each network link. Due to the application of aggregated travel zones, no information on the distribution of the regional indicators in the travel zones is provided. This shortcoming is particularly relevant for larger travel zones encompassing more than one city, given that these cities cannot be distinguished and separately modelled. All travel zones are represented by their respective centroids that are specific nodes in the network modelling graph. As all trips originate and end in these centroids, observed trips cannot be modelled one to one. In principle, this is an issue for all O/D relations, but it is particularly relevant for the modelling of trip estimates between neighbouring regions.

2.1.3.1 Limitation of accuracy

Figure 2.3 illustrates the loss of accuracy due to the application of aggregated travel zones. The left-hand image shows an observed commuting trip from the western district of the city of Karlsruhe (DE122) to an important industrial area in the region of Germersheim (DEB3E), which is located around 5 km outside Karlsruhe. The right-hand image shows how this trip is modelled in an aggregated transport model operating at NUTS-3 level. The trip starts in the city centre of Karlsruhe, which is the centroid of the origin travel zone. It passes by the feeding link and ends in the city centre of Germersheim, which is the centroid of the destination zone. The trip length is about 40 km. Due to the large distance between the two centroids, it can be expected that the transport model will significantly underestimate the trip demand between the two NUTS-3 regions Karlsruhe and Germersheim. The model accuracy can be improved, for instance, if the region of Germersheim is modelled by two travel zones. This example clearly points out the necessity to carefully determine the set of travel zones and centroids, taking into account the purpose and the scope of the transport model. If only long-distance trips over 100 km have to be modelled, for instance, NUTS-3-based travel zones are quite well suited.

2.1.3.2 Travel zones and centroids

The definition of travel zones often correlates with the definition of administrative units, given the availability of socio-economic, demographic and other indicators like population, GDP and the number of workplaces for these regions. The indicators are required to estimate the trip demand for each travel zone in terms of trip generation and trip endings within the trip generation model. In the transport model, each travel zone is represented by its centroid. If disaggregated information on the spatial distribution of the population and other relevant indicators is available, the location of the centroid C can be derived according to a generic approach which is outlined by (2.9):

$$C = \sum_k \omega_k S_k \qquad \text{with} \qquad \sum_k \omega_k = 1 \tag{2.9}$$

where S_k refers to the centre of gravity of a single indicator and ω_k to a weighting factor. If no disaggregated information on the spatial distribution of relevant indicators is available or if their distribution is too heterogeneous, e. g. if the travel zone covers several distant settlements, the centroid C has to be set manually, for instance, to the geographical centre of the largest settlement.

It has to be stressed that the geographic location of all centroids representing the travel zones, as well as their linkage to the network model, is of particular importance for the trip distribution model and the computation of trip estimates for each pair of travel zones. In the illustrated example (Fig. 2.3), the centroid representing the city of Karlsruhe is connected via one feeding link in an easterly direction to the network model. For this reason, the shortest routing from Karlsruhe to Germersheim first runs to the north before it turns to the west. Actually, most trips between these two travel zones use the southern route crossing the Rhine bridge between Wörth and Karlsruhe. The current transport model is therefore not suited for explaining the traffic loads on the Rhine bridge.[11] The model can be

[11] This is quite interesting to note, as the extension of the existing Rhine bridge, or the building of a second bridge, is an important transport infrastructure project, which has been under discussion for several years. Several case studies and cost-benefit analyses have been carried out (see, for example, Schick and Wieczorek (2008); Schulze and Waßmuth (2009); Ellinghaus et al. (2017)) and it can be expected that the applied transport models were much more detailed than the one outlined in Figure 2.3. However, for the assessment of this infrastructure project, the trans-European passenger

improved if the centroid of Karlsruhe is connected by a second feeding link in a westerly direction to the network model. In this case, the southern route between Karlsruhe and Germersheim would be faster.

2.1.3.3 Network model

A network modelling graph consists of nodes and links and can be compiled on the basis of cartographic maps, for instance, which show the rail and the road network infrastructure. The nodes are connected by links and can refer to crossings, starting and ending points, where each link provides specific attributes. For road network models, these attributes typically comprise travel impedance indicators such as information on number of lanes, link length and speed limit, as well as observed transport demand indicators like the so-called AADT value[12], and general information on the network link, e. g. the name of the motorway. The travel impedance indicators are important for computing the travel impedances and accordingly for deriving the generalised cost of travelling for each pair of travel zones. In general, several mode-specific network models, which can be connected by specific nodes, are considered in a transport model. These nodes can refer to inter-modal terminals, for instance.

2.1.3.4 Feeding links and feeding nodes

The feeding links and the feeding nodes identify specific elements of a network model, with each feeding link connecting the centroid of a travel zone to the so-called feeding node. For European rail and road transport network models, feeding nodes often refer to the most important motorway exit in a travel zone and to the most important railway station.[13] If possible, the feeding node should be located near the centroid, since the feeding link is also taken into consideration for the computation of the generalised cost of travelling between two regions. Frequently, feeding links are used in order to model the so-called access/egress (A/E) cost for accessing and egressing the main transport mode, e. g. the bus trip to the railway station. Feeding links are often derived based on a secondary network infrastructure model, but they can also be determined arbitrarily.

It is important to note that all originating and ending trips pass by the feeding link of a region. For this reason, the modelled traffic loads, which are computed by the assignment model, are commonly much overestimated in the feeding links and their adjacent links. The validation and interpretation of the assignment results are therefore often challenging and require a basic understanding of the transport model. The overestimation of modelled traffic loads for the feeding links can be corrected by following a so-called multi-feeder approach, which is also implemented in the European network-based transport model VACLAV (Schoch, 2004).

---

and freight transport flows also have to be considered due to possible route-changing effects. In regional transport models with a limited scope, changes of transit flows are commonly modelled exogenously based on assumptions, e. g. by applying growth factors. It should be emphasised that, besides these assumptions, the transport modeller can also have significant influence on the results of a study by determining the set of travel zones and centroids.

[12] The annual average daily traffic (AADT) value refers to the number of vehicles using the road segment on an average day. AADT values are often provided at a disaggregated level, e. g. several vehicle types are distinguished.

[13] In the right-hand image of Figure 2.3, the feeding link is outlined in yellow for the travel zone Karlsruhe connecting the city centre and the motorway junction on the east side.

2.1.3.5 Travel impedances

Travel impedance indicators are a crucial input for the two sub-models trip distribution and modal split. They are determined for each pair of travel zones and for each mode of transport on the basis of network modelling graphs and the application of shortest path algorithms. In most cases, only the shortest path between origin and destination is considered for deriving the travel impedances and, accordingly, the generalised cost of travelling. In transport literature, more sophisticated approaches for deriving generalised costs are also discussed. The approach outlined by Kato et al. (2003), for instance, for estimating the user's benefit from infrastructure investments takes into account any possible routing between two regions, and not only the shortest one. Their approach relies on the observation that the perceived costs of travelling between two regions are lower than the costs resulting from the best routing, provided that several options are available.

Travel impedances include cost indicators like travel time, travel distance, toll costs and A/E costs, but also Level of Service (LoS) indicators that are related to the quality of a transport mode, e. g. the frequency of train connections between two regions per day as well as the reliability and safety of a transport mode.

Particular attention has to be paid to the computation of the cost indicator travel time for road transport, since it varies depending on the level of congestion on the road network infrastructure. For the computation of the indicator travel time, so-called speed-flow functions (FSV, 1997) are applied at link level. The functions provide estimates on the current travelling speed, when given information on the current traffic load, the maximum link capacity and the maximum speed limit. For the modelling of long-distance trips at European scale, commonly only one set of travel impedances, referring to an average pre-loaded network model, is considered. In order to improve the accuracy of regional and urban transport models, it may be adequate to consider different datasets of travel impedances distinguishing between peak and off-peak hours. The transportation model MOSART[14] follows a similar approach. It distinguishes between several periods of the day, e. g. morning, afternoon and off-peak hours, to produce traffic predictions and assignments (Bonnafous et al., 2009).

Besides the relevance of travel impedance indicators in computing the generalised cost of travelling, they also facilitate the direct conversion of trip volumes into passenger-kilometres, measuring the transport performance, based on the O/D trip matrix, without having to apply the assignment model. It should be stressed, however, that the computed impedance indicator travel distance, which refers to the shortest path between the two centroids, is commonly higher than the actual average length of the trips that occur between two travel zones. The discrepancy between these two indicators is quite obvious for the illustrated example (Fig. 2.3), where travel distance is much higher for the modelled trip than for almost any other possible trip occurrences between the two travel zones.

[14] MOSART (French acronym for Modelling and Simulating Accessibility to Networks and Territories) is applied for the French agglomeration area "Grand Lyon" and supports decision makers in the planning of transport policies.

2.2 Application of European transport models

Passenger demand transport models following the four-step approach are basically computing a set of transport demand indicators including the trip matrices and the network assignment. They are sensitive to a set of input indicators like population, GDP and travel impedances, as well as to a set of modelling coefficients. Specific coefficients of this set reflect, for instance, the elasticity of travellers to changing travel costs. Through the variation of specific input indicators, different types of policy scenarios can be simulated. The transport model is then applied again, and the effects can be investigated and assessed. This facilitates, for instance, the investigation and prioritisation of planned infrastructure projects (Szimba, 2008) and the identification of critical infrastructures (Schulz, 2012). In addition, external effects of transport can be derived from modelled transport demand indicators (Schmedding, 2006).

The following sections briefly shed light on two exemplary European transport models that were intensively studied within this thesis:[15]

- the VACLAV model; and
- the TRANS-TOOLS model.

For both models the source code was available, making it easier to trace back model reactions to the level of equations and to gain in-depth expertise on interdependencies between input indicators, modelling coefficients and computed output indicators in the four-step modelling approach.

2.2.1 The VACLAV model

VACLAV is a European network-based passenger transport model operating at NUTS-3 level that has been developed since the late 1990s at KIT (Eberhard et al., 1998). It uses two complex network models for rail and road transport that can be modified to simulate network changes, pricing and other policies. VACLAV has been applied in many European projects and corridor studies. It has been updated and extended several times (e. g. to new zoning systems and base years, to new purposes and transport modes). Within the ETISplus project, it was updated to the base year 2010 (Szimba and Kraft, 2011). The model covers Europe and neighbouring countries with approximately 1600 travel zones.

The most recent research at KIT concerns the linkage of VACLAV with the high-level strategic transport model HIGH-TOOL in order to improve European transport policy assessment. Due to their different approaches, both models have different advantages and disadvantages. The HIGH-TOOL model, for instance, encompasses demographic, economic and freight models but operates only at NUTS-2 level. In addition, the transport infrastructure is modelled only at a very strategic level via the hypernet sub-model. In contrast, the VACLAV model relies on very detailed networks and operates at NUTS-3 level but lacks detailed demographic, economic and freight modelling. The current thesis work of Kraft (2017) aims to combine and integrate both models.

The VACLAV model largely follows the four-step approach in which trip generation and trip distribution are computed in a single step by a so-called direct demand model. The market shares of transport

[15] Discussion of their results is avoided. Their understanding requires deeper insights into the rather complex mathematical frameworks underlying both models.

modes are computed by an enhanced multinomial logit model in order to overcome unrealistic properties of the classical logit model. The methodological focus of VACLAV is the network assignment model, that follows an incremental approach as described in section 2.1.2. A major strength of the assignment model is to handle intra-zonal trips.[16] This is facilitated by subdividing each NUTS-3 region into several so-called access zones and by disaggregating the trip matrices to the level of the access zones. Hence, the network assignment model is virtually applied at the level of access zones. Besides intra-zonal trips, inter-zonal trips are also assigned at the level of access zones.
Freight transport is not modelled explicitly by VACLAV, but is covered by exogenous trip matrices, which are added for the network assignment step. This improves the quality of the road network assignment results, in particular as trucks make a significant contribution to congestion.

The runtime of VACLAV depends on the number of incremental steps and the number of available CPUs. The computation of the network assignment for rail and road passenger transport, on the basis of on pre-compiled trip matrices, for instance, took around three hours within the ETISplus project. In this specific case, VACLAV was applied on eight CPUs at the same time and was carried out in ten incremental steps. A specific limitation of VACLAV is the direct demand model, which follows a rather basic methodology. However, this methodology is not perfectly suited for the computation of the trip matrices from scratch. Instead, VACLAV is usually re-calibrated, based on a given trip matrix for a base year. It then computes the changes for a specific policy scenario in terms of trip generation, trip distribution and modal split. Based on this information, it derives the new trip matrices, which are assigned to the networks. In this way, VACLAV was applied for passenger demand modelling and impact assessment in many European projects, including:

- TRANS-TOOLS v1, in which it was combined with the ASTRA model;[17]
- TINA-Turkey, in which the trip matrices were computed by another model (TINA et al., 2007);
- ETISplus, in which it was complemented by the PAT model (see sec. 2.3.4); and
- Rail Baltica, in which the ETIS trip matrices were used (COWI et al., 2006).

2.2.2 The TRANS-TOOLS model

TRANS-TOOLS, which is an English acronym for TOOLS for TRansport Forecasting ANd Scenario testing, is a European network-based transport model. It has been developed and continuously improved since 2004, for application by the European Commission as a reference model for analysing policy questions and for impact assessment at regional level (NUTS-3). The first version was built upon several existing European models and calibrated to the base year 2000 (Burgess et al., 2008). The second version was calibrated to the base year 2005 (Petersen et al., 2009) and the third version is currently being updated, based on 2010 data (Berglund and Algers, 2016). Besides the base year, the methodology was continuously updated to extend the geographical scope, to implement additional features and to overcome identified shortcomings like the very long runtime (Ibañez-Rivas, 2010).

[16] This feature was developed within the thesis work of Schoch (2004), which was carried out at KIT.

[17] The development of the ASTRA model started initially at KIT (Schade, 2005). The model was updated several times and is recently applied by the European Commission for strategic policy analysis and impact assessment (Krail and Schade, 2014). Following a System Dynamics approach, and encompassing no modifiable network models, the methodology of ASTRA differs significantly from TRANS-TOOLS and VACLAV.

The model covers passenger and freight transport, detailed network models for all modes of transport, an economic and a trade model, an environmental model and impact models (Nielsen, 2009). The components were developed in different programming languages and integrated into the ArcGIS software platform. Through the ArcGIS user interface, the user can implement changes in the transport networks as well as in logistics and distribution systems. In addition, the user can execute the model, display and analyse the results.

Within this thesis, the passenger demand transport model that comes with version 2.1.9 of TRANS-TOOLS was analysed. It consists of a trip generation, a trip distribution and a modal split step that are implemented as stored procedures in an MS SQL database. The assignment step is complemented by another model that is implemented as an ArcGIS toolbox. The passenger model covers 42 countries and 1441 NUTS-3 regions. Short-distance trips (<100 km) and long-distance trips (>100 km) are treated differently. Short-distance trips are categorised as rail, bus, metro, car driver and car passenger, and long-distance trips as train, bus, air and car.

The trip generation step applies a multinomial logit model for estimating trip frequencies and accordingly the number of trips. The logit model is sensitive to economic and transport supply indicators and was estimated based on observed travel data from Denmark. Issues of the chosen approach are related to the behavioural interpretation of the logit model and the projection of Danish mobility behaviour to other countries (cf. Monzón et al., 2010). Trip distribution and modal split are modelled by a nested logit model. The runtime of the passenger model including the first three steps is 2h45.

The assignment model follows an excessively ambitious approach by computing a stochastic user equilibrium (SUE) that has several advantages compared to other approaches. However, SUE is very time demanding (the assignment takes 25 hours) and leads to the problem of stochastic variations of the results. This makes the comparison of different policy scenarios difficult (cf. Monzón et al., 2010).

2.2.3 Other transport models

Besides the VACLAV and the TRANS-TOOLS models, further transport models were investigated based on the provided model documentations in order to gain further experience. The so-called intercity generation-distribution (GENDIS) model should be mentioned in this context (cf. Gaudry et al. (1994a); Gaudry et al. (1994b)). The model relies on an innovative approach in which an autoregressive, contiguous distributed process was used for modelling the competition between travellers.[18] Within an early stage of this thesis, it was investigated whether the GENDIS model could be transferred to Europe. This was, however, not possible, as the autoregressive component underlying the GENDIS model relies on rather specific assumptions that were not documented in detail. For this reason, the competition between travellers will be modelled based on an accessibility measure in the IPAT and HIPAT models that will be discussed in section 3.3.

[18] Autoregressive (AR) models are a specific type of statistical model. They are often applied in time series analysis and in image data analysis in order to describe time-varying and space-varying patterns in the data. However, AR models require a certain degree of homogeneity of the data to be analysed. Their application in the context of the trip distribution model is therefore very challenging, given the large degree of heterogeneity of the O/D relations to be analysed for modelling the competition between travellers.

2.3 The ETISplus project and the PAT model

ETISplus was a research project from 2009 to 2012 funded by the European Commission (DG Mobility & Transport) under the 7th Framework Programme (Newton et al., 2013) in order to update and extend the existing European Transport Policy Information System (ETIS). This system was developed within the forerunner project ETIS-BASE to which KIT-IWW already contributed (NEA et al., 2005). The ETIS system provides a reference dataset covering transport supply and transport demand indicators at a regional level for member states of the European Union and for neighbouring countries.

KIT was involved in the development of the passenger demand data set and the coordination of the work package that dealt with socio-economic and supply indicators, networks, and external effects of transport. Within this thesis, the socio-economic dataset[19] collected within the ETISplus project was reviewed. All data gaps were closed in order to establish a complete and consistent input database for the base year 2010, covering socio-economic and demographic indicators at NUTS-3 level for member states of the European Union and for neighbouring countries. The compilation of this input database was a crucial prerequisite for the development of the PAT model that was applied for the computation of the ETISplus trip matrices covering inter-zonal rail and road passenger trips (Ihrig, 2012). Later on, the compiled input database was further required for the IPAT and the HIPAT models.

This chapter centres around the following topics: section 2.3.1 provides a general overview of the scope and the objectives of the ETISplus project. The subsequent section describes the process of constructing the supply indicator datasets including the network models, determinant and indicator data. Section 2.3.3 outlines the construction of the input database, which is required for the application of the passenger model, and section 2.3.4 describes the approach of the PAT model, its calibration and specific results. The last section then evaluates the PAT model with regard to the advantages and drawbacks of the implemented approach, and summarises the general findings, which are of particular relevance for the elaboration of the central research questions of this thesis.

2.3.1 Scope and objectives of ETISplus

The main objectives of the ETISplus project were the collection, consolidation and/or computation of data in order to establish a new reference data set for transport modelling and for European Transport Policy. Particular attention was paid to the support of the TRANS-TOOLS model and the user requirements as defined in several framework projects. The data covered belonged to the domains of socio-economic data, transport demand, transport networks, level-of-service and external impacts. Both passenger and freight transport were in the scope of ETISplus. Transport demand was dealt with at the level of O/D relations (NUTS-3 level) as well as at the level of network flows. The geographical scope of the databases encompassed – apart from EU countries – all accession countries, neighbouring countries and, for some data sets, links to the rest of the world. In addition to the approach followed within the forerunner ETIS project, several innovations were realised in ETISplus, such as the use of intelligent transport systems to provide data feeds, the development of a business model in order to make the system self-supporting and the introduction of a specific data validation procedure.[20]

[19] See: Siegele et al. (2012); Schimke et al. (2012).

[20] The project description is largely taken from the KIT's website, URL: `http://netze.econ.kit.edu/102_186.php` (Retrieved October 10, 2017).

The ETISplus project was subdivided into 13 work packages (WP) which can be grouped into three blocks, namely Communication and Management, Innovation and Design, and Building the Database (Newton, 2010, p. 59f). The first group, Communication and Management, dealt with the dissemination of results, the organisation of thematic workshops, the involvement of stakeholders, and overall project management activities.

Within the second group, Innovation and Design, several innovative data collection methods were investigated, including the GPS tracking of vehicles and travellers to gather detailed routing information, the exploitation of comprehensive data sources like OpenStreetMaps (OSM) and cost-effective methods for accessing and exploiting transport demand and supply indicators, notably in the non-EU member states. Several pilots were launched to evaluate these innovative concepts. Central outputs of Innovation and Design were determination of the database structures and the two zoning systems for the base years 2005 and 2010.

Work packages in the third group, Building the Database, dealt with the collection, computation and validation of transport demand and supply data for all modes of passenger and freight transport at regional level, in accordance with the predetermined database structures and the regions specified by the two zoning systems. KIT led WP7 and developed the PAT model which was applied for the computation of the G/A trip matrices in WP9.

2.3.2 Networks, determinant and indicator data

WP7 dealt with socio-economic data, transportation network data, passenger and freight level-of-service (LoS) data, impedance data and the assignment of external costs of transport. The datasets were generated for two base years, 2005 and 2010. KIT coordinated this work package and was responsible for the compilation of the socio-economic dataset, the network models, and the computation of LoS and impedance data for rail and road passenger transport. These datasets were of particular importance and fed into the development of the PAT model. Further tasks which were carried out at KIT included the assignment of external costs of transport and the compilation of a dataset comprising rail terminals for freight transport.

2.3.2.1 Socio-economic data set

The socio-economic data set covers approximately 1,600 regions in 42 countries and was compiled for two base years. For 2005, the regions correspond to the NUTS-2003 zoning system (Siegele et al., 2012) and for 2010 to the NUTS-2006 zoning system (Schimke et al., 2012).[21] For the non-EU member states not covered by the NUTS zoning system, regions were defined according to the classification that was already applied in the WORLDNET project (NEA et al., 2009). The main challenges to the compilation of the socio-economic dataset were related to the application of two different zoning systems for the base years 2005 and 2010. A particular challenge was the data collection of indicators for 2010, as these indicators had not been published completely by statistical offices during the project runtime. Missing indicators, i. e. data gaps, had to be filled by other means. The final socio-economic dataset comprises 65 indicators for 2005 and 66 indicators for 2010. These indicators are provided at different

[21] The NUTS classification distinguishes the four regional levels NUTS-0 to NUTS-3. NUTS-0 level refers to countries and NUTS-3 level refers to regional administrative units (e. g. cities).

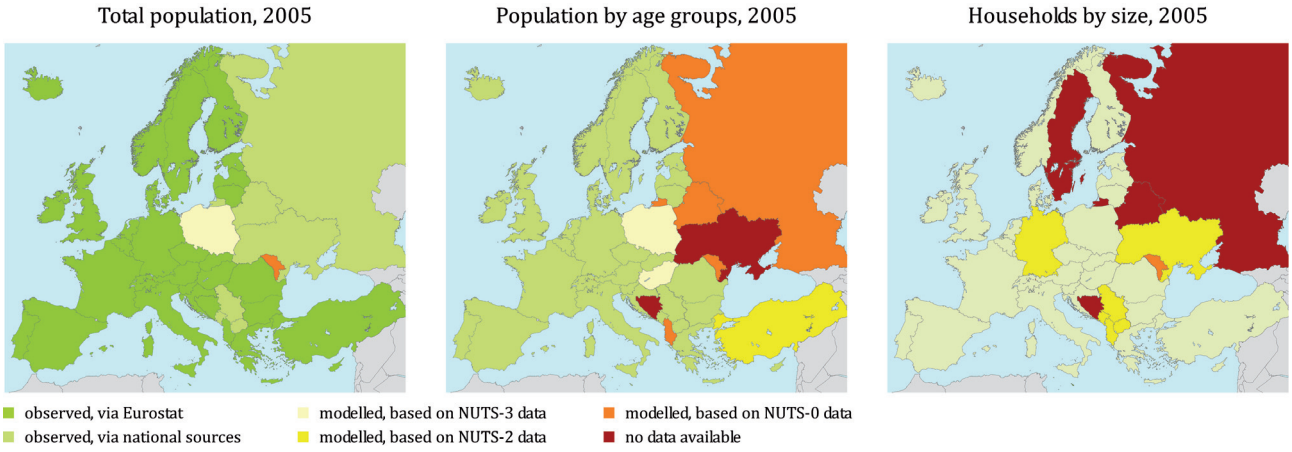

Figure 2.4: Primary source of three selected NUTS-3 indicators by country

levels of detail: for example, population is distinguished by age group and sex, while GDP is provided only in totals.

The regional database of the Statistical Office of the European Union (Eurostat) was used as a primary source. Data was also collected from national and regional statistical offices. Additional data sources were exploited in order to fill data gaps. In principle, the data availability was very heterogeneous and many NUTS-3 indicators, which are included in the ETISplus dataset, had to be modelled based on available data.

Figure 2.4 outlines the primary source of three selected indicators referring to the base year 2005, which are included in the ETISplus dataset.[22] For the aggregated indicator "total population by NUTS-3 region", data availability was given by Eurostat and by national sources for almost all countries. For Poland, the indicator depended on the "Population and Housing Census 2002", and for Moldova it could be exploited only at NUTS-0 level through the "National statistical office of Moldova". The indicator "population by age groups by NUTS-3 region" was not available at Eurostat for any country except Malta. It could be extracted from national sources for many countries but it had to be modelled as well, based on available NUTS-3, NUTS-2 or NUTS-0 data. For some countries it was not available. The indicator "households by size" was not available at NUTS-3 level. It could be modelled for most countries based on available census data from other years. For several countries the indicator was not available.

As a general rule it was found that disaggregated indicators are less easily available than aggregated indicators. Belgium, for instance, distinguished the indicator GVA by 60 economic sectors at national level, but only by three sectors at NUTS-3 level. This is very likely due to data protection regulations. It was also found that tables contained frequent data gaps at the regional level. However, data gaps also appear at the country level. A possible explanation for missing indicators is changes in the zoning system. For the federal state of Sachsen-Anhalt in Germany, for instance, indicators for 2005 could not be accessed in the required NUTS-2003 data format but only in the NUTS-2006 data format. In such cases, if an indicator could be exploited neither by Eurostat nor by national and regional sources,

[22]The information was extracted from Siegele et al. (2012), chap. 4, p. 18-41.

several methodical approaches for filling data gaps were applied including disaggregation, forecasting, population-weighted extrapolation and tailored approaches.

Disaggregation-based approaches were considered for the computation of indicators like GVA at NUTS-3 level that had to be broken down into the sectors agriculture, industry and services. Given the non-availability of this particular indicator, it was computed based on available information on the indicators of total GVA at NUTS-3 level and GVA by sector at NUTS-2 level according to (2.10) where $\hat{x}$ is the required indicator for sector s and region i, x the aggregated indicator at a regional level (e. g. total GVA at NUTS-3 level) and $\tilde{x}_s$ the detailed indicator for sector s at an aggregated level (e. g. GVA by sector at NUTS-2 level).[23]

$$\hat{x}_{i,s} = x_i \, \omega_s \qquad \text{with} \qquad \omega_s = \frac{\tilde{x}_s}{\sum\limits_s \tilde{x}_s} \tag{2.10}$$

Forecasting-based approaches were used for indicators that were not available for the required base year but for two other years. Given their availability for the two previous years t-1 and t-2, the required indicator $\hat{x}$ could be computed according to (2.11). It should be noted that forecasting relies on the assumption of constant framework conditions, which is an important prerequisite for the extrapolation of an observed trend to future years. Thus, the impact of events like the financial crisis in 2008 have to be carefully considered for forecasting approaches.

$$\hat{x}_t = x_{t\text{-}1} \, \frac{x_{t\text{-}1}}{x_{t\text{-}2}} \tag{2.11}$$

For missing indicators which were strongly correlated with the population, like the number of cars in a NUTS-3 region, population-weighted approaches were followed. In these cases, the indicator $\hat{x}_i$ for region i was computed according to (2.12) where pop_i is the population at regional level, and X is the required indicator at an aggregated spatial level, e. g. the number of cars at country level.

$$\hat{x}_i = X \, \omega_i \qquad \text{with} \qquad \omega_i = \frac{\mathrm{pop}_i}{\sum\limits_s \mathrm{pop}_i} \tag{2.12}$$

Tailored approaches were considered for filling data gaps related to the indicators “unemployment rate” and “the number of persons living in households” as well as for converting indicators between the two zoning systems NUTS-2003 and NUTS-2006. The final socio-economic dataset, which was produced within WP7, still comprises data gaps, since it was not always possible to elaborate a satisfactory approach for estimating a missing indicator. Thus, the activities to complete the dataset were continued in WP9 when compiling the input database for the passenger model.

2.3.2.2 Network data set

The second target of WP7 was the compilation of the new reference network models for all modes of transport including passenger and freight transport. To ensure their consistency and backwards-

[23] For some countries, GVA by sector was only available at NUTS-0 level.

compatibility, the structure developed in the TRANS-TOOLS and the WORLDNET projects was largely maintained. At KIT, the road and rail network models were updated. Both models relied on the corresponding TRANS-TOOLS v2 network models published in autumn 2009 (Szimba and Kraft, 2012b) by the Institute for Prospective Technological Studies (JRC-IPTS), which is a Joint Research Centre of the European Commission.

The main objectives were to improve these existing network models, to extend the list of link attributes provided, to improve the availability of traffic count data at the level of network links, and to improve the alignment of the network model with regard to the real infrastructure. For the rail network model, the implementation of new high-speed lines and the merging of the two separate network models dealing solely with rail passenger or rail freight transport were further objectives.

The update of the network models was carried out under consideration of a wide range of different data sources including the TRANS-TOOLS v2, TEN-T, and ETIS-BASE network models as well as UN/ECE and national traffic count data, OpenStreetMaps, and specific sources like the Maps of Railway Network by Büker[24] (Szimba and Kraft, 2012a). Most of the updates were implemented in the network models by hand in order to correct existing errors. For instance, one of the considered input network models showed non-existent level crossings between Alpine passes and tunnels. A possible explanation for this issue could be that an unobserved algorithm was applied when compiling this network model, inserting networks nodes for overlapping links (i. e. crossings). Another reason for the manual implementation of updates was related to the possibility of working out tailored solutions for some regions, e. g. for Paris the network density was optimised and a specific solution for connecting the railway stations was elaborated.

The final dataset which is provided by ETISplus consists of the road network model and the rail network model. The road network model covers 64 countries and comprises 53 indicators, and the rail network model covers 58 countries and comprises 35 indicators. Both network models were then implemented at KIT in the VACLAV model in order to derive the travel impedances for each O/D relation and to derive the traffic loads at the level of network links within the assignment step of the passenger model.

2.3.2.3 Level-of-Service and travel impedance data set

The Level-of-Service dataset comprises several impedance variables such as travel time, travel distance and user costs for each O/D relation. At KIT, the impedance matrices for rail and road transport were computed by the VACLAV model, based on the previously compiled network models. In the course of the implementation of the network models in VACLAV, the feeding nodes and the feeding links were determined and implemented in the rail and in the road network models.[25] The final impedance matrix comprises approximately 2.6 million O/D relations, 17 different impedance variables for road transport and 11 for rail transport.

The computation of the impedance variables was mainly carried by use of the Dijkstra algorithm for identifying the fastest path in an uncongested network for each O/D relation (Szimba and Kraft, 2012c). For the computation of the two impedance variables rail travel time and frequency of train

[24] URL: `www.bueker.net/trainspotting/maps.php` (Retrieved October 10, 2017).
[25] A feeding link connects the centroid of a travel zone to the feeding node in the network model (cf. sec. 2.1.3).

connections, and for the computation of the travel fares for rail passenger transport, timetable-based information from the "HaCon Fahrplan-Auskunfts-System" (HAFAS) software published by Deutsche Bahn was used (Szimba et al., 2012).

2.3.3 Compiling the input database for the PAT model

The availability of a comprehensive and consistent input database tailored to the needs of the PAT model was a crucial prerequisite for its development. The socio-economic dataset previously produced within WP7 was used as a basis, but given several data gaps further effort was necessary to produce this database. Moreover, the data format required by the PAT model was not in line with the format provided by the socio-economic dataset. The input database basically comprises mobility determinants at a national level (e. g. annual trips per capita), socio-economic and other indicators at a regional level, and travel impedances for each pair of travel zones.

During the construction of this input database several challenges had to be faced, including the definition of specific mobility determinants under consideration of the objectives and data requirements of the PAT model, the collection and harmonisation of mobility indicators from travel surveys and transport statistics in order to derive the required mobility determinants, and the completion of the socio-economic dataset. Further challenges were posed by the consideration of four different trip purposes (business, commuting, private and vacation), two different transport modes and two different zoning-systems, which had to be included in the input database.

The mobility determinants were derived at national level based on the evaluation and harmonisation of various sources, including national travel surveys like the National Travel Survey (NTS) of the UK (DfT, 2008) and the German Mobility Panel (MOP) (Zumkeller et al., 2007), transport statistics like EU energy transport in figures (EC, 2009), transport studies like iTREN-2030 (Fiorello et al., 2009) and Mobility in German cities (Ahrens et al., 2009) as well as other sources like the International Union of Railways (UIC) and the United Nations Economic Commission for Europe (UN/ECE). The objective of this comprehensive analysis was to obtain a general overview of the availability of mobility indicators in the EU countries. The two most important mobility indicators, which were collected and harmonised at country level, referred to the annual number of trips per capita and the annual mileage travelled per capita. These two indicators were further distinguished by trip purpose and by transport mode. In addition to these basic mobility indicators, specific coefficients were derived at the country level for modelling the relationship between the transport demand and the urbanisation degree of a region and the correlation between the income level and the transport demand. The mobility indicators and the coefficients were applied in the first three steps of the passenger model, i. e. trip generation, trip distribution, and modal split.

A general finding of this analysis was that only few EU countries provided comprehensive travel surveys that could be considered for the development of the passenger model. The surveys used often differed in terms of their definitions with regard to trip purposes, transport modes, and age groups. For this reason, the development of the input database had to be carried out in three steps. In the first step, available data sources were exploited. In the second step, harmonised indicators were derived according to the ETISplus definitions of transport modes and trip purposes, and in the last step, missing indicators were estimated, e. g. by transferring indicators from one country to another, taking into account the difference in per-capita income.

Besides mobility determinants, the input database also comprises indicators at a regional level and travel impedances for each pair of travel zones. The indicators were derived based on the corresponding ETISplus datasets which were constructed within WP7 (cf. sec. 2.3.2). The regional indicators included demographic indicators like population by age group and gender, economic indicators like GDP in purchasing-power standard (PPS), tourism indicators like the number of hotels, and other indicators like the area of a region. Travel impedances included indicators like travel time, travel distance and user cost for each O/D relation, as well as indicators at the national level like occupancy rates.

A specific challenge was related to the closing of existing data gaps, which was an important prerequisite for the application of the passenger model. Whenever it was not possible to further complement the data sources already exploited within WP7, proxy regions or proxy countries were defined for the estimation of missing indicators. This approach was, for instance, applied for filling gaps related to the indicators motorisation level and population by age group and gender.

2.3.4 Implementing the PAT model

The PAT model was developed in order to compute the trip matrices covering inter-zonal rail and road passenger trips at NUTS-3 level, which are part of the ETISplus passenger demand data set. This data set was generated within WP9. Its structure was developed based on the existing ETIS-BASE database, with increased emphasis on air transport, public transport, coach transport and short-distance transport (i. e. intra-zonal trips). The main features of the ETISplus passenger demand data set were the so-called O/D trip matrices covering road, rail and air passenger transport as well as the allocation of intra-zonal trip demand by distance band and transport mode.

Besides the PAT model (Ihrig, 2012), two other tailored models were developed. These models dealt specifically with the demand of intra-zonal trips (TRT, 2012) and the demand of air passenger trips (MKmetric, 2012).[26] In order to ensure the consistency between the three tailored models, the scope of the PAT model was augmented. In the first step, trip generation, PAT estimated the overall transport demand, including slow modes and motorised modes, intra-zonal and inter-zonal transport. In the subsequent step, the PAT model estimated the demand of intra-zonal trips and provided these results to the other model dealing with intra-zonal trips. The demand of air passenger flights was excluded from the PAT model's trip generation step.

2.3.4.1 General methodology

The methodology of the PAT model follows the classical four-step approach. The first two sub-models, trip generation and trip distribution, were developed from scratch within this thesis. For the computation of market shares, a tailored approach was developed based on the existing VACLAV model which was also applied for the computation of the network assignment. The main outcomes of the PAT model were the so-called G/A trip matrices by mode and purpose. These matrices were computed within the first three sub-models. Figure 2.5 outlines the structure of the PAT model including the executing order of the four sub-models, data exchange with the database and implemented feedback loops for calibrating the overall model.

[26] For intra-zonal trips additional transport modes were considered: bus and coach, tram and metro, cycling and walking. These trips were distinguished by four distance bands.

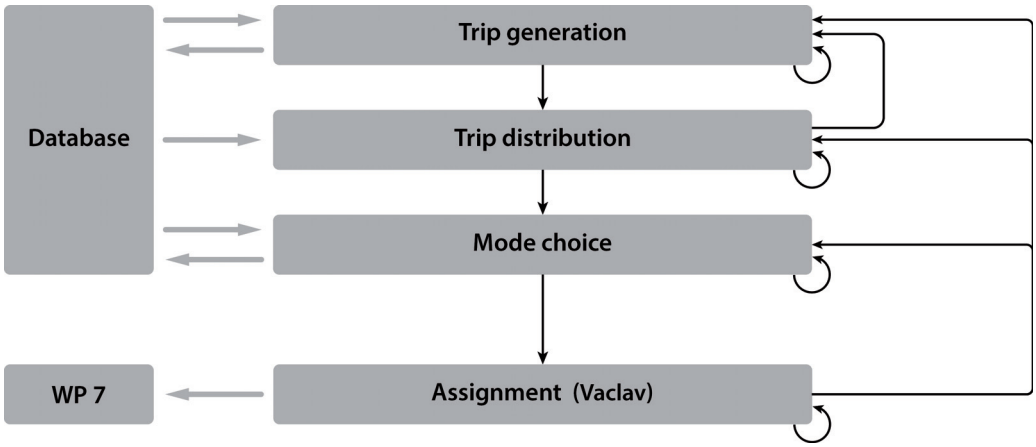

Figure 2.5: Structure of the PAT model and data flow

Additional outcomes included trip estimates for the intra-zonal demand, which were further processed by another model, and the network assignment results. In order to validate and to calibrate the PAT model, the computed network assignment was compared to observed traffic count data. Furthermore, the network assignment fed back into WP7 and was further processed for calculation of the external costs of transport. The key features of the PAT model were as follows:

- the computation of G/A trip matrices by mode and purpose;
- two modes: rail and road transport;
- four trip purposes: business, commuting, private and vacation;
- two base years: 2005 and 2010;
- 1603 travel zones for 2005 and 1622 for 2010;
- the computation of trip estimates for intra-zonal passenger demand including slow modes; and
- the computation of assignment results for road and rail passenger transport.

The development and calibration of the PAT model faced several challenges, mainly related to the heterogeneous size of the travel zones, which varied between a few and several million square kilometres[27], the application of the two different zoning systems NUTS-2003 and NUTS-2006 for the two base years 2005 and 2010[28], and the very limited availability of observed travel demand data.

2.3.4.2 Trip generation

The first sub-model, trip generation, had four objectives: the computation of the trip demand for rail and road passenger transport, the distinction between intra-zonal and inter-zonal trip demand, the estimation of trip endings, and the estimation of the trip demand by slow modes. All indicators were distinguished by purpose and computed for each travel zone.

[27]The smallest travel zone in the PAT model, Melilla (ES640), covers an area of 13.4 km^2. The largest, Norrbottens län (SE332), covers an area of 106,011.5 km^2.

[28]For this reason, the modelled trip matrices for 2005 and 2010 could not be compared directly.

Building on (2.13) to (2.15), the trip demand T_i by rail and road transport was computed for each travel zone i where a zone-specific trip rate factor $\hat{R}_i$ was multiplied by the population (pop_i) and a correction factor $\omega_{\text{ctry(i)}}$ according to (2.13). The correction factor was computed for each country (2.14). All zone-specific trip rate factors $\hat{R}_i$ rely on the respective country-specific, average trip rate factor $\bar{R}_{\text{ctry(i)}}$ previously determined based on travel survey data. Zone-specific coefficients were derived by a regression model $\Gamma(..)$ (2.15) taking into account the following zone-specific indicators ϕ_i:

- a cohort indicator for modelling the composition of a population by age group and gender;
- an economic indicator related to GDP and employment;
- a mobility indicator related to the degree of urbanisation and to the motorisation rate; and
- a spatial indicator related to the accessibility and centrality of a travel zone.

$$T_i = \omega_{\text{ctry}(i)} \cdot \hat{R}_i \cdot \text{pop}_i \tag{2.13}$$

$$\omega_{\text{ctry}(i)} = \frac{\sum_k \text{pop}_k \bar{R}_{\text{ctry}(i)}}{\sum_k \text{pop}_k \hat{R}_k} \qquad (\forall\ k \subseteq \text{ctry}(i))^{29} \tag{2.14}$$

$$\hat{R}_i = \bar{R}_{\text{ctry}(i)} \cdot \Gamma(\phi_i^{\text{coh}}, \phi_i^{\text{eco}}, \phi_i^{\text{mob}}, \phi_i^{\text{spa}}) \tag{2.15}$$

The distinction between the intra-zonal and inter-zonal trip demand was carried out on the basis of split factors $S_i \in (0, 1)$. The share $T_i\,(1 - S_i)$ is actually leaving zone i. Assuming that the shape of each zone is circular, split factors were computed by the function $\xi(..)$ based on the respective radii r_i of the zones and, in addition, based on observed country-specific trip length distributions (TLD) reporting the percentage of trip occurrences by distance bands (see sec. 4.3 and Fig. 4.1).[30] The computation of split factors can be outlined as follows:

$$S_i = \xi(\text{TLD}, r_i) \cdot F_i \cdot \omega_i \tag{2.16}$$

Besides the radius of a travel zone, the share of intra-zonal trips depends on zone-specific peculiarities like urbanisation degree, accessibility and employment rate. In order to model these peculiarities an aggregated factor F_i was derived for each travel zone.[31] The computation of split factors (2.16) is adjusted by explicit, zone-specific calibration factors ω_i which are initially set to 1. Each calibration factor is then updated based on the outcomes of the validation of the trip distribution step and the validation of the overall PAT model.[32] For this reason, several feedback loops are implemented according to the methodology of the PAT model outlined by Figure 2.5.

Further tasks of the trip generation step include the computation of trip endings $\tilde{D}_j$ and the computation of the trip demand by slow modes for each travel zone. The number of inter-zonal trips ending in destination zone j is estimated based on purpose-specific attraction indicators like the number of

[29] Index k counts all zones of a given country.

[30] The larger the radius of a travel zone, the lower the percentage of intra-zonal trips.

[31] The travel survey Mobility in Germany (infas and DLR, 2010b, p. 42), for instance, reports a negative correlation between trip length and the urbanisation degree of a region. Accordingly, the share of intra-zonal trips can be expected to be higher in the more urban region, given two regions of identical size.

[32] Calibration factors are required, as the share of intra-zonal trips depends on the accessibility of destinations (i. e. the trip distribution) which is only modelled in the trip distribution step.

workplaces attracting commuting trips and the number of accommodations attracting vacation trips. The trip demand by slow modes is computed based on national trip rate factors.

The main outcomes of the generation step are inter-zonal trip demands $\tilde{O}_i = T_i\,(1 - S_i)$ and attracted trip endings $\tilde{D}_j$ distinguished by trip purpose and travel zone where $\sum_i \tilde{O}_i = \sum_j \tilde{D}_j$. They are fed into the trip distribution model.

2.3.4.3 Trip distribution

The second sub-model, trip distribution, was developed for the computation of the G/A trip matrices covering only inter-zonal passenger trips by rail and road transport. For each trip purpose, a gravity model was applied calculating the number of trip estimates T_{ij} as follows:

$$T_{ij} \quad = \quad \alpha_{ij}\,\tilde{O}_i\,\tilde{D}_j\,f(c_{ij}) \qquad \text{with} \qquad \alpha_{ij} = a_i\,\nu_{ij}\,\omega_{ij} \tag{2.17}$$

where $\tilde{O}_i$ identifies the generated inter-zonal trip demand for origin i, $\tilde{D}_j$ the inter-zonal trip endings in destination j, c_{ij} the generalised cost of travelling, α_{ij} a set of model parameters, and $f(c_{ij})$ the deterrence function. The applied deterrence function largely followed the so-called EVA function which is discussed in more detail in section 4.3.2. The deterrence function contains several calibration parameters which were iteratively adjusted to ensure the consistency of the computed trip matrix with transport statistics in terms of mileage travelled and with the observed distributions of trip length.

Generalised cost of travelling c_{ij} was computed by a cost function (cf. sec. 3.2.1) by country as follows:

$$c_{ij} = \sum_m \varsigma_m\,(a_m t_{ijm} + b_m d_{ijm}) \qquad \text{with} \qquad m \in \{\text{rail}, \text{road}\} \tag{2.18}$$

where m identifies the transport mode, i the origin, and j the destination. For the computation of c_{ij}, two different cost components were considered: time-related travel cost t and distance-related travel cost d. These cost components were weighted by the factors a and b. The factor ς corresponds to the national market share of rail and road passenger transport demand in the respective country.

The computation of the model parameters α in the distribution model (2.17) and accordingly the computation of the purpose-specific trip matrices T was carried out in three steps under application of the Furness method.[33] In the first step, an initial trip matrix T_0 was computed with $a, \nu, \omega = 1$. Based on the origin trip constraint $\tilde{O}_i$, the parameter a_i was then updated as follows:

$$a_i = \sum_{j \neq i} T_0 / \tilde{O}_i \tag{2.19}$$

In the second step, a modified Furness algorithm was applied to determine the balancing factors ν_{ij} for each O/D relation. In comparison to its original form, which is discussed in section 4.1.1, the modified algorithm introduces a non-linear component p which can be described in simple terms by a root

[33] The Furness method is an iterative adaptation method. It is frequently applied in trip distribution modelling in order to adjust a precomputed trip matrix T_{ij} on the basis of two exogenously predetermined constraints referring to trip generation $\tilde{O}_i$ and trip endings $\tilde{D}_j$ by computing a balancing factor for each O/D relation or, to be more precise, two sets of balancing factors that are related to the origins and to the destinations (cf. sec. 4.1.1). The objectives are to ensure that the two constraints $\sum_j T_{ij} = \tilde{O}_i$ and $\sum_i T_{ij} = \tilde{D}_j$, referring to inter-zonal trips ($i \neq j$), are satisfied.

function with $0 < p < 1$. This reduces the effect of the balancing factors x_t, which are derived in each step t by the Furness algorithm $F(..)$ on the trip matrix, and also improves the convergence behaviour of the Furness algorithm (2.20). The method for deriving and updating the balancing factors $\nu_{ij_{t+1}}$ can be outlined as follows:

$$\nu_{ij_{t+1}} = \nu_{ij_t} \cdot \left(x_{ij_t}\right)^p \qquad \text{with} \qquad x_{ij_t} = F_t(a_i \cdot T_0,\, \nu_{ij_t}) \tag{2.20}$$

The iterative computation of ν_{ij} was stopped after the changes between ν_{t+1} and ν_t had become insignificant. The final set of factors was then validated against a priori defined rules. These rules included a general confidence range for ν_{ij} and a list of tolerated exceptions. For each ν_{ij} which violated the given criteria, the corresponding O/D relation was analysed further, taking into account various data sources. Following this approach, it emerged that for several origins the model which was applied for the computation of the intra-zonal share factors (2.16) was formulated too generally and could not cope with specific regional peculiarities. For instance, the original model underestimated the share of intra-zonal trips for the German capital Berlin (Ahrens et al., 2009). In such cases, the corresponding calibration factor ω_i was updated and the set of share factors S was computed again (2.16). The trip distribution model (2.17) was then applied again with the updated constraints. This procedure was repeated until the factors ν_{ij}, which were computed by the modified Furness algorithm, were inside a reasonable confidence range.

In the third step, the calibration factors ω_{ij} (2.17) that had initially been set to 1 were determined based on validation results for the overall PAT model. The main outcome of the trip distribution model, namely the purpose-specific G/A trip matrices, are fed into the mode choice model.

2.3.4.4 Modal split

The third sub-model, modal split, was applied for disaggregating the previously computed purpose-specific G/A trip matrices T_{ij} by transport mode m. In the first step, market shares s_m were derived for each O/D relation, and in the second step mode-specific trip matrices T_{ijm} were calculated as follows:

$$T_{ijm} = T_{ij} \cdot s_{ijm} \qquad \text{with} \qquad s_{ijm} = L_m(c_{ijm_1} + \omega_{ijm_1},\, c_{ijm_2} + \omega_{ijm_2}) \tag{2.21}$$

where $L_m(..)$ refers to the logit model[34] that is applied for mode $m \in \{\text{rail}; \text{road}\}$, c_{ijm} refers to generalised cost and ω_{ijm} is a calibration coefficient. The calibration coefficients were initially set to 0 and updated based on validation results for the modal split and the overall PAT model.

The main outcomes of the modal split model, namely the G/A trip matrices by mode and purpose, are fed into the assignment model.

2.3.4.5 Network assignment

The last sub-model, network assignment, was applied for the computation of traffic loads at the level of network links in terms of the number of vehicles per year and the number of passengers per year. For this purpose, the European network-based transport model VACLAV (see sec. 2.2.1) was applied.

[34] The logit model is discussed in sections 2.1.2 and 3.4.

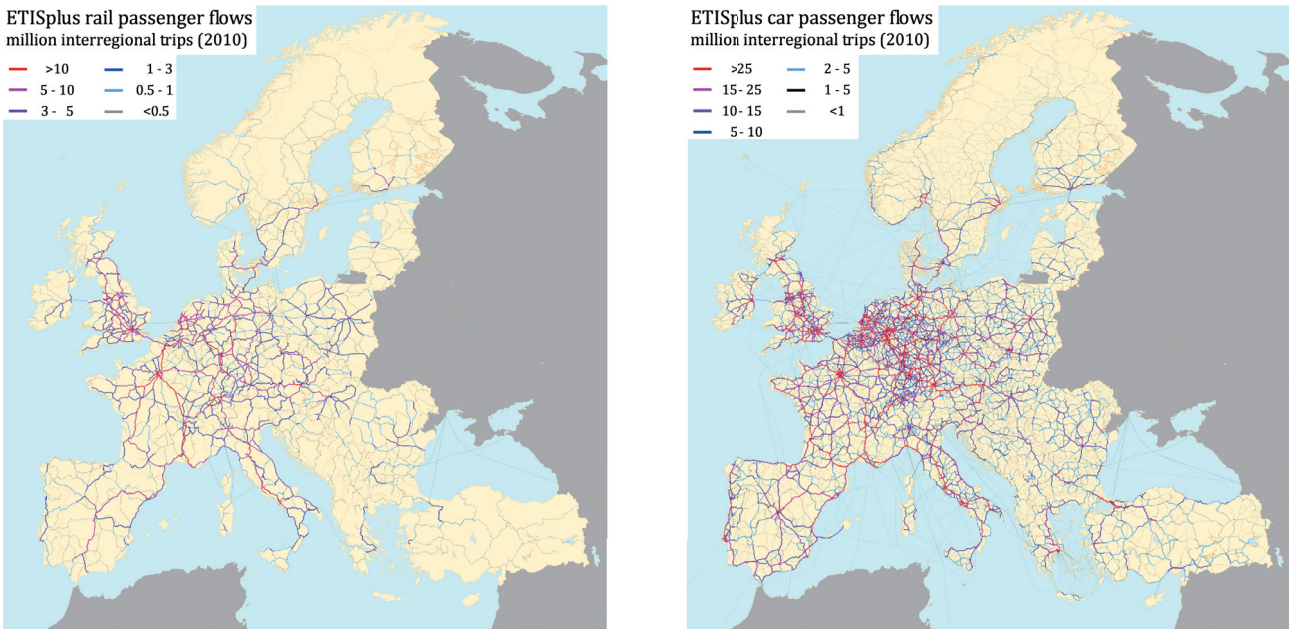

Figure 2.6: Modelled network loads for rail and road passenger transport

Figure 2.6 shows the main outcomes of the assignment model for rail and road passenger transport, which formed the basis for the validation of the overall PAT model. The assignment results fed back into WP7 and were used to calculate the external costs of transport.

2.3.4.6 Overall calibration and model output

The calibration process of the PAT model included three different mechanisms, which are outlined in Figure 2.5 by feedback loops. First of all, each sub-model was implicitly calibrated to meet given boundary conditions, e. g. concerning the trip volumes and the transport performance by mode and country. The second mechanism is the feedback loop between the trip distribution and the trip generation model. After the modified Furness algorithm was applied and the computed trip matrix was validated, the set of calibration parameters ω (2.16) related to the split factors was adjusted. The third calibration mechanism is the comparison of the modelled network loads at the level of network links with observed traffic counts. Based on the validation results, the three sets of calibration factors ω, which are related to the computation of the split factors (2.16), to the computation of the trip distribution (2.17), and to the computation of the market shares of transport modes (2.21), were updated.

The calibrated PAT model produced two important outputs: the purpose-specific G/A matrices covering inter-zonal trips for rail and road passenger transport, and the network assignment for rail and road passenger transport. The trip matrices are part of the passenger demand data set and can be accessed through the ETISplus online repository. The network assignment results for rail and road passenger transport were required for the computation of external costs of transport within WP7. Another output of the PAT model is the computation of the intra-zonal trip demand including rail and road passenger transport and slow modes. The corresponding dataset fed into another model, which was also developed and applied within WP9 (TRT, 2012).

2.3.5 Evaluating the PAT model

The development of the PAT model within ETISplus pursued several objectives. Its basic target was the computation of the O/D trip matrices for the four considered trip purposes and for the European NUTS-3 regions. However, the PAT model, focusing on inter-zonal rail and road passenger trips, was complemented by two other transport models. One of them dealt with intra-zonal trips by land-based transport modes and the other with air passenger trips between the NUTS-3 regions. In order to facilitate an independent development of the three transport models, three different transport demand segments were defined a priori. The first segment focused on inter-zonal rail and road passenger trips for the PAT model, the second segment on intra-zonal passenger trips by land-based transport modes, and the third segment on air passenger trips.

It has to be stressed that the a priori distinction of the three different demand segments and the application of three different transport models was a serious methodological restriction for the development of the PAT model. Modelling passenger transport demand as a whole and separating the three demand segments based on the outcomes of the transport model (i. e. the mode-specific trip matrices produced in third step of the four-step transport modelling approach) would clearly have been the better alternative in terms of consistency of the results and modelling approach.[35] For the PAT model, first of all, a non-integrated approach had to be chosen which immediately raised questions concerning the consistency of the overall modelling approach. Furthermore, the methodology of the four submodels trip generation, trip distribution, modal split and network assignment had to be tailored to overcome specific disadvantageous properties which stemmed from the restriction of the PAT model to inter-zonal trips. For instance, the methodology of the trip generation model had to be extended in order to compute estimates on the share of intra-zonal trips. In addition, the methodology of the trip distribution and assignment models had to be modified, as the complete trip demand was not considered but rather only inter-zonal rail and road passenger trips. It should be noted that three different cost functions are applied in the PAT model, namely in the trip distribution, the modal split and the assignment models. This is not ideal.

2.3.5.1 Trip generation

The trip generation model was applied for the computation of the transport demand in terms of originating rail and road passenger trips as well as the number of trip endings for each travel zone. Fulfilling specific requirements, which stemmed from the development and application of the three different models dealing with passenger transport, the methodology of the trip generation model had to be extended in order to produce two other output indicators. The first indicator was related to the trip demand by slow modes and the second indicator to the share of intra-zonal trips.

While the development of an integrated trip generation model covering all land-based transport modes improved the consistency between the two models dealing with intra-zonal and inter-zonal passenger trips, the computed split factors (2.16) for distinguishing between intra-zonal and inter-zonal trips were initially very inaccurate. Thus, they had to be refined several times based on the findings of the validation of the trip distribution and the overall PAT model. For this reason, the two models dealing with intra-zonal and inter-zonal trip demand had to be continuously harmonised.

[35] E. g. market shares of transport modes are computed differently for intra-zonal and inter-zonal trips.

It is important to stress that the inaccuracy of the computed split factors is related to the lack of information on the trip distribution and the market shares of transport modes within the trip generation step. For this reason, a better method would have been to develop one integrated passenger model covering all modes of transport and all trips.

2.3.5.2 Trip distribution

The trip distribution model was applied for each trip purpose for the computation of the O/D trip matrix covering inter-zonal rail and road passenger trips. It followed a gravity-based approach, which largely relied on the EVA deterrence function, for deriving the base matrices T_0 in the first step. In the second step, a modified Furness algorithm was applied in order to adjust the base matrix T_0 to the two given constraints on the number of inter-zonal trips originating and ending in each travel zone.

It should be pointed out that the Furness algorithm can only perform successfully if T_0 represents a valid starting point. This was not the case for many O/D relations, given that the number of intra-zonal trips (i. e. the share factors) had previously been estimated rather inaccurately during the trip generation step. As a consequence, the two derived constraints on the number of originating trips and on the number of trip endings were also inaccurate. Thus, the computed base matrix T_0 did not represent a perfect starting point for the successful application of the Furness algorithm. For this reason, the derived adjustment factors had to be validated against a predefined confidence range. In most cases where a factor violated the given confidence range, the cause could be traced back to an inaccurate share factor. For instance, the number of inter-zonal trips which was initially derived for the German capital Berlin was greatly overestimated. Accordingly the number of trip estimates from Berlin to its neighbouring destinations was greatly overestimated and the constraint on the number of trip endings was violated for these destinations.

In such cases, the model for estimating the share factors had to be updated together with the inter-zonal trip demand. The trip distribution model then had to be reapplied. This procedure had to be repeated several times until the given confidence range was no longer violated by any adjustment factor computed by the Furness algorithm, and the set of derived share factors had become stable. The procedure was further aggravated by the application of the complex EVA deterrence function, given its very challenging calibration (cf. sec. 4.3.2).

2.3.5.3 Modal split

The market shares of rail and road passenger transport were computed for each O/D relation by a logit model, taking into account mode-specific travel cost that had previously been computed by cost functions. During the development of the modal split model, the impedance dataset, which fed into the cost functions, was updated several times for the continuous implementation of new transport corridors and the refinement of existing corridors in the network models within WP7. The availability of different intermediate versions of the impedance dataset made it possible to evaluate the sensitivity of the logit model to changes of travel impedances and travel cost.

It was discovered, for instance, that the applied logit model was not very sensitive to cost changes for short-distance trips, i. e. for trips between neighbouring regions. The sensitivity increased as trip

distance rose. In order to overcome this specific model error, calibration coefficients were introduced, the ω_{ijm} of (2.21). These calibration coefficients were determined depending on the travel distance. This reveals a fundamental problem of the standard logit model, i. e. that it cannot simultaneously be applied for short-distance and long-distance trips. The logit model is therefore further investigated in section 3.4.

2.3.5.4 Network assignment

For the computation of traffic loads at the level of network links, the VACLAV model was applied. The calibration of the model and the interpretation of its results was rather challenging, given the lack of information on intra-zonal rail and road passenger trips. In addition, a systematic overestimation of traffic loads for the feeding links and their adjacent links could be observed, which is due to the specific relevance of these links in the network model. The degree of overestimation could be reduced for several feeding links by considering multiple feeding nodes for each travel zone for the computation of the network assignment according to Schoch (2004). For the above reasons, the network assignment results had to be analysed very carefully in order to derive the appropriate measures for improving the calibration of the overall PAT model including the four sub-models.

2.3.5.5 Overall findings

It was found out that the fully calibrated PAT model included several contradictory results. For instance, major discrepancies could be identified at the level of network links between the modelled and the observed traffic loads for several regions, in particular for the feeding links and their adjacent links. To a certain extent, these differences were inevitable and could be explained by the chosen modelling approach.

A general obstacle disabling the computation of more accurate results was the application of travel zones that were too large and heterogeneous, as the zones were defined according to the European NUTS-3 regions. This specific issue can be best described for the computation of the trip demand between neighbouring regions. In this case, the accuracy of the modelled trip demand solely depends on the location of the two centroids in the two travel zones. If the centroids are located in adjacent corners, i. e. very close to each other, the trip distribution model is very likely to overestimate the trip demand between the two regions. By contrast, if the centroids are located in distant corners of the two travel zones, the trip distribution model is very likely to underestimate the trip demand between the two regions. It should be noted that regional trips below 100 km could in general not be modelled very accurately by the PAT model, since the diameter of more than half of the applied NUTS-3 regions was over 50 km.

2.4 General conclusions and ways forward

This chapter provided a brief introduction into transport modelling and discussed several basic modelling concepts and transport models. It presented the PAT model which was developed during the ETISplus project for updating the trip matrices in the European transport information system. Due to specific requirements (e. g. intra-zonal trips are not modelled, and the four sub-models trip generation, trip distribution, modal split and assignment were developed and applied independently), the PAT model followed a very tailored approach that led to several methodological shortcomings. These shortcomings could be counterbalanced by the introduction of additional calibration factors, but the process of calibrating the PAT model was challenging and time consuming. In particular, the trip distribution model required three subsequent calibration steps before the produced trip matrix was acceptable. For these reasons, the PAT model was no longer considered for passenger transport modelling within this thesis apart from its basic result, the ETISplus trip matrices.

It should be emphasised that a major obstacle for European transport modelling could be solved, namely the generation of a complete European modelling database covering all required regional and mobility indicators at NUTS-3 or NUTS-0 level. This database is further required for the IPAT and the HIPAT models. A second important output relates to the expertise built up during the development and/ or analyses of the PAT, the VACLAV and the TRANS-TOOLS models. This experience facilitated the process of developing the methodology of the IPAT and the HIPAT models taking the following approach:

1. Modelling the whole European passenger trip demand by the trip distribution model overcomes the necessity to distinguish between intra-zonal and inter-zonal trips based on split factors (S_i) and calibration factors (ω_i) within the trip generation model (2.16). Moreover, this contributes to the consistency of the trip distribution model, as intra-zonal and inter-zonal trips are processed in the same way.

2. Fundamental revision of the trip distribution model by consideration of additional explanatory factors, such as competition of travellers, overcomes the necessity to use the Furness method for massaging the trip matrix through the use of balancing factors.

3. Substantial revision of the deterrence function simplifies the difficult calibration process related to the EVA function and the trip distribution model, while the explanatory power of the deterrence function is simultaneously improved.

4. Integration of the trip distribution and the modal split models as well as application of one consistent cost function in both models increases the consistency of the overall modelling approach.

5. Inspection of the logit model in order to reveal the reasons why this model responds differently to changes in travel costs for short-distance and long-distance trips, and elaboration of an approach to solve this deficiency.

6. Substantial downsizing of the diameter of NUTS-3-based travel zones from about 50 km to 5 km in order to enable an accurate modelling of regional trips between 10 and 100 km.[36]

[36] Trips between two adjacent travel zones cannot be computed accurately, as they are modelled by one "representative" travel distance. In reality, however, travel distances of trips between two adjacent average-sized NUTS-3 regions can vary between a few metres and almost 100 km.

The first three issues require a fundamental revision of the trip distribution model. In the PAT model, these issues were tackled by introducing additional calibration and balancing factors. However, we should bear in mind that such factors do not contribute to the explanatory power of a transport model but only compensate shortcomings of the chosen methodology. In particular, the Furness method massaging the trip matrix based on two given constraints might even reduce the explanatory power of a transport model. If the values of the computed balancing factors are not reviewed against a reasonable confidence range, this method is likely to whitewash existing model errors. In a very early stage of the PAT model, for instance, the Furness method computed very large (>10) and small (<0.1) balancing factors for many O/D relations for the German capital Berlin. This model error could be traced back to the overestimation of the inter-zonal trip demand for Berlin. Thus, the respective split factors were corrected and, in consequence, the balancing factors subsequently computed by the Furness method improved significantly.

However, the balancing factors that were finally applied for adjusting the trip matrix still differed markedly from 1 due to the non-consideration of important explanatory factors in the trip distribution model.[37] One of these explanatory factors is the competition between travellers for available destination opportunities. In the PAT model, for instance, the principle "first come, first served" is not implemented, but rather "all come, all served", given that the number of available destination opportunities (D_j) is kept identical for any traveller irrespective of the actual demand for destination j in (2.17). Hence, the trip demand to those destinations that are accessible for many travellers is overestimated (and vice versa) and has to be counterbalanced a posteriori by balancing factors that were computed by the Furness method. To avoid application of this method, the level of competition of travellers for destination opportunities has to be modelled, for instance, by use of an appropriate accessibility indicator as discussed in section 3.3.

Another shortcoming of the trip distribution model underlying the PAT model is the rather weak concept of a deterrence function (2.17), which is only sensitive to travel costs (c_{ij}) and not to other important explanatory factors, such as the "quality" of destination opportunities (e. g. the wage level of jobs).[38] For this reason, the trip distribution model and the deterrence function are substantially revised in chapter 4. In addition, the concept of a deterrence model is developed in chapter 5 which considers more explanatory factors than only the deterrence function. Besides the trip distribution model, also the logit model underlying the modal split model is also intensively reviewed (cf. sec. 3.4), firstly to correct the identified distance-related error and, secondly, to develop a cost indicator for integrating the trip distribution and modal split models (issues 4 and 5).[39]

The methodological revisions (issues 1 to 5) leading to the development of the IPAT model are not tackled and solved independently one after another. They are carried out in chapters 3 to 5 starting with an in-depth analysis of the standard theory. The downscaling of travel zones (issue 6) is then discussed in chapter 7 and leads to the development of the HIPAT model in which European transport flows are modelled at LAU-2 level, based on more than 100,000 travel zones.

[37] If sufficient explanatory factors were considered, no additional balancing factors would be required.
[38] Employees tend to accept higher travel costs for commuting to jobs proposing higher salaries.
[39] It should be emphasised that the applied logit model is in line with standard theory in transport modelling.

Part II

Methodological revisions

Chapter 3

Economic and spatial concepts

This chapter summarises several important modelling concepts that are relevant with regard to the four step-approach in transport modelling. The concepts discussed contribute to the consistency, the explanatory power and the transferability of a transport model. Although the mathematical theories have been known for a relatively long time, these concepts are implemented in rather problematic ways in some transport models.[1] It is therefore necessary to discuss some important transport modelling concepts in detail and to reveal their specific characteristics in order to understand the fundamental ideas behind the methodological revisions elaborated within this thesis. This led to the development of the integrated trip distribution and modal split model, which is implemented in the IPAT model.

This chapter starts with a brief review of the concept of price elasticities. Price elasticities are an important concept for describing the sensitivity of a transport model to changes of travel costs. Furthermore, elasticity coefficients are used to determine the behaviour of modelling components, including cost functions and deterrence functions. Section 3.2 discusses the concept of generalised costs that are computed by cost functions taking into account relevant cost components. It further explains the differences between the two most common cost units, monetary terms and time units. In section 3.3 the concept of accessibility is studied, which is frequently applied in urban planning for developing transport infrastructures. According to Crozet (2012) this concept has to be taken into account in order to predict the mobility behaviour of city residents to infrastructure changes. For this reason, an improved accessibility indicator is developed, forming the basis for the further development of the trip distribution model within this thesis.

Section 3.4 thoroughly investigates the mathematical properties of the logit model, given that the application of the PAT model indicated a suspicious behaviour when computing market shares for rail and road passenger transport with the logit model for short-distance and long-distance trips. According to the relevant academic literature, this issue can be eliminated quite easily. Finally, the logit model underlying the IPAT model is briefly introduced, which can consistently be applied for more than two transport modes as well as for any travel distance. Originating from the logit model, the concept of expected minimum cost is discussed in section 3.4.4, facilitating the integration of the trip distribution and the modal split models implemented in the IPAT model.

[1] In the PAT model, for instance, following basic transport modelling theory, additional calibration factors had to be introduced in order to circumvent identified methodological issues. This limits transferability of the PAT to the investigation of transport policy scenarios if calibration factors are kept constant.

3.1 Price elasticities

In transport modelling, the concept of price elasticities is commonly used in order to describe the behaviour of a transport model and the four sub-models to changes of travel costs, while a fundamental differentiation is made between direct price elasticities and cross price elasticities. Based on the value of the derived elasticity indicator E_D, the behaviour of the respective models can be grouped into several categories according to Table 3.1. For the evaluation of the IPAT model E_D was computed

Table 3.1: Classification of elasticity values

Category	Value	Descriptive Term
1	$E_D \to -\infty$	Perfectly elastic demand
2	$-\infty < E_D < -1$	Elastic demand
3	$E_D = -1$	Unit elastic demand
4	$-1 < E_D < 0$	Inelastic demand
5	$E_D = 0$	Perfectly inelastic demand
6	$E_D > 0$	Abnormal demand

according to the expressions introduced below ((3.1) and (3.6)). E_D was then compared to observed price elasticities indicating the response of travellers to changes of travel costs (see sec. 6.3.3). For example, inelastic response of travellers can be observed for compulsory trips that are related to the trip purposes of commuting and shopping for goods for everyday consumption, while elastic demand can be observed for non-mandatory trips that are related to the trip purposes of leisure or vacation.[2] Besides the validation of a model, such information on observed price elasticities is also an important input for the development of a transport model or, to be more precise, for the specification of the cost functions (see sec. 3.2), the logit model (see sec. 3.4) and the deterrence functions (see sec. 4.3) amongst others.

3.1.1 Direct price elasticities

Direct price elasticity indicates the percentage change of the demand in relation to the percentage change of the price. If the former demand Q_1 for the price P_1 is known, as well as the current demand Q_2 for the price P_2, the price elasticity E_D can be computed by the method of mean values according to Samuelson and Nordhaus (2007, p. 107):

$$E_D = \frac{\Delta Q}{0.5 * (Q_1 + Q_2)} \Bigg/ \frac{\Delta P}{0.5 * (P_1 + P_2)} \tag{3.1}$$

where $\Delta Q = Q_1 - Q_2$ is the absolute change in the demand, and $\Delta P = P_1 - P_2$ the absolute change in the price. If changes are infinitesimally small ($\Delta Q, \Delta P \to 0$) Equation (3.1) can be transformed as follows:

$$E_D = \frac{\Delta Q}{Q} \Bigg/ \frac{\Delta P}{P} = \frac{\Delta Q}{\Delta P} \cdot \frac{P}{Q} \tag{3.2}$$

[2]The strong increase in air passenger transport, for instance, can be explained by the emergence of low-cost carriers, leading to an increase in leisure and vacation trips amongst others.

If the demand function Q: $x(p)$ and the price function P: $p(x)$ are known for some areas, and if both functions are defined and $p(x)$ can be differentiated, the marginal elasticity function $\eta_{x,p}$ is:

$$\eta_{x,p} = \frac{dQ}{dp} \cdot \frac{p}{Q} = \frac{x'(p)}{x(p)} \cdot p \tag{3.3}$$

Generally, the elasticity of a demand function is different for each point. Demand functions in which the elasticity is constant are called isoelastic. These functions are of particular importance for modelling changes in a dependent variable y in relation to changes in an explanatory variable x, since the whole model behaviour can then be described by one constant elasticity coefficient. The power function $y = x^\alpha$ with $x > 0$ is isoelastic (3.4):

$$\eta_{x,y} = \frac{y'(x)}{y(x)} \cdot x = \frac{\alpha \cdot x^{\alpha-1}}{x^\alpha} \cdot x = \alpha \tag{3.4}$$

Isoelastic models are frequently used as deterrence functions in gravity-based trip distribution models (cf. sec. 4.3) with the elasticity coefficient α adjusted to a specific demand segment. It can be assumed, for instance, that students respond rather more elastically to price changes than business travellers. Thus, both demand segments are commonly modelled independently, and different elasticity coefficients are chosen in the deterrence functions.[3] It should be noted that in the IPAT model a more complex deterrence function is used to better meet the requirements of a European transport model.

In the context of cost functions (see sec. 3.2.1), the Box-Cox transformation is discussed, which is very similar to a power function and therefore also isoelastic. Another important model relying on an isoelastic function is the so-called pivot point method. Pivot point methods facilitate the extrapolation of a dependent variable D, knowing the change rate $\dot{C}$ of the explanatory variable V between two time steps t_1 and t_2 as well as the elasticity coefficient ε. The extrapolated variable $\hat{D}_{t_2}$ can then be computed as follows:

$$\hat{D}_{t_2} = D_{t_1} \cdot \left(\dot{C}\right)^\varepsilon \qquad \text{with} \qquad \dot{C} = \frac{V_{t_2}}{V_{t_1}} \tag{3.5}$$

Pivot point methods have been applied in several research projects, e. g. in the TINA Turkey project for forecasting the freight demand up to the year 2020 (TINA et al., 2007) and in the TEN-STAC project for computing modal shift effects with the VACLAV model (NEA et al., 2004).

3.1.2 Cross price elasticities

Cross price elasticities are considered for describing the behaviour of the demand if a traveller can choose between several comparable goods and the price changes for precisely one good. Given a choice set of two alternative goods {A, B}, for instance, the cross elasticity $E^{\mathrm{Q(A)}}_{\mathrm{P(B)}}$ indicates how the level of demand Q for product A changes if the price P for product B changes. Analogous to the direct price

[3] In the IPAT model four trip purposes are distinguished that are modelled by different deterrence functions and, accordingly, different elasticity coefficients.

elasticity (3.1), the cross elasticity of the demand can be computed as follows:

$$E_{\mathrm{P(B)}}^{\mathrm{Q(A)}} = \frac{\Delta Q(A)}{0.5 * \big(Q_1(A) + Q_2(A)\big)} \Bigg/ \frac{\Delta P(B)}{0.5 * \big(P_1(B) + P_2(B)\big)} \tag{3.6}$$

where $\Delta Q(A) = Q_1(A) - Q_2(A)$ refers to the absolute change in the demand, and $\Delta P(B) = P_1(B) - P_2(B)$ to the absolute change in the price.

Cross elasticities are commonly considered in the context of discrete choice models for describing the changes in the choice probabilities in relation to changes in an explanatory variable. In transport modelling, for instance, modal shift effects between the transport modes can be described by cross elasticities, e. g. changes in the demand of coach passengers in relation to price changes of rail tariffs.

3.2 Generalised Costs

Generalised costs are an important concept in transport modelling, which combines a variety of possible cost components in a single cost indicator. These components comprise disincentives as well as incentives for travelling, such as travel time and monetary travel costs related to a trip. Generalised costs are computed by so-called cost functions (see sec. 3.2.1). They are commonly measured either in a monetary or a time-related unit. The advantages and disadvantages of the two cost units are discussed in sections 3.2.2 and 3.2.3.

3.2.1 Cost functions

Cost functions provide a general framework for deriving generalised cost of travelling between two travel zones under consideration of a variety of cost components. In transport literature, the terms "utility function" or "disutility function" are also used as synonyms. Cost functions are often defined by purpose and by transport mode, since different cost components are relevant for different demand segments. These cost components are commonly measured in a monetary unit like EUR or a time-related unit like hours. The conversion between these two cost units is facilitated by the so-called Value-of-time (VoT) indicator, which is given in EUR per hour, and is outlined by (3.9) and (3.10).

The most common cost components are time- and monetary travel cost that are directly measured in hours and EUR. However, Level-of-Service (LoS) indicators, like the frequency of train connections per day and indicators related to the quality of travelling[4], can also be considered for the computation of generalised costs. These qualitative variables are measured on an ordinal scale and must therefore be transformed in order to be incorporated into the arithmetic model of the cost functions.

Equation (3.7) drafts a linear model for computing generalised cost c_{ij} of travelling from origin i to destination j.[5] The variables x_t refer to time-based cost indicators and x_d to distance-based cost indicators, and $f_l(..)$ refers to specific functions for transforming LoS and other quality-related indicators x_l for their application within the cost functions. The β coefficients refer to weighting factors and have to be estimated. It should be noted that all components of the cost function (3.7) are given in a monetary cost unit, i. e. time-related variables x_t are converted to EUR by the VoT indicator.

$$c_{ij} = \beta_0 + \sum_t \beta_t \cdot x_{tij} \cdot \text{VoT} + \sum_d \beta_d \cdot x_{dij} + \sum_l \beta_l \cdot f_l(x_{lij}) \tag{3.7}$$

Within this thesis, the following time-based cost indicators are of particular importance:

- the indicator net time, referring to the travel time with the main transport mode, e. g. the flight time; and
- the indicator A/E time, referring to the travel time for accessing and egressing the main transport mode, e. g. the travel time for arriving at the railway station and the travel time from the railway station to the real destination.

[4] Business class, for instance, commonly provides more comfort and allows travellers to work during their journey.
[5] Although not indicated by indexes, trip purposes and transport modes are implicitly distinguished.

The following distance-based cost indicators, which are distinguished by private and public transport modes, are considered within this thesis:

- for private transport modes, the indicators fuel, toll and fixed costs[6] for the trip with the main transport mode;
- for public transport modes, the indicator travel fares for the trip with the main transport mode, e. g. the railway ticket;
- the indicator A/E cost, referring to the travel cost for accessing and egressing the main transport mode, e. g. a tram ticket to the railway station.

The third category of cost components in the linear model (3.7) comprises LoS indicators which are related to the quality of a travel opportunity such as the frequency of train connections per day, safety, reliability and comfort aspects. Due to the difficulty of measuring and converting qualitative indicators, only relative changes in a single, aggregated LoS indicator are considered within this thesis. The indicator is set to zero for the base year when calibrating the cost functions and determining the values of the β-coefficients. It can be modified for the simulation of transport policies (cf. Van Grol et al., 2016, p. 67).

The linear model (3.7) has several disadvantages which are discussed in detail in Gaudry (1993) and Mandel et al. (1997) among others. In order to overcome these disadvantages, the Box-Cox transformation can be applied to a cost indicator (Box and Cox, 1964). The transformation is defined as follows:

$$x^{(\lambda)} = \begin{cases} \dfrac{x^{\lambda} - 1}{\lambda} & \text{for } \lambda \neq 0 \\ \ln x & \text{else} \end{cases} \tag{3.8}$$

where x is the linear cost indicator, $x^{(\lambda)}$ the non-linear indicator, and λ the parameter of the Box-Cox transformation. The Box-Cox transformation should be taken into account if a non-linear response of travellers to changing travel cost can be observed. A prominent example is the frequency of business trips in relation to the travel time. As two out of three business trips are one-day trips (Schneider, 2008), the frequency of trip occurrences declines strongly between two and three hours. This can be explained by the fact that the "waste" of half a working day for travelling to the client and back is just about acceptable for a one day business trip.

3.2.2 Generalised travel times and generalised travel costs

Generalised costs are computed by cost functions and may be given in monetary terms or as a time-related unit. If generalised costs are given in monetary terms, they are also known as generalised travel costs (GTC), while they are known as generalised travel times (GTT), if they refer to a time-related unit.[7] Even though GTT and GTC can be applied almost equivalently, they provide specific advantages. The concept of GTC is particularly suited for measuring the consumer surplus and may

[6] Fixed costs include components which are given on a yearly basis, e. g. car insurance, taxes, purchasing and repairing costs. These components can be disaggregated based on assumptions about the average annual mileage of a vehicle.

[7] The term GTC is often used ambiguously in transport literature and can refer to both cost units.

be preferred for cost-benefit analyses, e. g. for the planning of new infrastructure, while the concept of GTT is more appropriate as a cost indicator with regard to the trip distribution model. Both indicators can be converted into each other according to (3.9) and (3.10).

$$\text{GTT} = \text{time costs} + \text{monetary costs/VoT} \tag{3.9}$$

$$\text{GTC} = \text{time costs} * \text{VoT} + \text{monetary costs} \tag{3.10}$$

Within this thesis, VoT indicators are distinguished by purpose and country, enabling the consistent application of a generalised cost function for different demand segments. In a wider sense, the VoT indicator can be understood as the travellers' willingness to pay for higher travel speed. It can therefore explain different mode-choice decisions of different demand segments. For instance, business travellers, who have a higher willingness to pay than students, frequently choose faster and more expensive transport modes than students, who prefer to travel cheaply. The same holds true for travellers who live in different countries with different wage levels.

It should be stressed that different values for the VoT indicators can be found in transport literature. The HEATCO study, for instance, distinguishes the VoT indicator or the so-called Value of travel time Savings (VTTS) indicator by income, journey length, journey purpose, and transport purpose. For the EU25, the average estimated VTTS values vary between 5.11 and 32.80 EUR (Bickel et al., 2005, p. 11-13). The study recommends the estimation of VoT values for future years in relation to GDP growth, by application of an elasticity coefficient of 0.7. In contrast, the passenger demand module of the HIGH-TOOL model distinguishes the VoT indicators by purpose and country. Depending on the trip purpose, the average value for the EU28 varies between 5.89 and 12.05 EUR (Kiel et al., 2016b). Given the large discrepancies, VoT indicators from different data sources should not be combined. Within this thesis, VoT indicators were therefore solely taken from the ETISplus dataset, which refers to the base year 2010.

3.2.3 Advantages of generalised travel times

The basic advantage of choosing GTT as a cost indicator within a transport model is the consistent application of the trip distribution model. The trip distribution model takes account of the decreasing likelihood of travelling with increasing travel cost for different demand segments and, accordingly, for different values of the VoT indicator.[8] Table 3.2 compares GTT and GTC, which have been computed for the same O/D relation according to (3.9) and (3.10), for different willingness to pay values. While the GTT indicator decreases with increasing willingness to pay, the GTC indicator

Table 3.2: Comparison of GTT and GTC for different willingness to pay values

Willingness to pay	Value of time [1]	Time costs [2]	Monetary costs [3]	GTT [2]	GTC [3]
low	5	60	10	180	15
moderate	10	60	10	120	20
high	20	60	10	90	30

[1] [EUR/h] [2] [min] [3] [EUR]

[8] This relationship is modelled by the deterrence function in the gravity trip distribution model (cf. sec. 4.3).

increases. Accordingly, a trip distribution model, which relies on the GTT indicator, can be applied consistently for different demand segments. For example, if we assume a high willingness to pay of business travellers, their travel costs in terms of GTT are 90. Assuming a low willingness to pay of students, their travel costs in terms of GTT are 180. Given the lower travel costs for business travellers, a trip distribution model relying on the GTT indicator will derive a higher likelihood of travelling to the destination for business travellers than for students. For this reason, the concept of GTT should be favoured as a cost indicator for transport models which are applied for different countries and for different demand segments. The trip distribution model can then be estimated at an average European VoT level and can be transferred to different countries and different demand segments.

By contrast, a trip distribution model relying on the GTC indicator cannot be applied consistently for different countries and different demand segments. If GTC is considered as a cost indicator within the trip distribution model, it computes a higher likelihood of travelling to the destination for students than for business travellers.[9] Hence, transferability is not given. If GTC is used, a trip distribution model has to be estimated independently for each country and each demand segment, which instantly raises questions about the consistency of the overall transport model.

[9] Travel costs in terms of GTC are 15 for students and 30 for business travellers.

3.3 Accessibility

The accessibility concept is important for transport modelling and transport planning issues. It is commonly related to a region and can be defined from the perspective of departing or arriving trips. The origin-based accessibility indicator summarises information about the number of accessible opportunities in all destinations, where each destination is weighted by a factor. This factor is defined in relation to the utility of travelling to a destination, and decreases with increasing travel costs. It reaches zero for infinite travel costs. The accessibility is commonly defined by trip purpose, taking into account different opportunities of attraction. Workplaces, for instance, are frequently considered for the trip purpose of commuting, and accommodation for the trip purpose of vacation. The origin-based accessibility indicator is of particular importance with regard to the trip distribution model. It facilitates the computation of the most likely distribution of the generated trip demand to the destinations.

Vice versa, a destination-based accessibility indicator summarises the information about the number of travellers who could travel to a certain destination. This indicator can also be defined by trip purpose. Destination-based accessibility indicators are frequently considered, for instance, for location planning of shopping malls, distribution centres and labour-intensive enterprises, as well as for urban planning with regard to the development of the transport infrastructure (the more people can reach a location with reasonable effort, the better the location).

Taking into account different types of accessibility indicators, the chapter is structured as follows: the basic concept of accessibility is explained in section 3.3.1 and a revised indicator, which is also sensitive to competition effects, is introduced in section 3.3.2. Section 3.3.3 concludes the main findings. It should be noted that the index $\gamma = \{2, 3\}$ of the accessibility indicator $A^{(\gamma)}$ specifies whether second or third order competition effects are also considered.[10]

3.3.1 Basic accessibility indicators

The gravity accessibility indicator was introduced by Hansen (1959). Referring to the trip purpose of commuting, the origin-based accessibility indicator A_i for origin i can be defined as follows:

$$A_i = \sum_{j=1}^{n} W_j e^{-\beta c_{ij}} \tag{3.11}$$

where W_j refers to the number of workplaces in destination j, and c_{ij} to the generalised cost of travelling from i to j.[11] The exponential function is a classical deterrence function, which is frequently considered in gravity-based trip distribution models. It models the disincentive of travelling between two regions in relation to travel costs where the coefficient $\beta > 0$ determines the sensitivity of the deterrence function.[12] Based on the accessibility indicator (3.11), the likelihood L_{ij} of people living in the origin region i and travelling to destination j can be computed as follows:

$$L_{ij} = \frac{A_{i|j}}{A_i} \qquad \text{with} \qquad A_{i|j} = W_j e^{-\beta c_{ij}} \tag{3.12}$$

[10] Accordingly, $A^{(5)}$ would indicate that fifth-order competition effects are also considered.
[11] In the following, i with $i = \{1 \dots m\}$ is used for origins and j, l with $j, l = \{1 \dots n\}$ for destinations ($m = n$).
[12] Deterrence functions are further discussed in section 4.3.

In contrast to the origin-based accessibility indicator (3.11), the destination-based accessibility indicator $\hat{A}_j$, which is specifically related to commuting trips, provides information on the number of employees who are willing to travel from any region to destination j, taking into account generalised travel cost c_{ij}. $\hat{A}_j$ is computed according to (3.13), where E_i is the number of employees (i. e. labour force) in region i.

$$\hat{A}_j = \sum_{i=1}^{m} E_i e^{-\beta c_{ij}} \tag{3.13}$$

Besides their consideration within the trip distribution model, accessibility indicators can also be applied for disaggregating observed transport demand indicators, which are only provided at a regional level, to the level of O/D relations. For instance, German transport statistics provide the number of in- and out-commuters at a regional level (Destatis, 2007). In this case, the origin-based accessibility indicator (3.11) can be applied for estimating the demand of out-commuters at the level of O/D relations and the destination-based accessibility indicator (3.13) for estimating the demand of in-commuters at the level of O/D relations.

3.3.2 Competition-enhanced accessibility indicator

An important shortcoming of the previously discussed accessibility indicators is the non-consideration of competition effects. For instance, an origin-based accessibility indicator formulated according to (3.11), which is specifically related to vacation trips, disregards the number of tourists who are competing for available accommodations. In consequence, the same accessibility value is derived irrespectively of whether only two tourists or one hundred tourists are competing for ten available accommodations in a destination. Vice versa, a destination-based accessibility indicator formulated according to (3.13), which is specifically related to shopping trips, disregards the competition between shopping malls which are attracting the same clients. Since the demand of the clients is distributed among all accessible malls, a more realistic estimation of the potential demand $\tilde{A}$ of the clients could be computed according to (3.14), in which the number of potential clients can be derived according to (3.13).

$$\tilde{A} \sim \frac{\text{number of potential clients}}{\text{number of competing malls}} \tag{3.14}$$

The limitations of the gravity accessibility indicator, which disregards the effects of competition, and the implications for transport modelling, have only been discussed peripherally in the literature (see, for example, Crozet et al. (2012); Shen (1998); and Wang (2001)). In order to improve the origin-based accessibility indicator, the authors recommend dividing the number of opportunities provided by a destination by the number of travellers who can travel to this destination. The competition-weighted accessibility indicator $A_i^{(2)}$ can then be computed as follows:

$$A_i^{(2)} = \sum_{j=1}^{n} \frac{W_j}{V_j} e^{-\beta c_{ij}} \qquad \text{with} \qquad V_j = \sum_{k=1}^{m} E_k e^{-\beta c_{kj}} \tag{3.15}$$

If $A_i^{(2)}$ is formulated for the trip purpose of commuting, W_j refers to the number of workplaces in destination j, c_{ij} to the generalised cost of travelling from i to j, and V_j to the number of employees E_k living in region k who can travel to the destination j. The factor W_j/V_j would identify the expected level of competition for the workplaces in region j if all employees commuted to this region. However, it is obvious that their demand is split among all accessible destinations. This leads to a further extension of the accessibility indicator $A_i^{(2)}$ and the introduction of the competition-enhanced accessibility indicator $A_i^{(3)}$. Formulated for the trip purpose of commuting, $A_i^{(3)}$ satisfies the following equation:

$$A_i^{(3)} = \sum_{j=1}^{n} \frac{W_j}{V_j^{(2)}} e^{-\beta c_{ij}} \qquad \text{with} \qquad V_j^{(2)} = \sum_{k=1}^{m} \frac{E_k}{J_k} e^{-\beta c_{kj}} \qquad \text{and} \qquad J_k = \sum_{l=1}^{n} W_l e^{-\beta c_{kl}} \tag{3.16}$$

where $V_j^{(2)}$ indicates the number of employees E travelling to a workplace in region j taking into account the fact that each employee has several options.

3.3.3 Summary

The basic accessibility indicator A_i (3.11), which is specifically related to commuting trips, determines the number of workplaces which can be reached from origin i, taking into account the utility of travelling to the destinations. However, A_i does not reveal the real number of *available* workplaces due to the non-consideration of competition between employees, given that many employees from different regions can travel to the same workplace. Its application within accessibility-based gravity trip distribution models, which will be discussed in section 4.2, is therefore limited to a few specific cases. In contrast, the competition-enhanced indicator $A_i^{(3)}$ (3.16), which explicitly models these competition effects by combining three individual accessibility indicators, provides more realistic estimates of the number of available workplaces for a region. It can therefore be applied within the trip distribution model almost without limitations.

It should be stressed, however, that the competition-enhanced accessibility indicator $A_i^{(3)}$ disregards indirect competition effects of a higher order, such as the indirect competition between employees living in distant cities like Frankfurt and Stuttgart. Consider the following hypothetical example. If we assume that employees are not willing to travel more than 100 km, only employees living in neighbouring cities like Frankfurt and Mannheim, Mannheim and Karlsruhe, Karlsruhe and Stuttgart, would directly compete for accessible workplaces in the respective regions. In contrast, the travel distance between distant cities like Frankfurt and Stuttgart is larger than 100 km. Hence, if the demand of commuters were computed based on $A_i^{(3)}$, the choice of employees living in Frankfurt would have no impact on the choice of employees living in Stuttgart. However, following the competition chain, Frankfurt - Mannheim - Karlsruhe - Stuttgart, the destination choice of employees living in Frankfurt certainly has an impact on the destination choice of employees living in Stuttgart.[13] It must therefore be concluded that all travellers indirectly compete for all available destination opportunities. Thus, it could be investigated whether a further revision of $A_i^{(3)}$ towards a more complex accessibility indicator promises more accurate results with regard to the trip distribution model.

[13] If many employees from Frankfurt decide to work in Mannheim, employees from Karlsruhe are likely not to work in Mannheim but rather in Stuttgart, thus changing the choice set of employees living in Stuttgart.

3.4 Logit Model

The logit model originates from discrete choice theory. In the context of the four-step transport model it can be applied with different objectives, including travellers' choices of destination, transport mode and routing. Within this thesis, the logit model is applied in the context of a macroscopic transport model, solely for the computation of the market shares of transport modes (i. e. choices of transport modes) in relation to absolute differences between mode-specific travel costs. Travel costs and differences are, however, not uniform in European transport models applied for short-distance and long-distance trips, i. e. for travel costs ranging from a few to several hundred Euro. This is an important difference to the standard application of a logit model at a single cost level, in which the scale parameter μ is kept constant (3.19) and the behaviour of the logit model is only described by the weighting factors β which are part of the mode-specific cost functions (3.17). The fact that μ was also kept constant in the PAT model explains the suspicious behaviour of the modal split model that responded differently to cost changes for short-distance and long-distance trips (cf. sec. 2.3.5).

The described issue is better known as scale dependency. It was circumvented in the PAT model by introducing additional calibration coefficients determined in relation to travel distances and, accordingly, to travel costs. It has to be emphasised that the same issue is also relevant for other transport models analysed within this thesis. As the scale dependency of the logit model was not identified in these models, even more questionable workarounds were applied than in the PAT model in order to improve the behaviour of the mode choice model.

For the above reasons, the mathematical theory underlying the logit model is briefly reviewed in the following sections and an extension is introduced, facilitating the application of the logit model in the context of a European transport model. Section 3.4.1 briefly introduces the concept of random utility and illustrates the derivation of the logit model and the EMU indicator. Section 3.4.2 investigates the scale dependency of the logit model that can be explained by means of the properties of the Gumbel distribution (see sec. 3.4.3). Section 3.4.4 discusses the EMC indicator that is used in this thesis for the consistent integration of the trip distribution and modal split models. In section 3.4.5, the nested multinomial logit model is briefly introduced, which will be applied in the IPAT model for the computation of market shares of transport modes.

3.4.1 Derivation of the Multinomial Logit model

According to the concept of random utility, a traveller decides between several options on the basis of rational criteria. Some of these criteria are transparent and can be measured, but some criteria cannot be revealed and are therefore unknown. For instance, a traveller may decide which transport mode m to use on the basis of generalised costs c_m and unknown, individual preferences ϵ_m which are modelled by a random term. Following the principle of utility maximisation or, to be more precise, the principle of cost minimisation, the traveller chooses the transport mode that promises him the highest utility U according to:

$$U = \max_m \{u_m + \epsilon_m\} \qquad \text{with} \qquad u_m = -c_m \qquad \text{and} \qquad c_m = \sum_k \beta_{km} x_{km} \tag{3.17}$$

where x refers to a cost component and β to a weighting factor in the cost function (cf. sec. 3.2). If the random term ϵ_m is statistically independent and Gumbel-distributed with the scale parameter $\mu > 0$, U is also Gumbel-distributed with the same scale parameter μ. The expected value of U, which is better known as the expected maximum utility (EMU), can then be computed as follows:

$$\text{EMU} = \frac{1}{\mu} \cdot \ln \sum_m e^{\mu u_m} + \text{cnst} \tag{3.18}$$

where cnst is a model constant. The probability of choosing the transport mode m with utility u_m can be computed according to (3.19), which is better known as the multinomial logit (MNL) model.

$$P(u_m) = \frac{e^{\mu u_m}}{\sum_\nu e^{\mu u_\nu}} \tag{3.19}$$

MNL and EMU can be derived relatively straightforwardly on the basis of the properties of the Gumbel distribution, which are discussed in more detail in Ben-Akiva and Lerman (1985) and Maier and Weiss (1990) amongst others.

3.4.2 Scale dependency of the logit model

The MNL is often applied within the third step of the four-step transport model for computing the market shares of the transport modes. A frequent approach used to determine the logit model (3.19) is to estimate the weighting factors β (3.17) that are applied in the cost functions while keeping the scale parameter constant, e. g. $\mu = 1$. This, however, represents a crucial model restriction, which leads to incorrect results in many cases, given that the scale parameter μ determines the Gumbel distribution that is used to model the error term ϵ.

Ben-Akiva and Lerman (1985) have paid some attention to this issue. Even they recommended setting the scale parameter to a convenient value such as one if it is not identifiable when deriving the MNL model (3.19) and if the prerequisite of homoscedastic disturbances is given. However, homoscedastic disturbances imply that a traveller responds equally to a price difference of 10 EUR between two transport modes whether he is travelling only a few or several hundred kilometres. Obviously, a traveller travelling only a few kilometres is not likely to accept a price difference of 10 EUR if the cheapest acceptable transport mode costs only 1 EUR. In contrast, he is likely not to notice a difference of 10 EUR if the cheapest acceptable transport mode costs more than 100 EUR. The assumption of homoscedastic disturbances is therefore not given in the context of a European transport model.

In order to overcome this issue, which is related to the scale dependency of differences[14], and to facilitate the application of a single logit model at different price levels, the scale parameter μ should be adapted according to the corresponding price level. The necessity of modelling μ by a function in relation to the price level exists not only for European transport models, which deal simultaneously with short-distance, regional and long-distance trips, but also for transport models operating at a regional and at an urban level.

[14]The scale dependency of differences is a potential source of errors in statistical analysis (Stelzl, 1982).

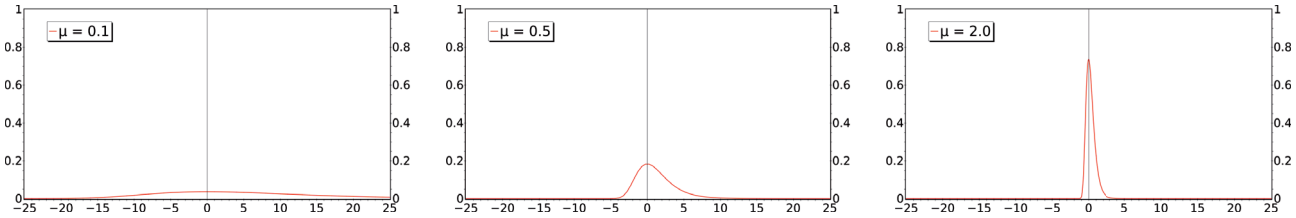

Figure 3.1: Probability density function of the Gumbel distribution for different values of μ

3.4.3 Properties of the Gumbel distribution

It is important to understand that the error term and, accordingly, the Gumbel distribution is an implicit part of the MNL model, though not explicitly stated in (3.19). For a better understanding of how to solve the issue of the scale dependency of the logit model on the cost level, the properties of the Gumbel probability distribution are briefly discussed in the following. The distribution function of a Gumbel-distributed random variable ϵ is $F(\epsilon)$ (3.20). Its density function is $f(\epsilon)$ (3.21).

$$F_{\eta,\mu}(\epsilon) = \exp[-e^{-\mu(\epsilon-\eta)}] \tag{3.20}$$

$$f_{\eta,\mu}(\epsilon) = \mu e^{-\mu(\epsilon-\eta)} \exp[e^{-\mu(\epsilon-\eta)}] \tag{3.21}$$

The Gumbel distribution is determined by the location parameter η and the scale parameter μ. Its mode[15] is μ, its standard deviation is $\sigma = \pi/\sqrt{6}\mu$ and its expected value is $\eta + \gamma/\mu$.[16]

Figure 3.1 shows the density function $f_{\eta,\mu}(\epsilon)$ (3.21) of the Gumbel distribution with $\eta = 0$ for three different values of the scale parameter, $\mu = \{0.1, 0.5, 2\}$, where the y-axis shows the value of $f(..)$ and the x-axis ϵ. For $\mu \to 0$, the standard deviation becomes infinite ($\sigma \to \infty$). Within this limit, the density function follows the x-axis and the random variable ϵ is uniformly distributed and can take on any value. With regard to the model of random utility (3.17), this can be understood as follows: the traveller decides completely randomly which transport mode to take and pays no attention to the generalised cost of travelling. Vice versa, for $\mu \to \infty$, the standard deviation becomes zero ($\sigma \to 0$) and the random variable becomes deterministic ($\epsilon = 0$). In this case, the traveller decides solely on the basis of deterministic criteria, and reacts to any cost differences, even if they are negligibly small.

The three examples outlined by Figure 3.1 relate to long-distance, regional and short-distance trips. With $\mu = 0.1$, the Gumbel distribution produces random values approximately in the range of -15 to 25.[17] This example clearly relates to long-distance trips and the error term can explain why one traveller chooses the cheaper transport mode, paying 100 EUR, and another traveller chooses a more expensive transport mode, paying 130 EUR for the same trip. For obvious reasons $\mu = 0.1$ cannot be used to model the mode-choice behaviour for short-distance trips related to travel costs of about 5 EUR. For such trips the value of $\mu = 2.0$ producing random values approximately in the range of -1 to 2 is more appropriate.

[15] In statistics, the mode is the value that appears most often in a set of data. Since the Gumbel distribution is a skewed probability distribution, the mode is not equal to the expected value and the median of the distribution.

[16] Here, $\gamma \sim 0.577$ is Euler's constant.

[17] Values like -100 and 100 are also possible, but rather unlikely.

For a similar reason, the MNL model should not be applied in a homogeneous manner for a group of travellers if their mode-choice behaviour is heterogeneous. For example, if the majority chooses the cheapest transport mode (e. g. $\mu = 2.0$) while some travellers do not care about travel costs (e. g. $\mu = 0.1$), different scaling parameters have to be chosen in order to explain the different mode-choice behaviours. This might be the case if travellers from different demand segments were modelled together, such as students and business travellers. For this reason, four demand segments are distinguished within this thesis when computing market shares of transport modes by the logit model, including business, commuting, private and vacation trips.

3.4.4 Expected Minimum Cost

The concepts of expected minimum cost (EMC) and expected maximum utility (EMU) can be applied almost equivalently. They both relate to the logit model and the logsum measure. As travel costs are the central focus of the transport models developed within this thesis, the concept of expected minimum cost is used. Following the definition of EMU in (3.18), EMC can be computed according to (3.22):

$$\text{EMC} = \frac{1}{-\mu} \cdot \ln \sum_m e^{-\mu c_m} + \text{cnst} \quad > 0 \tag{3.22}$$

where $\mu > 0$ is the heterogeneity parameter of the logit model which relates to the scale parameter of the Gumbel distribution determining the error term, c_m are travel costs related to transport mode m and cnst is the model constant. Bearing in mind the discussion on the scale dependency of the logit model, μ has to be chosen differently for each O/D relation depending on the level of travel costs. The computation of EMC is similar if an NMNL model is used (cf. Equations (3.23), (3.24), (3.25)).

Within this thesis, EMC are used as cost indicator in the deterrence function. Given that EMC are derived based on the specific market shares of transport modes computed by the logit model, this facilitates the consistent integration of the trip distribution and the modal split models. The integration of these two modelling steps is implemented in the IPAT model, improving the consistency of results and simplifying the model calibration.

3.4.5 Nested Multinomial Logit model

The nested multinomial logit (NMNL) model is an extension of the MNL model which is based on a mathematical property of the Gumbel function, namely that the expected maximum of two or more Gumbel-distributed random variables with the scale parameter μ is also Gumbel-distributed with the scale parameter μ. This specific property of the Gumbel function facilitates the consistent modelling of hierarchical decision processes, where each decision step is modelled by an MNL model.

In the IPAT model, for instance, the mode choice of travellers is modelled by a hierarchical decision process according to Figure 3.2. The first level is related to the general decision of preferring to travel with a land-based transport mode or to take the plane. The second level distinguishes between public transport modes and private road transport and the third level between rail and coach transport. The decision problem is, however, solved in reverse order. Firstly, expected minimum cost of travelling with a public transport mode N_{pbl} are computed on the basis of the generalised costs c_m of travelling with

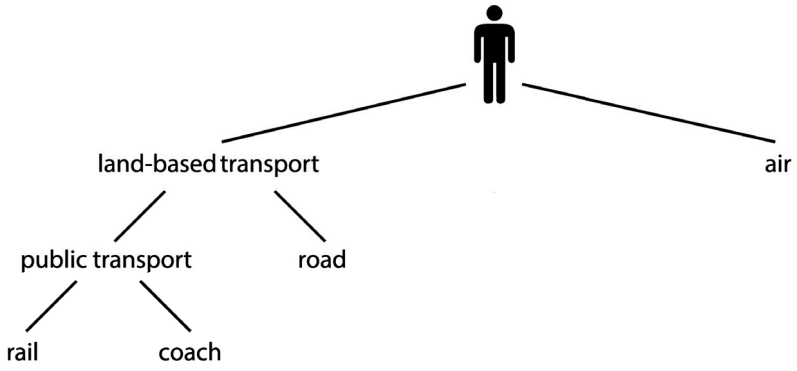

Figure 3.2: Nested multinomial logit model for mode choice

the corresponding transport modes m (3.23).[18] Secondly, the expected minimum costs of travelling N_{lnd} with land-based transport modes are computed (3.24). In the last step, expected minimum costs of travelling with all transport modes are derived (3.25). The EMC indicator further facilitates integration of the trip distribution and the mode choice models (see sec. 3.4.4).

$$N_{\text{pbl}} = \frac{1}{-\mu_1} \cdot \ln\left(e^{-\mu_1 c_{\text{rail}}} + e^{-\mu_1 c_{\text{coach}}}\right) + \text{cnst}_1 \tag{3.23}$$

$$N_{\text{lnd}} = \frac{1}{-\mu_2} \cdot \ln\left(e^{-\mu_2 c_{\text{road}}} + e^{-\mu_2 N_{\text{pbl}}}\right) + \text{cnst}_2 \tag{3.24}$$

$$\text{EMC} = \frac{1}{-\mu_3} \cdot \ln\left(e^{-\mu_3 c_{\text{air}}} + e^{-\mu_3 N_{\text{lnd}}}\right) + \text{cnst}_3 \tag{3.25}$$

The market shares of the transport modes can then be computed by the logit model (3.19) as follows:

$$P_{\text{air}} = e^{-\mu_3 c_{\text{air}}} \left(e^{-\mu_3 c_{\text{air}}} + e^{-\mu_3 N_{\text{lnd}}}\right)^{-1} \tag{3.26}$$

$$P_{\text{road}} = (1 - P_{\text{air}}) \cdot e^{-\mu_2 c_{\text{road}}} \left(e^{-\mu_2 c_{\text{road}}} + e^{-\mu_2 N_{\text{pbl}}}\right)^{-1} \tag{3.27}$$

$$P_{\text{rail}} = (1 - P_{\text{air}} - P_{\text{road}}) \cdot e^{-\mu_1 c_{\text{rail}}} \left(e^{-\mu_1 c_{\text{rail}}} + e^{-\mu_1 c_{\text{coach}}}\right)^{-1} \tag{3.28}$$

$$P_{\text{coach}} = (1 - P_{\text{air}} - P_{\text{road}} - P_{\text{rail}}) \tag{3.29}$$

Within this thesis, the NMNL model was preferred over the MNL model for the computation of market shares of transport modes, due to a fundamental limitation of the MNL model originating from the so-called "independence from irrelevant alternatives" (IIA) property of the logit model. As a result of this IIA property, the demand ratio between two mode alternatives is not affected by price changes of other mode alternatives. For example, if coach services become unavailable and if market shares are computed by the MNL model, the initial demand for coach trips distributes equally among all alternative transport modes in relation to their market shares. This is not realistic, as a stronger increase for rail transport can be expected than for road transport (some travellers have no car).

[18] Intermediate levels of a decision tree are frequently called "nests". For this reason, two different parameter names are used in the Equations (3.23) to (3.29). N is used for expected minimum cost computed at an intermediate level of a decision tree, whereas EMC is used solely for expected minimum cost computed at the top level of this tree.

The issue of the IIA property relates to the limitation of equal cross price elasticities in the MNL model, i. e. if travel costs change for one transport model, all competing modes respond in a similar way (cf. sec. 3.1). In the NMNL model, cross price elasticities can be chosen differently, for example by modelling competing public transport modes in a separate "nest" (cf. Fig. 3.2). In this case, rail transport benefits first if coach services become unavailable (and vice versa).[19] The NMNL model is therefore preferred over the MNL model if different cross price elasticities have to be modelled.

Besides the mode choice problem, the transport modelling problem itself can also be understood as a hierarchical decision problem including decisions on the destination, the transport mode and the routing as well as a general decision on whether to travel or not. For instance, an NMNL model covering the four-step transport modelling approach as a whole was recently introduced by Schulz (2012) for the identification of critical parts of the road infrastructure in the German state of Baden-Württemberg.

Understanding the four-step model as a hierarchical decision problem and formulating an NMNL model might be advantageous in some cases, as the EMC indicator or, to be more precise, the EMU indicator that is computed at the top level of this NMNL model, can be converted into monetary units in order to be used as an evaluation measure (see, for example, de Jong et al. (2005); de Jong et al. (2007)). It should be emphasised, however, that for the application at European scale the logit model has several specific properties which may be disadvantageous with regard to solving the trip distribution problem (see, for example, Lohse et al. (1997); Lohse et al. (2006)) and the routing problem (see, for example, Cascetta et al. (1996); Kato et al. (2003)). However, this issue is not relevant within this thesis, as the NMNL model is only applied for computing the market shares of transport modes within the third step of the four-step transport model.

[19] In transport literature, the problem of IIA is frequently illustrated based on the "red bus, blue bus paradox" (see, for example, Train (2009); Koppelman and Bhat (2006)).

Chapter 4

Trip distribution modelling revisions

Following the four-step approach in transport modelling, the trip distribution model is applied in the second step. It produces the best possible prediction of travellers' destination choices in terms of trip estimates for each pair of travel zones, based on information on the generated trip demand and the number of attracted trip endings for each travel zone as well as information on the generalised cost of travelling. The trip distribution model computes trip estimates for each pair of travel zones and stores this information in the so-called O/D trip matrix. Due to the quadratic dependency of the trip matrix on the number of travel zones, the methodology of the trip distribution model determines the overall complexity of the transport model to a large extent.

For the computation of the O/D trip matrix, transport literature suggests various distribution models of different complexity. A good overview of the different approaches can be found, for instance, in Cascetta et al. (2007) and Cascetta (2009). For this thesis, gravity-based approaches are particularly important and therefore discussed in detail with a specific focus on the deterrence function. In addition, growth-factor methods are described.

This chapter is structured as follows: in the first section, the advantages and limitations of the less complex growth-factor methods are investigated. Section 4.2 introduces the gravity-based approach, providing a more general framework for the computation of the trip matrix. This approach is further elaborated, and an improved model is introduced relying on the competition-enhanced accessibility indicator. Section 4.3 investigates the role of the deterrence function, which is a key component of the gravity model, and introduces a revised functional form. The last section summarises the methodology of the revised trip distribution model.

4.1 Growth-factor methods

Growth-factor methods rely on rather simple rules and enable the forecast of a future trip matrix from a given trip matrix by applying growth factors. These factors can be derived from the changes of the trip demand and supply indicators between the base year and the forecast year. It is, for instance, very likely that the generated trip demand in a region will increase between the base year and the forecast year if the population of this region and their purchasing power increase between the two years. The method can be outlined by the following equation:

$$\hat{t}_{ij} = t_{ij} g_{ij} \tag{4.1}$$

where $\hat{t}_{ij}$ is the estimated trip demand in the forecast year, t_{ij} the known trip demand between origin i and destination j in the base year, and g_{ij} the applied growth factor.

Rather simple models consider a uniform growth factor for all O/D relations. The factor can be derived, for instance, based on changes in the population. Such models are limited and will inevitably fail if people have moved from one region to another one between the base year and the forecast year but the overall population stays constant. For this reason, improved models are often preferred which are sensitive to changes at regional level. These models include the singly and doubly constrained growth-factor method in which one or two constraints must be determined for the forecast year. These constraints refer to estimates on trip generation and trip endings for each region. While for singly constrained methods growth factors can be derived straightforwardly, for doubly constrained methods these factors have to be iteratively determined, e. g. by applying the Furness algorithm.

In principle, growth-factor methods imply two important limitations. Firstly, they depend on a given trip matrix for the base year and, secondly, the forecast scenario has to be consistent with the base year, notably with regard to the definition of travel zones. Growth-factor methods are therefore more appropriate for computing short-term forecasts than for computing long-term forecasts.

4.1.1 Furness method

The Furness method (Furness, 1965) belongs to the group of doubly constrained growth-factor methods. It derives the trip matrix $\hat{T}_{ij}$ for the forecast year following an iterative approach by sequentially applying two balancing factors a_i and α_j to the matrix of the given base year T_{ij}. As a result, the given matrix is progressively adjusted to meet the two given constraints for the forecast year on trip generation $\hat{O}_i$ in the origin zone i (4.2) and trip endings $\hat{D}_j$ in the destination zone j (4.3):

$$\sum_{j=1}^{n} \hat{T}_{ij} = \hat{O}_i \tag{4.2}$$

$$\sum_{i=1}^{m} \hat{T}_{ij} = \hat{D}_j \tag{4.3}$$

The Furness method stops after K iterations. It can be described as follows:

$$\hat{T}_{ij} := \left(t_{ij}^{(\mathrm{K})}\right) \tag{4.4}$$

$$t_{ij}^{(\mathrm{k})} = \tau_{ij}^{(\mathrm{k})} a_i^{(\mathrm{k})} \tag{4.5}$$

$$\tau_{ij}^{(\mathrm{k})} = t_{ij}^{(\mathrm{k\text{-}1})} \alpha_j^{(\mathrm{k})} \tag{4.6}$$

$$\left(t_{ij}^{(0)}\right) := T_{ij} \tag{4.7}$$

where $t_{ij}^{(\mathrm{k})}$ (4.5) identifies the derived trip estimates from region i to j in iteration (k), and $t_{ij}^{(0)}$ (4.7) the given trip demand for the base year. In each iteration k, the balancing factors a_i and α_j are updated, based on the comparison of the two constraints ($\hat{O}_i$, $\hat{D}_j$) with the totals ($\tilde{O}_i^{(\mathrm{k})}$, $\tilde{D}_j^{(\mathrm{k\text{-}1})}$) of the most recent version of the matrix. The computation of the two balancing factors satisfies the following equations:

$$\alpha_j^{(\mathrm{k})} = \frac{\hat{D}_j}{\tilde{D}_j^{(\mathrm{k\text{-}1})}} \quad \text{with} \quad \tilde{D}_j^{(\mathrm{k\text{-}1})} = \sum_{i=1}^{n} t_{ij}^{(\mathrm{k\text{-}1})} \tag{4.8}$$

$$a_i^{(\mathrm{k})} = \frac{\hat{O}_i}{\tilde{O}_i^{(\mathrm{k})}} \quad \text{with} \quad \tilde{O}_i^{(\mathrm{k})} = \sum_{j=1}^{n} \tau_{ij}^{(\mathrm{k})} \tag{4.9}$$

More briefly, the Furness algorithm can be illustrated as follows: it starts with a given trip matrix T_{ij} for the base year and two given constraints for the forecast year in terms of originating trips $\hat{O}_i$ and trip endings $\hat{D}_j$ that are commonly computed by the trip generation model. The algorithm then continuously computes new balancing factors, a_i and α_j, and alternately applies these factors to the most recent version of the trip matrix. This progressively adjusts the matrix to meet both constraints. The algorithm stops after K iterations if the matrix has been perfectly adjusted, i. e. the new balancing factors have converged to one. However, the convergence of the Furness algorithm is not guaranteed (Ortúzar and Willumsen, 2001, p. 185).

Recalling that the Furness method relies on a given trip matrix, it is evident that the algorithm cannot cope with any changes in the definition of the zoning system. Its application is further limited to scenarios with slight differences between the base and the forecast year in terms of generated trip demand and trip endings in each travel zone as well as in terms of observed and expected travel patterns. The application of the Furness method for forecasting a given trip matrix is particularly inadequate in the case of comprehensive infrastructure changes such as high-speed rail corridor upgrades. For instance, the German city of Montabaur was only connected to the high-speed rail network in 2002 (Demuth, 2004). Since then, the two large cities of Frankfurt and Köln can be reached within one hour by train. Given the significantly improved accessibility of Montabaur, many service companies have settled in this city in the meantime. Obviously, the spatial structure and the settlement structure have changed significantly between the years 2000 and 2010 in this particular town and, in consequence, the travel patterns of the population living in this region have also changed.[1] It is, for instance, very likely that a former trip matrix referring to the year 2000 reports zero business and commuting trips for some O/D relations which are very important today. Following a multiplicative approach, however,

[1] In contrast to ordinary geographical maps displaying the distance between two locations, so-called time maps can be produced, based on the travel time between two locations. Given the development of the high-speed rail network, Europe appears to have shrunk since the appearance of the first TGV in France (Spiekermann and Wegener, 1993, p. 29). Accordingly, larger distances can be bridged within one hour than in the past and mobility patterns have changed.

the Furness algorithm cannot cope with zero values in the base matrix, as these would also remain as zeros in the extrapolated matrix for the forecast year.

Transport literature shows awareness of this issue and suggests the application of the so-called seeding strategy. Empty O/D relations in the trip matrix of a base year reporting a trip demand of zero can be seeded with ones which depend, however, on the transport modeller's best guess. This seeding strategy should be applied with caution, since small updates of the base matrix generally have much larger impacts on the structure of the forecast matrix. If a base matrix is seeded in different ways, the respective forecast matrices are likely to be very different. This not limited to changes of zeros into ones but also to other changes. This issue relates to the criterion of numerical stability which identifies a desired property of numerical algorithms. Numerical stability is fulfilled if an algorithm produces similar results given small changes of the input data (Stewart, 1973, p. 76). According to this definition, the seeding strategy and the Furness method are not numerically stable methods.

It should be noted that the Furness algorithm is also relevant with regard to the doubly constrained gravity trip distribution model (see sec. 4.2). In this case, the Furness algorithm is often considered to be a necessary but rather questionable method that improves the quality of the modelled matrix. However, the application of the algorithm does not improve the explanatory power of the trip distribution model and may even reduce it. This is the case, for instance, if inconsistencies in the modelled trip matrix are caused by an improper calibration of the trip distribution model or by following a simplified methodology lacking important explanatory factors. In this case, the better way would obviously be to carefully review and to update the methodology of the trip distribution model rather than to apply the Furness algorithm.

4.1.2 Limitations of growth-factor methods

Growth-factor methods can only be assigned in a broader sense to the group of trip distribution models. Their main advantage is the simplicity of the forecasting method, which enables the extrapolation of the trip matrix of a known base year to a forecast year given updated constraints on trip generation and trip endings. For this reason, growth-factor methods are not sensitive to any changes in the transport system and changes in the travel behaviour of people unless captured by the growth factors.

Growth-factor methods may be adequate for short-term planning if the changes between the base year and forecast year are not too significant. Otherwise, their application should be well thought-out. For instance, if the population of a travel zone has increased significantly between the base year and the forecast year, but the capacity of the infrastructure was not improved accordingly, congestion may increase. In consequence, travellers are stimulated to adapt their behaviour with regard to destination, mode and route choice.

A general limiting property of growth-factor methods is the prerequisite of an invariable zoning system, i. e. travel zones must not be modified between the base year and the forecast year. Given frequent changes of the European NUTS zones and, accordingly, breaks in region-based time series data, the application of growth-factor methods can be very difficult and is not always reasonable. Although it would be possible to disaggregate region-based indicators, such as population and GDP, between different sets of travel zones, this would open the door to another problem, namely the so-called modifiable areal unit problem which is investigated more in detail in section 7.1.

4.2 Gravity-based approaches

A common approach for the computation of the O/D trip matrix within the trip distribution model relies on the "theory of gravity" and the concept of a "deterrence function" which models peoples' disincentive to travel due to travel cost. Gravity-based approaches are therefore sensitive to infrastructure changes. Such changes can be simulated in the network models, cost indicators can be updated and the gravity model can be reapplied. For the same reason gravity models enable the derivation of the O/D trip matrix from scratch, which is a major advantage compared to growth-factor methods building on a given matrix for a base year.

The gravity model was rigorously used for the first time in the middle of the 20th century.[2] It is still the subject of current research, as the standard gravity model has to be improved in several directions. For instance, the standard model disregards competition effects originating from the heterogeneous spatial distribution of demand and supply[3] as well as congestion effects originating from capacity constraints of the transport system.[4] Amongst others, the standard gravity model is discussed in de Grange et al. (2010), providing a brief comparison of three suggested model improvements. However, these improvements only relate to the introduction of additional modelling coefficients[5] but not a general methodological revision of the gravity model. As the coefficients cannot be transferred from one region to another one, the discussed models are not suitable for application at European scale and a different approach is taken in this thesis. First, a general rule for modelling competition effects is developed based on the competition-enhanced accessibility indicator that was introduced in section 3.3.2. Secondly, an incremental approach is developed in order to model congestion effects, which is implemented in the HIPAT model (see sec. 9.2.6).

In the next sections, the concept of the standard gravity trip distribution model and several extensions are discussed, including the elasticity-based gravity model and the doubly constrained model. Subsequently, accessibility-based approaches are discussed and an improved trip distribution model is introduced which is implemented in the IPAT and the HIPAT models.

4.2.1 Gravity model

Gravity-based modelling approaches provide a more general concept for the computation of the O/D trip matrix than growth-factor methods. They require information on:

- the generated trip demand for each travel zone;
- the attraction value or, to be more precise, the number of trip endings for each travel zone; and
- the generalised cost of travelling for each pair of travel zones.

[2] See Ortúzar and Willumsen (2001) p. 182.

[3] In agglomerations travellers can choose between many competing destinations, while in sparsely populated regions their choice set is limited.

[4] In aggregated transport models, travel costs are often computed only once, based on pre-loaded networks. This approach is reasonable in many cases but it is not particularly accurate if traffic loads (congestion) change significantly within the same day. For this reason, the MOSART model that is applied for the agglomeration area "Grand Lyon" distinguishes different load scenarios, including a night-time, an off-peak and a peak-hour scenario (Bonnafous et al., 2009).

[5] Modelling coefficients must be determined by empirical analysis based on observed data.

Gravity-based trip distribution models follow Newton's law of universal gravitation, outlined by (4.10), stating that a force F can be observed between two planets. This force is directly proportional to their masses m_1 and m_2, and inversely proportional to the square of the distance r between the centre of both masses:

$$F = G\,\frac{m_1 m_2}{r^2} \tag{4.10}$$

where G is a gravitational constant. For the trip distribution in transport modelling, instead of a gravitational force F, the number of trip estimates T_{ij} from origin i to destination j is computed (4.11), where the two masses are replaced by information on generated trip demand O_i and the number of trip endings D_j.[6] The distance indicator is replaced by generalised cost c_{ij}. In its simplest form, the gravity trip distribution model satisfies:

$$T_{ij} = \alpha O_i D_j f(c_{ij}) \tag{4.11}$$

where α is a model constant and f identifies the deterrence function which models travellers' sensitivity to varying travel cost. The exact value of α is commonly determined during the model calibration.

The model (4.11) as such is not restricted by any constraints (e. g. on the number of generated trips and trip endings for each travel zone) and is called the zero-constrained gravity model. The existence of only one model constant (α) for the whole trip distribution model, however, is an important limitation for its transferability, taking into account that this single parameter has to cover all kinds of heterogeneities, such as heterogeneous levels of trip generation, trip endings and competition effects. Given that any two O/D relations are generally not uniform in terms of these three boundary conditions, trip estimates can be expected to be more accurate for some relations and less accurate for others. Moreover, if the model (4.11) is transferred spatially or temporally, the boundary conditions commonly change, which would require adjustment of the parameter α.

Some drawbacks of the zero-constrained gravity model can be shown easily if we assume a doubling of trip generation and trip endings and if we apply the model again. We also would expect a doubling of computed trip estimates, but the model predicts a quadruplication according to (4.12).

$$\alpha(2O_i)(2D_j)f(c_{ij}) = 4T_{ij} \tag{4.12}$$

The zero-constrained gravity model (4.11) can be improved slightly by introducing two elasticity coefficients (λ_1, λ_2) which are typically chosen in the range of 0.5 to 1.0. The improved model (4.13) is called the elasticity-based gravity model and satisfies the following equation:

$$T_{ij} = \alpha O_i^{\lambda_1} D_j^{\lambda_2} f(c_{ij}) \quad \text{with} \quad \lambda_1, \lambda_2 > 0 \tag{4.13}$$

where α is again a model constant. However, the general transferability of the elasticity-based gravity model is not given as well, for the reason mentioned above, i. e. heterogeneities of travel zones and O/D relations. Hence, α should be chosen differently for each O/D relation.

[6] The indicators O_i and D_j refer to the total trip demand, unlike in section 2.3.4 in which $\tilde{O}_i$ and $\tilde{D}_j$ refer to inter-zonal trips only.

The transferability of the gravity model can be improved, for instance, by introducing additional constraints for each travel zone referring to the number of originating trips and trip endings. These constraints can be derived from transport statistics, if available. Lenormand et al. (2012a) and Lenormand et al. (2012b) therefore apply statistics depicting the number of in-commuters and out-commuters per region for the modelling of commuting networks (i. e. trip flows). Intra-zonal commuting trips are not considered in the presented model. If detailed commuter statistics are not available, these two constraints can be defined, for instance, based on the number of economically active people in a region and the number of workplaces. This approach is much more flexible. It also includes intra-zonal trips and can be transferred to other trip purposes.

These types of models, which have to fulfil additional boundary conditions on trip generation and trip endings, include the singly and the doubly constrained gravity model. At first glance, they look very similar to the previously discussed growth-factor methods (cf. sec. 4.1) but they are superior for two reasons. Firstly, gravity models do not require a given base matrix and, secondly, they are more sensitive to changing travel costs due to the application of a deterrence function. According to Ortúzar and Willumsen (2001), the doubly constrained gravity model can be described as follows:

$$T_{ij} = x_i O_i y_j D_j f(c_{ij}) \tag{4.14}$$

$$x_i^{-1} = \sum\nolimits_j y_j D_j f(c_{ij}) \tag{4.15}$$

$$y_j^{-1} = \sum\nolimits_i x_i O_i f(c_{ij}) \tag{4.16}$$

where x_i and y_j are correction factors that are computed following the Furness algorithm starting with $x_i, y_j = 1$ (cf. sec. 4.1.1).

It should be stressed that the doubly constrained model should be applied with caution, since the quality of the derived matrix (4.14) depends on the reliability of the initial O/D trip matrix which is computed with $x_i, y_j = 1$. It is therefore reasonable to allow only a few iterations of the Furness algorithm and then to validate the correction factors x_i and y_j. If they differ much from 1, the initially computed trip matrix must be adjusted significantly in order to meet the two constraints on trip generation and trip endings. This indicates that the applied distribution model disregards important explanatory factors and that the methodology of the trip distribution model should be improved, e. g. by revising the assumptions on trip generation and trip endings, by revising the deterrence function, and by introducing additional explanatory factors.

4.2.2 Accessibility-based gravity model

In comparison to the ordinary gravity trip distribution model (4.11), accessibility-based gravity models consider an additional explanatory component measuring the accessibility of a region. Depending on the trip purpose, different accessibility indicators can be considered (cf. sec. 3.3). For private and business trips, Hansen's gravitational accessibility (3.11) is more suitable, while for commuting and vacation trips, the competition-enhanced accessibility indicator (3.16) promises better results. This leads to the following two different trip distribution models $\mathrm{M_H}$ (4.17) and $\mathrm{M_{CMP}}$ (4.18) with the well-known parameters O, D and f from the gravity model (4.11). Here, all indices i, j, k, l refer to travel zones. The parameter γ_j (4.19) refers to the level of competition. It is computed based on $N_k^{(\mathrm{o})}$ and

$N_l^{(\mathrm{d})}$, namely the generated trip demand in origin k and the number of trip endings in destination l.

$$\mathrm{M_H}: \quad T_{ij} = O_i \frac{D_j f(c_{ij})}{A_i} \qquad \text{with} \qquad A_i = \sum_{j=1}^{n} D_j f(c_{ij}) \tag{4.17}$$

$$\mathrm{M_{CMP}}: \quad T_{ij} = O_i \frac{D_j f(c_{ij})}{\gamma_j A_i^{(3)}} \qquad \text{with} \qquad A_i^{(3)} = \sum_{j=1}^{n} \frac{1}{\gamma_j} D_j f(c_{ij}) \tag{4.18}$$

$$\gamma_j = \sum_{k=1}^{m} \frac{1}{\mu_k} N_k^{(\mathrm{o})} f(c_{kj}) \qquad \text{with} \qquad \mu_k = \sum_{l=1}^{n} N_l^{(\mathrm{d})} f(c_{kl}) \tag{4.19}$$

The formulae (4.17) and (4.18) can be interpreted as follows: both models distribute the generated trip demand to all destinations, based on the accessibility of each destination in relation to the total accessibility of all destinations ($D_j f(c_{ij})/A_i$). Their basic difference is that $\mathrm{M_H}$ assumes the absence of competition between travellers for destination opportunities, while $\mathrm{M_{CMP}}$ explicitly models this competition through the additional explanatory factor γ_j. Hence, $\mathrm{M_{CMP}}$ is the more powerful model. It should be preferred if the number of trip endings in a region can be estimated accurately based on known supply indicators (e. g. workplaces for commuting trips or accommodations for vacation trips). It can also be applied for business and private trips.

The methodological revision of the gravity model, by introducing an additional explanatory factor modelling the accessibility of a region, improves the explanatory power of the trip distribution model. This eliminates the necessity of constraining the model, and additionally ensures its spatial and temporal transferability. This transferability is not guaranteed for the ordinary gravity model (4.11). For instance, if we assume that better-paid jobs are more attractive and if we further assume that the incentive of commuters to spend time and money for travelling correlates with the expected income, we can expect a higher level of competition for more attractive jobs. In this case, the level of competition γ_j (4.18) is greater than one (demand is higher than supply) and the accessibility-based trip distribution model $\mathrm{M_{CMP}}$ adjusts the demand accordingly. However, if we apply a zero-constrained model (4.11) ignoring the factor γ_j, we overestimate the trip demand to those regions providing more attractive jobs. At the same time, we underestimate the trip demand to those regions providing less attractive jobs where $\gamma_j < 1$.

To resolve this issue and to establish the transferability of the gravity model, the doubly constrained model (4.14) was introduced. Though the accessibility-based model (4.18) and the doubly constrained model appear to be similar, the two models are quite different. The latter ensures its spatial and temporal transferability by two sets of correction factors which are derived during an iterative counterbalancing process. However, this also whitewashes other systematic shortcomings in the chosen modelling approach, like potential lack of important explanatory factors, the application of an unsuitable deterrence function and travel zones that are too heterogeneous.

The accessibility-based trip distribution model $\mathrm{M_{CMP}}$ (4.18) applies no correction factors. If the computed trip matrix differs from expected trip endings, the reasons can be traced back to shortcomings in the chosen modelling approach. For this reason, the model $\mathrm{M_{CMP}}$ is chosen as a basis for the further methodological revision of the trip distribution model in the following sections.

4.3 Deterrence functions

The frequency of trip occurrences decreases with increasing travel cost. This relationship is tackled by the deterrence function in the gravity trip distribution model. Typically, strictly monotonic decreasing functions with $0 \leqslant f \leqslant 1$ are applied to distribute the generated trip demand to the accessible destinations in relation to the generalised cost of travelling. The functional shape of these functions can be determined on the basis of the so-called trip length distribution (TLD), which is generated from observed travel demand data, and which reports the frequency of trip occurrences by distance band. Depending on the trip purpose and the scope of the model, deterrence functions of different complexity can be considered for modelling the distribution of trips.

The chapter is structured as follows. The first section investigates and assesses three classical deterrence functions which are frequently used. Section 4.3.2 introduces the more complex EVA function and discusses its advantages in comparison to a classical deterrence function, but also the major challenge related to its calibration. The discussion on the definition of an optimal function in the first two sections finally leads to the development of the composite deterrence function in section 4.3.3, combining the simplicity of classical deterrence functions and the advantages of the EVA function.

4.3.1 Classical deterrence functions

Table 4.1 shows definitions and implicit elasticity curves[7] for three classical deterrence functions that are often applied in trip distribution modelling (cf., for example, Ortúzar and Willumsen (2001)): the exponential, the power, and the combined function. The coefficients $(\alpha, \beta, \gamma, \delta > 0)$ are chosen so that

Table 4.1: Classical deterrence functions

Descriptive term	Equation	Elasticity
Power function	$f(c) = c^{-\alpha}$	$-\alpha$
Exponential function	$f(c) = e^{-\beta c}$	$-\beta c$
Combined function	$f(c) = c^{\gamma} e^{-\delta c}$	$\gamma - \delta c$

the deterrence function comes close to the given TLD. The TLD can be compiled based on observed travel demand data that is provided, for instance, by national travel surveys. It should be noted that in the following examples discussing different deterrence functions the variable c does not refer to generalised cost but rather to the travel distance. This simplification does not represent a restriction for the findings.

The following analysis is carried out based on the frequency of trip occurrences by distance bands in Germany. It relies on travel diary data for 44,581 motorised trips surveyed by the German mobility panel (MOP) in 2015 (Weiß et al., 2016). The data reports trip lengths between 0.1 and 1000 km with an average trip length of 16.9 km. Figure 4.1 shows the shape of the derived TLD and the three classical deterrence functions in relation to the travel distance (in kilometres) in six different graphs, in which deterrence functions are drafted in red and the TLD, compiled based on distance bands, is

[7] Elasticity curves are computed according to (3.4) (cf. sec. 3.1.1).

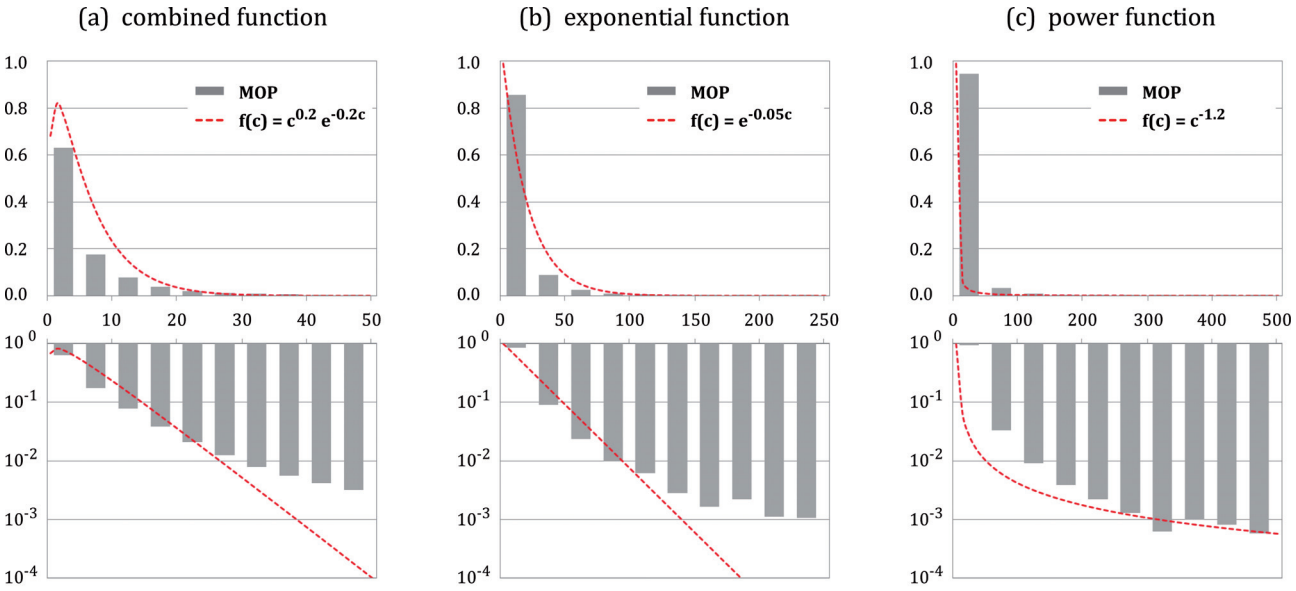

Figure 4.1: Modelling TLD with different deterrence functions

outlined in grey bars.[8] From brief analysis of the TLD it can be concluded that the majority of the surveyed trips have short travel distances, and the frequency of trip occurrences quickly becomes zero for larger travel distances. This pattern has to be reproduced by the deterrence function.

In the top row data is displayed in the usual scale and in the bottom row in the logarithmic scale to the basis of 10 (y-axis). In addition, limits of the x-axis are chosen differently, so that the data can be displayed, compared and analysed in a reasonable way. For instance, the MOP reports the occurrence of 28,198 motorised trips in the distance band (0; 5] km accounting for about 60% of all reported trips (left-hand graph in the top row) but only 26 trips in the distance band (450; 500] km accounting for about 0.06% (right-hand graph in the bottom row).

The graphs on the left-hand side focus on short-distance trips and the TLD is displayed in distance bands with a width of 5 km up to 50 km. The graphs in the centre display the TLD in distance bands with a width of 25 km up to 250 km and the graphs on the right-hand side display the TLD in distance bands with a width of 50 km up to 500 km. Overall, the values of the TLD decrease monotonically with increasing travel distance with the exception of the distance bands (150; 175] km in the central figure and (300; 350] km in the right-hand figure.

Short-distance trips can be best modelled by the combined function (Fig. 4.1a) as the factor c^{γ} reproduces the peak value at the first distance band up to 5 km. The level of the TLD then drops almost constantly in the logarithmic representation which can be reproduced by the factor $e^{-\delta c}$. In the combined function, however, a comparable high value has to be chosen for the decay coefficient δ in order to counterbalance the effect of c^{γ}. Hence, the combined function becomes progressively worse for distances larger than 30 km, as the exponential function dominates the behaviour of the combined function for large values and runs to zero.

[8] For reasons of presentation, the widths of all bars were reduced in order to improve the distinction between different distance bands, e. g. the first bar (left-hand graph in the top row) refers to the distance band (0; 5] km.

Considering trip distances between 25 and 150 km, the shape of the TLD drops almost linearly in the logarithmic scale as shown in the central figures. This linear pattern can be reproduced very accurately by the simple exponential function $e^{-\beta c}$ (Fig. 4.1b). It should be noted, however, that the outlined exponential function with $\beta = 0.05$ tends to overestimate the demand for trips between 25 and 100 km, while it underestimates the demand between 100 km and 150 km.

Considering long-distance trips, the frequency of trip occurrences drops only slightly with increasing travel distance. This pattern can be reproduced by the power function $c^{-\alpha}$ (Fig. 4.1c) which is also suitable for distances larger than 500 km. It should be stressed that the decay coefficient α should not be smaller than 1.2 in order to improve the fit of the power function for shorter travel distances. It would simultaneously lead to an overestimation of long-distance trips which cannot be counterbalanced.

It can be concluded that each of the discussed deterrence functions (Tab. 4.1) provides advantageous properties which are relevant for the modelling of either short, medium, or long-distance trips. The power function provides a constant elasticity term and is suitable for long-distance trips. The exponential function provides a linear elasticity function and can therefore successfully reproduce the pattern of the TLD for medium-distance trips. The combined function also provides a linear elasticity function. Due to its factor c^{γ}, it even increases for very small travel distances. Hence, the combined function is better suited to modelling short-distance trips than the exponential function. However, none of the classical deterrence functions are capable of reproducing the whole TLD with sufficient accuracy.

4.3.2 Deterrence function of the EVA model

The deterrence function, which models disincentive to travel as a function of travel costs, is probably the most important component of a gravity-based trip distribution model. As shown above, the three investigated functions are only appropriate either for modelling short-distance, medium-distance or long-distance trips (cf. Fig. 4.1), but none of the three classical functions are capable of reproducing the observed TLD for all travel distances with sufficient accuracy. Moreover, Lohse et al. (1997) stress that travellers respond differently to cost changes at different cost levels. This pattern cannot be modelled by the constant and the linear elastic function underlying the power, the exponential and the combined deterrence functions (Tab. 4.1).

In order to overcome these limitations, different types of functions have been introduced. In this thesis, the EVA function (Lohse et al., 1997, p. 84) is used, which is already applied in the PAT model with minor adjustments (cf. sec. 2.3.4). The EVA function is capable of reproducing complex elasticity trajectories and therefore enables the simultaneous modelling of short, medium and long-distance trips. Following the ideas behind its definition, the concept of a composite deterrence function was developed within this thesis and applied in the IPAT and the HIPAT models. It is introduced below in section 4.3.3.

The EVA function consists of three calibration parameters $E, F, G > 0$, and satisfies (4.20).

$$f(c) = (1 + c)^{-\phi(c)} \tag{4.20}$$

$$\phi(c) = \frac{E}{1+\exp(F-Gc)} \tag{4.21}$$

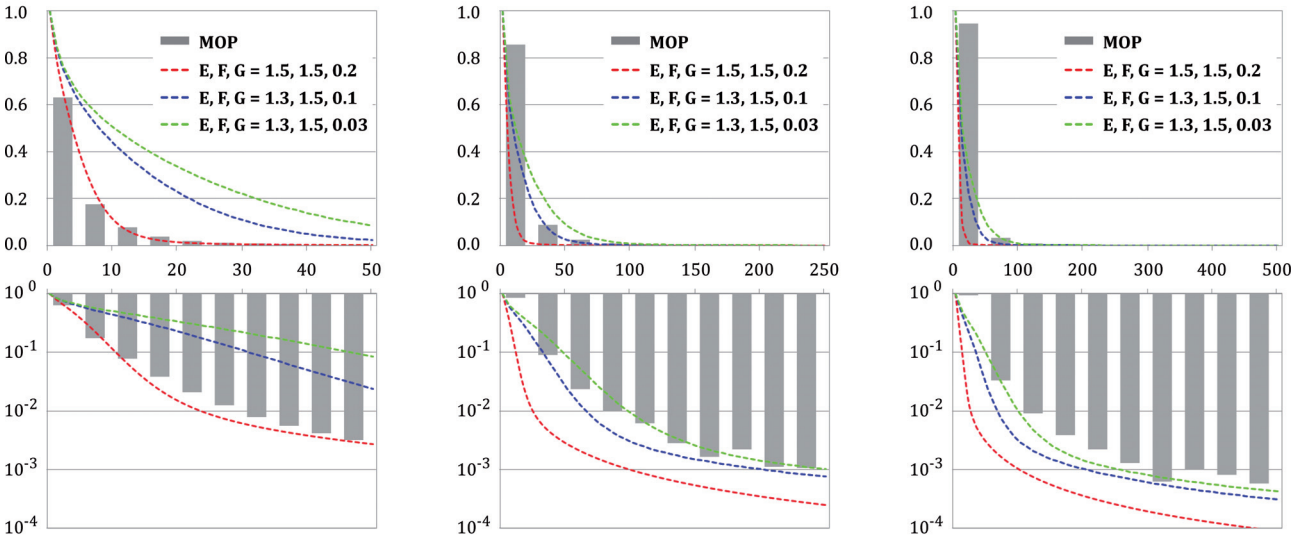

Figure 4.2: Calibrating the EVA function for different distance segments

where c are travel costs (travel distances in this example) and E determines the shape of the EVA function with regard to long-distance trips, since:

$$\lim_{c \to \infty} \exp(F - Gc) = 0 \quad \Rightarrow \quad \lim_{c \to \infty} \phi(c) = E \tag{4.22}$$

For short-distance trips, the value of the EVA function is almost constant, since:

$$\lim_{c \to 0} \phi(c) = \frac{E}{1 + \exp(F)} \quad \Rightarrow \quad \lim_{c \to 0} f(c) = 1 \tag{4.23}$$

Hence, the function $\phi(c)$ (4.21) is almost constant for short and long-distance trips. The range in between, where the value of $\phi(c)$ changes between $E/(1 + \exp(F))$ and E, is rather narrow and the calibration of the EVA function to medium-distance trips is therefore challenging.

Figure 4.2 shows the EVA function for different values of the model parameters E, F, G and outlines the three different distance segments similar to Figure 4.1. One of the major disadvantages of the EVA function is its closed representation, since the function cannot be independently calibrated to short, medium, and long-distance trips. While E is closely related to long distances, the parameters F and G, which calibrate the EVA function to short and medium-distance trips, cannot be independently optimised. To say the least, the larger the value of G, the more volatile the behaviour of the function changes, since the denominator of $\phi(c)$ (4.21) changes abruptly between infinity and one. Calibration of the EVA function is therefore very challenging and should be carried out manually.

4.3.3 Composite deterrence function

European transport models operating at the level of NUTS-2 or NUTS-3 regions have to cope with a large degree of heterogeneity, including the inhomogeneous size, shape and weight of the NUTS regions. As a consequence, the estimated values of the calibration parameters of deterrence functions (e. g. E, F, G for the EVA function) are not optimal for each O/D relation. In order to improve the accuracy

of the deterrence function, the set of calibration parameters should be revised, e. g. origin-specific and destination-specific parameters could be introduced. These additional calibration parameters could be determined following an iterative approach, for instance by comparing a modelled trip matrix with expected output indicators, like the distribution of trip length by origin and the travelled mileage by origin, as well as by updating the additional calibration parameters.

An iterative calibration approach can be implemented with reasonable effort for simple deterrence functions which only depend on one parameter, like the power and the exponential function. For the more complex EVA function depending on three parameters, the calibration identifies a multi-dimensional problem. Accordingly, the procedure for updating E, F, G and for adjusting the EVA function to the data is not clear. Changing the parameter E, for instance, adjusts the trajectory of the function for long-distance trips but simultaneously changes its trajectory for short and medium-distance trips as well.

If we want to combine both, i. e. the simplicity of classical deterrence functions and the explanatory power of the EVA function, we can follow the principle of superposition by approximating the complex trajectory of the EVA function by more simpler functional terms. This leads to the composite deterrence function satisfying (4.24) with $\sum_z \omega_z = 1$, $\omega_z \geqslant 0$ for $z \in \{S, M, L\}$:

$$f(c_{ij}) = \omega_S f_S(c_{ij}) + \omega_M f_M(c_{ij}) + \omega_L f_L(c_{ij}) \tag{4.24}$$

$$f_S(c_{ij}) = {c_{ij}}^{\gamma} e^{-\delta c_{ij}} \tag{4.25}$$

$$f_M(c_{ij}) = e^{-\beta c_{ij}} \tag{4.26}$$

$$f_L(c_{ij}) = {c_{ij}}^{-\alpha} \tag{4.27}$$

where the combined function f_S is calibrated to short-distance trips, the exponential function f_M to medium-distance trips and the power function f_L to long-distance trips.

The calibration of the composite deterrence function follows an iterative approach. In the first step, the three weighting factors (ω) must be determined. Reasonable values relying on practical experience from calibrating the IPAT model within the HIGH-TOOL project are: $\omega_L \in (0, 0.05)$, $\omega_M \in (0.08, 0.12)$, and $\omega_S \in (0.85, 0.9)$. Having determined the weighting factors, f_L can be calibrated almost independently to long-distance trips, then f_M to medium-distance trips and finally f_S to short-distance trips. In the second step, the model must be applied and the outcomes must be validated. Given an almost independent calibration of f_L, f_M and f_S, interdependencies between the three functions are relatively small. For this reason, the calibration factors generally only require a slight update and the calibration process requires few iterations. This is of particular importance with regard to the implementation of a European transport model, which has to cope with a large degree of heterogeneity. The composite deterrence function can be calibrated straightforwardly following the outlined method and is therefore considered for the methodological revision of the trip distribution model and the development of the IPAT and the HIPAT models.

It should be noted that the calibration of the composite deterrence function is still challenging. For example, the exact shape of the TLD depends on the chosen scale. Figure 4.3 outlines this issue in which the TLD is shown in grey bars and the composite deterrence function satisfying (4.28) in green.

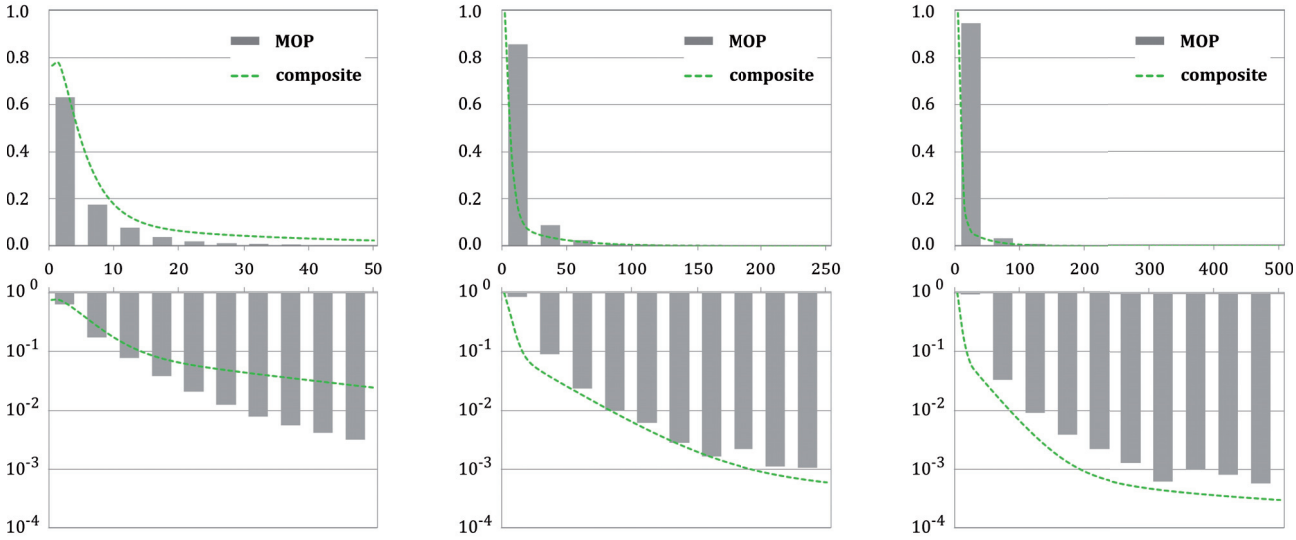

Figure 4.3: Modelling TLD with composite deterrence function

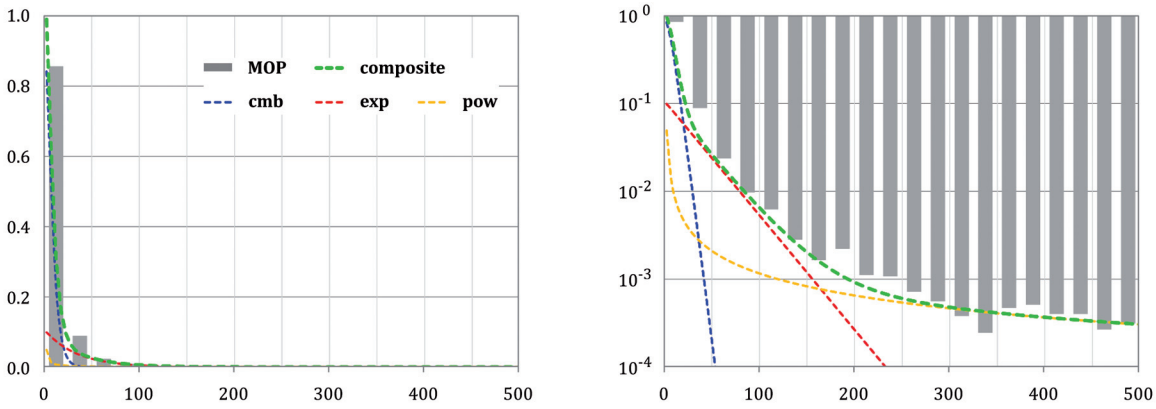

Figure 4.4: Principle of superposition for composite deterrence function

Considering distance bands of 5 km in the left-hand graphs, the composite deterrence function seems to overestimate the TLD for short distances. If we choose distance bands of 25 or 50 km in the other graphs, the function seems to underestimate the TLD for short distances. However, a quick analysis of the shape of the composite deterrence function reveals that it reproduces the pattern of the TLD relatively well for all travel distances. At the same time, it is much easier to calibrate than the EVA function (cf. sec. 4.3.2). For this reason, the outlined issue seemed to be negligible at this point and was not further investigated within this thesis.

Figure 4.4 outlines the principle of superposition for the following composite deterrence function:

$$f(c_{ij}) = 0.85\, {c_{ij}}^{0.3} e^{-0.3\, c_{ij}} + 0.1\, e^{-0.03\, c_{ij}} + 0.05\, {c_{ij}}^{-0.82} \tag{4.28}$$

In the left-hand graph, data is displayed in the ordinary scale and in the right-hand graph in the logarithmic scale. The composite function drafted in green is built on three classical functions where the combined function (cmb) has a share of 85%, the exponential function (exp) a share of 10%, and the power function (pwr) a share of 5%. The TLD is shown in grey bars.

4.4 The revised trip distribution model

This chapter discussed several basic concepts of trip distribution modelling. It started with simple growth-factor methods and concluded that these types of models are of limited use, as they rely on a given trip matrix for a base year. Subsequently, the gravity model was discussed, which has some unfavourable properties if it is applied in its most basic form (4.11). Thus, in order to overcome these drawbacks, additional boundary conditions and calibration factors must be introduced. However, neither of these two measures effectively benefits the methodology of the gravity model.

Finally, the advantages of accessibility-based gravity trip distribution models were described. Two models were presented relying on different accessibility indicators. The first model (4.17) builds on Hansen's gravitational accessibility but is only suitable for modelling private and business trips. The second model (4.18) builds on a competition-enhanced accessibility indicator and can also cope with commuting and vacation trips.

The role of deterrence functions in gravity trip distribution models was then investigated and the properties of different types of functions were compared. It was revealed that applying the complex EVA function (4.20) instead of one of the classical deterrence functions provides major advantages, but its calibration is very challenging. To combine the simplicity of the classical deterrence functions and the explanatory power of the EVA function, the composite deterrence function (4.24) was introduced following the principle of superposition.

In total, this led to the revised trip distribution model:

$$T_{ij} = O_i \frac{D_j f(c_{ij})}{\gamma_j A_i^{(3)}} \tag{4.29}$$

where $A_i^{(3)}$ is the competition-enhanced accessibility indicator, γ_j the level of competition of travellers for opportunities in destination j, f the composite deterrence function, and c_{ij} generalised travel cost.

The revised trip distribution model (4.29) provides an important property in comparison to classical gravity models. It is enriched by the additional explanatory factor γ_j, which models the level of competition for opportunities D_j in destination j. Given that the level of competition depends on the heterogeneous spatial distribution of demand and supply, γ_j is generally different for each region. Hence, consideration of γ_j improves the transferability of the revised trip distribution model from one European region to another. For this reason, the revised accessibility-based gravity model is superior to the classical gravity model. It is therefore used in the IPAT and the HIPAT models.

Methodological revisions of the trip distribution model or, to be more precise, the deterrence function are continued in the following chapter 5. While testing the revised trip distribution model under ideal boundary conditions (identical travel zones and travel costs), the model (4.29) did not produce the expected output concerning distribution of trip lengths in the produced trip matrix. This issue results from a major difference between the concepts of deterrence function and TLD, limiting the spatial transferability of the deterrence function. This is an important issue for European transport modelling.

Chapter 5

Improved deterrence modelling

The deterrence function is a crucial component of the gravity trip distribution model. It models the disincentive to travel between two travel zones in relation to travel costs and is commonly calibrated based on a known trip length distribution (TLD). The TLD reports the frequency of trip occurrences by distance bands and can be derived from observed travel demand data. This concept is not particularly satisfying, as the disincentive to travel depends on more explanatory factors than just travel costs. As these factors generally differ for all O/D relations, the deterrence function (4.24) and, accordingly, the revised trip distribution model (4.29) introduced in section 4.4 do not provide the same accuracy for all O/D relations.

The limited transferability of travel demand models is a well-known problem in transport literature (see, for example, Karasmaa, 2003). It was also a problem in the ETISplus project, in which the developed PAT model had to be improved by additional calibration factors. However, in order to solve this problem of limited transferability additional explanatory factors should be introduced instead of additional calibration factors. One of these additional explanatory factors, for instance, should relate to the "quality" of a destination (e. g. by modelling different wage levels in different regions). This quality-based factor can explain the willingness of commuters to accept higher travel costs when travelling to more "attractive" destinations.[1] Furthermore, complementary factors must be considered in order to counterbalance differences stemming from heterogeneous sizes of travel zones and, accordingly, heterogeneities of O/D relations.

In this chapter, the concept of modelling the disincentive to travel between two regions is studied in more depth. Four explanatory factors are introduced in order to complement the deterrence function in the trip distribution model. Section 5.1 reveals a basic discrepancy between the concept of deterrence function and the TLD, originating from their different means of application. While the definition of the TLD relies on distance bands, the deterrence function is applied at the level of O/D relations. It is shown that this discrepancy can be resolved by introducing a spatial conversion factor following a generic rule. Section 5.2 reveals another shortcoming of the deterrence function, namely its non-applicability

[1] For instance, rail travel times from the city of Freiburg, located in the south-west of Germany, to the cities of Basel, located in Switzerland, and to Offenburg, located in Germany, are about 30 minutes for both destinations. Hence, the values computed by the deterrence function, which models the disincentive to travel in relation to travel costs, are almost identical for both destinations. However, commuters from Freiburg have a higher incentive to travel to Basel due to the higher income level. For this reason, an additional explanatory factor is required in the trip distribution model or, to be more precise, in the deterrence model besides the deterrence function.

to intra-zonal O/D relations. This issue is well-known and concerns the lack of representative travel impedances and travel costs for intra-zonal relations. It can be dealt with by a correction factor which is defined based on the radius of the travel zone.

Section 5.3 introduces two further important explanatory factors in order to improve the transferability of the deterrence function in a European transport model. The first one addresses specific barriers between countries, such as language, cultural and institutional barriers. It is determined based on an empirical study. The second factor addresses economic and other differences between origin and destination and is defined in relation to attraction indicators specific to the trip purpose (e. g. the price level in a region). The finalised deterrence model is presented in section 5.4. It contains four explanatory factors complementing the deterrence function (4.24) that was introduced in section 4.3.3. This model is already implemented in the IPAT model, which is successfully applied as part of the European HIGH-TOOL model (see chapter 6).

5.1 Accessible destinations by travel distance

An important difference between the deterrence function and the trip length distribution becomes apparent if one recalls how the TLD is derived from observed travel demand data. While the TLD is an aggregated distribution that relies on the definition of distance bands, the deterrence function is applied for each O/D relation. As several O/D relations often refer to the same distance band, the deterrence function is applied more than once for the same distance band. Since the number of accessible destinations is constantly growing with increasing travel distance, this particular difference between the TLD and the deterrence function cannot be resolved by applying a constant scaling factor. In this case, a scaling function has to be introduced in order to establish the compatibility of the deterrence function and the TLD. The enhanced deterrence model $M_{ij}^{(1)}$ computing the disincentive to travel from origin i to destination j satisfies the following equation:

$$M_{ij}^{(1)} = \rho_d(x_{ij}) f(c_{ij}) \tag{5.1}$$

where $f(..)$ identifies the deterrence function, c_{ij} generalised cost of travelling, and $\rho_d(..)$ the scaling function that is defined in relation to the travel distance x_{ij} and to the average diameter d of the travel zones.

In the following, the mathematical formula of $\rho_d(x)$ is derived. Figure 5.1 shows the interrelation between the TLD, the travel zones and the deterrence model. The deterrence function $f(c)$ approximating the TLD is outlined in light grey. For the purpose of simplification, we assume homogeneously distributed quadratic travel zones of identical size.[2] Following this approach, we can easily derive the respective conversion factors for each distance band b by counting the number of accessible destinations which are located within this distance band. For intra-zonal trips, the conversion factor for establishing the compatibility between the TLD and the deterrence function is 1 as the deterrence function is applied for only one O/D relation. For trips to neighbouring travel zones, the conversion factor is $1/8$ given that eight destinations are accessible within this specific distance band (cf. Fig. 5.1b).

[2] It is not important that the diameter of the travel zones is chosen in line with the width of the distance bands.

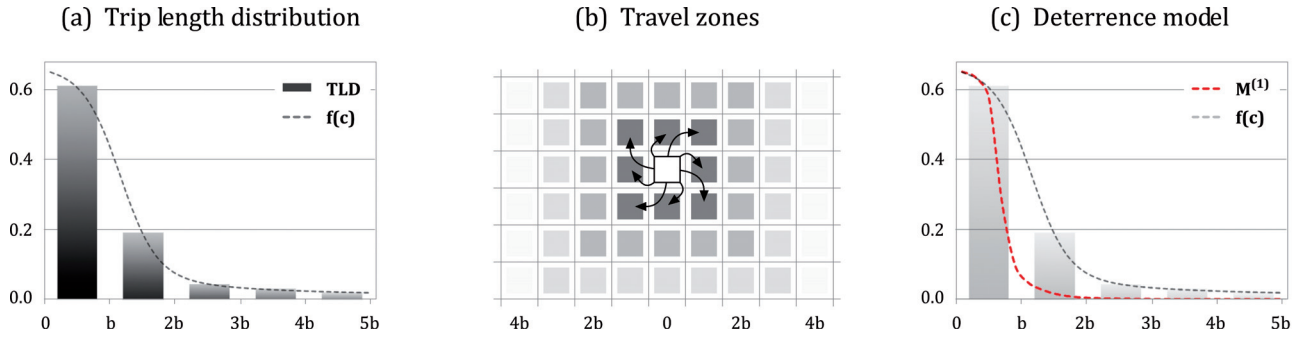

Figure 5.1: Interrelation between trip length distribution, travel zones and deterrence function

The number of accessible destinations $\nu(b)$ which are located within the distance band $b \in \{1, 2, \ldots\}$ follows the general rule (5.3). It depends on the area $A(b)$ (5.2) of the distance band b and the area $A(0)$ of the origin, where a is the edge length of the quadratic travel zones.

$$\begin{aligned} A(0) &= a^2 \\ A(1) &= [(2+1)a]^2 - [(2-1)a]^2 = 8a^2 \\ A(b) &= [(2b+1)a]^2 - [(2b-1)a]^2 = 8ba^2 \qquad (5.2) \\ \nu(b) &= \frac{A(b)}{A(0)} = 8b \qquad (5.3) \end{aligned}$$

With each additional distance band b, the number of accessible destinations increases by a factor of eight. For the definition of distance bands the concept of quadratic zones is not optimal, given that the distance from the centre of the origin zone to destinations located in the same distance band depends on the exact location of a destination.[3] Instead of quadratic zones one can also consider distance bands based on concentric circles.[4] This, however, leads to the same rule (5.5) for the computation of the number of accessible destinations $\nu(b)$ located within the distance band b, where r is the radius of a travel zone.

$$\begin{aligned} A(0) &= \pi r^2 \\ A(1) &= \pi[(2+1)r]^2 - \pi[(2-1)r]^2 = 8\pi r^2 \\ A(b) &= \pi[(2b+1)r]^2 - \pi[(2b-1)a]^2 = 8b\pi r^2 \qquad (5.4) \\ \nu(b) &= \frac{A(b)}{A(0)} = 8b \qquad (5.5) \end{aligned}$$

It should be pointed out that the conversion factor $\rho(b)$ that must be applied for a specific distance band (b) is inversely proportional to the number of accessible destinations $\nu(b)$:

$$\rho(b) \approx 1/\nu(b) \qquad (5.6)$$

[3] Distances from the origin to the eight destinations located in the first distance band are either b or $b\sqrt{2}$ (Fig. 5.1b).

[4] Concentric circles have the same centre but their radii are different. In our case, the origin zone is a circle with the radius r. All destinations in the first distance band are then located around the origin in a concentric circle with the radius $2r + 1$. Destination zones are not circles, but their areas are identical to those of the origin (πr^2).

Consideration of $\rho(b)$ ensures the compatibility between the TLD and the deterrence function (cf. Fig. 5.1c). For example, for the distance band $(b;2b]$ the value of the applied deterrence model $M^{(1)}$ is eight times lower than the deterrence function $f(c)$ due to the fact that $f(c)$ is applied eight times (eight destinations are located in this distance band). It is important to stress that the deterrence function is a crucial determinant of the trip distribution model. It should be calibrated based on a given TLD that provides information on the response of travellers to changing travel cost.[5] For this reason, consideration of $\rho(b)$ in the deterrence model and, accordingly, in the trip distribution model is crucial in order to ensure that the computed O/D trip matrix is in line with the given TLD.

Due to the heterogeneity of travel zones in a European transport model, the definition of accurate distance bands is not possible. For this reason, a more general rule is derived in the following, in order to count the number of accessible destinations $\nu_d(x) \in \{1;8;16;...\}$. This rule relies on the travel distance x between origin and destination as well as on the average diameter d of the travel zones (5.7).

$$\nu_d(x) \quad = \quad \begin{cases} 1 & \text{for } x < \frac{d}{2} \\ 8\,\lfloor x/d + \frac{1}{2} \rfloor & \text{else} \end{cases} \tag{5.7}$$

$$\big(\rho_d(x)\big)^{-1} \quad = \quad \begin{cases} 1 & \text{for } x < \frac{d}{2} \\ 8\,(x/d) & \text{else} \end{cases} \tag{5.8}$$

With regard to the heterogeneous sizes of the NUTS regions, which often dictate the specification of travel zones, it is advisable not to restrict the factor $\big(1/\rho_d\big)$ to being an integer when computing the conversion factor. Hence, the conversion factor, which establishes the compatibility of the deterrence function and TLD, is computed by a slightly modified rule (5.8). Its range is $0 < \rho_d \leqslant 1$.

Consideration of the conversion factor is important for the proper specification of the deterrence function in a European transport model. This property of ρ_d is revealed in the following example, in which the exponential function $\big(\exp(-\beta c)\big)$ is chosen as the deterrence function and travel cost c_{ij} is approximated by the travel distance x_{ij}. In this case, the enhanced deterrence model (5.1) is as follows:

$$M_{ij}^{(1)} = \rho_d(x_{ij})\, e^{-\beta c_{ij}} \tag{5.9}$$

If we assume travel cost c, computed by a linear model according to (5.10), and if we further assume a high correlation between the two explanatory variables of travel distance x and travel time t, we can specify x in relation to travel cost c (5.12). Thus, we can also specify the conversion factor ρ_d (5.8) in relation to travel cost c and we can approximate the deterrence model by (5.13).

$$c_{ij} \approx a_0 t_{ij} + a_1 x_{ij}\,, \qquad a_0, a_1 > 0 \tag{5.10}$$

$$t_{ij} \approx a_2 x_{ij}\,, \qquad a_2 > 0 \tag{5.11}$$

$$c_{ij} \approx a x_{ij} \quad \Leftrightarrow \quad x_{ij} \approx \frac{c_{ij}}{a}\,, \qquad a > 0 \tag{5.12}$$

$$M_{ij}^{(1)} \approx \frac{\alpha}{c_{ij}} e^{-\beta c_{ij}}\,, \qquad \alpha > 0 \tag{5.13}$$

[5] E. g. short-distance trips are cheaper than long-distance trips and therefore more likely.

For long-distance travelling, where $c_{ij} > \alpha$, we can define a strictly decreasing monotonic function $g \in (0, 1)$ as follows:

$$g(c_{ij}) := \alpha / c_{ij} \tag{5.14}$$

The significance of g or, to be more precise, of ρ_d, for the proper specification of the deterrence function now becomes apparent. Application of the exponential deterrence function alone without regarding g, would lead to a systematic overestimation of the deterrence value for long-distance trips. This undesired property has to be counterbalanced by choosing a smaller value for the β coefficient in the exponential function, to avoid an overestimation of long-distance trips. However, the lower the chosen value of β, the faster the exponential function disappears (5.15), which in turn leads to a systematic underestimation of long-distance trips, since the deterrence model (5.13) is dominated by the exponential function.

$$\lim_{c \to \infty} \frac{\alpha}{c_{ij}} e^{-\beta c_{ij}} = \lim_{c \to \infty} e^{-\beta c_{ij}} = 0 \tag{5.15}$$

For this reason, so-called heavy-tailed distributions following the power law should be considered as a deterrence function for the modelling of long-distance trips. However, if we replace the exponential function by a power function in the enhanced deterrence model (5.9), the conversion factor ρ_d, which can be approximated by the function g (5.14), should also be explicitly modelled.

At this point it has to be stressed that the conversion factor is of particular importance for the simultaneous application of the deterrence model at different spatial levels (e. g. NUTS-2 and NUTS-3) since average diameters d of these zone types are different and, in consequence, ρ_d will also vary. In contrast to the concept of a conventional deterrence function, which is implicitly restricted to a single aggregation level, spatial transferability is given for the deterrence model (5.1). Hence, awareness of the conversion factor ρ_d has been a crucial prerequisite for the development of the HIPAT modelling approach discussed in chapter 9, in which the trip distribution model is simultaneously applied at different aggregation levels.

It can be concluded that the conversion factor modelling the number of accessible destinations in relation to the travel distance is required for two reasons. First of all, the conversion factor establishes the compatibility of deterrence function and TLD. Secondly, it enables the consistent application of the deterrence model at different spatial levels. However, it should be noted that the general rule (5.8) for computing the conversion factor relies on the assumption of homogeneous travel zones. This condition is not fulfilled if travel zones are defined on the basis of administrative units (NUTS regions). Accordingly, the error term, which can be related to the application of the conversion factors for each O/D relation, correlates with the size of the two respective travel zones. If the diameter of a destination is twice as large as the average, this destination encompasses two distance bands (i. e. four average-sized travel zones). This is neither in line with the concept of quadratic travel zones, which is outlined by Figure 5.1, nor the concept of concentric circles. In order to reduce the error term, which originates from the application of the conversion factor, it would therefore be necessary to review the set of travel zones and to improve their homogeneity with regard to the size of the travel zones.

5.2 Intra-zonal travelling

Intra-zonal trips have to be dealt with separately by a transport model, since these trips originate and end in the same travel zone and, accordingly, in the same centroid. For this reason, travel impedances cannot be determined and the concept of a deterrence function cannot be applied. A simple approach, which was implemented in the PAT model (cf. sec. 2.3.4), is to estimate the share of intra-zonal trips within the trip generation model, based on the size of a travel zone, and to apply the trip distribution model (i. e. the deterrence model) only for inter-zonal O/D relations. However, as was illustrated in the description of the PAT model, this approach requires additional correction factors in order to adjust the precomputed share factors, which were of poor quality. Moreover, separate modelling of intra-zonal and inter-zonal trips calls into question the very concept of the accessibility-based gravity trip distribution model developed in section 4.2.2. According to this concept, the generated trip demand for a travel zone is distributed among all destinations, based on the accessibility of a certain destination in relation to the accessibility of all destinations, including the origin itself.

For these reasons, the deterrence model should also be applied for intra-zonal trips. One way to overcome the lack of travel impedances for intra-zonal O/D relations could be based upon the common understanding that the average intra-zonal trip length increases with the size of a travel zone. However, the straightforward implementation of this approach leads to an inconsistent model behaviour, as the deterrence function is commonly defined as a strictly monotonically decreasing function with $c_1 > c_2 \Leftrightarrow f(c_1) < f(c_2)$. This behaviour makes sense for inter-zonal O/D relations: the greater the travel costs, the greater the disincentive to travel to this destination and, accordingly, the lower the computed trip demand to this destination. For intra-zonal O/D relations this behaviour makes no sense, since it can be expected that with increasing size of a travel zone more and more trips are carried out intra-zonally.

This undesired model property could be avoided with considerable effort by disaggregating the large travel zones into homogeneous micro cells of equal size. In this case, travel impedances for intra-zonal O/D relations have to be estimated as well, but they are similar for all micro cells. While for regional transport models, which deal with a small number of rather disaggregated and homogeneous travel zones, a trade-off could possibly be made between the homogeneity and the number of micro cells, this is not an option for European transport models like VACLAV and TRANS-TOOLS, due to already existing complexity and runtime problems (cf. chap. 2.2).

Within this thesis, intra-zonal travel impedances are estimated based on the distribution of the destination opportunities in a travel zone and its size. Furthermore, a set of correction factors is introduced in order to rectify the undesired property of the deterrence function for intra-zonal trips that computes lower values for larger travel zones. The previously discussed deterrence model (5.1) is extended as follows:

$$M_{ij}^{(2)} = M_{ij}^{(1)}\delta_{ij} \quad \text{with} \quad \delta_{ij} = \begin{cases} \Omega(r_i) & \text{for } i = j \\ 1 & \text{else} \end{cases} \tag{5.16}$$

$$\Omega(r_i) = \log_{\upsilon}(1 + r_i) + \mathrm{o} \quad \text{with} \quad \upsilon, o > 1 \tag{5.17}$$

where r refers to the radius of a travel zone, and υ, o are modelling coefficients. The correction factors $\Omega > 1$ are computed for each origin by a monotonically increasing logarithmic model (5.17) with $r_1 > r_2 \Leftrightarrow \Omega(r_1) > \Omega(r_2)$. The application of Ω counterbalances the undesired property of the deterrence function for intra-zonal trips.

It should be noted that the correction term (5.17) could be avoided in the deterrence model if the integral of the deterrence function were computed, given that:

$$r_1 > r_2 \Leftrightarrow \int_0^{r_1} f(c(r))\,\mathrm{d}r > \int_0^{r_2} f(c(r))dr \tag{5.18}$$

where $f(c(r))$ is the deterrence function, $c(r)$ are generalised costs, and r is the radius of a travel zone. Understanding the frequency of trip occurrences by distance bands reported by the TLD as integral might serve as an incentive to apply the integral of the deterrence function in the gravity trip distribution model. Within this thesis, however, no subsequent analysis was carried out to develop a new concept for the computation of disincentives to travel between any two travel zones in relation to travel cost based on the integral. Given that the HIPAT transport modelling approach to be developed relies on disaggregated travel zones that are similar in size, correction factors for intra-zonal trips are not particularly relevant in the HIPAT model that will be discussed in chapter 9.

5.3 Economic and border effects

The concept of a deterrence function relies on the observation that the frequency of trip occurrences decreases with increasing trip distance and accordingly with increasing travel cost. However, besides travel costs, further explanatory factors have to be taken into account in the deterrence model. These factors are related to specific incentives and disincentives to travel, like economic differences between origin and destination as well as specific country barriers (e. g. language, cultural and institutional barriers). They are therefore of particular importance with regard to international trips.

The factor related to country barriers, for instance, can be derived by comparing passenger trip volumes in relation to travel cost for national and international O/D relations. While this analysis is hardly possible for trips with private transport modes, due to the non-availability of a comprehensive dataset, it can be carried out for trips with public transport modes, given the availability of data on the number of ticket sales as well as on trip volumes by O/D relation. Another approach to modelling country barriers, i. e. specific travel cost related to border crossings, is to introduce virtual cost rates into the network model which are then considered for the computation of travel cost. These virtual cost rates can be related, for instance, to specific waiting times at the border.

In order to reveal and to quantify the exact value of the additional deterrence factor related to border-crossing trips, a study has been carried out in France. This study analyses time series data on trip volumes for national and international O/D relations and covers rail and air passenger transport (Bonnafous et al., 1989). In a first step, the authors formulated a gravity trip distribution model and estimated the modelling coefficients, based on travel demand data for French O/D relations. In a second step, they applied their estimated gravity model for international O/D relations and compared the modelled trip volumes with observed ones. The authors noticed a significant overestimation of

modelled volumes for international O/D relations, except for two long-distance relations to Italy. While the degree of overestimation varied between 1.5 and 7.0[6] for different country pairs, it stayed almost constant over time for each country pair. The overestimation can be explained by the lack of a specific variable in the model for explaining the additional disincentive related to border-crossing trips. Accordingly, generalised costs of travelling abroad are underestimated and the applied gravity model overestimates the trip volumes for international O/D relations.

It should be noted that the authors considered a simple cost indicator referring to the linear distance between origin and destination. Furthermore, the applied data basis of French O/D relations for estimating the gravity model was dominated by rail passenger trips while the passenger demand for international O/D relations was dominated by air trips. However, the study clearly highlighted the necessity of enriching the gravity trip distribution model, or rather the deterrence model, by additional explanatory variables for modelling specific incentives and disincentives related to border-crossing trips. An obvious disincentive to travel abroad refers to language, cultural and institutional barriers. In contrast, economic differences between origin and destination in terms of living cost and wage levels commonly constitute both an incentive and a disincentive: travellers who benefit from economic differences are encouraged to travel abroad, while travellers who are disadvantaged by working or shopping abroad are discouraged. Economic differences between two countries can therefore lead to asymmetric travel demand patterns on international O/D relations and should therefore be modelled explicitly by an additional factor in the deterrence model. Moreover, economic differences between two regions should also be considered for national O/D relations.

A study of the European Commission estimates that 2.5 million Europeans commute to another Member State (COM, 2002, p. 55). The study mentions two notable examples of commuters from neighbouring countries: Luxembourg and the south-eastern part of the Netherlands. Further examples of relevant demand for border-crossing commuting trips can be found for:

- Copenhagen, the Danish capital, with commuters from Sweden (see, for example, Knudsen and J. (2013); Nauwelaers et al. (2013));
- Germany, with commuters from neighbouring countries (see, for example, Heining and Möller (2009));
- Lothringen, a French region, with commuters to Germany (see, for example, Dechoux (2013)); and
- Switzerland, with commuters from neighbouring countries (see, for example, BFS (2017); Walter et al. (2016)).

Besides daily commuters on international O/D relations, it is also necessary to understand the phenomenon of weekend and long-distance commuters. According to a German study (DGB, 2016), their number has increased significantly in the last ten years. An important mobility driver for long-distance commuting trips is related to different wage levels. Hence, the travel patterns of commuters cannot be explained solely by generalised cost of travelling and the number of workplaces in a region. Incentives for travelling long distances have to be modelled explicitly, e. g. by an indicator describing the "quality" of workplaces.

[6]Source data see Table "EFFETS FRONTIERE PAR PAYS" Bonnafous et al. (1989, p. 77).

Within this thesis, two additional explanatory factors are considered in the deterrence model for modelling specific disincentives of travelling between origin and destination. The first factor B_{IJ} is particularly related to country barriers between the countries I and J. The second factor E_{ij} models specific differences between origin i and destination j, i. e. incentives as well as disincentives to travel. The further improved deterrence model, building on (5.16), satisfies the following equation:

$$M^{(3)}(c_{ij}) \quad = \quad B_{IJ}E_{ij}M^{(2)}(c_{ij}) \quad \text{with} \quad 0.2 \leqslant B_{IJ} \leqslant 1.0\,, \quad (i \in I,\, j \in J) \tag{5.19}$$

Both factors are distinguished by purpose. B_{IJ} refers to a static modelling parameter which is chosen in the range $[0.2, 1]$ according to the findings of Bonnafous et al. for each pair of countries. Setting $B_{IJ} := 1$, which signifies the non-existence of a specific country barrier, could be reasonable if the same language is spoken in the countries I and J, and if no other barriers exist. It can also be reasonable for trips related to the purpose of vacation, given that language, cultural and institutional barriers are frequently unimportant for this specific demand segment. Setting $B_{IJ} := 0.2$ indicates the existence of a large barrier between the two countries. Within this thesis the factor B_{IJ} is kept constant for all O/D relations between two countries. Taking into account cross-border regions where most people speak both national languages fluently, it might be reasonable to choose different values.

The factor E_{ij} is defined according to (5.20):

$$E_{ij} \quad = \quad \left(\frac{R_j}{R_i}\right)^p\,, \quad (p \geqslant 0) \tag{5.20}$$

where R is a rating indicator related to the travel zone i. The power coefficient p determines the variance of the computed factors E_{ij} and, accordingly, the sensitivity of the deterrence model (5.19) to specific differences between origin and destination. The purpose-specific rating indicators R can be defined according to Table 5.1. They rely on quality-based indicators which describe the quality

Table 5.1: Rating indicators by purpose

Purpose	Indicator R_i
business	GDP by workplace
commuting	Wage level in the region
private	Price level in the region
vacation	Number of sunny days per year

of the destination opportunities in a region. It should be mentioned that the indicators can also be defined in different ways. Furthermore, the purpose-specific values of the exponent p (5.20) can be set to zero if quality differences between two regions should not be modelled for a specific trip purpose.

It should be stressed that the additional factors (B, E) in the deterrence model (5.20) only influence the computation of the deterrence value. Neither factor influences the number of trip endings attracted by a travel zone that are computed in the first step, namely the trip generation model. This is a crucial prerequisite in order to determine the competition level for each destination and to apply the accessibility-based gravity trip distribution model which was introduced in section 4.2.2. The more incentives exist for travelling to a destination, like high wage levels, the greater people's willingness to

pay higher travel costs, and the higher the derived deterrence value. By contrast, the more disincentives exist, like language and cultural barriers, the lower the computed deterrence value.

Besides the two discussed factors, B and E, a multitude of other additional factors can be considered for improving the deterrence model and the trip distribution model. For instance, the HIGH-TOOL passenger demand module (Van Grol et al., 2016) introduces an additional parameter for modelling capital effects, i. e. a particularly high transport demand to capitals, due to their importance as economic and tourist centres of a country.

5.4 The finalised trip distribution model

The deterrence function, which models the disincentive of travellers in relation to generalised costs, is a crucial component of a gravity-based trip distribution model. Its trajectory is commonly determined on the basis of a known TLD, but one has to be aware of the basic discrepancy between the deterrence function and the TLD. Their compatibility can be established by considering an additional conversion factor in the deterrence model and, accordingly, in the distribution model. In addition, a correction factor is required which enables the application of the deterrence function for intra-zonal O/D relations. However, besides generalised costs, other factors have to be taken into account for explaining travellers' destination choice. These factors refer to country barriers, economic and other differences between origin and destination, and also have to be considered by the trip distribution model. For these reasons, a number of improvements were discussed within the last sections and the classical concept of applying a sole deterrence function has been revised several times, leading to the development of a more sophisticated deterrence model. The deterrence model, which is further considered within this thesis for trip distribution modelling, is as follows:

$$M_d(c_{ij}) \quad := \quad B_{IJ} E_{ij} \rho_d(c_{ij}) f(c_{ij}) \delta_{ij} \,, \quad (i \in I,\, j \in J) \tag{5.21}$$

where f is the composite deterrence function (4.24), which can be understood as an approximation of the EVA function by less complex functions, and c generalised travel cost. The two factors B_{IJ} (5.19) and E_{ij} (5.20) model country barriers, economic and other differences between origin and destination. The correction factor δ_{ij} (5.16) is only relevant for intra-zonal O/D relations and the scaling function ρ_d (5.8) establishes the comparability between the deterrence function and the TLD.

The finalised trip distribution model relies on the competition-enhanced accessibility indicator $A_i^{(3)}$ discussed in section 3.3.2 and satisfies (5.22):

$$T_{ij} = O_i \frac{D_j M_d(c_{ij})}{\gamma_j A_i^{(3)}} \qquad \text{with} \qquad A_i^{(3)} = \sum_{j=1}^{n} \frac{1}{\gamma_j} D_j M_d(c_{ij}) \tag{5.22}$$

$$\gamma_j = \sum_{k=1}^{m} \frac{1}{\mu_k} N_k^{(\mathrm{o})} M_d(c_{kj}) \qquad \text{with} \qquad \mu_k = \sum_{l=1}^{n} N_l^{(\mathrm{d})} M_d(c_{kl}) \tag{5.23}$$

It has to be emphasised that the definition of the scaling function ρ_d follows a generic rule relying on the average diameter d of the travel zones. This ensures the transferability of the deterrence model (5.21) and, accordingly, the trip distribution model between different zoning systems. Moreover, the deterrence model can be applied simultaneously at different aggregation levels, which is the major prerequisite for the development of the hierarchical modelling approach in chapter 9.

With regard to the trip distribution model, in principle it would be desirable to apply more homogeneous travel zones. This would reduce the necessity to consider the correction factor δ_{ij} for intra-zonal travelling in the deterrence model and establish the full compatibility between the deterrence function and the TLD.

Chapter 6

The IPAT modelling concept

The methodological revisions carried out in the last chapters led to the development of the IPAT modelling concept. The revisions particularly address fundamental shortcomings in the methodologies of the trip distribution and the modal split models that were revealed in the PAT model (cf. sec. 2.4). The IPAT model was realised at NUTS-2 level. The IPAT model is published and successfully applied as part of the larger HIGH-TOOL policy assessment model that was developed for the European Commission. In the HIGH-TOOL context the IPAT model interacts with several other models, providing forecasts on demographic and economic indicators, for instance. In this context it was also intensively tested and validated. The methodology of the IPAT model and its validation are described in Ihrig (2014) and in the deliverables D4.3, D8.1, D8.2 of the HIGH-TOOL project (see Van Grol et al. (2016); van Meijeren et al. (2016); Kiel et al. (2016a)).

This chapter starts with an introduction of the HIGH-TOOL model, paying attention to the specific role of the IPAT model. Section 6.2 then focuses on the IPAT model, briefly summarising the main methodological revisions, the overall methodology and the calibration process. Section 6.3 describes a series of checks that were carried out in order to test and validate the IPAT model. The last section summarises the main findings of this chapter with regard to the development of the HIPAT model discussed in chapter 9, which is built on the IPAT model.

6.1 The HIGH-TOOL policy assessment model

HIGH-TOOL was a research project carried out between 2013 and 2016 and funded by the European Commission (DG Mobility & Transport) under the 7th Framework Programme. It was coordinated by the KIT (Dr. Szimba, Network Economics), which was in charge of the development of the passenger demand module and involved in many other tasks. The HIGH-TOOL project was recently awarded the German Mobility Prize 2017 (KIT, 2017). Its key objectives were the development of a user-friendly, open source, high-level strategic transport model with a runtime of a few hours for use by the European Commission for transport policy assessment, which is called the HIGH-TOOL policy assessment model (cf., for example, Szimba et al., 2018).

The HIGH-TOOL model was developed based on existing tools and models, and on widely approved data sources, as well as in close cooperation with the targeted model users, to ensure an optimal coverage of user requirements. It was developed in three steps based on feedback by the potential model users, including a prototype, a pre-final and the final version of the HIGH-TOOL model. In order to receive more feedback, the methodology of the HIGH-TOOL model was already discussed during the development period at several international transport conferences, e. g. the European Transport Conference 2015 (Szimba et al., 2015), the Transport Research Arena 2016 (Szimba and Ihrig, 2016) and the World Conference on Transport Research 2016 (Szimba et al., 2017).

The following sections briefly discuss the final version of the HIGH-TOOL policy assessment model.[1] Section 6.1.1 gives an overview of the objectives and the key model features, section 6.1.2 describes the functional modules that are part of the HIGH-TOOL model, section 6.1.3 discusses the interdependencies between the modules and section 6.1.4 finally discusses the model application and the results. The specific role of the IPAT model in the HIGH-TOOL model is described in the last section.

6.1.1 Overview and key model features

The HIGH-TOOL model provides forecasts on European passenger and freight transport in 5-year steps from 2010 to 2050 and assesses the economic, social and environmental impacts of transport policies. It covers the EU28, Norway and Switzerland at NUTS-2 level and neighbouring countries at NUTS-0 level. Countries outside Europe are grouped into so-called country bundle regions. The model is calibrated to the EU Reference Scenario 2013 (EC, 2013), providing forecasts at country level in 5-year steps from 2010 to 2050. These forecasts include indicators such as passenger and freight transport demand (pkm, tkm), energy demand, composition of vehicle fleet, GDP and population. The forecasts are in line with EU statistics for the year 2010. The HIGH-TOOL model reproduces the EU Reference Scenario 2013 if no specific transport policy is investigated. The results of each model run are summarised in a policy assessment report, providing an overview of the investigated transport policy as well as a predefined set of relevant output indicators for transport policy assessment. For instance, the report shows tables and graphs summarising the expected savings of CO_2 emissions and other pollutants in relation to the reference scenario. The key features of the HIGH-TOOL model are summarised in Table 6.1.

[1] The information provided in this chapter mainly relies on publicly available project deliverables including: Final Report (Szimba et al., 2016); Elasticities and equations (Van Grol et al., 2016); Updated input database for the HIGH-TOOL model (Kiel et al., 2016b); Final version of the HIGH-TOOL model (Larrea, 2016); User guide (Larrea and Szimba, 2016); Validating the HIGH-TOOL model (Kiel et al., 2016a); Final structure of the HIGH-TOOL model (Mandel et al., 2016).

Table 6.1: Key model features of the HIGH-TOOL policy assessment model[1]

Feature	Description
Model type	Strategic high-level model derived from existing tools, models, equations and elasticities; where necessary enriched by new models
Scope & zoning	EU28, Norway and Switzerland: NUTS-2 level; neighbouring countries: NUTS-0 level; other countries worldwide: country bundles
Timeline	5-year steps from 2010 to 2050
Modes	Passenger: air, rail, road (passenger car and powered 2-wheelers), long-distance coach, urban public transport, slow modes; further differentiation by vehicle technologies Freight: air, rail, road, inland waterways, maritime transport; further differentiation by vehicle technologies
Transport types	Passenger by trip purpose (business, private, vacation, commuter). Freight transport by NST2 aggregation
Distance bands	0–300 km, 300–1000 km, 1000+ km (based on distances between NUTS-2 centroids)
Model sensitivity	The dependent variables of a module are sensitive to a variety of independent variables to cope with transport policy measures
Validation	EU Reference Scenario 2013, EU transport in figures/ Statistical Pocketbook, ETISplus
Baseline	EU Reference Scenario 2013
Model complexity	Runtime: <3 hours; database size: 15 gigabytes

[1] Source: Transport Research Arena (TRA) 2016, poster session (Szimba and Ihrig, 2016)

6.1.2 Brief description of the functional modules

Figure 6.1 outlines the general structure of the HIGH-TOOL model including the data exchange between its functional modules. It consists of three main elements: the core modules, the "Data Stock" and the user modules. The core modules represent the modelling framework of the HIGH-TOOL model and comprise the following seven functional modules:

- **Economy & Resources (ECR)** simulates the impacts of transport policies on the economy by capturing changes in passenger and freight transport, the type and the number of vehicles purchased as well as changes in population and labour force. It produces estimates for economic indicators like GDP, gross value added, inter-regional trade-flows as well as environmental and material use indicators. The general drivers for computing the economic baseline trajectory 2010-2050 are defined exogenously by the EU Reference Scenario 2013.

- **Demography (DEM)** produces estimates on the regional population and the labour force for each travel zone. For the EU28, Norway, and Switzerland the assumptions on fertility rates, life expectancy at birth and net migration, underlying the EU Reference Scenario 2013, are reflected. Estimates for other countries worldwide rely on UN projections.

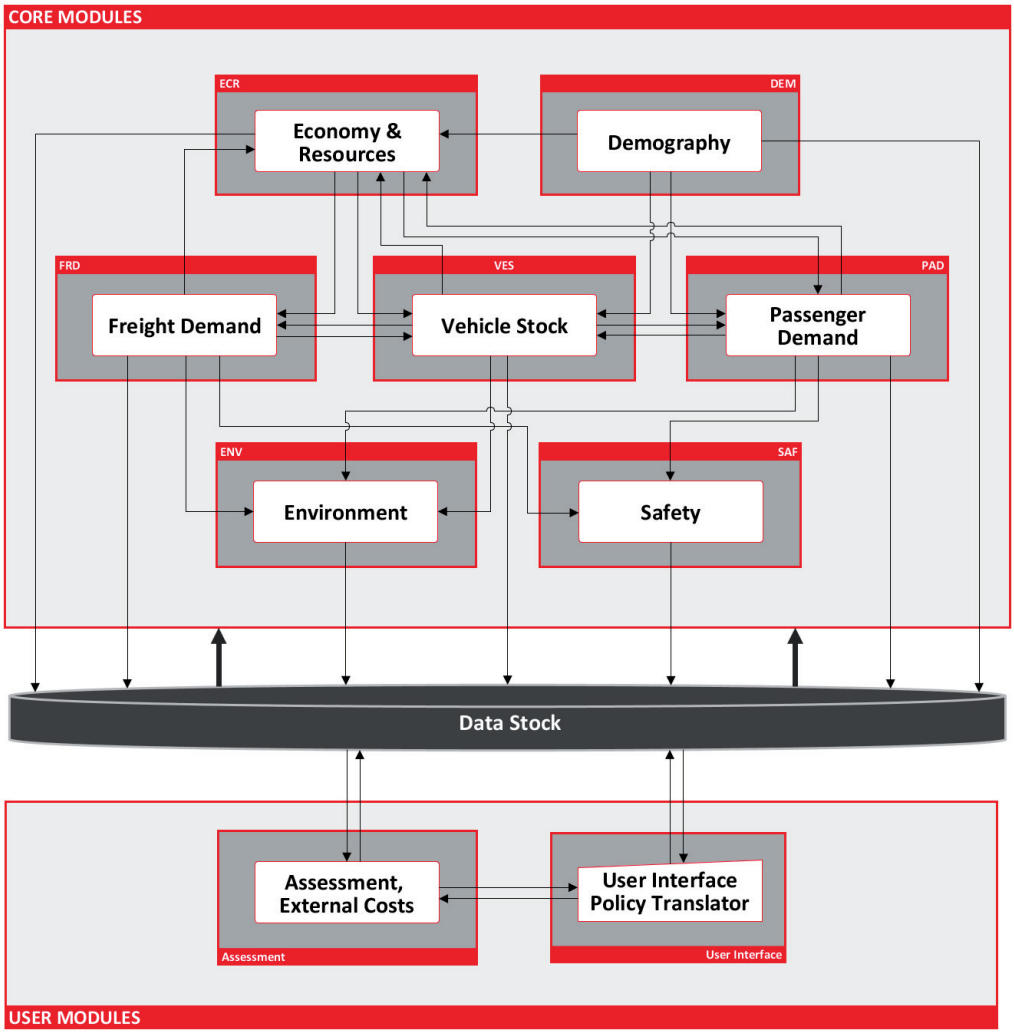

Figure 6.1: Structure of the HIGH-TOOL model (source: Mandel et al., 2016)

- **Freight Demand (FRD)** deals with the mobility of goods under consideration of estimates on inter-regional trade-flows provided by ECR, vehicle fleet data provided by VES, available infrastructure and transport modes as well as transport-related policies. Its main output comprises transport performance indicators in terms of tkm and vkm that are of particular importance for the assessment of transport policies and the calculation of external costs. The methodology of FRD relies on existing models including TRANS-TOOLS v2 and ETISplus. For the projection of mobility to 2050, FRD follows the general trends on trade volumes provided by ECR.

- **Vehicle Stock (VES)** computes changes in the fleet composition over time by a cohort-based modelling approach. The vehicles are distinguished by vehicle type, fuel type and propulsion technology. For each time step, the number of new vehicles is computed based on information on the fleet composition of the previous time step, passenger and freight transport demand, population and applied transport policies. In addition, VES provides forecasts on average fixed and variable costs by vehicle type and on total tax revenues at country level. The vehicle type segmentation relies on the TRACCS project database, the TREMOVE and the MOVEET model.[2]

[2]See: Papadimitriou et al. (2013); De Ceuster et al. (2007); Purwanto et al. (2017).

- **Passenger Demand (PAD)**[3] deals with the mobility of the population under consideration of mobility drivers like GDP and household income per capita provided by ECR, mode-specific cost indicators provided by VES, available infrastructure and transport modes as well as transport-related policies. Its main output comprises the O/D trip matrices by transport mode at NUTS-2 level as well as derived travel performance indicators in terms of pkm and vkm. These derived indicators are of particular importance for the assessment of transport policies and the calculation of external costs. For the projection of mobility to the year 2050, passenger demand follows the general trends that are identified by the EU Reference Scenario 2013.

- **Environment (ENV)** computes the emission of air pollutants like CO_2, $PM_{2.5}$ and NO_x that are caused by passenger and freight transport, based on information on the vehicle fleet composition and the given emission factors by vehicle type. The computed results are of particular importance in the assessment of the environmental impacts of transport policies.

- **Safety (SAF)** assesses the effects of transport policies on the number of fatalities, injuries and associated social costs caused by passenger and freight transport.

The other functional modules of the HIGH-TOOL model are:

- **Data Stock** manages the data exchange between the HIGH-TOOL modules. The stored data comprise input and output indicators, modelling parameters and calibration coefficients of the seven core modules. In addition, the Data Stock stores parameters related to the transport policy measures that can be changed by the user through the User Interface.

- **Assessment, External Costs** assesses the transport policy defined by the user with regard to its economic, social and environmental impacts. The module generates the policy assessment report that includes tables and figures for a predefined set of assessment variables for both scenarios, the reference scenario and the user-defined policy scenario.

- **User Interface, Policy Translator** organises the interaction between the user and the HIGH-TOOL model. This includes the definition of the transport policy to be investigated, triggering the model run as well as access to the produced policy assessment report. In addition, the User Interface also provides access to the tables that are stored in the Data Stock. The Policy Translator translates the user-defined transport policy into specific policy levers triggering the core modules of the modelling framework and the HIGH-TOOL model run.

6.1.3 Application order of the core modules

The seven modules of the core modelling framework are applied in consecutive order, precisely once for each time step t.[4] For this reason, some modules use the results of other modules referring to the previous time step t-1. For instance, ECR is executed before PAD and uses estimates on passenger transport referring to t-1. For the base year 2010, the output indicators of each module are predefined by the Data Stock and are not changed during a model run.

[3] The PAD module is an implementation of the IPAT modelling concept described in section 6.2.
[4] A modelling time step is five years.

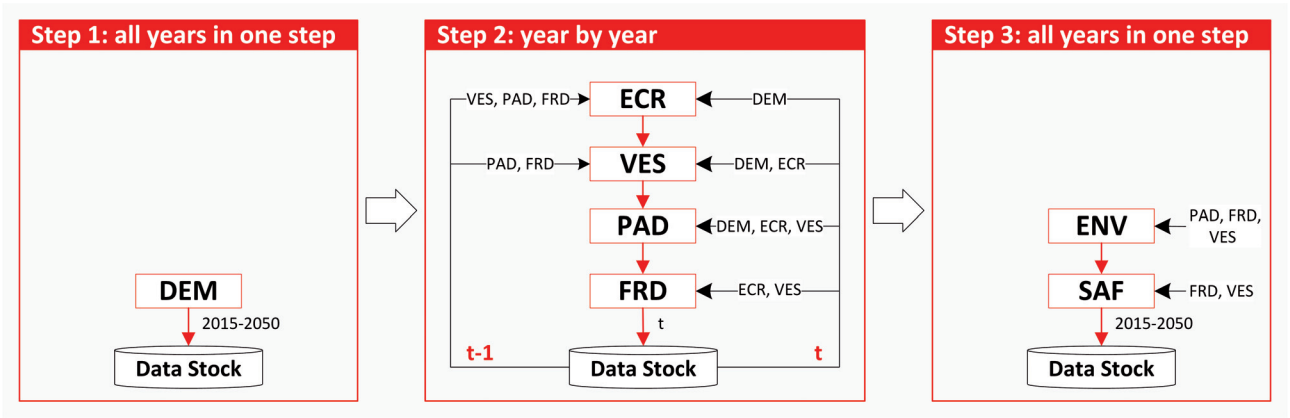

Figure 6.2: Application order of the seven modules

Figure 6.2 shows the application order of the modules. In the first step, DEM is executed, producing forecasts on demographic indicators for each travel zone for the whole time period from 2015 to 2050. The four modules ECR, VES, PAD and FRD are then executed in consecutive order for each time step from 2015 to 2050. In the last step, ENV and SAF are executed for the whole time period from 2015 to 2050 in order to assess environmental and safety impacts.

6.1.4 Analysing transport policies

A policy analysis with the HIGH-TOOL model is carried out in the following steps:

1. defining the policy scenario;
2. running the simulation; and
3. accessing and evaluating the outputs.

In the first step, the user defines the policy scenario via the Policy Translator. It can be chosen among a set of predefined single transport policy measures (TPM) or a combined TPM package, or a customised policy package can be defined. Each TPM is modelled by specific policy lever variables that are provided by the Policy Translator. For instance, one lever variable is related to the TPM "CO_2 certificate system for road transport" (Fig. 6.3a), which can be changed by the user within a reasonable range. In addition, the intensity can be varied - spatially and over time (Fig. 6.3b).

For the customised policy package, the user has full access to each variable that is provided by the User Interface. Besides the definition of the policy scenario through changes in the provided policy lever variables, an "expert mode" is available in which the user can access and change any value that is stored in the Data Stock. This expert mode is also required for implementing network-based transport policies, such as the simulation of corridor improvements for an important high-speed rail (HSR) corridor in central Europe (cf. sec. 6.3.4). All changes are stored in the Data Stock and later processed during the HIGH-TOOL model run.

In the second step, the user starts the simulation and the HIGH-TOOL model computes output indicators for the seven core modules including passenger and freight transport, economic, social and environmental indicators. In the last step, the user can download the policy assessment report (Fig. 6.3c),

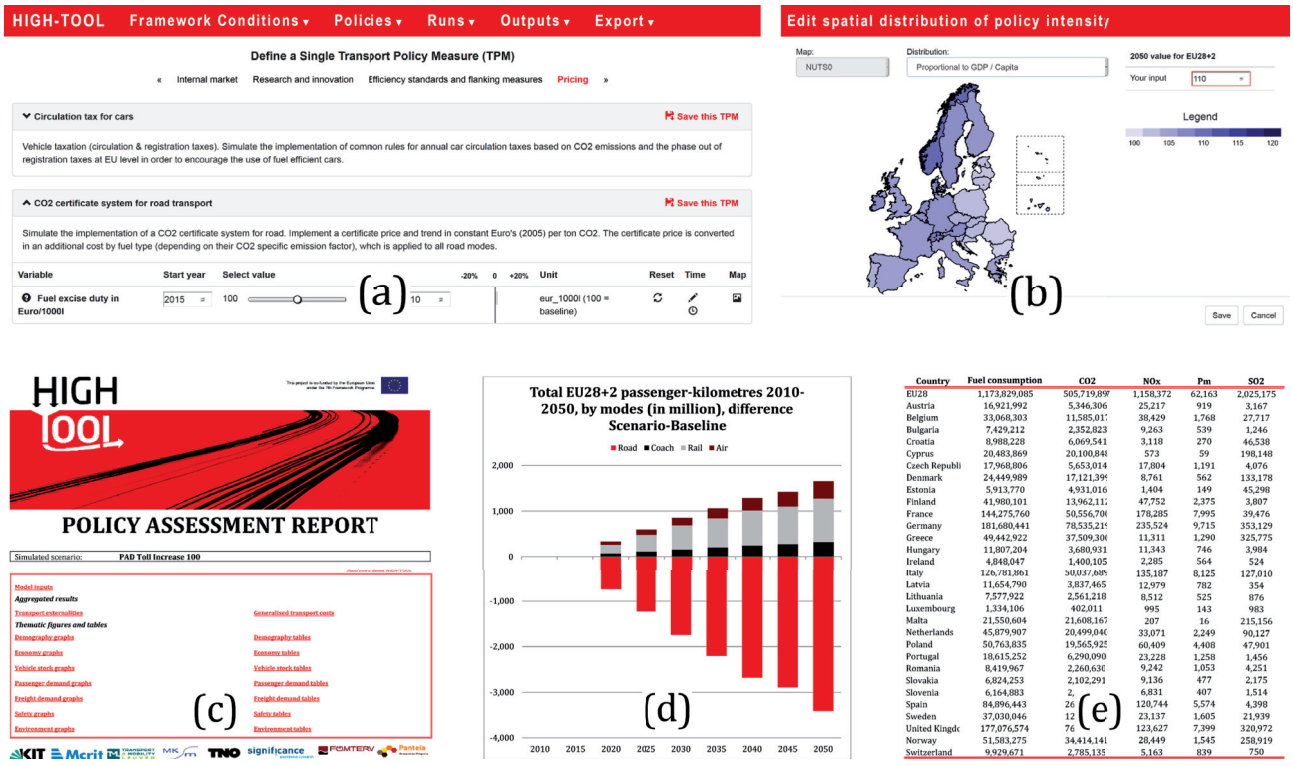

Figure 6.3: Screenshots from the HIGH-TOOL User Interface

which is generated by the system. It summarises the settings of the investigated policy scenario and provides precompiled graphs and tables for specific output indicators (Fig. 6.3d-e). Based on this information, the user can assess the economic, social and environmental impacts of the applied transport policy. Through the User Interface, he can also obtain full access to each produced output indicator at a more detailed level.

6.1.5 Role of the IPAT model

The PAD module, which computes the effects of transport policies on passenger transport demand, is one of the two central modules of the HIGH-TOOL model. It is a special and slightly modified implementation of the IPAT model and it is referred to as the "PAD model" later in the text. The models, IPAT and PAD, are calibrated to the base year 2010 and are in line with EU transport statistics provided by the EU Reference Scenario 2013.

The EU Reference Scenario 2013 focuses on the energy, transport and climate dimensions of EU developments and the various interactions among policies, including also specific sections on emission trends not related to energy (EC, 2013, p. 10). It provides forecasts for the EU and for each of its 28 member states at country level from 2010 to 2050 in 5-year steps. Important output indicators which are particularly related to passenger demand modelling include forecasts on the transport demand in terms of pkm by mode of transport, and on travel costs in terms of Euro per vkm by mode of transport, as well as projections of GDP and population.

In order to simulate the effects of EU policies on passenger transport for the timeline of 2015 to 2050, the methodology of the PAD model was slightly extended. In comparison to the IPAT model, the PAD model introduces additional calibration factors that are related to the forecast years from 2015 to 2050 as well as additional "policy lever variables" that are required for the simulation of transport policies. Calibration factors related to the forecast years were derived based on the EU Reference Scenario 2013, while the basic calibration parameters related to the base year 2010 were kept constant.

For the base year 2010, the results of both models are almost identical. For 2015 to 2050, the results are slightly different. Given that both models rely on the same cost functions, they share equal sensitivity characteristics concerning changes of input parameters in the cost functions. For instance, if travel times are decreased by 10% for rail passenger transport or if fuel costs are increased by 25% for private cars, both models respond in a similar way. These similarities enabled a comprehensive validation of the IPAT model within the framework of the HIGH-TOOL project that was carried out by a broader audience, including transport modelling specialists and transport policy makers.

6.2 Description of the IPAT model

This section describes the implementation of specific methodological improvements of the IPAT model that were discussed in the chapters 3, 4 and 5. First, a general overview of the key features is provided, followed by a discussion of the methodology and the calibration process.

6.2.1 Overview and key model features

The IPAT model provides estimates on European passenger transport at the level of NUTS-2 regions inside the EU28, Norway and Switzerland. Neighbouring countries are covered at NUTS-0 level in order to extend the geographical scope. Key output indicators include the G/A passenger transport demand matrices by mode and purpose in terms of trips, passenger-kilometres and vehicle-kilometres. The model applies ETISplus data and is calibrated to the base year 2010. In contrast to the PAT model developed within ETISplus, the IPAT model covers additional transport modes and intra-zonal trip demand.

The IPAT model is implemented in the programming language `Java` and uses so-called "worker threads" (i. e. multi-threading) to benefit from the high performance of modern computers and the multi-core architecture of modern CPUs.[5] This reduces its runtime to less than 30 seconds including data exchange with the file system. In addition to the implemented `Java` source code, a modelling framework was realised in `Microsoft Excel` for calibrating and for testing the IPAT model. Table 6.2 gives an overview of the key features of the IPAT model.

Table 6.2: Key features of the IPAT passenger demand transport model

Feature	Description
Methodology	Tailored four-step approach
Scope & zoning	EU28, Norway and Switzerland: NUTS-2 level; neighbouring countries: NUTS-0 level
Modes	Rail, road, air, rail, long-distance coach
Purposes[1]	Business, commuting, private, vacation
Demand segments	Male and female; age groups: 0-14, 15-24, 25-64, 65+
Reference data	ETISplus 2010
Software	Model implementation: `Java`; calibration procedure: `MS Excel`
Model runtime	30 seconds including data exchange with the file system

[1] ETISplus definition of trip purposes is applied.

6.2.2 Methodology

The methodology of the IPAT model mainly follows the classical four-step approach of passenger transport demand modelling. It consists of the four sub-models of trip generation, trip distribution, modal split and conversion[6] (cf. Fig. 6.4). In the IPAT model, the trip generation model is implemented

[5] For more information on parallel computing see, for example, Ullenboom (2014); Goetz (2008); Bal and Haines (1998).
[6] The conversion model replaces the assignment model and computes passenger-kilometres and vehicle-kilometres.

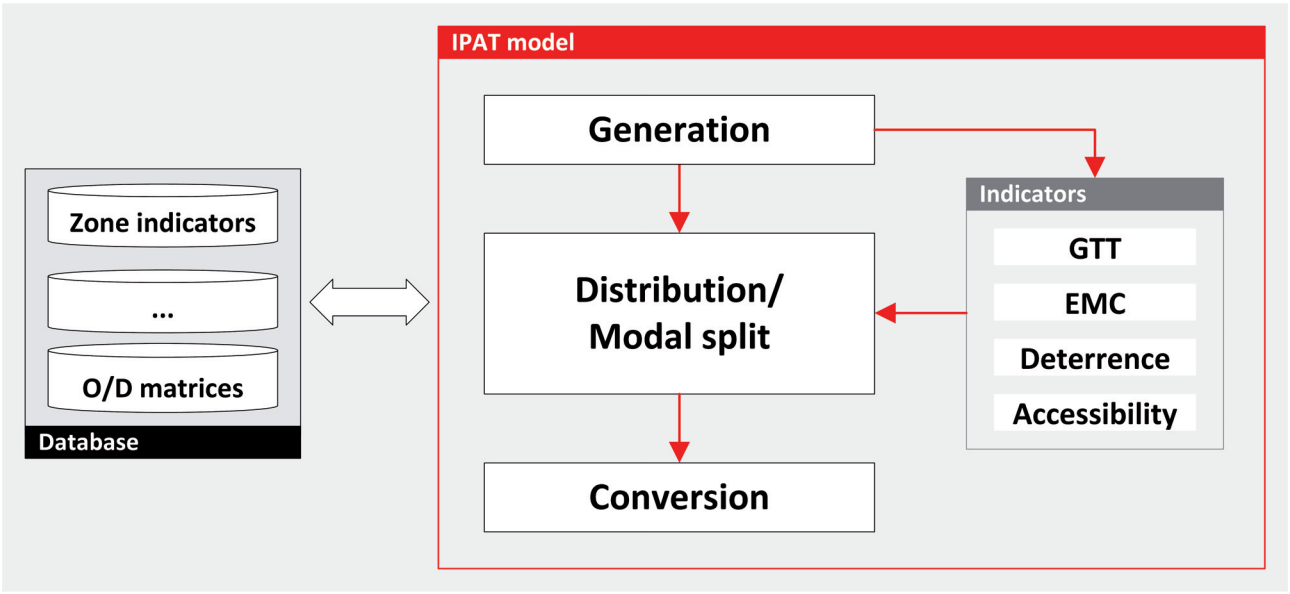

Figure 6.4: Structure of the integrated passenger transport model (IPAT)

straightforwardly and the two sub-models dealing with destination and mode choice are integrated. Integration of these two sub-models contributes to the consistency of the overall IPAT model and was implemented by using the EMC (expected minimum cost) indicator in both sub-models.[7]

Figure 6.4 outlines the interaction in the IPAT model, including the four-step approach, four important indicators and data flows. The main focus of developing the IPAT model was on the methodological improvement of the distribution and the modal split models. In order to reduce the overall model complexity, the network assignment model was replaced by a more lightweight conversion model. During this conversion step, trip volumes are translated into passenger-kilometres and vehicle-kilometres, based on information on the travelled distance and on vehicle occupancy rates.

6.2.2.1 Trip generation model

The trip generation model computes the number of trips $T_{i,p}$ that are generated in each origin i by a trip production function (6.1) for each trip purpose p (see Tab. 6.2). The function applies the population (pop) by age group and gender (ag) and country-specific (ci) trip rate factors (R). Trip rate factors are distinguished by purpose and were determined at a very early stage during this thesis within the ETISplus project (cf. sec. 2.3).

$$T_{i,p} = \sum_{ag} R_{ci,ag,p} \cdot \mathrm{pop}_{i,ag} \tag{6.1}$$

6.2.2.2 Integrating distribution and modal split model

In the IPAT model, destination and mode choice are integrated. This improves the consistency of the overall transport model. The overall methodology of the integrated trip distribution and modal split model encompasses several modelling steps. While the methodology of the two steps referring

[7] The EMC indicator is an output of the logit model, which computes market shares of transport modes (cf. sec. 3.4.4).

to trip distribution and modal split are straightforward, the methodology of the steps referring to the four subsequently derived indicators, GTT (generalised travel times), EMC, deterrence value and EMC accessibility, is more complex.[8] These four indicators and their equations have already been investigated in chapters 3, 4 and 5. In the following, only extensions and changes are discussed in greater detail.

During the **integrated trip distribution and modal split** step, trip matrices $T_{ij,p,m}$ are computed for each transport mode m and trip purpose p as follows:

$$T_{ij,p,m} = T_{ij,p} \cdot \mathrm{PM}_{ij,p,m} \tag{6.2}$$

$$T_{ij,p} = T_{i,p} \cdot \mathrm{PA}_{ij,p} \tag{6.3}$$

where PM identifies the market share of the respective transport mode, and PA the share of trips that are generated in origin i and end in destination j. The share factor PM is derived during the computation of the EMC indicator, and PA during the computation of the EMC accessibility indicator. This will now be explained.

The **EMC accessibility** indicator $A^{(3)}_{i,p}$ (6.6) is computed by purpose p for each origin and quantifies the accessibility of destination opportunities. In the IPAT model, the competition-enhanced accessibility indicator is applied. It was introduced in section 3.3.2 and is the basis of the trip distribution model that was developed and continuously improved in chapters 4 and 5 of this thesis. The final trip distribution model is documented by (5.22) in section 5.4, where the variable O_i now has to be replaced by $T_{i,p}$ from (6.3). The share factor $\mathrm{PA}_{ij,p}$ applied in the trip distribution model (6.3) is computed as follows:

$$\mathrm{PA}_{ij,p} = \tilde{A}_{ij,p} \left(A^{(3)}_{i,p} \right)^{-1} \tag{6.4}$$

$$\tilde{A}_{ij,p} = (\gamma_{j,p})^{-1} \left(D_{j,p} M(\tilde{c}_{ij,p}) \right) \omega^{(\mathrm{opp})}_{ij,p} \tag{6.5}$$

$$A^{(3)}_{i,p} = \sum_j \tilde{A}_{ij,p} \tag{6.6}$$

where $\gamma_{j,p}$ quantifies the expected level of competition of travellers for opportunities located in destination j. Estimated trip endings $D_{j,p}$ are computed based on attractiveness indicators summarised in Table 6.3 where $\sum_i T_{i,p} = \sum_j D_{j,p}$. M identifies the deterrence model, and $\tilde{c}$ the applied cost indicator that relies on the EMC indicator.

Table 6.3: Attractiveness indicator by purpose

Purpose	Indicator by region
Business	Total GDP produced
Commuting	Number of workplaces
Private	Population
Vacation	Number of beds in accommodation facilities

[8] GTT are given in a time unit and relate to the generalised cost of travelling (cf. sec. 3.2).

The parameter $\omega_{ij,p}^{(\text{opp})}$ in (6.5) is a calibration coefficient that captures adverse effects stemming from the heterogeneous size of the origin travel zones for measuring the accessibility of destination opportunities of inter-zonal O/D relations ($i \neq j$). It is defined according to (6.7), where the values of $a_{i,p}$ are distributed around 1 and determined during the model calibration.

$$\omega_{ij,p}^{(\text{opp})} = \begin{cases} 1, & \text{if } i = j \\ a_{i,p} \in \mathbb{R}^+, & \text{else} \end{cases} \tag{6.7}$$

The **deterrence model** $M_{d,p}$ (6.8) computes the disincentive to travel between two travel zones in relation to travel costs $\tilde{c}_{ij,p}$ and other explanatory variables. The methodology of the deterrence model was elaborated and intensively revised in chapter 5 and is documented by (5.21). Besides the deterrence function f, the deterrence model considers the parameters $B_{ci,cj,p}$ and $E_{ij,p}$ that are related to possible barriers between the two countries ci and cj as well as to economic and other differences between origin i and destination j. The consistent application of $M_{d,p}$ for various trip distances and in particular for intra-zonal trips is ensured by the two factors ρ_d and δ. In the IPAT model, the deterrence model is applied for each trip purpose p and satisfies:

$$M_{d,p}(\tilde{c}_{ij,p}) \quad := \quad B_{ci,cj,p} E_{ij,p} \rho_d(\tilde{c}_{ij,p}) f(\tilde{c}_{ij,p}) \delta_{ij,p}\,, \quad (i \in ci,\, j \in cj) \tag{6.8}$$

At this point, it is necessary to draw attention to an important and required property of the cost indicator $\tilde{c}_{ij,p}$ that is applied in (6.8). In principle, $\tilde{c}_{ij,p}$ relies on the EMC indicator $c_{ij,p}$ that is given in the "virtual utility scale", but as the deterrence model is defined in the "ordinary distance scale" referring to travel distances, EMC values have to be translated. This is carried out by a linear scale transformation that does not change existing characteristics between different EMC values. The scale transformation is defined by (6.9) in which $r, s > 0$ are two regression coefficients.

$$\tilde{c}_{ij,p} = r \cdot c_{ij,p} + s \tag{6.9}$$

Amongst other things, this translation is required to determine the deterrence function based on a known TLD given in distance bands (cf. sec. 4.3), and to derive the parameter $\rho_d(\tilde{c}_{ij,p})$, which is defined in relation to the average diameter d of the applied travel zones and to the trip distance (cf. sec. 5.1).

The **EMC indicator** is an integrated cost indicator that refers to the expected minimum cost of travelling between two travel zones taking into account all available transport modes. It is computed for each purpose, following the hierarchical structure of the NMNL model outlined in Figure 3.2 (in section 3.4.5). Equation (6.10) outlines the mathematical model for computing the expected value $E(N_{ij,p})$ of a "nest" given the expected values of the lower-level "nests" $E(\tilde{N}_{ij,p})$:

$$E(N_{ij,p}) \quad = \quad \frac{1}{-\mu_{N,p}(x)} \cdot \ln \sum_m e^{-\mu_{N,p}(x) E(\tilde{N}_{ij,p,m})} + \text{cnst}_{N,p} \tag{6.10}$$

$$\mu_{N,p}(x) \quad = \quad \frac{1}{(s \cdot x)^\alpha} \quad \text{with} \quad 0 \leqslant s, \alpha \leqslant 1 \tag{6.11}$$

where $\mu > 0$ refers to the heterogeneity or, to be more precise, the scaling parameter, and cnst to the constant term. At the bottom of the tree structure (cf. Fig. 3.2, sec. 3.4.5), the expected value of a nest refers to the given mode-specific generalised travel cost, and at the top, the expected value is the EMC indicator. At this point, it is important to draw attention to the scaling parameter μ of the model (6.11), which is defined in relation to the trip distance x as the IPAT model is applied for a broad range of travel distances and price levels. The parameter is distinguished by nest and by purpose.

During the computation of the EMC indicator, the second share factor PM that is applied in the integrated destination and mode choice model (6.2) can be determined according to Equations (3.26)-(3.29) that are documented in section 3.4.5.

The **GTT indicator** is an integrated, mode-specific cost indicator that refers to the generalised cost of travelling between two travel zones. It is computed by a cost function which applies travel impedance indicators computed by the VACLAV model. The cost function applied in the IPAT model satisfies the following equations:

$$c_{ij,p,m} = \omega_p^{(\text{time})} \cdot \left(\beta_{ci,m}^{(\text{time})} \cdot c_{ij,p,m}^{(\text{time})}\right) + (1 - \omega_p^{(\text{time})}) \cdot \left(\beta_{ci,m}^{(\text{dist})} \cdot c_{ij,p,m}^{(\text{dist})}\right) + \text{cnst}_{ci,p,m} \quad (6.12)$$

$$c_{ij,p,m}^{(\text{time})} = (t_{ij,p,m}^{(\text{net})})^{\lambda_m^{(\text{time})}} + t_{ij,p,m}^{(\text{ae})} \quad (6.13)$$

$$c_{ij,p,m}^{(\text{dist})} = \frac{60}{\text{VoT}_{ci,p}} \cdot \left((d_{ij,p,m}^{(\text{net})})^{\lambda_m^{(\text{dist})}} + d_{ij,p,m}^{(\text{ae})}\right) \quad (6.14)$$

where c identifies the GTT indicator, ω, β, and λ are modelling coefficients, and cnst is the constant term. The computation of time-related travel costs (6.13) and distance-related travel costs (6.14) relies on the following cost indicators:

- $t_{ij,pm}^{(\text{net})}$: network travel time with the main transport mode;
- $t_{ij,pm}^{(\text{ae})}$: time for accessing and egressing the main transport mode;
- $d_{ij,pm}^{(\text{net})}$: monetary travel cost of the trip with the main transport mode; and
- $d_{ij,pm}^{(\text{ae})}$: monetary travel cost of the A/E trip.

For converting monetary cost indicators into minutes, country and purpose-specific VoT indicators given in the unit [EUR/minute], which were taken from ETISplus data, are applied. It should be noted that the β coefficients and the cnst terms are not distinguished by country when the cost function (6.12) is calibrated in the first step at a European scale. This ensures the general transferability of the GTT indicator and, accordingly, the IPAT model to all European countries. The β coefficients and cnst terms are only adjusted in the second step when the calibration is revised for each country.

6.2.2.3 Conversion step

The conversion model translates the G/A trip matrices into passenger-kilometres and vehicle-kilometres. The two output indicators are computed as follows:

$$\text{pkm}_{ij,p,m} = T_{ij,p,m} \cdot \text{dist}_{ij,m} \quad (6.15)$$

$$\text{vkm}_{ij,p,m} = \text{pkm}_{ij,p,m} \cdot \text{occ}_{ci,p,m} \quad (6.16)$$

where the trip distance (dist) is distinguished by the O/D relation ij and mode m. The occupancy rate factor (occ) is distinguished by country ci, mode m and purpose p. These factors are given in the unit [passengers/vehicle] and rely on ETISplus data.

In addition, the conversion model computes several aggregates of the three output indicators of trips, pkm and vkm, e. g. the aggregated indicator "transport flows by mode at country level" in pkm. This indicator is particularly important for validating the IPAT model based on national transport statistics. It is computed as follows:

$$\text{pkm}_{ci,m} \quad = \quad 2 \cdot \sum_{ij,p} \delta_{ci,p}(i,j) \cdot \text{pkm}_{ij,p,m} \quad \text{with} \quad 0 \leqslant \delta_{ci}(i,j) \leqslant 1 \tag{6.17}$$

where $\delta_{ci,p}$ identifies the share of a trip for relation ij that takes place in country ci.[9] The factor is commonly determined based on routing information of the trips and is computed during the network assignment step. For the conversion step of the IPAT model, $\delta_{ci,p}$ was precalculated by the VACLAV model for each O/D relation based on the ETISplus network models.

It is important to note that most of the disaggregated output indicators of the IPAT model, such as the G/A matrices, only reflect half of the transport demand (they only refer to the outgoing trip demand). In contrast, most aggregated output indicators, such as pkm at country level, reflect the full demand, since the returning trips were implicitly added during their computations (this explains the factor 2 in (6.17)).

6.2.3 Calibration

The calibration process of the IPAT model follows an iterative approach in which the calibration parameters are continuously improved by comparing model outputs with transport demand indicators from statistics. Table 6.4 provides a rough overview of the calibration parameters.

Table 6.4: Calibration parameters of the IPAT model

Category	Parameters
GTT indicator	β, λ coefficients, cnst terms
Deterrence model	ω parameters, decay factors
EMC indicator	Heterogeneity parameter μ, cnst terms
Accessibility	$\omega_{ij,p}^{(\text{opp})}$, inter-zonal opportunities

A particular challenge stems from the integrated destination and mode choice model in which all calibration parameters interact. If the parameters of the deterrence model are updated, for instance, the trip matrix changes. As a result, overall market shares of transport modes at country level change and, accordingly, the EMC indicator changes as well. For this reason, the parametrisation of the IPAT model has to be revised parameter by parameter and incrementally in several intermediate steps. In each step, the differences between the IPAT model and transport statistics are evaluated as follows:

[9] For instance, the road travel distance between the cities of Karlsruhe (Germany) and Lyon (France) is about 550 km, where $\delta = 30\%$ can be allocated to Germany.

- the level of agreement between estimated and reported passenger-kilometres;
- the level of agreement between estimated and reported market shares of transport modes; and
- the alignment of the estimated deterrence function and the reported TLD.

Based on the outcome of this evaluation, the parameters of the IPAT model are slightly updated and the differences between model output and transport statistics are re-evaluated.

The overall calibration process is carried out at three levels. The model is initially calibrated at a European level, then at a national level, and finally at a regional level. This three-step approach ensures the transferability of the IPAT model between EU countries and regions, since the set of calibration parameters, which is derived during the first step, is valid for all countries. In addition, parameters of the GTT indicator are not revised when improving the model calibration at national and regional levels.

During the second step, the parameters of the deterrence model and the EMC indicator are revised country by country in order to improve the alignment of the IPAT model to transport statistics. In the third step, the calibration parameter $\omega_{ij,p}^{(\mathrm{OPP})}$ (6.7) is determined for each travel zone and purpose. Its values are distributed around 1 as the parameter primarily catches effects stemming from the heterogeneous size of the travel zones in a country. Subsequently, the level of agreement of the IPAT model with transport statistics is evaluated. If the discrepancy is small, all calibration parameters are accepted. Otherwise, the whole calibration process is re-started, following a slightly different strategy for updating the calibration parameters.

Thus the IPAT calibration was carried out in several incremental steps. Due to high interaction between several calibration parameters, the parameters were slightly updated in each step, and the computed output indicators of the IPAT model were then evaluated. For instance, if the decay factors related to the deterrence model (Tab. 6.4) are increased, the demand for trips is shifted towards shorter trip lengths. Accordingly, the market shares of transport modes also change[10] and the calibration of the modal split model computing the EMC indicator must be updated.

Questions concerning robust criteria for parameter optimisation arise for many multi-dimensional problems. If dimensions are correlated and cannot be decorrelated, parameter optimisation should be carried out in incremental steps in most cases. This conclusion is confirmed by experience obtained during the calibration of the PAT model (cf. sec. 2.3), the development of algorithms for classifying patterns in multi-dimensional image data (Ihrig, 2009), and in the calibration of the IPAT model. The elaboration of an automated approach for calibrating the IPAT model could be a subject of further studies.

[10] Air transport is particularly relevant for long-distance trips, while the market share of road transport is very high for short-distance trips (trams are not considered in the model but only regional trains).

6.3 Testing and validation of the IPAT model

The IPAT model was successfully tested and reviewed by a broader audience within the HIGH-TOOL project (see, for example, van Meijeren et al. (2016); Kiel et al. (2016a)). Within this thesis additional tests are carried out but with a greater emphasis on the methodology of the IPAT model. Outcomes of these tests are described in the following sections, starting with a brief discussion of the general test conditions. In the first assessment, the alignment of the IPAT model to the EU Reference Scenario 2013, which provides forecasts on transport statistics, is inspected for the year 2010. Sensitivities of the trip distribution and the modal split models to changes in travel cost indicators are then investigated for the year 2010 and the respective elasticity coefficients are discussed. Forecast capabilities of the IPAT model for the year 2030 are inspected by simulating several transport policy scenarios. In the final section, the main findings are summarised and conclusions are drawn that could be relevant for further methodological revisions of the IPAT model.

6.3.1 Test conditions

The testing and validation of the IPAT model were carried out by comparing the computed output indicator "annual pkm by transport mode" for each country to transport statistics provided by the EU Reference Scenario 2013, subsequently referred to as "Reference Scenario" and "EU Ref". The Reference Scenario provides important indicators at country level in 5-year steps from 2010 to 2050, including population and GDP as well as transport demand in terms of pkm by mode. For 2010 these indicators are in line with EU statistics.

For its application, the IPAT model is connected to the Data Stock that is published with the final version of the HIGH-TOOL model (Kiel et al., 2016b). The Data Stock encompasses modelled output indicators that are produced by the HIGH-TOOL modules. These indicators are largely in line with the Reference Scenario. However, given an almost linear dependency between population and generated transport demand, discrepancies for the indicator of population between the Data Stock and the Reference Scenario were counterbalanced as follows:

$$\mathrm{pkm}_{ci} = \tilde{\mathrm{p}}\mathrm{km}_{ci} \cdot \mathrm{cf}_{ci} \tag{6.18}$$

$$\mathrm{cf}_{ci} = \mathrm{pop}_{ci}^{(\mathrm{ref.})} \cdot \left(\mathrm{pop}_{ci}^{(\mathrm{mod.})}\right)^{-1} \tag{6.19}$$

where pkm_{ci} is the corrected demand in country ci after application of the respective correction factor cf_{ci}, $\tilde{\mathrm{p}}\mathrm{km}_{ci}$ is the demand computed by the IPAT model, $\mathrm{pop}_{ci}^{(\mathrm{mod.})}$ is the modelled population provided by the Data Stock and $\mathrm{pop}_{ci}^{(\mathrm{ref.})}$ the population provided by the Reference Scenario.[11] Consideration of the correction factors improves the results of the IPAT model concerning the comparison of modelled pkm to transport statistics. The factors were only applied for the base year 2010.

It should be noted that the results of the current validation process are not entirely consistent with those obtained for the HIGH-TOOL model. This applies especially for the case study "stimulating the competitiveness of high-speed rail" that is briefly described below. Discrepancies can be explained as follows: within this thesis the case study is simulated by directly modifying the cost function in the

[11] Relevant differences for 2010 exist only for two countries. The French population is overestimated by 3% and the Croatian population is underestimated by 2.8% (van Meijeren et al., 2016, p. 22).

IPAT model, while it is simulated in the HIGH-TOOL model by modifying input variables of the Vehicle Stock module.[12] The captured effects on transport demand in the HIGH-TOOL study are therefore strictly speaking second order effects.[13]

6.3.2 Comparing model results for 2010 with transport statistics

In this section, the results of the IPAT model computed for the base year 2010 are compared to transport statistics provided by the Reference Scenario. For this analysis, O/D-based output indicators "passenger transport performance by mode" were aggregated to NUTS-0 level and then adjusted according to (6.18). In addition, the market shares of transport modes were computed. The derived indicators of the IPAT model ($\hat{x}$) were then compared to those provided by the Reference Scenario (x) based on absolute errors (e_a) and relative errors (e_r). These error terms are computed as follows:

$$e_a = \hat{x} - x \tag{6.20}$$

$$e_r = \hat{x}/x - 1 \tag{6.21}$$

Table 6.5 summarises the pkm values by country and mode, while Figure 6.5 shows the relative model error by country.

Analysing the values of the indicator "total passenger transport performance" at NUTS-0 level in general reveals a good alignment of the IPAT model with the transport statistics provided by the Reference Scenario (EU Ref). For the majority of EU countries, except Cyprus, Ireland and Portugal, the discrepancies are below 2%. In the case of Cyprus (CY), the model overestimates the transport demand by 3.8%. Amongst other things, this can be explained by the specific structure of the transport market on Cyprus, which has no railway. Given that the IPAT model was calibrated to an average EU level encompassing four competing transport modes in the first step (cf. sec. 6.2.3), the model is not optimally suited for countries in which the structure of the transport market is different. However, for Malta (MT), which has no railway either, the goodness of fit of the IPAT model is better.

It is interesting to note that for both countries, Cyprus and Malta, both the IPAT model and the Reference Scenario show a small demand for rail passenger kilometres, different from zero. For the IPAT model, the generated rail pkm are related to intra-zonal trips. This can be explained by the application of a generic approach for estimating the travel impedance indicators for intra-zonal trips based on the region size.[14] Hence, the computation of an almost negligible market share for rail passenger transport on Cyprus and Malta could only be ensured by pushing the calibration of the applied mode choice model to its limits. A better approach would probably have been to set the rail travel time and the rail travel distance manually to costs of 9,999 for both countries.[15] For the Reference Scenario, the reasons are unknown and cannot be further investigated, given that detailed information on the applied modelling approach is not available.

[12] The methodology of this module, including the assumptions for computing the composition of the vehicle fleet and for deriving average travel cost by transport mode, is documented in Van Grol et al. (2016).

[13] Interpretation of the results of the HIGH-TOOL study is difficult, given that an increase of taxes on jet fuel may stimulate transport operators to use more fuel-efficient aircraft and, in consequence, air flights could even become cheaper.

[14] Travel impedance indicators cannot be derived based on shortest path information, as intra-zonal trips start and end in the same node of the network model.

[15] 9,999 is the default value for inter-zonal O/D relations if no connection exists between two travel zones.

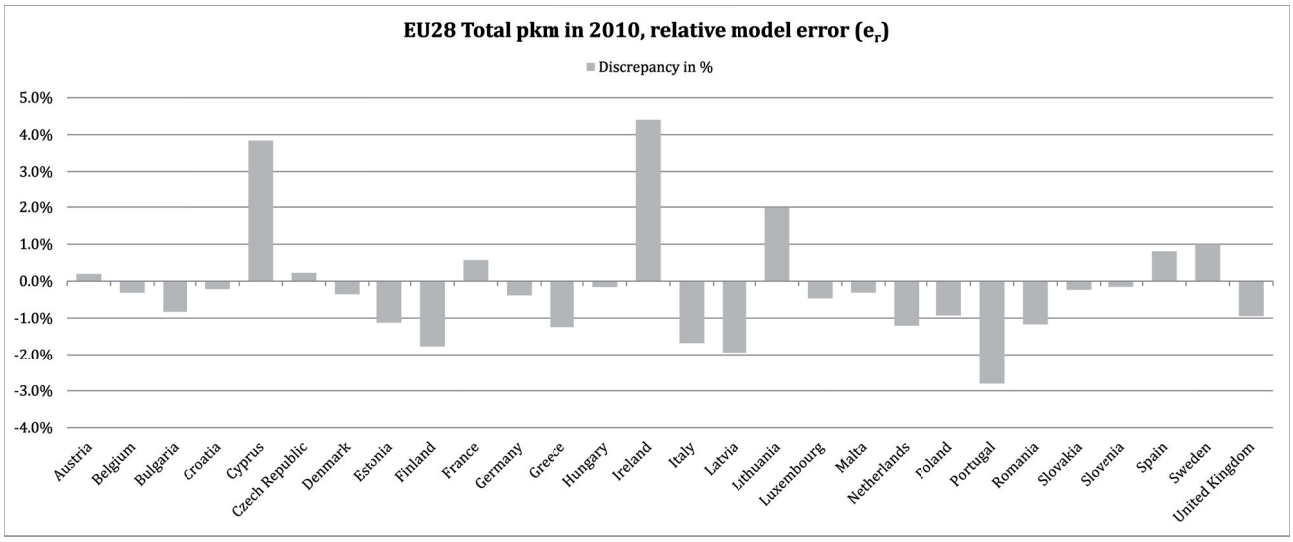

Figure 6.5: Validation of total passenger transport performance in 2010

Table 6.5: IPAT results vs. EU Ref, total pkm in 2010 by mode (Gpkm)

	road		rail		coach		air	
	IPAT	EU Ref	IPAT	EU Ref	IPAT	EU Ref	IPAT	EU Ref
EU28	4867.5	4893.4	405.4	405.5	364.1	365.1	525.1	525.6
EU15	4175.5	4197.4	359.6	359.5	284.9	285.7	476.2	476.7
EU13	692.0	696.0	45.8	46.0	79.2	79.4	48.9	48.9
AT	74.8	74.6	10.8	10.7	7.0	7.0	7.1	7.1
BE	110.1	110.5	10.0	10.0	14.3	14.4	8.6	8.6
BG	47.5	47.9	2.1	2.1	7.6	7.7	3.9	3.9
HR	26.0	26.0	1.7	1.7	3.0	3.0	3.1	3.1
CY	6.5	6.0	0.0	0.0	1.1	1.1	7.2	7.2
CZ	67.3	67.1	6.6	6.6	15.5	15.5	8.5	8.5
DK	51.4	51.6	6.3	6.3	4.5	4.5	10.2	10.2
EE	10.2	10.3	0.2	0.2	1.4	1.4	0.6	0.6
FI	64.4	65.7	3.9	4.0	4.4	4.5	8.5	8.5
FR	745.7	741.2	86.3	85.9	32.0	31.7	65.9	65.9
DE	900.0	904.7	83.4	83.0	44.7	44.6	58.5	58.5
EL	103.8	105.4	1.3	1.3	13.4	13.5	23.8	23.8
HU	53.8	53.8	7.7	7.7	13.7	13.8	3.6	3.6
IE	49.1	46.5	1.7	1.7	4.3	4.2	10.3	10.3
IT	726.5	740.5	46.8	47.3	75.9	76.5	49.2	49.4
LV	16.6	16.9	0.7	0.7	1.3	1.3	2.0	2.0
LT	31.2	30.6	0.4	0.4	2.3	2.3	1.2	1.2
LU	6.6	6.7	0.3	0.3	0.4	0.4	0.6	0.6
MT	2.2	2.3	0.0	0.0	0.4	0.4	2.6	2.6
NL	151.9	154.0	15.3	15.4	7.7	7.7	13.9	13.9
PL	300.2	303.3	17.8	17.9	16.3	16.4	8.4	8.4
PT	82.5	85.5	4.0	4.1	8.8	9.0	18.1	18.1
RO	77.3	78.3	5.4	5.4	10.2	10.3	6.6	6.6
SK	27.4	27.4	2.3	2.3	4.5	4.5	0.9	0.9
SI	25.9	26.0	0.8	0.8	1.8	1.8	0.3	0.3
ES	356.2	351.8	22.4	22.4	36.0	36.3	106.2	106.2
SE	101.3	100.1	11.3	11.2	4.9	4.8	14.2	14.2
UK	651.2	658.8	55.8	55.8	26.7	26.6	81.2	81.3

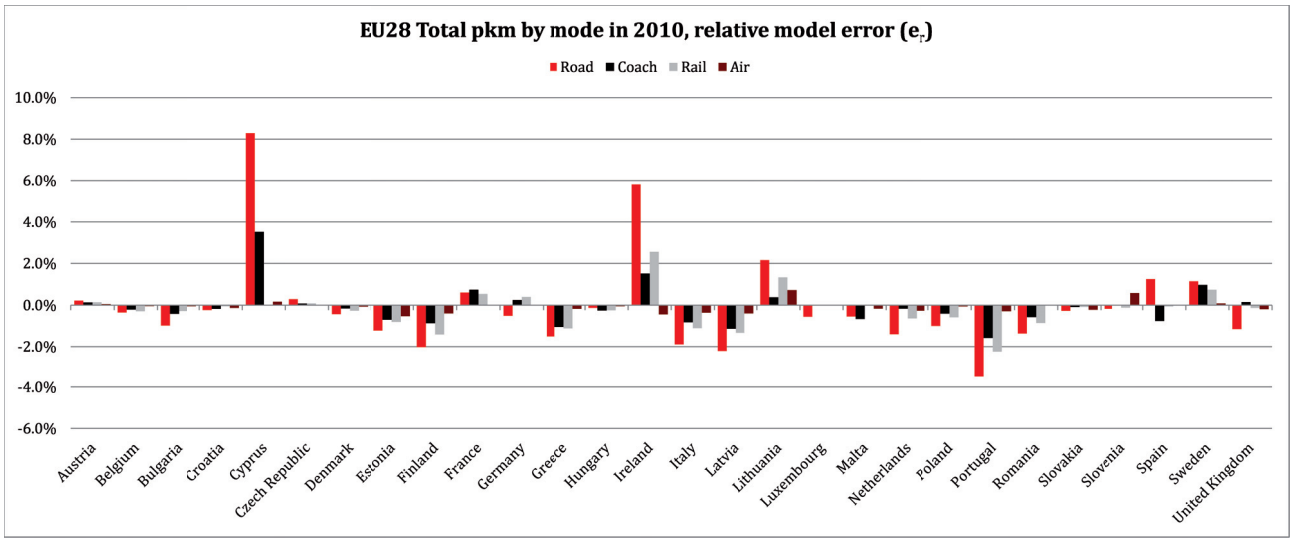

Figure 6.6: Validation of passenger transport performance by mode in 2010

In the case of Ireland (IE), the IPAT model overestimates the transport demand by 4.4% (Fig. 6.5). This cannot be explained at first glance but rather requires deeper analysis of the transport demand indicators summarised in Table 6.5. Compared to central EU countries, a key distinction of Ireland is the relevance of the transport mode of air for travelling.[16] While the EU average market share provided by the Reference Scenario for air passenger transport is only 8.5%, it is twice as high in Ireland with 16.4%. In the case of Portugal (PT), the IPAT model underestimates the transport demand by 2.8%. In this country, the market share of air transport is 15.5% and thus also about twice as high as the EU average.

Given the different relevance of the air transport mode in EU countries for travelling, particular attention was paid to this transport mode when calibrating the IPAT model. The discrepancy is therefore less than 0.5% for most of the countries[17] and the IPAT model is slightly overfitted for the air transport mode. This has adverse effects on the goodness of fit for the other transport modes.

Figure 6.6 shows the differences between modelled and reported pkm by transport mode and country.[18] In general, the alignment of the IPAT model is good except for the three countries already mentioned. The largest discrepancies between modelled pkm and the Reference Scenario can be observed for road transport in Cyprus (+8.3%) and Ireland (+5.8%). At the level of market shares, the differences are significantly lower (see below).

[16] According to the Reference Scenario, the high demand for air passenger kilometres applies particularly to long-distance flights over 1000 km.

[17] Except Lithuania: +0.7%; Slovenia: +0.6%; Estonia: -0.5%.

[18] For Cyprus and Malta, the values for rail passenger transport were set to zero (otherwise e_r would attach infinity).

In addition to the indicator "total passenger transport performance", the IPAT model was validated against the Reference Scenario concerning the output indicator "modal split based on pkm" (Tab. 6.6).

Table 6.6: IPAT results vs. EU Ref, modal split in 2010 based on pkm (%)

	road		rail		coach		air	
	IPAT	EU Ref	IPAT	EU Ref	IPAT	EU Ref	IPAT	EU Ref
EU28	79.0	79.1	6.6	6.6	5.9	5.9	8.5	8.5
EU15	78.8	78.9	6.8	6.8	5.4	5.4	9.0	9.0
EU13	79.9	80.0	5.3	5.3	9.1	9.1	5.6	5.6
AT	75.1	75.1	10.8	10.8	7.0	7.0	7.1	7.1
BE	76.9	77.0	7.0	7.0	10.0	10.0	6.0	6.0
BG	77.7	77.9	3.4	3.4	12.5	12.4	6.3	6.3
HR	76.7	76.7	5.1	5.1	8.9	8.9	9.3	9.3
CY	44.0	42.2	0.0	0.0	7.3	7.4	48.7	50.4
CZ	68.8	68.7	6.7	6.8	15.8	15.8	8.7	8.7
DK	71.0	71.1	8.7	8.7	6.2	6.1	14.1	14.1
EE	82.1	82.1	2.0	2.0	11.4	11.4	4.5	4.5
FI	79.3	79.5	4.8	4.8	5.5	5.4	10.4	10.3
FR	80.2	80.2	9.3	9.3	3.4	3.4	7.1	7.1
DE	82.8	82.9	7.7	7.6	4.1	4.1	5.4	5.4
EL	73.0	73.2	0.9	0.9	9.4	9.4	16.7	16.6
HU	68.3	68.2	9.7	9.7	17.4	17.5	4.6	4.6
IE	75.1	74.1	2.6	2.7	6.6	6.8	15.7	16.4
IT	80.9	81.0	5.2	5.2	8.4	8.4	5.5	5.4
LV	80.4	80.7	3.6	3.6	6.3	6.3	9.7	9.5
LT	88.7	88.6	1.1	1.1	6.6	6.8	3.5	3.6
LU	83.9	83.9	4.4	4.4	4.5	4.5	7.2	7.2
MT	42.9	43.0	0.1	0.0	7.6	7.6	49.5	49.4
NL	80.5	80.6	8.1	8.1	4.1	4.1	7.3	7.3
PL	87.6	87.7	5.2	5.2	4.8	4.7	2.4	2.4
PT	72.7	73.3	3.5	3.5	7.8	7.7	15.9	15.5
RO	77.7	77.8	5.4	5.4	10.3	10.2	6.6	6.6
SK	78.0	78.1	6.6	6.6	12.8	12.8	2.6	2.6
SI	89.7	89.7	2.8	2.8	6.3	6.3	1.2	1.2
ES	68.4	68.1	4.3	4.3	6.9	7.0	20.4	20.5
SE	76.9	76.8	8.6	8.6	3.7	3.7	10.8	10.9
UK	79.9	80.1	6.8	6.8	3.3	3.2	10.0	9.9

Figure 6.7 shows the differences between modelled and reported modal split values at NUTS-0 level for all transport modes. Analysing the discrepancies reveals an overall good alignment of the IPAT model to transport statistics. As might be expected, the maximum discrepancies can be observed for three countries already mentioned (CY, IE, PT) and follow directly from the differences between modelled and observed passenger kilometres.

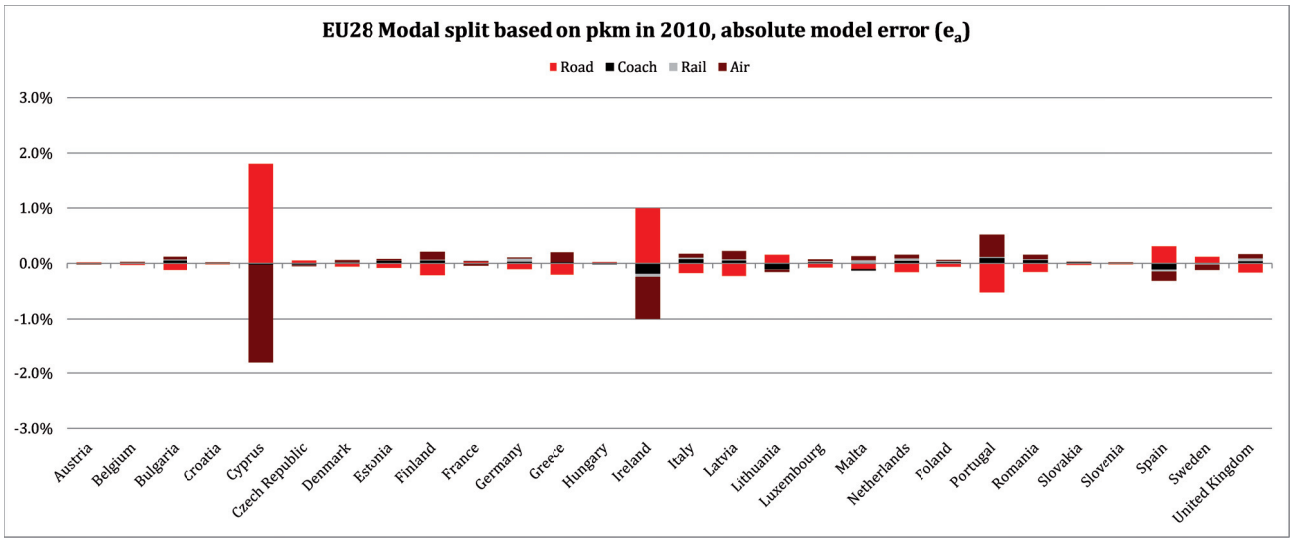

Figure 6.7: Validation of modal split in 2010

6.3.2.1 Summary

Comparing the IPAT results with pkm values of the Reference Scenario reveals an overall good alignment of the model for most countries. The highest differences can be observed for the three countries, Cyprus, Ireland and Portugal. They can partly be explained by the specific structure of the transport market in these countries, in which the market share of air transport is above the European average. In addition, a railway system is not present in Cyprus.

6.3.3 Model sensitivity to cost changes

In this section, the sensitivity of the IPAT model to changing travel costs is evaluated and in particular the sensitivity of the integrated destination and mode choice step. The tests were carried out with the final version of the HIGH-TOOL model (Kiel et al., 2016b) by modifying specific modelling coefficients of the IPAT model and by reproducing the passenger demand forecasts for the base year 2010. The procedure for testing the sensitivity of the IPAT model largely follows the procedure that was applied for validating the pre-final version of the HIGH-TOOL model (van Meijeren et al., 2016).

6.3.3.1 Defining testing scenarios and collecting results

In order to enable a largely independent evaluation of the elasticity coefficients for the trip distribution model and the mode choice model, two different testing methodologies were applied. For the trip distribution model, the following cost variable:

- EMC indicator described in section 6.2.2.2 by (6.10)

was modified by -5%, -2%, +2%, and +5%. These changes were carried out by adjusting the modelling coefficient r in (6.9). Table 6.7 shows the forecast passenger demand values and the values for the baseline run at an aggregated level, comparing old member states (EU15) and new member states (EU13) of the EU.

Table 6.7: Impact of cost changes (EMC) on passenger demand in 2010 (Gpkm)

	EU15				EU13			
	road	rail	coach	air	road	rail	coach	air
Baseline	4175.5	359.6	284.9	476.2	692.0	45.8	79.2	48.9
EMC +2%	4156.4	356.3	283.9	465.8	685.7	45.1	78.8	47.6
EMC +5%	4128.8	351.6	282.5	451.0	676.7	44.0	78.2	45.8
EMC -2%	4195.2	362.9	286.0	487.0	698.5	46.5	79.6	50.3
EMC -5%	4225.9	368.2	287.6	504.0	708.8	47.7	80.3	52.4

In order to evaluate the sensitivity of the modal split model, the following cost variables:

- monetary travel cost;
- time-related travel cost; and
- both cost variables

were modified for each individual transport mode independently by -5%, -2%, +2%, and +5%. These changes were carried out by modifying the modelling coefficients $\beta_{ci,m}^{(\text{time})}$ and $\beta_{ci,m}^{(\text{dist})}$ in (6.12). In total, 48 model runs were conducted. Table 6.8 shows the forecast passenger demand values in relation to monetary travel cost and the values for the baseline run at an aggregated level for the EU15 and the EU13, Table 6.9 shows the forecast values in relation to time-related travel cost, and Table 6.10 in relation to modifications of both cost variables.

Table 6.8: Impact of monetary cost changes on passenger demand in 2010 (Gpkm)

	EU15				EU13			
	road	rail	coach	air	road	rail	coach	air
Baseline	4175.5	359.6	284.9	476.2	692.0	45.8	79.2	48.9
road +2%	4157.1	363.9	287.1	478.1	687.6	46.4	79.8	49.2
road +5%	4129.9	370.4	290.3	481.0	681.2	47.5	80.8	49.5
road -2%	4194.1	355.3	282.8	474.2	696.4	45.1	78.6	48.7
road -5%	4222.4	349.0	279.7	471.2	703.1	44.2	77.7	48.3
rail +2%	4180.1	352.1	285.1	476.5	692.6	44.7	79.3	48.9
rail +5%	4186.8	341.5	285.5	476.9	693.5	43.1	79.4	49.0
rail -2%	4170.7	367.2	284.7	475.9	691.3	46.9	79.2	48.9
rail -5%	4163.4	379.2	284.3	475.4	690.3	48.7	79.1	48.8
coach +2%	4178.2	359.8	281.1	476.3	692.7	45.8	78.2	48.9
coach +5%	4182.1	360.1	275.6	476.4	693.7	45.9	76.6	49.0
coach -2%	4172.7	359.3	288.9	476.1	691.2	45.7	80.3	48.9
coach -5%	4168.4	358.9	295.0	475.9	690.1	45.6	82.1	48.9
air +2%	4176.4	359.7	285.0	472.3	692.1	45.8	79.2	48.5
air +5%	4177.8	360.0	285.0	466.4	692.3	45.8	79.2	47.8
air -2%	4174.5	359.4	284.9	480.1	691.9	45.8	79.2	49.4
air -5%	4173.1	359.2	284.8	486.2	691.7	45.8	79.2	50.1

Table 6.9: Impact of time cost changes on passenger demand in 2010 (Gpkm)

	EU15				EU13			
	road	rail	coach	air	road	rail	coach	air
Baseline	4175.5	359.6	284.9	476.2	692.0	45.8	79.2	48.9
road +2%	4162.7	363.0	286.5	477.6	688.4	46.3	79.7	49.1
road +5%	4143.6	368.2	288.9	479.8	683.2	47.1	80.4	49.4
road -2%	4188.4	356.1	283.3	474.7	695.6	45.3	78.8	48.7
road -5%	4208.0	351.1	281.0	472.5	701.0	44.5	78.1	48.5
rail +2%	4179.2	354.2	285.1	476.4	692.6	44.7	79.3	48.9
rail +5%	4184.6	346.2	285.4	476.7	693.5	43.2	79.4	48.9
rail -2%	4171.8	365.0	284.7	476.0	691.3	46.9	79.2	48.9
rail -5%	4166.0	373.5	284.5	475.6	690.3	48.6	79.1	48.9
coach +2%	4177.2	359.7	282.3	476.3	692.5	45.8	78.3	48.9
coach +5%	4179.7	360.0	278.6	476.4	693.4	45.9	77.0	48.9
coach -2%	4173.7	359.4	287.6	476.1	691.4	45.7	80.2	48.9
coach -5%	4171.0	359.1	291.7	476.0	690.5	45.7	81.7	48.9
air +2%	4176.1	359.7	284.9	474.1	692.0	45.8	79.2	48.6
air +5%	4176.9	359.9	285.0	471.0	692.1	45.8	79.2	48.2
air -2%	4174.9	359.4	284.9	478.3	691.9	45.8	79.2	49.2
air -5%	4174.0	359.3	284.9	481.5	691.8	45.8	79.2	49.6

Table 6.10: Impact of time and monetary cost changes on passenger demand in 2010 (Gpkm)

	EU15				EU13			
	road	rail	coach	air	road	rail	coach	air
Baseline	4175.5	359.6	284.9	476.2	692.0	45.8	79.2	48.9
road +2%	4144.4	367.3	288.7	479.6	684.1	47.0	80.3	49.3
road +5%	4098.6	379.2	294.4	484.6	672.7	48.8	82.0	50.0
road -2%	4207.1	351.9	281.2	472.7	700.0	44.6	78.2	48.5
road -5%	4255.6	340.8	275.8	467.5	712.3	42.9	76.6	47.8
rail +2%	4183.7	346.9	285.3	476.7	693.2	43.7	79.3	48.9
rail +5%	4195.5	328.8	285.9	477.4	695.0	40.7	79.5	49.0
rail -2%	4166.9	372.8	284.5	475.7	690.7	48.0	79.1	48.9
rail -5%	4153.6	393.8	283.9	474.8	688.6	51.7	78.9	48.8
coach +2%	4179.9	360.0	278.6	476.4	693.3	45.9	77.3	48.9
coach +5%	4186.1	360.6	269.7	476.6	695.0	46.0	74.6	49.0
coach -2%	4170.9	359.1	291.6	476.0	690.6	45.7	81.3	48.9
coach -5%	4163.6	358.4	302.3	475.7	688.5	45.5	84.7	48.8
air +2%	4177.0	359.8	285.0	470.2	692.2	45.8	79.2	48.2
air +5%	4179.2	360.3	285.1	461.3	692.4	45.8	79.3	47.1
air -2%	4174.0	359.3	284.9	482.3	691.8	45.8	79.2	49.6
air -5%	4173.1	359.2	284.8	486.2	691.7	45.8	79.2	50.1

6.3.3.2 Sensitivity of the trip distribution model

The trip distribution model computes the G/A trip matrix. It relies on the deterrence model that estimates disincentives to travel in relation to travel costs (i. e. EMC values). If the EMC values are increased, it is expected that the distribution of trip lengths reported by the G/A trip matrix will shift towards short-distance trips and accordingly the overall demand in terms of passenger kilometres will decrease. If travel costs are decreased, the opposite is expected. Table 6.11 gives an overview of the derived elasticity values by transport mode if the EMC values are modified in a percentage change.[19] In addition, the average elasticity values ($\bar{x}_{\text{EMC}}$) are included in the table.

Table 6.11: Model elasticities in 2010 in relation to changes of EMC

	EU15				EU13			
	road	rail	coach	air	road	rail	coach	air
$\bar{x}_{\text{EMC}}$	-0.232	-0.463	-0.178	-1.111	-0.464	-0.806	-0.263	-1.345
EMC +2%	-0.232	-0.462	-0.177	-1.112	-0.461	-0.802	-0.263	-1.342
EMC +5%	-0.231	-0.461	-0.176	-1.114	-0.458	-0.796	-0.262	-1.338
EMC - 2%	-0.233	-0.464	-0.178	-1.109	-0.466	-0.810	-0.264	-1.348
EMC - 5%	-0.234	-0.465	-0.179	-1.106	-0.470	-0.816	-0.265	-1.351

Given that only EMC values are modified for this analysis, but not the mode-specific travel costs underlying the EMC indicator, the modal split values remain constant for all O/D relations. However, the sensitivity values differ by transport mode, as the structure of the trip matrix changes slightly and, accordingly, the overall market shares of transport modes changes also.

The obtained elasticity coefficients are plausible and meet the expectations in general. The highest sensitivity can be observed for air transport demand which is particularly related to non-mandatory and long-distance trips (e. g. vacation trips). The sensitivity of rail and road transport demand is significantly lower and both transport modes are also relevant for compulsory trips (e. g. commuting). For coach transport, the IPAT model tends to respond too inelastically. In principle, the IPAT model responds more sensitively for new member states (EU13), which have joined the EU since 2004. Compared to the elasticity values found in the literature, the IPAT model responds less sensitively to cost changes.[20] This is still within acceptable limits, taking into account the high standard deviations of the mean values for the collected elasticities.

The comparatively low sensitivity of the IPAT model has several causes. First of all, the effects of induced transport are not considered by the current analysis, as cost changes have no impact on trip generation. Hence, if travel costs (i. e. EMC values) are decreased, short-distance trips are replaced by longer trips but "new" demand like weekend trips is not created. Another reason is the large diameter of the NUTS-2 regions, which is over 125 km. Thus, a very high share of the transport demand is carried out by intra-zonal trips that respond relatively inelastically to cost changes due to the modelling

[19] The elasticity values are computed according to (3.1) in section 3.1.

[20] As part of the HIGH-TOOL validation process (see Van Grol et al., 2016, Tables 36 and 37, page 166), elasticities were collected and analysed by category. Concerning changes of the overall passenger transport demand (pkm), the mean value of the direct cost elasticity is given as -0.75 and that of the direct time elasticity as -0.94. The corresponding standard deviations are 0.33 and 0.71.

approach. This mitigates the effects of changing travel costs on the distribution of trip length and, accordingly, the computed transport demand in terms of pkm.

6.3.3.3 Sensitivity of the modal split model

The IPAT model applies an NMNL model for computing the market shares of transport modes and for deriving the EMC values that feed into the deterrence model (see above). If a cost variable for a specific transport mode is increased, a modal shift in the demand towards the competing transport modes is expected. If a cost variable for a specific transport mode is decreased, the opposite is expected. Table 6.12 gives an overview of the mean elasticity and cross-elasticity values.[21]

Table 6.12: Model elasticities in 2010 in relation to changes of mode cost indicators

	EU15				EU13			
	road	rail	coach	air	road	rail	coach	air
monetary travel costs								
$\bar{x}_{\text{road}}$	-0.221	0.594	0.374	0.206	-0.316	0.715	0.383	0.261
$\bar{x}_{\text{rail}}$	0.056	-1.048	0.040	0.031	0.046	-1.215	0.036	0.024
$\bar{x}_{\text{coach}}$	0.033	0.034	-0.681	0.011	0.053	0.068	-0.689	0.020
$\bar{x}_{\text{air}}$	0.011	0.022	0.008	-0.415	0.009	0.015	0.006	-0.465
time-related travel costs								
$\bar{x}_{\text{road}}$	-0.154	0.476	0.280	0.153	-0.257	0.573	0.300	0.186
$\bar{x}_{\text{rail}}$	0.044	-0.756	0.032	0.022	0.046	-1.174	0.038	0.018
$\bar{x}_{\text{coach}}$	0.021	0.026	-0.461	0.009	0.042	0.055	-0.584	0.016
$\bar{x}_{\text{air}}$	0.007	0.017	0.004	-0.221	0.005	0.007	0.003	-0.279
monetary and time-related travel costs								
$\bar{x}_{\text{road}}$	-0.376	1.070	0.653	0.359	-0.573	1.287	0.682	0.447
$\bar{x}_{\text{rail}}$	0.100	-1.804	0.072	0.053	0.092	-2.387	0.074	0.042
$\bar{x}_{\text{coach}}$	0.054	0.060	-1.142	0.020	0.095	0.123	-1.273	0.036
$\bar{x}_{\text{air}}$	0.016	0.035	0.011	-0.581	0.012	0.020	0.007	-0.676

For instance, if the monetary costs of road transport increase by 1%, the demand decreases for road transport by 0.2% and increases for competing transport modes due to modal shift effects. If travel costs for other transport modes change, the effects are very similar. Accordingly, all elasticity values have a negative sign and all cross-elasticity values have a positive sign.

Overall, the derived elasticity values show very similar patterns for the EU15 and the EU13, and the model responds less sensitively to modifications in time-related travel costs than in monetary travel costs. This model behaviour is driven to a large extent by the assumption that the demand for private and vacation trips is more sensitive to changing monetary cost than to changing time-related costs.[22] Moreover, the model responds less to cost changes for private road transport than for public transport

[21] Cross elasticities are computed according to (3.6) in section 3.1.2 for the four cost changes -5%, -2%, +2%, +5% (cf. Tab. 6.8 to 6.10). The variables $\bar{x}_{\text{mode}}$ report the mean values of these elasticities over the four cost changes.

[22] The model behaviour is determined by the parameter ω (6.12) weighting monetary and time-related travel costs.

modes. This is in line with the expectations and can be explained by the fact that a large majority of private road transport demand is related to mandatory trips (e. g. commuting and private trips) and intra-zonal trips. By contrast, trips with public transport modes, and in particular flights, are often optional and the respective demand is therefore more sensitive to cost changes. In addition, the sensitivity analysis reveals that the response of the IPAT model is stronger for the new member states (EU13). This can be explained by the lower income level in these countries and, accordingly, by a lower "mobility budget". Hence, the demand is more sensitive to changing travel costs.

It should be noted that the elasticity values (Tab. 6.12) do not exclusively refer to modal shift effects but also to changes in the trip distribution. If travel costs for a single transport mode change, overall travel costs (i. e. EMC values) also change and, accordingly, the trip distribution shifts. Induced trips are not considered, as the trip generation model is not sensitive to changing travel costs.

6.3.3.4 Summary

Testing the sensitivity of the integrated destination and mode choice step of the IPAT model generally reveals reasonable model reactions. If a mode-specific travel cost variable is changed, modal shift effects can be observed. If overall travel costs are changed, the distribution of trips changes and, accordingly, the distribution of trip lengths. The sensitivity of the IPAT model is lower than might be expected. This is due to the large size of the travel zones and, accordingly, a very high share of intra-zonal trips.[23] In addition, and though it often occurs in reality, if travel costs decrease, induced transport demand is not modelled.

6.3.4 Model validation for 2030 by policy scenarios

This section analyses several transport policy scenarios and quantifies their effects on traffic demand forecasts for the year 2030. The defined scenarios largely refer to the case studies that were conducted in order to test and to validate the overall HIGH-TOOL model with regard to transport policy assessment (Kiel et al., 2016a). In contrast to the case studies conducted with the HIGH-TOOL model that often encompass several policy assumptions and that are simulated for quantifying the social, economic and environmental effects for the whole timeline from 2015-2050, this analysis quantifies only the effects on passenger transport demand for the year 2030. Moreover, precisely one policy assumption is made per scenario and only direct input indicators of the IPAT model are changed. It is further assumed that all policies are only enacted in 2030. This avoids the occurrence of second order effects and effects of higher order that are introduced through feedback loops between the HIGH-TOOL modules.[24]

The required traffic demand indicators for 2030, like population and GDP, are provided by the HIGH-TOOL baseline run and reflect the Reference Scenario. The following policy scenarios are investigated with the IPAT model:

[23] Given an average diameter of over 125 km of the NUTS-2 regions, the IPAT model is particularly suited to long-distance trips. This issue is further discussed in Part III of this thesis, leading to the development of the HIPAT model.

[24] A second order effect is an effect that is not directly induced by the transport policy measure (TPM). For instance, if high-speed rail (HSR) is subsidised for a single time period (e. g. between 2015 and 2019), a modal shift towards rail passenger transport can be expected and, in addition, a higher demand due to induced transport. This can have stimulating effects on economic growth - and economic growth in turn on transport demand. Thus, for the next time period a higher demand can be expected even if HSR is no longer subsidised.

- benchmark scenario;
- corridor improvement for rail passenger transport; and
- stimulating the competitiveness of high-speed rail.

6.3.4.1 Benchmark scenario

The benchmark scenario reproduces the traffic demand forecasts that are provided by the HIGH-TOOL model for 2030, as no policy assumption is made. The results of this scenario are compared to the forecasts provided by the Reference Scenario. Table 6.13 summarises the pkm values, and Figures 6.8 and 6.9 show the relative model errors concerning total pkm and pkm by mode.

Table 6.13: IPAT results vs. EU Ref, total pkm in 2030 by mode (Gpkm)

	road		rail		coach		air	
	IPAT	EU Ref	IPAT	EU Ref	IPAT	EU Ref	IPAT	EU Ref
EU28	5650.7	5713.7	639.8	606.5	412.4	411.3	876.8	887.5
EU15	4488.8	4526.2	521.6	489.5	301.4	299.3	566.8	573.0
EU13	1161.9	1187.5	118.2	117.0	111.0	112.0	309.9	314.4
AT	85.8	85.4	15.7	14.3	7.9	7.8	13.0	13.2
BE	127.8	129.4	15.9	14.6	16.6	16.7	14.3	14.5
BG	51.3	51.8	3.2	3.3	8.0	7.9	8.2	8.7
HR	30.7	34.1	2.1	2.4	3.6	3.9	5.8	6.7
CY	8.5	7.7	0.0	0.0	1.3	1.2	15.0	15.0
CZ	88.2	88.4	12.4	12.5	20.0	19.8	16.2	16.5
DK	57.2	56.4	8.3	7.9	5.0	4.9	17.9	17.9
EE	12.3	12.8	0.4	0.4	1.6	1.7	1.2	1.3
FI	70.8	71.7	5.7	5.2	4.6	4.6	14.9	15.1
FR	877.7	867.2	124.7	121.0	38.8	37.4	104.0	102.4
DE	927.8	942.3	143.3	122.8	48.5	48.4	92.4	96.5
EL	112.7	111.9	1.7	1.7	14.6	14.2	44.4	44.3
HU	67.3	69.0	12.2	11.6	16.1	16.0	7.5	7.7
IE	61.9	61.1	2.1	2.1	4.5	4.5	16.7	16.6
IT	803.8	818.0	66.9	65.5	80.4	80.7	81.7	83.9
LV	19.6	19.8	1.2	1.2	1.6	1.6	4.8	4.8
LT	34.2	35.7	0.6	0.6	2.6	2.6	2.7	2.9
LU	8.5	8.5	0.5	0.5	0.4	0.4	1.0	1.0
MT	2.4	2.4	0.0	0.0	0.4	0.4	5.0	5.0
NL	166.6	169.2	23.4	20.2	8.2	8.2	22.8	23.3
PL	406.3	410.7	35.0	35.4	16.0	15.9	17.2	17.6
PT	96.9	100.0	6.9	7.0	10.8	10.9	28.5	28.8
RO	107.9	111.1	12.8	12.5	13.0	13.0	17.0	17.5
SK	41.3	42.2	4.3	4.2	6.8	6.8	2.0	2.0
SI	32.3	33.2	2.1	2.1	1.7	1.7	0.7	0.7
ES	469.5	483.2	47.6	50.7	41.5	42.4	168.8	170.9
SE	119.0	118.1	15.2	14.5	5.5	5.5	25.6	25.3
UK	762.4	772.3	75.5	72.4	32.3	32.2	127.4	127.3

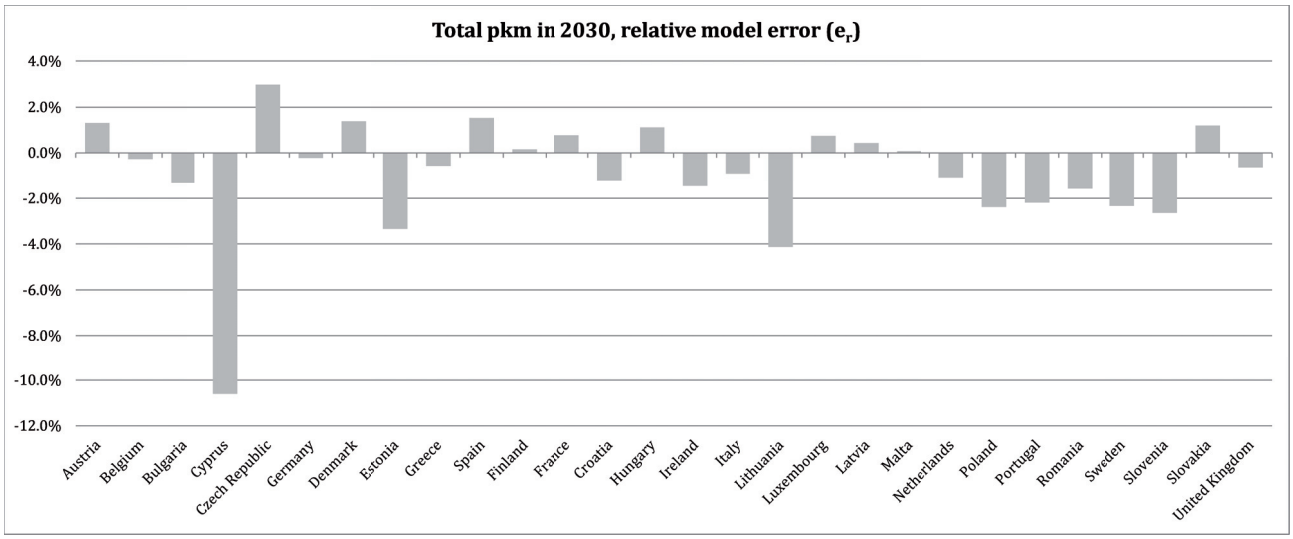

Figure 6.8: Validation of total passenger transport performance in 2030

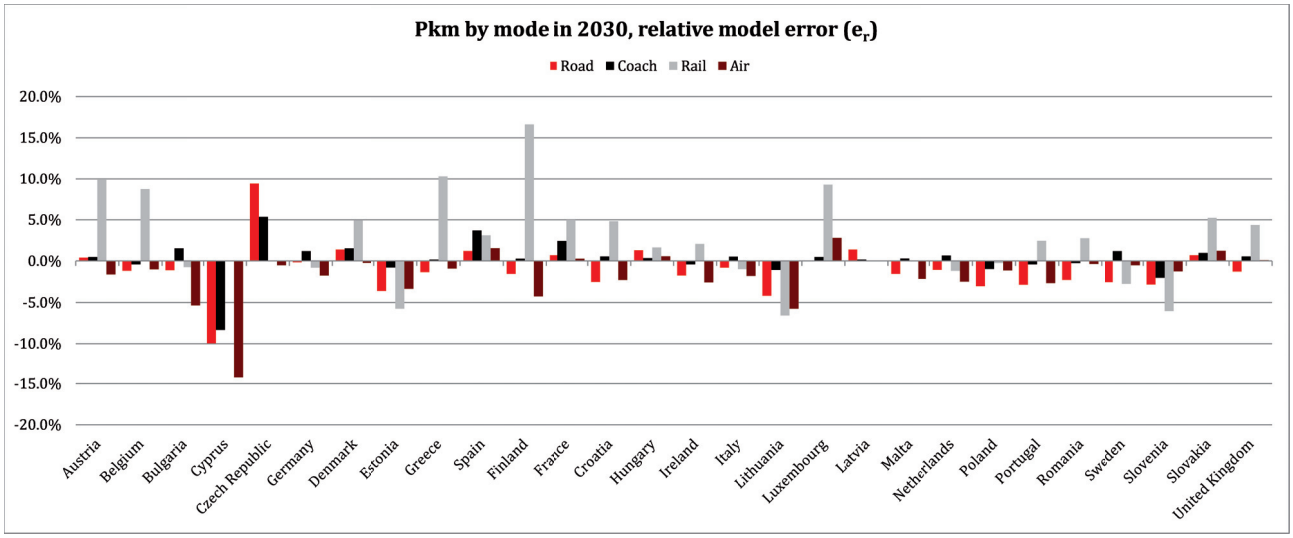

Figure 6.9: Validation of passenger transport performance by mode in 2030

Overall, the modelled demand for 2030 is much in line with the transport demand forecasts provided by the Reference Scenario. For most countries, the relative model error e_r is lower than 2% in terms of total pkm and pkm by mode. The largest discrepancies can be observed for Cyprus (total pkm: -10.6%, air pkm: -14.3%) and Finland (rail pkm: 16.6%). For Cyprus, this can be explained by the stronger growth of the air transport demand between 2010 and 2030 in comparison to the EU average (148% compared to 129%). As the model was calibrated to an average EU level, it is not optimally suited for countries with significantly different growth rates.

For Finland, the growth rates are very similar to the EU average and thus the relative model error for rail transport cannot be explained in simple terms. This would require deeper analysis of the general discrepancies between the HIGH-TOOL baseline run and the assumptions underlying the EU Reference Scenario for 2030 in terms of GDP, population and cost indicators, for instance. Furthermore, it has to be emphasised that the IPAT model has only been calibrated to the base year 2010. It can therefore be concluded that the differences for the forecast year 2030 are still within an acceptable range.

6.3.4.2 Corridor improvement for rail passenger transport

This case study illustrates the potential of a network-based transport model to evaluate the implementation of a new trans-European high-speed rail (HSR) corridor. In HIGH-TOOL, such analyses are made possible by the hypernet sub-module which complements the IPAT model. The hypernet relies on two simplified network models for rail and road passenger transport in which neighbouring NUTS-2 regions are connected by precisely one link, and only if these regions are also connected in the ETISplus networks. Hence, the hypernet sub-module can simulate network effects of transport policies but unlike the VACLAV model (cf. sec. 2.2.1), it is not suited for the modelling of traffic loads at link level.

The current study refers to a hypothetical investment scenario, where the "Magistrale" rail corridor Paris–Strasbourg–Karlsruhe–Munich–Vienna–Bratislava is upgraded. It is assumed that travel times decrease by 10% for all rail links along the corridor for 2030 and beyond. Based on these modifications, the hypernet updates the set of travel impedances that feed into the cost functions of the IPAT model, computing the effects on O/D-based passenger transport flows.

In the following, the impact of this policy on transport demand is investigated and quantified. For the countries France, Austria, Germany and Slovakia, larger changes are expected than in countries which do not directly benefit from this corridor upgrade. Table 6.14 compares the forecast passenger demand values by mode at NUTS-0 level for the policy scenario with those of the benchmark scenario for the year 2030. Figure 6.10 illustrates the changes in the demand by mode.

Overall, the forecast values for the policy scenario show reasonable patterns that are in line with expectations. The hypothetical investment in the "Magistrale" corridor stimulates an increase in rail passenger transport stemming almost equally from modal shift effects and from induced traffic demand.[25] Although the effect of induced traffic demand seems to be exceptionally high, this pattern is much in line with observations made for the French HSR line Paris-Lyon (Vickerman, 1996). The large increase can be explained by direct and indirect effects, such as modifications of travellers' mobility behaviour and new activities due to the reduced travel time (Cascetta and Coppola, 2014).

The strongest effects of the hypothetic investment scenario can be observed in Germany, followed by Austria, France and Slovakia that are the main beneficiaries. Smaller effects can be observed in countries like the Czech Republic, Hungary and Romania that are located around the corridor or in its extensions. Some effects can also be observed in distant countries like Spain and the United Kingdom. This is due to the modelling approach. For instance, the decrease in travel costs for rail transport between Ile de France (FR10) and Champagne-Ardenne (FR21) has a slight impact on the travel costs for travelling from both countries to France, Germany, etc. due to network effects. However, the percentage changes for Spain and the United Kingdom are negligible (see Tab. 6.14).

In principle, the changes in relation to the total passenger demand are very small, in particular for air transport. This can firstly be explained by the relatively limited geographical scope of the assumed infrastructure upgrade in the concerned countries. In addition, the assumed travel time savings affect inter-zonal, long-distance rail passenger transport flows, which represent a comparatively small share of the overall transport market. Large modal shift effects from air to rail transport cannot be observed, given that flights are particularly related to trips of very long travel distances. These long-distance trips are, however, hardly influenced by the investigated policy scenario.

[25] In Figure 6.10 the induced travel demand can be read as the difference between all positive and negative changes.

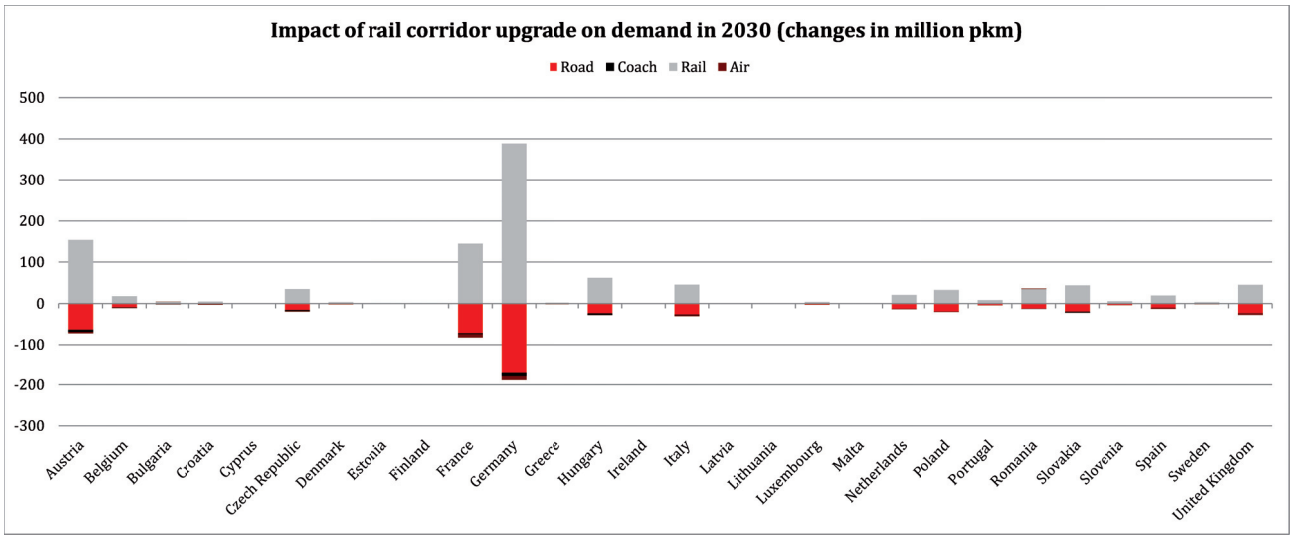

Figure 6.10: Impact of rail corridor upgrade on demand in 2030

Table 6.14: Impact of rail corridor upgrade on demand in 2030 [1]

	road		rail		coach		air	
	pkm	%	pkm	%	pkm	%	pkm	%
EU28	-500.5	-0.01	1,090.4	0.01	-29.6	0.00	-35.1	0.00
EU15	-424.6	-0.01	919.7	0.01	-22.7	0.00	-29.5	0.00
EU13	-75.9	-0.01	170.7	0.01	-6.9	0.00	-5.6	0.00
AT	-63.4	-0.10	154.1	0.12	-5.4	-0.01	-3.9	-0.01
BE	-9.6	-0.01	17.9	0.01	-0.9	0.00	-0.7	0.00
BG	-1.7	-0.01	4.5	0.01	-0.1	0.00	0.2	0.00
HR	-2.3	-0.01	4.8	0.01	-0.2	0.00	-0.1	0.00
CY	0.0	0.00	0.0	0.00	0.0	0.00	-0.1	0.00
CZ	-15.7	-0.02	35.4	0.02	-2.8	0.00	-1.0	0.00
DK	-1.8	0.00	3.6	0.00	-0.1	0.00	-0.3	0.00
EE	-0.1	0.00	0.3	0.00	0.0	0.00	0.0	0.00
FI	-0.6	0.00	1.5	0.00	0.0	0.00	-0.2	0.00
FR	-71.8	-0.01	144.9	0.01	-2.5	0.00	-8.4	0.00
DE	-171.3	-0.03	388.5	0.03	-8.2	0.00	-9.1	0.00
EL	-1.2	0.00	2.2	0.00	0.0	0.00	-0.3	0.00
HU	-23.9	-0.05	62.6	0.06	-3.7	-0.01	-0.1	0.00
IE	-0.5	0.00	1.1	0.00	0.0	0.00	-0.2	0.00
IT	-26.7	0.00	45.9	0.00	-1.5	0.00	-2.8	0.00
LV	-0.5	0.00	1.1	0.00	0.0	0.00	0.0	0.00
LT	-0.2	0.00	0.4	0.00	0.0	0.00	0.0	0.00
LU	-2.5	-0.03	3.9	0.04	-0.1	0.00	-0.1	0.00
MT	0.0	0.00	0.0	0.00	0.0	0.00	0.0	0.00
NL	-12.4	-0.01	21.2	0.01	-0.3	0.00	-1.3	0.00
PL	-19.8	-0.01	33.3	0.01	-0.4	0.00	-0.3	0.00
PT	-4.0	0.00	8.4	0.01	-0.1	0.00	-0.5	0.00
RO	-12.5	-0.02	35.6	0.02	-0.4	0.00	1.3	0.00
SK	-19.6	-0.07	44.2	0.08	-2.3	-0.01	-0.2	0.00
SI	-3.5	-0.02	6.0	0.02	-0.1	0.00	0.0	0.00
ES	-9.6	0.00	19.9	0.00	-0.3	0.00	-2.9	0.00
SE	-1.5	0.00	3.4	0.00	0.0	0.00	-0.3	0.00
UK	-23.8	0.00	45.8	0.00	-0.1	0.00	-3.7	0.00

[1] Difference Scenario-Baseline (in million pkm)

6.3.4.3 Stimulating the competitiveness of high-speed rail

This scenario investigates two policies attempting to improve the competitiveness of HSR at the expense of air passenger trips by subsidising rail ticket costs or by raising taxes on air tickets in the year 2030. The first policy is modelled by decreasing the coefficient $\beta_{ci,m}^{(\text{dist})}$ in (6.12) by 10% for rail trips in the distance range of 300-1000 km and the second policy is modelled by increasing the respective coefficient by 10% for air trips. Tables 6.15 and 6.16 summarise the impact of the two investigated policies on the demand in 2030, and Figures 6.11 and 6.12 illustrate the changes by mode.

Overall, the changes for 2030 look plausible. For both policies, modal shift effects can be observed as well as induced transport demand. If rail trips are subsidised, the demand increases at the expense of the other transport modes and the overall demand is higher compared to the benchmark scenario (cf. Tab. 6.13). Equally, if taxes on air trips are raised, a modal shift to the other transport modes can be observed and the overall transport demand is lower compared to the benchmark scenario. The observed effects, i. e. modal shift towards the cheaper transport mode(s) as well as induced demand, to changes in travel costs are largely in line with observations and can be explained by the fact that travellers tend to minimise their individual travel costs.

However, changes in market shares for the first investigated policy are lower than expected taking into account the elasticities of the modal split model that were previously derived for the base year 2010. If HSR trips are subsidised by 10% in the year 2030, rail demand increases on average by 2.96% for the EU15, and 1.98% for the EU13 countries (Tab. 6.15), and elasticities are -0.3 and -0.2 respectively. In contrast, elasticities derived for the logit model for the base year 2010 are -1.05 for the EU15 and -1.22 for the EU13 (cf Tab. 6.12). This can be explained by the fact that for the investigated policy only rail trips in the distance range of 300-1000 km are subsidised. Thus, conventional rail trips in the distance range of 0-300 km are not affected by the policy. In contrast, rail travel costs were reduced for all trip distances when elasticities for the logit model for the base year 2010 were derived. This explains the differences between the derived elasticity values for the years 2030 and 2010.

For the second investigated policy, changes in market shares are more in line with expectations. If air ticket costs are increased by 10% in the year 2030, the demand decreases on average by 3.31% for the EU15 and 2.97% for the EU13 countries (Tab. 6.15), and elasticities are -0.33 and -0.30 respectively. These elasticities are more consistent with the previously determined elasticities for the base year 2010 for the logit model, i. e. -0.42 for the EU15 and -0.47 for the EU13 (cf. Tab. 6.12). The existing differences between the derived elasticity values for the 2030 policy scenario and the 2010 base year scenario can be explained by the fact that the IPAT model was calibrated to the base year 2010. The calibration was not revised for 2030 but adjustment factors were introduced in order to align the IPAT model to the Reference Scenario. Hence, the IPAT calibration is slightly different for all years, explaining the differences between the elasticity values derived for the years 2030 and 2010.

The derived modal shift effects between rail and air passenger transport are relatively small for both investigated policies in relation to those between rail and road transport (Fig. 6.11) or air and road transport (Fig. 6.12). This can be explained by the fact that the HSR transport mode is not modelled explicitly. It therefore has to be concluded that a policy scenario attempting to improve the competitiveness of HSR and air transport cannot be investigated with the IPAT model.

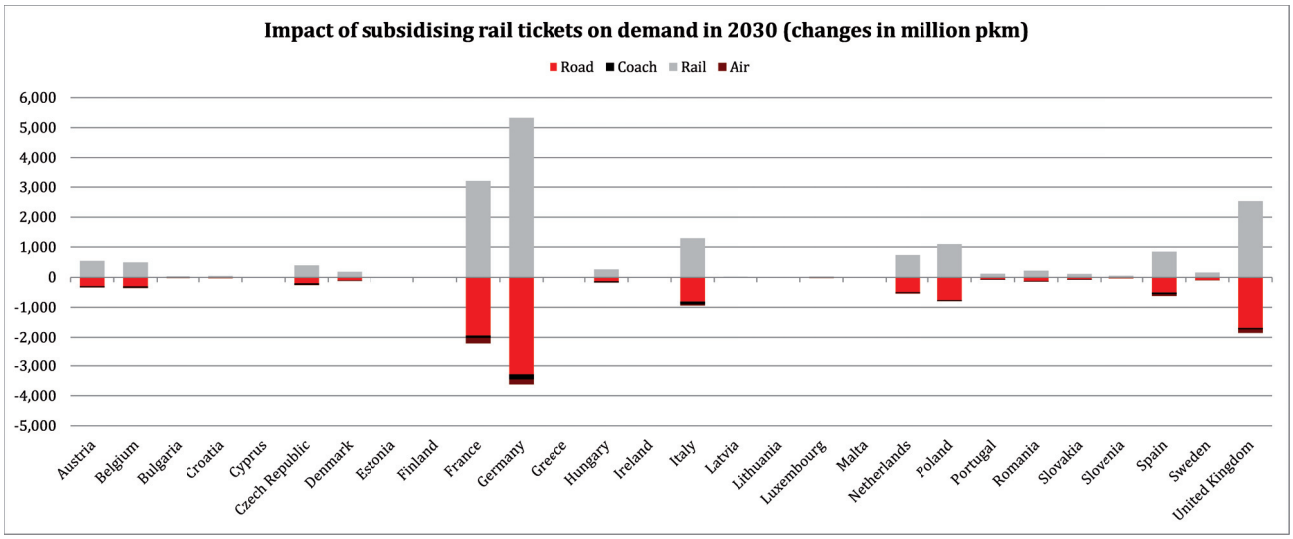

Figure 6.11: Impact of subsidising rail tickets on demand in 2030

Table 6.15: Impact of subsidising rail tickets on demand in 2030 [1]

	road		rail		coach		air	
	pkm	%	pkm	%	pkm	%	pkm	%
EU28	-10,994	-0.19	17,763	2.78	-688	-0.17	-724	-0.08
EU15	-9,589	-0.21	15,427	2.96	-557	-0.18	-599	-0.11
EU13	-1,405	-0.12	2,336	1.98	-132	-0.12	-125	-0.04
AT	-296	-0.35	541	3.45	-29	-0.36	-17	-0.13
BE	-302	-0.24	492	3.10	-39	-0.23	-17	-0.12
BG	-18	-0.04	37	1.14	-3	-0.03	-2	-0.02
HR	-25	-0.08	45	2.16	-3	-0.10	-2	-0.03
CY	0	0.00	0	0.00	0	0.00	-1	-0.01
CZ	-195	-0.22	390	3.14	-51	-0.25	-12	-0.08
DK	-99	-0.17	176	2.13	-10	-0.20	-12	-0.07
EE	-5	-0.04	8	2.10	0	-0.02	0	-0.01
FI	-9	-0.01	14	0.24	0	-0.01	-2	-0.01
FR	-1,939	-0.22	3,224	2.58	-82	-0.21	-179	-0.17
DE	-3,286	-0.35	5,328	3.72	-165	-0.34	-170	-0.18
EL	-12	-0.01	19	1.07	-2	-0.01	-5	-0.01
HU	-132	-0.20	259	2.12	-32	-0.20	-7	-0.10
IE	-7	-0.01	12	0.56	0	-0.01	-2	-0.01
IT	-808	-0.10	1,297	1.94	-93	-0.12	-44	-0.05
LV	-15	-0.08	23	1.86	-1	-0.06	-1	-0.03
LT	-3	-0.01	4	0.76	0	-0.01	0	-0.01
LU	-20	-0.24	31	6.04	-1	-0.19	-1	-0.09
MT	0	-0.01	0	0.00	0	0.00	0	-0.01
NL	-492	-0.30	735	3.14	-19	-0.23	-30	-0.13
PL	-764	-0.19	1,096	3.13	-25	-0.16	-9	-0.05
PT	-65	-0.07	123	1.77	-8	-0.08	-7	-0.02
RO	-124	-0.11	213	1.66	-14	-0.11	-7	-0.04
SK	-63	-0.15	114	2.67	-11	-0.16	-1	-0.05
SI	-32	-0.10	54	2.60	-2	-0.10	0	-0.04
ES	-508	-0.11	850	1.78	-52	-0.12	-61	-0.04
SE	-86	-0.07	149	0.98	-3	-0.05	-11	-0.04
UK	-1,688	-0.22	2,529	3.35	-43	-0.13	-122	-0.10

[1] Difference Scenario-Baseline (in million pkm)

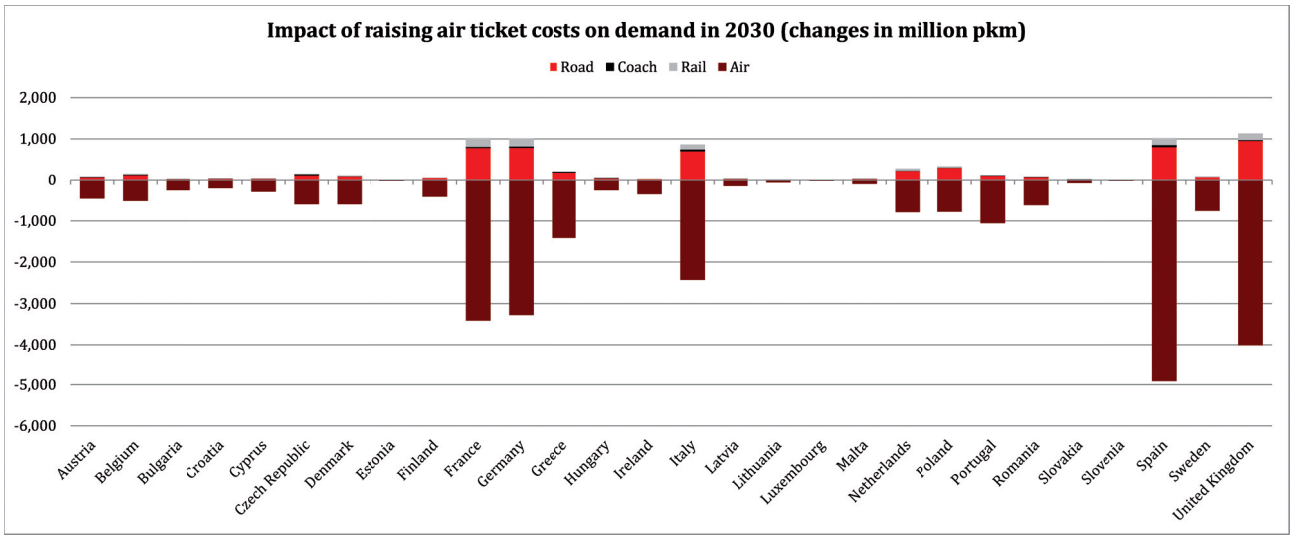

Figure 6.12: Impact of raising air ticket costs on demand in 2030

Table 6.16: Impact of raising air ticket costs on demand in 2030 [1]

	road		rail		coach		air	
	pkm	%	pkm	%	pkm	%	pkm	%
EU28	5,620	0.10	1,142	0.18	293	0.07	-27,960	-3.19
EU15	4,211	0.09	854	0.16	200	0.07	-18,767	-3.31
EU13	1,409	0.12	288	0.24	93	0.08	-9,193	-2.97
AT	66	0.08	14	0.09	6	0.07	-458	-3.52
BE	114	0.09	27	0.17	12	0.07	-516	-3.61
BG	24	0.05	1	0.05	4	0.04	-256	-3.12
HR	34	0.11	3	0.13	3	0.09	-207	-3.59
CY	32	0.37	0	0.07	4	0.29	-295	-1.97
CZ	111	0.13	20	0.16	23	0.11	-599	-3.69
DK	85	0.15	20	0.24	7	0.13	-598	-3.35
EE	9	0.07	1	0.16	0	0.02	-38	-3.03
FI	50	0.07	16	0.27	1	0.02	-415	-2.78
FR	777	0.09	220	0.18	24	0.06	-3,417	-3.29
DE	780	0.08	196	0.14	32	0.07	-3,281	-3.55
EL	177	0.16	3	0.16	20	0.14	-1,416	-3.19
HU	42	0.06	10	0.08	8	0.05	-260	-3.46
IE	29	0.05	3	0.15	1	0.02	-351	-2.10
IT	694	0.09	125	0.19	46	0.06	-2,429	-2.97
LV	30	0.15	3	0.21	1	0.09	-154	-3.25
LT	16	0.05	1	0.16	1	0.02	-70	-2.55
LU	8	0.09	1	0.10	0	0.06	-36	-3.63
MT	29	1.17	0	0.03	2	0.42	-107	-2.12
NL	214	0.13	59	0.25	6	0.07	-789	-3.46
PL	291	0.07	38	0.11	4	0.02	-780	-4.54
PT	100	0.10	10	0.14	9	0.08	-1,058	-3.71
RO	69	0.06	7	0.06	7	0.05	-619	-3.64
SK	16	0.04	2	0.04	2	0.03	-82	-4.10
SI	8	0.03	1	0.04	0	0.01	-32	-4.66
ES	797	0.17	175	0.37	52	0.13	-4,920	-2.92
SE	67	0.06	24	0.16	1	0.03	-760	-2.96
UK	952	0.12	165	0.22	18	0.05	-4,017	-3.15

[1] Difference Scenario-Baseline (in million pkm)

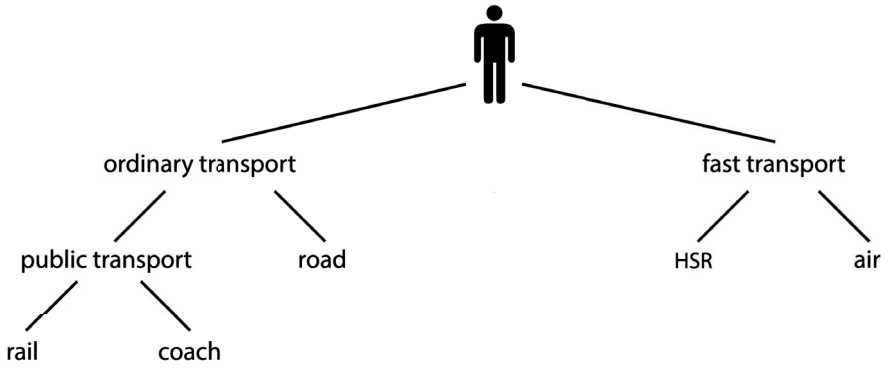

Figure 6.13: Revised nested multinomial logit model including the high-speed rail transport mode

In general, the reactions of the IPAT model are reasonable, as the frequency of trip occurrences decreases with the trip distance in a nonlinear manner. Hence, competition between "regional rail transport" and road transport has a bigger effect on the results than competition between HSR and air transport for long-distance trips. In order to improve the IPAT model, HSR could be introduced as an additional transport mode and the structure of the nested multinomial logit model computing market shares of transport modes could be adapted accordingly (cf. Fig. 3.2, sec. 3.4.5). It is, however, not evident how the structure can be best adjusted. For instance, HSR could be introduced explicitly as a competitor to air transport according to Figure 6.13. The suggested structure is quite simple but it is perhaps inappropriate to limit the competition between HSR and air transport to the distance band 300-1000 km. The further revision of the structure of the NMNL model requires deeper investigation and could be subject of further studies.

6.3.5 Summary

In the previous sections, the findings of a series of checks were discussed that were carried out in order to test and validate the IPAT model. These checks were performed with the Data Stock of the final version of the HIGH-TOOL model (Kiel et al., 2016b), including

- comparing model results with transport statistics provided by the EU Reference Scenario 2013;
- testing and evaluating the model sensitivity; and
- validating the model by specific policy scenarios.

In the first step, the IPAT model is validated for the year 2010 against the Reference Scenario. The obtained results reveal good overall consistency of the IPAT model. For the majority of EU countries, the discrepancy is less than 2% for the indicator "total passenger transport performance". However, the model overestimates the demand by 3.8% for Cyprus, 4.4% for Ireland and it underestimates the demand by 2.8% for Portugal. As the IPAT model is basically calibrated to an average EU level, these discrepancies can largely be explained by the specific structure of the transport markets in the three countries. Cyprus has no railways and the market share of air is almost twice as high in Ireland and Portugal compared to the EU average.

The second series of checks demonstrates reasonable reactions to cost changes of the trip distribution and the mode choice models underlying the IPAT model. If the values of the integrated cost indicator EMC are increased in a percentage range, the computed transport demand decreases slightly, and vice versa. Given the particular relevance of flights for long-distance trips, the largest effect can be observed for air transport, as long-distance trips are substituted by shorter trips with other transport modes if travel costs are increased, and vice versa. This behaviour is reflected by the determined elasticities related to the EMC cost indicator that are summarised in Table 6.11. If cost indicators of a single transport mode are modified, a modal shift of the demand to competing transport modes can be observed. More specifically, the analysis was carried out based on elasticity and cross-elasticity values that were determined for each transport mode in relation to changes in monetary travel costs, time-related travel costs and both cost variables. Overall, the obtained elasticity values summarised in Table 6.12 demonstrate a reasonable behaviour of the mode choice model implemented in the IPAT model that computes the market shares of transport modes in relation to mode-specific travel costs.

The model's reactions to variations in the cost variables, and in particular to variations of EMC values, tend to be relatively small. This can be explained by the comparatively small share of inter-zonal passenger flows between the NUTS-2 regions, given their average diameter of over 125 km. Thus, changes in travellers' destination choices can only be modelled for long-distance trips if the structure of the O/D trip matrix changes.

In the last step, the forecasting capability of the IPAT model is evaluated by analysing three policy scenarios influencing demand in the year 2030. The first one is the benchmark scenario in which no policy is applied. The modelled demand for 2030 is well aligned to the forecasts provided by the Reference Scenario. The differences are larger than for 2010 but still within an acceptable range. The second scenario examines network effects on transport demand by simulating a hypothetical investment in the "Magistrale" rail corridor via the hypernet sub-module. As might expected, the largest effects can be observed for the countries France, Germany and Austria, which are located along the corridor, but the overall changes are relatively small.

The last scenario examines two policies attempting to improve the competitiveness of HSR by subsidising rail trips within the distance range of 300-1000 km or by raising taxes on flights. Basically, the results obtained for these two policies look plausible and can be explained by the cost elasticities that were determined already for 2010. In both cases, modal shift effects between competing transport modes can be observed. The largest effects do not occur between rail and air transport, as might be expected, but between road and rail transport, or between air and road transport. However, the two policies attempting to improve the competitiveness of HSR vis-a-vis air cannot be investigated adequately since HSR transport is not explicitly considered as a transport mode. This is a limitation of the applied mode choice model, and it might be reasonable to introduce HSR as an additional transport mode in order to distinguish between conventional rail and HSR transport.

6.4 General conclusions

This chapter discussed the IPAT model that is published and successfully applied as part of the HIGH-TOOL policy assessment model. The IPAT model is calibrated to the base year 2010. It provides estimates on passenger transport at NUTS-2 level for the European countries. For its application within the HIGH-TOOL model, the methodology of the IPAT model was slightly extended in order to provide insights on changes in European passenger transport demand for the timeline 2010-2050. Compared to the PAT model, which was applied in the ETISplus project, the IPAT model covers all relevant motorised transport modes as well as intra-zonal and inter-zonal trip demand. Modelling the whole transport demand sector with a single transport model enabled the implementation of a more consistent approach with the main focus on:

- improving the methodology of the distribution model;
- improving the methodology of the modal split model; and
- integrating the destination and the modal split model

based on the theory discussed in chapters 3 to 5. In contrast to the rather complex methodology of the integrated trip distribution and mode choice step in the IPAT model, the trip generation step follows a straightforward approach relying on trip rate factors that were derived within ETISplus. The network assignment step was replaced by a more lightweight conversion step.

6.4.1 Methodological revisions introduced by the IPAT model

The two most important methodological revisions introduced by the IPAT model relate to the trip distribution model as a whole and to the logit model. For the trip distribution model a series of improvements had to be realised in order to overcome the necessity to apply the rather questionable Furness method as before in the PAT model. An accessibility-based gravity trip distribution model was first introduced, then the composite deterrence function, and finally the deterrence model establishing the consistency between the deterrence function and the trip length distribution (TLD).[26] For the logit model used in the PAT model for computing market shares of transport modes, a distance-dependent bias was revealed. This issue was solved in the IPAT model by defining the heterogeneity parameter μ of the logit model in relation to the travel distance (6.11).

6.4.2 Evaluation of the IPAT model

The IPAT model was successfully tested and validated by a series of checks demonstrating a very good alignment to transport statistics and forecasts provided by the EU Reference Scenario 2013. In addition, the model's reactions to modifications of several cost variables are reasonable. If the overall travel costs in terms of EMC values are changed, a response of the demand can be observed that is different for each transport mode. If travel costs are modified for a single transport mode, modal shift effects can be observed as well as changes of the O/D trip matrix (i. e. induced transport demand). It can therefore be concluded that the integration of the trip distribution and of the modal split model is successfully implemented and that the revised IPAT model operates correctly.

[26] The TLD is compiled based on observed data and therefore used as an empirical basis to calibrate the deterrence function.

However, it has to be emphasised that the sensitivity of the IPAT model to modifications in the cost variables is limited due to the modelling of the passenger demand at the level of NUTS-2 regions. Given their average diameter of over 125 km, the travel zones are far too large for capturing effects of travel cost changes on short-distance and regional passenger transport relating, to a large extent, to intra-zonal trips. Accordingly, the IPAT model is much more sensitive to changes in the inter-zonal trip demand, representing a relatively small share in the overall transport demand, than to changes in the intra-zonal trip demand, which represents the majority of the overall demand.

For the reasons stated above, the transport demand forecasts provided by the IPAT model at NUTS-2 level should be generally interpreted at a more strategic level, and in particular not at a regional level. For instance, short-distance passenger transport between the two German cities of Ulm and Neu Ulm that are located on both sides of the river Danube are not reflected in a satisfactory way in the trip matrices because the representative travel distance between the two corresponding NUTS-2 regions is over 150 km.[27] For the same reasons, computing a network assignment at NUTS-2 level would make no sense. However, modelling traffic loads at link level represents an excellent means to assess network effects of transport policies and to validate the results of a transport model in more detail. For instance, the PAT model computing the ETISplus trip matrices was validated and calibrated based on assignment results (see sec. 2.3.4).

6.4.3 Key findings for further revising the model

In brief explanation, the IPAT model operates at NUTS-2 level. As a consequence, the majority of the passenger transport demand (short-distance and regional trips) is modelled by intra-zonal trips. This limits the overall sensitivity of the model to cost changes and to transport policies. In order to augment the scope of the model to regional and short-distance trips, a systematic reduction of the size of the travel zones is required. This reduces the share of intra-zonal trips. However, if we downsized the travel zones to an edge length of 10 km, we would end up with 44 thousand zones to model the area of all 28 member states of the European Union.[28] It is quite evident that this large increase in the number of travel zones inevitably raises complexity issues concerning the number of O/D relations and the computability of the trip matrix. Moreover, travel zones with an average diameter of 10 km are still not small enough for an accurate modelling of trips below 20 km.[29]

Another reason for downsizing the travel zones relates to the computation of traffic load estimates at the level of network links by the network assignment model, for instance, in order to investigate the impact of transport policies on congestion. This requires detailed information on the spatial distribution of regional trips below 100 km that is not given if O/D trips are modelled at NUTS-2 level. The necessity to decrease the size of the travel zones while simultaneously increasing their number finally led to the development of the HIPAT model. The HIPAT model follows a hierarchical approach and is discussed in Part III of this thesis.

[27] Ulm and Neu Ulm are located in the NUTS-2 regions DE23 and DE31. The corresponding centroids are the city centres of Tübingen and Augsburg.

[28] The area of the EU28 member countries is 4.4 million km^2 (EC, 2015, p. 9).

[29] Trip estimates between neighbouring travel zones depend to a very large extent on the location of the two centroids. The diameter of the travel zones should therefore be at least half the length of the shortest trip to be modelled.

Part III

Hierarchical transport modelling

Chapter 7

Spatial disaggregation of travel zones

The most important decision in the conception of a transport model is the determination of the travel zones. As a general rule, the travel zones chosen should be homogeneous in size and at least one order of magnitude smaller than the required spatial resolution of the computed output indicators. If this is the case, the required output indicators can be aggregated from more disaggregated indicators, which reduces the error terms. For instance, the PAT model operating at NUTS-3 level is built on about 1,600 travel zones with an average diameter of over 50 km (cf. sec. 2.3.4). The trip length between two neighbouring travel zones can therefore vary between a few and one hundred kilometres, and it is quite evident that the NUTS-3 zones have to be further disaggregated in order to enable a more accurate modelling of regional trips. However, the spatial disaggregation of the travel zones immediately raises the question of the availability of socio-economic and demographic indicators, since these are, in general, not given below NUTS-3 level. Moreover, halving the diameter of the travel zones increases their number by a factor of 4 and the complexity of the O/D trip matrix by a factor of 16.[1]

In this chapter, a method is discussed for defining homogeneous travel zones at a more disaggregated level than NUTS-3. This requires a deeper understanding of the concept behind the travel zones and possible restrictions which are provided in section 7.1. In order to solve the non-availability of regional indicators for any travel zone at any level, and in order to define homogeneous travel zones at any level, so-called downscaling procedures must be implemented, based on georeferenced data and geographic information systems (GIS). Section 7.2 therefore provides a brief introduction to GIS and describes the most relevant datasets that are used in this thesis. Section 7.3 discusses the applied downscaling procedures.

[1] This issue is tackled by a hierarchical modelling approach that is discussed in chapter 8.

7.1 Determining the travel zones

This chapter deals with the determination of optimal travel zones, taking into account several homogeneity criteria that are discussed in section 7.1.1. Section 7.1.2 introduces the modifiable areal unit problem (MAUP), identifying a source of error in transport models relying on aggregated travel zones and region-based indicators. The problem cannot be avoided but can only be reduced, e. g. by defining a consistent zoning system that is described in section 7.1.3. Section 7.1.4 discusses two examples illustrating the advantages of disaggregated travel zones for the modelling of regional trips. The last section summarises the most important findings of this chapter.

7.1.1 Requirements and restrictions

Travel zones and their respective centroids are one of the key elements of a transport model. For this reason, the definition of travel zones should be carried out under consideration of several criteria. The travel zones should be consistent in size and weight (in terms of, for example, population and workplaces). Moreover, only one city should be located in each travel zone, as a single centroid cannot represent two or more cities with sufficient accuracy. Due to the low availability of regional indicators, the criterion of data availability frequently restricts the definition of travel zones. Another restriction is the complexity of a transport model that increases quadratically with the number of travel zones. For these reasons, the definition of the travel zones is often sub-optimal.

Recent European transport models operating at NUTS-3 level, for instance, have to face the fact that these regions are inhomogeneous in size, shape and weight. Frequently, more than one city is located in those regions. The application of inhomogeneous travel zones, however, can have certain adverse effects on the computation of transport demand indicators, like the O/D trip matrix, which is computed by the trip distribution model. In order to improve the accuracy of transport models, travel zones should be defined more in line with theoretical homogeneity criteria (cf., for example, Guo et al., 2000), including the following:

- **Intra-homogeneity of land use type**: travel zones should be characterised by a dominant land use type. In particular, they should not cover a mixture of several land use type clusters referring to different cities or different industrial areas. If we separated settlement areas and industrial areas, for instance, we could reduce the issue of computing intra-zonal commuting trips. For trip purposes like visiting friends, however, the issue of intra-zonal trips can hardly be improved, given that these trips can also start and end in the same settlement area.

- **Compactness**: the criterion of compactness describes the concurrence between the actual shape of a travel zone and its convex hull or, to be more precise, its bounding circle. Moreover, it forbids the definition of a travel zone that is completely surrounded by another travel zone.[2] The definition of compact travel zones is particularly important with regard to the computation of intra-zonal trips, as intra-zonal trip lengths are estimated based on the radii of the travel zones (i. e. it is assumed that all travel zones have circular shapes and their centroids are located in the centres).

[2] According to the NUTS-2006 zoning system, the region Baden-Baden (DE121) is completely surrounded by the region Rastatt (DE124). This has an adverse effect on the computation of the trip estimates between these two travel zones, since inter-zonal trips from Rastatt to Baden-Baden can be shorter than intra-zonal trips within Rastatt.

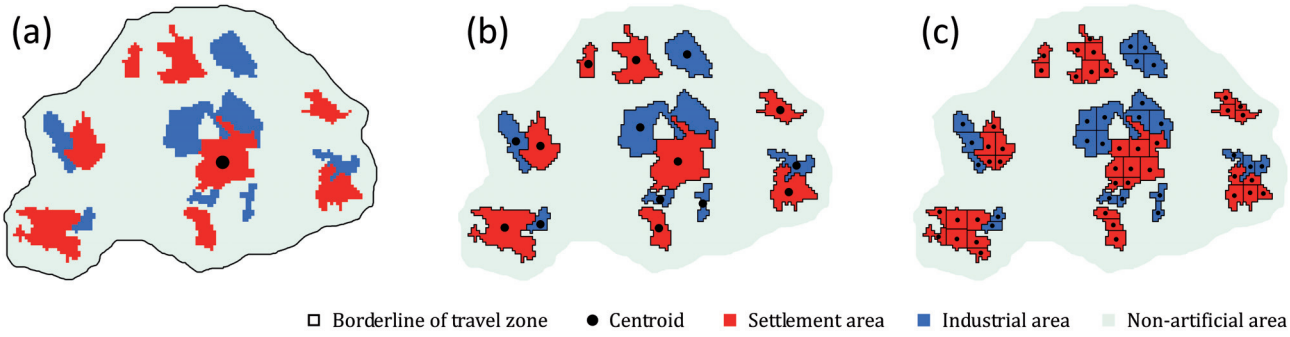

Figure 7.1: Generation of homogeneous travel zones

- **Inter-homogeneity of shape and size**: travel zones should be defined homogeneously with regard to their shape and size, as the applied deterrence model or, to be more precise, the scaling function $\rho_d(x)$, is sensitive to the average diameter of the travel zones (cf. sec. 5.1). The scaling function was introduced in this thesis in order to solve an existing inconsistency between the TLD and the deterrence function.
- **Level of disaggregation**: in principle, the travel zones chosen should be as small as possible in order to minimise the problems of intra-zonal trips and trips between neighbouring regions. This is particularly relevant with regard to the computation of the travel impedances and the trip matrix as well as the network assignment.

Figure 7.1 illustrates the approach of generating homogeneous travel zones that are more in line with the discussed criteria by successively dividing very heterogeneous zones into smaller units. The first image outlines one aggregated travel zone encompassing several settlement-, industrial- and non-artificial areas. Thus, the representativeness of the centroid set to the centre of the largest settlement area can be assumed to be rather weak. The second image shows the disaggregated subregions and their centroids, where each region refers to precisely one settlement- or industrial area. The third image outlines further disaggregated subregions that fulfil the criteria of compactness and of equal size more coherently.

It should be emphasised that the spatial disaggregation of large travel zones into homogeneous units reduces the issue of computing intra-zonal trips, given that settlement areas and industrial areas always refer to different travel zones. Areas referring to non-artificial land use types, like agricultural areas, forests and semi-natural areas, are often not relevant for transport modelling. These areas should not be added to the previously defined travel zones as they influence the computation of density-based indicators like inhabitants per square kilometre.

7.1.2 The modifiable areal unit problem (MAUP)

The modifiable areal unit problem (MAUP) identifies a source of statistical bias that can explain discrepancies in results when analysing data at different aggregation levels (Openshaw and Taylor, 1979). The problem is described in detail by Openshaw (1984). In transport modelling, MAUP has to be investigated with regard to the definitions of the travel zones and the centroids.

If travel zones are defined at an aggregated level as in Figure 7.1a, for instance, the average population density for these travel zones is rather low, due to the large share of non-artificial areas (i. e. unsettled

areas). Hence, aggregated travel zones are likely to be classified as rural areas. In contrast, if travel zones are defined at a more disaggregated level as in Figure 7.1b, it is possible to distinguish between densely populated and sparsely populated settlement areas (i. e. cities and rural areas). Given that residents of cities tend to travel less than residents of rural regions (cf. Fig. 2.1, sec. 2.1.2), the transport demands computed by a model relying on aggregated travel zones and by a model relying on disaggregated travel zones are likely to be different. This is an instance of MAUP.

The design of disaggregated travel zones according to Figure 7.1b is still likely to be sub-optimal with regard to the uneven distribution of wealth among the population and the strong correlation between the indicators of average income per capita and generated trip demand per capita (cf. Fig. 2.1, sec. 2.1.2). Hence, the definition of very disaggregated travel zones according to the districts outlined by Figure 7.1c provides the best basis for the computation of density-based indicators and for transport modelling, as long as these districts show a high intra-homogeneity and a high inter-heterogeneity (e. g. exclusive and poor residential areas should not be assigned to the same district).

Besides the computation of density-based indicators, MAUP must also be taken into account in the determination of representative centroids for the travel zones. These centroids are commonly determined based on the respective geographical centre of a zone or the centre of the largest city in a zone. The precise locations of the centroids are particularly important for the computation of travel distances and trip demands between neighbouring travel zones. In Figure 7.1a, for example, the conclusion must be drawn that the large settlement area located in the lower left corner is not accurately represented by the chosen centroid. In consequence, for this specific settlement area, the travel distances that are computed based on the centroids of the aggregated travel zone and of the other zones (that are not shown) are rather imprecise.

It should be stressed at this point that changes in the geographical locations of the centroids have major impacts on the computation of travel distances and therefore on the O/D trip matrix, given a non-linear correlation between the travel distances and the frequency of trip occurrences (cf. sec. 4.3). The smaller the travel zones, the better the representativeness of the respective centroids. Hence, the definition of very disaggregated travel zones according to the districts outlined by Figure 7.1c provides the best basis for the computation of the travel distances and the trip matrix in this example.

It should be noted that a well-known occurrence of MAUP refers to "gerrymandering" which is the deliberate manipulation of electoral districts in favour of a particular party.[3] In transport modelling MAUP refers to the determination of the travel zones and the respective centroids, given that there is no common definition. A modeller therefore has many options to influence the output of the transport model. For instance, if he decides to set the centroid to the settlement area located in the lower left corner (cf. Fig. 7.1a), the transport demands between the aggregated travel zone and the neighbouring zones located near the lower left corner are overestimated, while the transport demands to those zones located near the upper right corner are underestimated. The decision of the modeller might be correct, for instance, if the only railway station is located in this settlement area. However, the definition of the travel zones and the centroids underlying a transport model should be carried out in a transparent

[3]Gerrymandering was first identified in the Massachusetts state election of 1812 (Martis, 2008), but incidents are reported even today. For instance, the parliamentary elections in Hungary on 6. April 2014 were criticised "for lacking transparency, independence and consultation, and allegations of gerrymandering were widespread" (OSCE, 2014, p. 7).

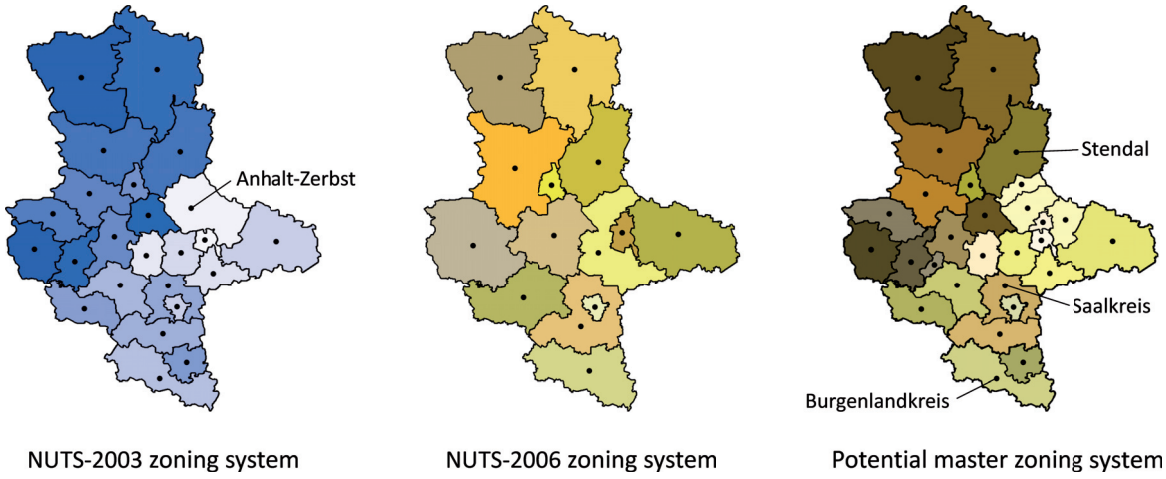

Figure 7.2: Changes in the NUTS-3 regions for Sachsen-Anhalt

way. Moreover, the definition of the travel zones and the centroids has to be taken into account when analysing the computed model output.

For the examples discussed it can be concluded that MAUP should be taken into account when defining the travel zones and the centroids. Hence, the step from highly aggregated travel zones (Fig. 7.1a) towards disaggregated districts (Fig. 7.1c), reducing the issues caused by MAUP, is a promising step forward in European transport modelling.

7.1.3 Requirements for a flexible zoning system

An important task of European transport models is the computation of transport demand indicators such as the O/D trip matrices that correspond to the administrative units defined by the NUTS classification maintained by Eurostat, the statistical bureau of the European Commission. Although this classification changes frequently, travel zones are commonly defined based on these NUTS regions. This is an obstacle to transferring a transport model from one base year to another, as the NUTS classification (i. e. the zoning system) might have changed. For example, the requirement to consider two different zoning systems, the NUTS-2003 and the NUTS-2006 zoning systems, was a specific challenge in the development of the PAT model during the ETISplus project (cf. sec. 2.3.4). Besides the additional work of compiling two different input databases providing regional indicators at NUTS-3 level for both zoning systems, the PAT model, which was calibrated for the first base year 2005, could not be instantly applied for the second base year 2010, as the calibration factors had to be updated. The transferability of the PAT model between these two zoning systems would have been possible if MAUP had been avoided, e. g. by defining a master zoning system underlying the NUTS-2003 and the NUTS-2006 zoning systems.

The concept of a master zoning system relying on disaggregated travel zones that remain constant over time is illustrated by Figure 7.2 showing the changes in the administrative units in Sachsen-Anhalt, a federal state in Germany, due to a reform of the administrative units ("Kreisgebietsreform 2007"). The left-hand image refers to the NUTS-2003 zoning system, where Sachsen-Anhalt comprises 24 NUTS-3 regions, and the second picture to the NUTS-2006 zoning system, where Sachsen-Anhalt

comprises 14 regions. As a result of the Kreisgebietsreform, for instance, the region Anhalt-Zerbst was dissolved and assigned to four different regions. For such reasons, the NUTS-2003 zoning system cannot be directly converted into the NUTS-2006 zoning system. This issue can be solved if a master zoning system is applied that can be converted into both the NUTS-2003 and the NUTS-2006 zoning systems. The right-hand image outlines a potential master zoning system comprising 28 regions for Sachsen-Anhalt in which the region Anhalt-Zerbst (NUTS-2003) is represented by four regions. Based on this master zoning system, all regions of the NUTS-2003 and NUTS-2006 zoning systems can be generated without loss.

It should be noted that the outlined master zoning system should be further improved. First of all, larger regions like Stendal should be subdivided into smaller units in order to homogenise the size of the travel zones modelling Sachsen-Anhalt. In addition, regions like Saalkreis and Burgenlandkreis, which almost completely surround another region, should be subdivided into several compact travel zones.

Changes of county boundaries like in Sachsen-Anhalt are not uncommon. They occur due to political changes and special considerations in other countries as well. In principle, all hierarchical levels are affected, including also changes in local administrative units defined at LAU-2 level.[4] Besides these changes, a fundamental issue of the NUTS classification is related to the fact that LAU-2 regions are not defined for many European cities, e. g. the German capital Berlin (DE3) is actually only defined at NUTS-1 level. A reliable master zoning system that remains stable over time must therefore be defined on the basis of artificial travel zones that are consistent in terms of size and weight.

The discussion on the advantages of a potential master zoning system relying on the very disaggregated LAU-2 regions as well as on artificial travel zones instantly raises questions concerning the availability of regional indicators. The availability of the most important indicators for passenger transport modelling is ensured through the ETIS database, but only at NUTS-3 level. At LAU-2 level, a European dataset worth mentioning is census data which is provided by Eurostat (Eurostat, 2015a). To overcome the non-availability of regional indicators at LAU-2 level, so-called downscaling methods (cf. sec. 7.3) have to be applied in order to disaggregate the indicators provided by the ETIS database to this level.

One could certainly say that the spatial disaggregation of travel zones and the downscaling of regional indicators should be seen as a special case of MAUP (Gallego, 2007). Within this thesis, MAUP can be minimised as the applied downscaling methods rely on very detailed datasets. For instance, the spatial resolution of the European dataset on the population density is one square kilometre (Bloch et al., 2012) and the spatial resolution of the Corine land cover (CLC) database providing information on land use types is only one hectare (EEA, 2007).

7.1.4 Revealing the advantages of smaller travel zones

Having discussed the general requirements and restrictions for the definition of optimal travel zones with regard to the application of different criteria, and having investigated the implications of MAUP on the processing of spatial indicators, two examples are presented in the following. These examples

[4] NUTS-2003: 112,108 (Eurostat, 2004), NUTS-2006: 113,098 (Eurostat, 2007), NUTS-2010: 111,195 (Eurostat, 2011), NUTS-2013: 109,499 (Eurostat, 2015b); all numbers of LAU-2 units refer to the geographical scope of the EU25.

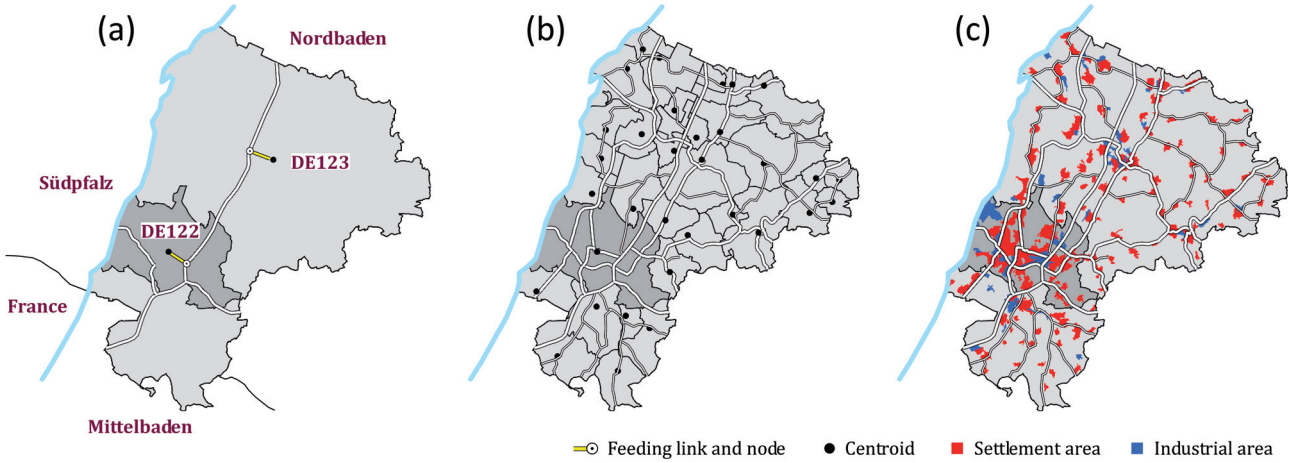

Figure 7.3: Definition of travel zones at different spatial levels

emphasise the necessity of applying smaller and less heterogeneous travel zones for modelling transport demand more accurately. The first example is related to the modelling of inter-zonal trips between two NUTS-3 regions, and outlines the potential for improvement if the travel zones are disaggregated. In the second example, a trip distribution model is applied at different spatial levels and the discrepancies between the computed trip matrices are quantified.

The first example outlined by Figure 7.3 shows the two neighbouring NUTS-3 regions of Karlsruhe Stadtkreis (DE122) and Karlsruhe Landkreis (DE123) at different spatial levels. The two regions are located in the south-western part of Germany near the river Rhine and close to France. The regions are bordered to the north by the region of Nordbaden, to the south by the region of Mittelbaden and to the west by the region of Südpfalz. The left-hand image shows both NUTS-3 regions, the corresponding centroids, feeding links and feeding nodes as well as the two major road corridors. The image in the centre shows both regions at LAU-2 level and the regional road network complementing the two major road corridors. The image on the right complements land use data, facilitating the identification of the location of the settlements and the industrial areas in both NUTS-3 regions.

European transport models operating at NUTS-3 level are likely to underestimate the trip demand from Karlsruhe Landkreis to Mittelbaden (Fig. 7.3a). The feeding node representing Karlsruhe Landkreis is located in the north and is therefore not representative of the population living in the south of this region. An underestimation of the trip demand from Karlsruhe Stadtkreis to Südpfalz can also be expected, given that the centroid of Karlsruhe Stadtkreis is connected to the motorway situated east of the city. In principle, the definition of the travel zone DE123 is in conflict with the criterion of compactness, given that it actually consists of two different travel zones - one located in the north of Karlsruhe Stadtkreis and one located in the south.

Besides the averse effects on the modelling of regional trips, the application of travel zones that are too large also affects the modelling of traffic loads at link level by the network assignment model. In the drafted example (Fig. 7.3a), trips between Karlsruhe Stadtkreis and Karlsruhe Landkreis are only assigned to the road segments between the two centroids. Alternative routings are not considered. If we increase the spatial resolution and define the travel zones in accordance with the LAU-2 regions

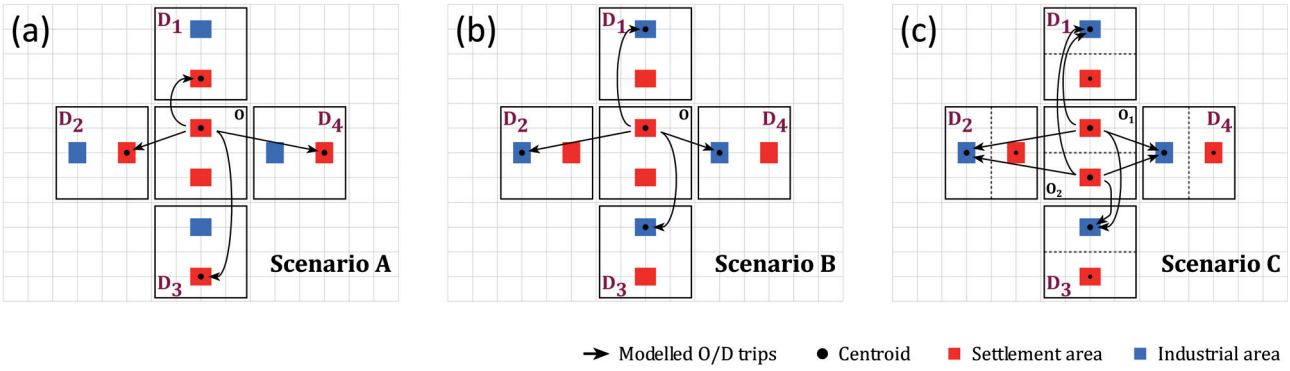

Figure 7.4: Computation of trip estimates between neighbouring travel zones

(Fig. 7.3b), Karlsruhe Landkreis is subdivided into 32 units but Karlsruhe Stadtkreis is not subdivided. This is an issue with regard to the criterion of inter-homogeneity of shape and size. It could be solved by defining artificial travel zones based on land use data (Fig. 7.3c) for each LAU-2 region as well as for Karlsruhe Stadtkreis. This would facilitate a more accurate modelling of regional trips.

The second example investigates the adverse effects of MAUP on the computation of commuting trips by varying the location of the centroids of the travel zones and by virtually disaggregating the travel zones into two units. The example is outlined by Figure 7.4 and distinguishes between three scenarios. We consider the case that people living in "O" want to go to the industrial areas of D_1 to D_4. Scenarios A and B differ in the definitions of the centroids. In scenario C, each of the five travel zones have been split into two.

For the computation of the trip distribution, an accessibility-based gravity trip distribution model is applied which is defined as follows:

$$T(O \to D_j) = O \frac{D_j \omega_j^{-1.5}}{\sum_n D_n \omega_n^{-1.5}} \quad \text{with} \quad j \in \{1, 2, 3, 4\}$$

where O refers to the generated trip demand in the origin travel zone and D_j refers to the number of workplaces provided by destination j. For the purpose of simplification, it is assumed that the two settlements located in the origin are of equal weight and the four industrial areas located in the four destinations are also of equal weight. Furthermore, linear travel costs (ω) are assumed. The basic model parameters and the outcomes of the three scenarios are summarised in Table 7.1.[5]

It is quite evident that the third scenario (Fig. 7.4c) provides the highest accuracy, since each travel zone and, accordingly, each centroid refers to exactly one settlement or one industrial area. This ensures a high representativeness of the centroids. It is also clear that the first scenario (Fig. 7.4a) provides the lowest accuracy, since each centroid represents two areas. Moreover, the industrial areas located in the destinations are represented rather weakly by the centroids that are set to the settlements. Hence, the first scenario shows large discrepancies from the results computed by the third scenario, e. g. 51 % compared to 17 % ($O \to D_1$) and 10 % compared to 38 % ($O \to D_3$). These

[5] For scenario C, which distinguishes between two origins, this notation needs to be expanded accordingly, i. e. $O_k \to D_j$ with $k \in \{1, 2\}$ instead of $O \to D_j$.

Table 7.1: Comparison of trip estimates for three scenarios

	Trip estimates			Travel costs		
O/D relation	Scen. A	Scen. B	Scen. C [1]	Scen. A	Scen. B	Scen. C [2]
$O \rightarrow D_1$	51 %	24 %	17 %	2.0	4.0	{4.0; 6.0}
$O \rightarrow D_2$	26 %	17 %	15 %	3.2	5.1	{5.1; 5.1}
$O \rightarrow D_3$	10 %	24 %	38 %	6.0	4.0	{4.0; 2.0}
$O \rightarrow D_4$	13 %	25 %	30 %	5.1	3.2	{3.2; 3.2}

[1] The values refer to the computed trip estimates for the two disaggregated relations.
[2] Different travel costs are considered for the two settlement areas.

discrepancies can be explained by the low representativeness of the centroids in the first scenario. Accordingly, the derived travel costs are inaccurate and differ strongly from the travel costs derived for the third scenario, e. g. 2.0 to {4.0; 6.0} ($O \rightarrow D_1$), and 6.0 to {4.0; 2.0} ($O \rightarrow D_3$). The results of the second scenario (Fig. 7.4b) are intermediate. The centroids of the destinations are also set to the industrial areas but the two different settlement areas located in the origin travel zone are represented by the same centroid.

The two examples discussed clearly reveal the advantages of smaller travel zones for the computation of trips between neighbouring regions. For instance, the trips between the two NUTS-3 regions of Karlsruhe Stadtkreis and Karlsruhe Landkreis could be computed more accurately if both regions were subdivided into smaller units and homogeneous travel zones were defined based on spatial data indicating the land use types. The need to disaggregate travel zones is further underlined by the second example, demonstrating the magnitude of the adverse effects of MAUP on the modelled output indicators when changing the travel zones and their centroids. These changes have an impact on the computation of the travel costs and accordingly on the computation of the trip distribution.[6] For the first O/D relation, the results provided by the first and the third scenarios differ by a factor of three (cf. Tab. 7.1). For the third O/D relation the discrepancy is even larger. The findings of the two examples are in line with the findings of other researchers who investigated the correlation between the size of the travel zones and the accuracy of the transport model (cf., for example, Schoch (2004); Ding (1998); Wildermuth et al. (1972)). Whenever possible, smaller travel zones should be chosen.

7.1.5 Summary

The consideration of homogeneous and disaggregated travel zones is an important prerequisite for transport modelling. It facilitates a more accurate computation of the trip estimates between travel zones by the trip distribution model, due to their dependency on the travel costs and therefore on the geographic location of the centroids representing the travel zones. The smaller, more compact and more homogeneous the travel zones, the higher the representativeness of the centroids.
Although the calculation in the second example only refers to five travel zones, the basic findings are transferable to a full-featured transport model encompassing several hundred travel zones. In order to reduce the adverse effects of a single, improperly chosen centroid on the accuracy of the modelled O/D trip matrix, large travel zones should be subdivided into smaller units. In general, this improves

[6] Adverse effects with respect to the computation of density-based indicators are not considered.

the representativeness of the centroids. Taking into account the discussion on the changes of regional boundaries in Sachsen-Anhalt due to the Kreisgebietsreform 2007 and frequent updates of the European NUTS classification, the subdivided travel zones should be defined according to a master zoning system. The master system should remain constant over time in order to avoid the necessity of converting spatially aggregated indicators between different zoning systems that is a potential source of error stemming from the modifiable areal unit problem (MAUP).

It is important to bear in mind that, in addition to the trip distribution model, the assignment model also depends on the geographic location of the centroids, given the fact that each trip originates and ends in a centroid. For this reason, modelled network loads are often significantly overestimated for the feeding links and their adjacent links. In contrast, other existing routes can be almost empty. Thus, the interpretation of the modelled network assignment can be very challenging. This issue could be significantly improved by considering less aggregated travel zones where each centroid only represents a small area. In consequence, regions referring to large cities of several hundred thousand inhabitants and regions referring to a patchwork of different settlement-, industrial- and non-artificial areas should be disaggregated.

This concept is already partly implemented in the European network-based transport model VACLAV that considers artificially subdivided NUTS-3 regions for the computation of the network assignment (Schoch, 2004). However, the model uses a given NUTS-3 trip matrix as input limiting the advantages of disaggregated travel zones for the computation of the network assignment. For this reason, disaggregated travel zones should be applied consistently for the whole process of transport modelling including trip generation, trip distribution, modal split, and network assignment.

7.2 Georeferenced data

Spatially disaggregated data, like socio economic and demographic indicators, is frequently not available at European scale, representing an outstanding challenge for transport modelling. During the ETISplus project, for instance, regional indicators were collected from Eurostat as well as from national and regional statistical offices. The collected indicators were harmonised, data gaps were closed and a consistent database covering all European NUTS-3 regions was compiled (cf. sec. 2.3). In order to ensure the availability of regional indicators at LAU-2 level, this database is disaggregated within this thesis using geographic information systems (GIS) and georeferenced data, e. g. the Corine land cover database which provides information on land use types in Europe at a spatial resolution of one hectare, and the GEOSTAT grid dataset which provides information on the population in Europe at a spatial resolution of one square kilometre. In addition, a spatial algorithm using GIS is developed for the generation of artificial travel zones at LAU-2 level.

The use of geographic information systems represents an outstanding opportunity to significantly improve European transport modelling, as GIS datasets providing georeferenced data for a variety of different topics are available at almost any spatial level. However, even the GIS datasets are not fully compatible and cannot always be easily combined. The reason is that GIS datasets come in various coordinate reference systems, and in the background looms the basic problem of generating planar maps due to the curvature of the Earth's surface. The accuracy of any map is only ensured for a limited geographical scope.[7] In order to avoid errors when processing GIS datasets it is therefore necessary to understand the basic mathematical theory behind data and the limitations of GIS.

The chapter is structured as follows. The first two sections provide a brief introduction to cartography and geographic information systems. In addition, the process of creating maps is discussed, paying specific attention to the multitude of existing coordinate reference systems underlying these maps. Section 7.2.3 discusses the differences between the two most common GIS data types, vector- and raster data, and section 7.2.4 describes the application of geographic information systems in transport modelling. The last section then briefly introduces the most relevant GIS datasets used in this thesis.

7.2.1 Geographic information systems and cartography

Cartography is the science dedicated to the production of two-dimensional maps of our three-dimensional terrestrial globe. For the production of these maps GIS software can be used. Based on the fact that undistorted mappings of the terrestrial globe onto a flat surface do not exist, all maps are distorted in some way. Depending on the purpose and the geographic focus of a map, different projection systems and coordinate reference systems (CRS) can be applied.[8]

However, GIS should not only be understood as a simple means to compile cartographic maps but rather as a comprehensive tool allowing the visualisation of georeferenced data as well as the implementation of spatial analyses. It should be noted that the term GIS is not consistently defined in literature with regard to the specific components which have to be part of the system. According to the National Geographic Society:

[7] This is important for the computation of travel distances.

[8] For instance, for the production of nautical charts angle-preserving (isogonic) projection systems are frequently preferred, while for the production of maps showing land-use patterns area-preserving projections and for road maps length-preserving (isometric) projections are used.

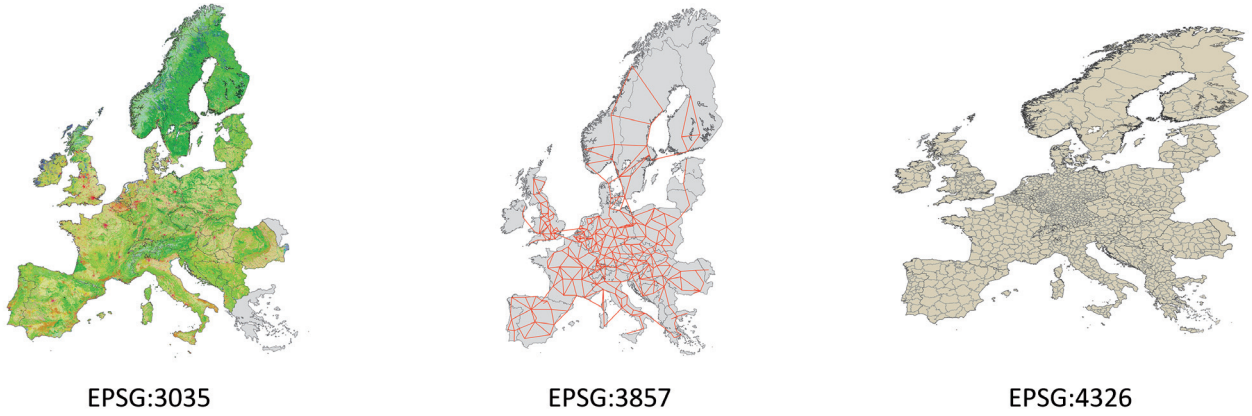

Figure 7.5: Representation of Europe in different coordinate reference systems

> "A geographic information system is a computer system for capturing, storing, checking, and displaying data related to positions on Earth's surface. GIS systems can show many different kinds of data on one map. This enables people to more easily see, analyse, and understand patterns and relationships."[9]

A central property of a geographic information system is the possibility to combine several maps or, to be more precise, several layers, providing information on different space-related indicators. In transport modelling, for instance, three types of layers are frequently combined that provide information on the boundaries of the travel zones, on the networks and on land use. If these layers rely on different CRS, a default coordinate system has to be defined and the GIS software tries to adjusts these layers as well as possible in order to link and congruently display those data elements which refer to the same position on the Earth's surface. This important property of geographic information systems provides the basis for carrying out sophisticated spatial analyses (cf., for example, Obe and Hsu (2015); Anselin and Rey (2010)). For instance, the potential for improving parcel transport in urban areas could be investigated by combining several GIS layers that provide information on the urban road network and on daily congestion patterns as well as on the distribution of parcel shops including the expected volumes of parcels to be delivered per day.

It should be emphasised at this point that a lossless conversion of data between different coordinate reference systems is often not possible. This is of particular importance in European transport modelling given the large extent of the respective maps. Furthermore, additional information from analogue or digital maps is often used to improve the data basis of a transport model. While free and comprehensive GIS layers given in one consistent CRS are often not available, analogue or digital maps for daily use (e. g. tourists maps) may refer to various CRSs or the CRS is unknown. Another potential source of error is the measurement of distances if they are computed based on two-dimensional map coordinates rather than three-dimensional global coordinates.

The discrepancies between different CRSs that were frequently used in this thesis are illustrated by Figure 7.5 which shows Europe in three projection systems. The first map on the left showing land use

[9] URL: `http://education.nationalgeographic.com/encyclopedia/geographic-information-system-gis/` (Retrieved November 14, 2014).

data is given in an area-preserving projection (EPSG:3035) that is a standard system used by the European Commission. The second map outlining the rail hypernet[10] is given the WGS-84 Pseudo-Mercator projection (EPSG:3857) that is a applied in many Web mapping and visualisation applications, including OpenStreetMap and Google Maps. The third map drafting the boundaries of the European NUTS-3 regions is given in the World Geodetic System WGS-84 (EPSG:4326) that is used, for instance, by GPS satellite navigations systems.

7.2.2 Coordinate reference systems

Taking into account the discrepancies between the different coordinate reference systems outlined in Figure 7.5, it is important to understand the basic theory behind a CRS, as well as the basic advantages and disadvantages, in order to avoid errors when processing GIS data. In principle, the projection of the Earth's surface onto a two-dimensional map by applying a CRS can be carried out according to the following steps:

1. recording, measuring and georeferencing of locations on the Earth's surface;
2. mapping the recorded locations to the surface of a reference body; and
3. projecting the reference surface onto a planar map.

None of these three steps can be performed without loss. Measurement uncertainties occur when recording the locations on the Earth's surface in the first step, mapping errors occur due to the consideration of a simplified reference body in the second step and projection errors occur due to the projection of the three-dimensional reference surface onto a planar map in the third step. For these reasons, GIS layers given in different CRS cannot be merged without loss.

Nowadays satellite systems are used to record locations on the Earth's surface, e. g. the global positioning system (GPS) and the global navigation satellite system (GNSS) Galileo. Recorded locations are georeferenced in the so-called ECEF (Earth-Centered, Earth-Fixed) coordinate system referring to a Cartesian coordinate system originating from the centre of mass of the Earth (cf. Fig. 7.6a). In step two and three, the ECEF coordinates are then converted into map coordinates under application of a chosen coordinate reference system (CRS).[11] The system defines the reference body that is applied for describing the surface of our terrestrial globe with a simplified mathematical model, as well as the projection system that is applied for projecting the three-dimensional surface of the reference body onto a planar map. In principle, there are three competing mathematical models for describing the reference body which are in ascending complexity: sphere, ellipsoid, and geoid. There are also three basic systems for projecting the surface of the chosen reference body onto a two-dimensional map. These systems can be distinguished based on the applied projection surface; the main categories are cylindrical-, conical- and azimuthal projections (cf. Fig. 7.6b to 7.6d).

For the cylindrical projection (Fig. 7.6b), the sphere is rolled up into a cylindrical surface and projected onto it. For the default projection, the cylinder is oriented along the Earth's axis and the

[10] The rail hypernet is a condensed network model that was used in HIGH-TOOL in order to simulate corridor effects in passenger transport (cf. Van Grol et al., 2016).

[11] Widely used reference systems include the World Geodetic System (WGS-84) modelling the whole terrestrial globe with satisfactory accuracy, the European Terrestrial Reference System (ETRS-89) focusing specifically on Europe and the North American Datum (NAD-83) focusing specifically on North America.

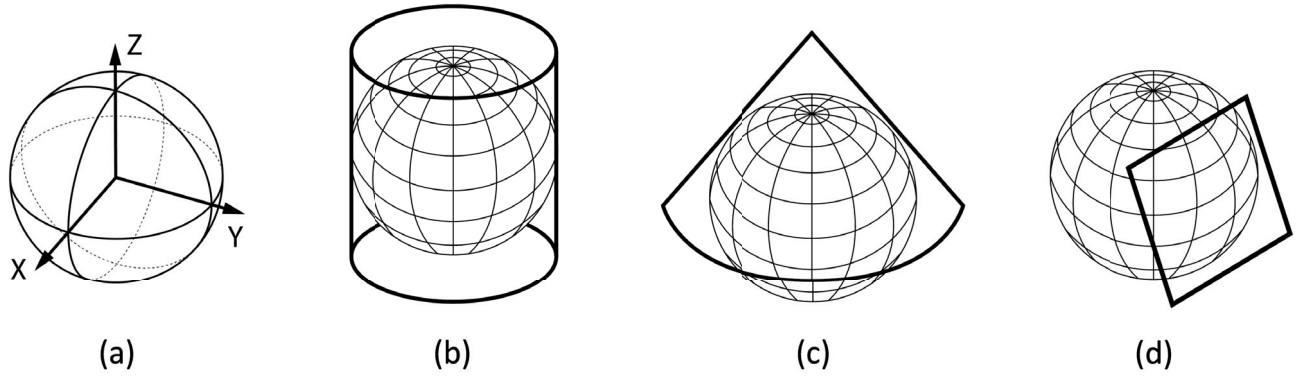

Figure 7.6: ECEF coordinate system, cylindrical, conical and planar projection

equator touches the cylindrical surface. For this reason, the equator is projected without distortion while in particular the polar regions (Arctic and Antarctic) are extremely distorted.

For the conical projection (Fig. 7.6c), the sphere is rolled up into a conical surface and projected onto it. For the default projection, the cone is oriented along the Earth's axis and touches the sphere in one circle of latitudes. It is convenient for the production of maps comprising regions with a large east-west extension on the northern or the southern hemisphere. It is not convenient for regions with a large north-south extension and in particular not for the whole terrestrial globe due to the increasing distortion in relation to the distance from the touching circle(s).

For the azimuthal (planar) projection (Fig. 7.6d), the sphere is touched by a planar surface. The degree of distortion grows rapidly with increasing distance from the centre of the projection due to the curvature of the terrestrial globe. Azimuthal projections are therefore appropriate for the production of maps of small regions with similar north-south and east-west extension. Besides the cylindrical-, conical- and azimuthal projections there are numerous other projections that cannot be clearly attributed to one of the three main categories. In order to produce angle-, distance- or area-preserving maps for specific regions of the Earth, each projection can be adjusted, e. g. by varying the orientation of the cylinder, by varying the cone angle, or by applying stretching-factors.

It should be stressed that a multitude of different map projections and, accordingly, different coordinate reference systems exist (see, for example, Kohlstock (2014); Annoni et al. (2001); Hennermann and Woltering (2014); Flacke (2010)). For the merging of maps and GIS layers relying on different CRS more or less complex coordinate transformations are required. Nowadays, these coordinate transformations as well as the different coordinate reference systems are covered by the EPSG Geodetic Parameter Dataset[12] (OGP, 2012) and implemented in most relevant GIS software environments. On the one hand, this simplifies the processing of GIS data significantly, as no sound understanding of the mathematical theory of GIS is required (e. g. for the merging of different data sets). On the other hand, this knowledge is required in order to recognise potential sources of error.

In this thesis, such GIS datasets relying on the ETRS-89 reference system were used in preference when building the data basis covering travel zones and regional indicators for the HIPAT model discussed

[12]The dataset was elaborated by the European Petroleum Survey Group (EPSG) and is provided by the International Association of Oil & Gas Producers (IOGP).

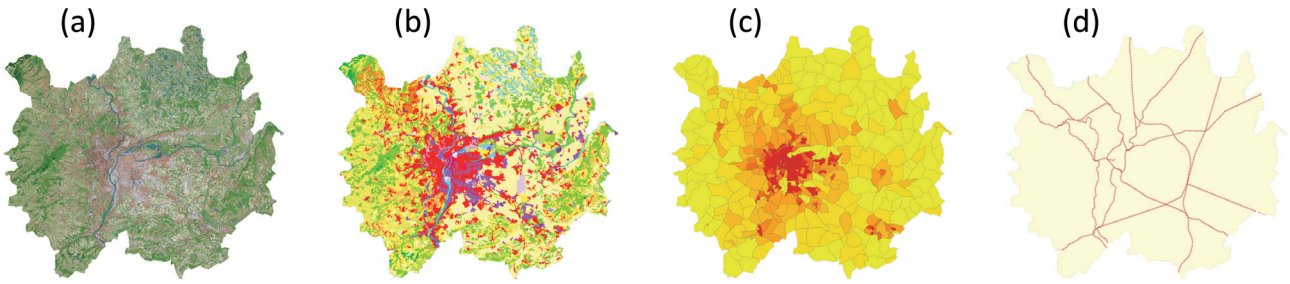

Figure 7.7: GIS raster and vector data layers for the French agglomeration of Grand Lyon

in chapter 9. Furthermore, an area-preserving coordinate reference system was chosen in order to facilitate the processing of space-related indicators, such as population density per square kilometre and land use type per hectare. In addition, particular attention was paid to the correct computation of travel distances. Given the Earth's curvature, the distance between two points was computed by specific GIS routines based on the corresponding three-dimensional coordinates rather than by measuring the linear distance between these points on a planar map.

7.2.3 Vector and raster data type

There are basically two different data types for storing georeferenced data in a GIS layer: the vector and the raster type. Vector data rely on geometric elements (i. e. geometric features) like points, polygon lines and multi-polygon lines that are described by mathematical functions. Hence, GIS data provided in the vector format can be analysed at any spatial level without loss of information. The vector format is commonly chosen if accurate information on spatial structures (topologies) needs to be stored in a compact data format. Typical examples include the definition of borderlines, rail- and road networks. Commonly, vector data are combined with attribute tables providing additional information on the topologies, e. g. the number of inhabitants for regions and speed limits for road links.

In contrast to vector data relying on geometric features, raster data rely on pixels that are arranged in a grid. Raster data may be preferred over vector data, for instance, in order to describe the spatial distribution of land use types in a region and the concentration of air pollutants in a city as well as to store true colour satellite images. Depending on the number of dimensions that are related to each pixel, raster data can store information about one or more attributes. Due to the close relation between raster data and ordinary image data, GIS layers provided in the raster format are frequently given in an area-preserving map projection where all pixels and, accordingly, all grid raster cells refer to areas of equal size on the map. Given that raster layers do not provide information on continuous surfaces, like regions and networks, they are often combined with vector layers containing information on region boundaries and on transport networks.

Figure 7.7 outlines four GIS layers for the French agglomeration of Grand Lyon that were used for a preparatory analysis and for testing several features of geographic information systems within this thesis. The two layers on the left (a, b) showing a satellite image and land use data are given in the raster format. The two images on the right (c, d) showing the boundaries of the city blocks[13] and the ETISplus rail network model are given in the vector format.

[13]The city blocks refer to the IRIS classification (French: Îlots regroupés pour l'information statistique) that are provided by the French National Institute of Statistics and Economic Studies.

7.2.4 GIS software environments and transport modelling

Given the application of georeferenced indicators in transport modelling, GIS software can be used for different purposes, including:

- the processing of GIS datasets and the compilation of the input dataset;
- the development and implementation of the transport model; and
- the processing of modelled output indicators and the preparation of thematic maps.

The elaboration of a complete and consistent input dataset is an important prerequisite for the development of a transport model. Typical tasks include the definition of travel zones and the compilation of mode-specific network models. In addition, centroids, feeding nodes and feeding links have to be defined in order to link the travel zones with the network models. Furthermore, data sources have to be exploited to extract the required regional indicators and the required link attributes (e. g. population, speed limits, etc.). Required indicators that are not available have to be complemented by other means. Missing regional indicators, for instance, can be estimated based on downscaling. In this case, two or more GIS layers are combined. The first layer provides the required indicator for an aggregated region while the other layers provide relevant information in order to estimate the spatial distribution of the required indicator in the aggregated region. For instance, the spatial distribution of the population in a NUTS-3 region can be estimated based on information on the location of the settlement areas that can be extracted from satellite images and land use data. Missing indicators for network links, like the number of lanes and speed limits, can be determined based on high-resolution satellite images. It is also possible to estimate these indicators based on information on the road category.

Having ensured the availability of a complete and consistent input database, existing GIS software can be used to solve specific problems related to the four-step transport modelling approach as well as for the implementation of the whole transport model.[14] However, for the implementation of the whole four-step model it is often more convenient to use specific GIS software environments that are tailored to the needs of transport modelling, like the software VISUM, which provides an implemented four-step transport model (PTV, 2015).[15]

It should be mentioned that the extensive use of a comprehensive GIS software environment certainly speeds up the process of developing, testing and validating a new transport model, i. e. it facilitates rapid prototyping. Given that such comprehensive GIS software environments are commonly not designed to solve complex spatial problems in an efficient way, transport models to be applied at large scale should be implemented in a programming language in order to improve the runtime. For instance, the passenger transport model of the TRANS-TOOLS v2.1.9 model, that is embedded in ArcGIS, has a runtime of more than two days (Ibañez-Rivas, 2010), while the HIPAT prototype model, that is implemented in `Java`, has a runtime of about two minutes (cf. Tab. 9.2, sec. 9.3.1).

[14] The assignment model of TRANS-TOOLS v2 (Traffic Analyst), for instance, is embedded into the proprietary GIS software environment ArcGIS (Monzón et al., 2010).

[15] URL: `http://vision-traffic.ptvgroup.com/de/produkte/ptv-visum/funktionen` (Retrieved January 12, 2018).

Besides the processing of input indicators and the implementation of a transport model, GIS software can also be used for the processing of modelled output indicators. Typical tasks include the computation of derived indicators, like the travelled mileage per capita per year and the preparation of thematic maps to present specific output indicators in a user-friendly way according to the idiom: "a picture is worth a thousand words". It should be emphasised that thematic maps are also important with regard to the validation process, including each component and the overall transport model.

Within this thesis, available GIS software was only used to process input and output indicators while the developed transport models are implemented in common programming languages. For instance, the IPAT model is implemented in `Java` by about 20,000 lines of source code (cf. sec. 6.2.1). The majority of the source code is related to efficient data processing while only a small proportion actually represents the mathematical equations underlying the model. It is therefore evident that the IPAT model could have been implemented with a fraction of the effort in a GIS software environment that manages the data processing itself.

The decision to favour `Java` for the development of the IPAT model was necessary due to already existing runtime problems at NUTS-2 level encompassing only 300 travel zones. In addition, experience showed that the performance of GIS software environments collapses earlier compared to programming languages like `Java` and `C++` if the number of travel zones underlying a transport model is increased. The main reason is that programming languages provide a high degree of flexibility to structure and to tailor the implementation of a transport model. This flexibility was required to comply with two contradictory requirements during the development of the HIPAT model within this thesis. On the one hand, the number of travel zones needed to be increased from 1,500 NUTS-3 regions to more than 100,000 LAU-2 regions. On the other hand, the runtime of the model needed to be reduced from more than two days to several hours at maximum. The HIPAT model is introduced in chapter 9.

7.2.5 Data availability

The setup of European transport models often demands the availability of a comprehensive database providing all required input indicators. Within this thesis, a consistent input database was developed at different aggregation levels in order to meet the requirements of the PAT, IPAT and HIPAT models on the basis of the ETISplus datasets (cf. sec. 2.3). For this purpose several GIS datasets providing georeferenced indicators had to be combined. The input database was first compiled at NUTS-3 level, then aggregated to NUTS-2 level and finally disaggregated to LAU-2 level. While the aggregation step could be simply carried out by summarising the regional indicators, for the disaggregation step different GIS datasets had to be combined and spatial models (i. e. algorithms) had to be developed. The applied datasets are discussed in the following.

7.2.5.1 Nomenclature of Territorial Units for Statistics (NUTS)

The NUTS-2006 classification (Eurostat, 2007) provides the primary basis for defining the travel zones. It basically distinguishes between four regional levels, NUTS-0 to NUTS-3, of which the NUTS-3 level is the most detailed. Besides the four NUTS levels, two local levels, LAU-1 and LAU-2, are defined.[16]

[16] The PAT model only uses NUTS-3-based travel zones and the IPAT model only NUTS-2-based travel zones. In contrast, the HIPAT model uses all travel zones defined by the NUTS-2006 classification and additional, artificially defined zones (cf. Table 9.4 on page 191 summarising the travel zones used in the HIPAT prototype model).

Each hierarchical level comprises more or less disaggregated regions which strongly correlate with the administrative structures in the corresponding countries. The official dataset comprises only the definition of the boundaries (of the regions) for the four NUTS-levels. It is provided in a GIS vector format and can be accessed through the Geographic Information System of the Commission (GISCO) of Eurostat.[17] The NUTS classification was developed by Eurostat in order to ensure harmonised standards for the production of statistics at regional level. These statistics are also provided by Eurostat. However, it has to be mentioned that the NUTS classification is updated every few years. At the time of writing, the most current classification is NUTS-2013.

7.2.5.2 ETISplus socio-economic dataset

The ETISplus socio-economic dataset comprises 66 regional indicators for approximately 1,600 regions in 42 countries (cf. sec. 2.3.2).[18] It is provided as an attribute table[19] and it can be merged, for instance, with the EBM dataset that is briefly introduced below. The regional indicators refer to the base year 2010 and the regions to the NUTS-2006 classification. Besides the member states of the EU, countries of the European Free Trade Association (EFTA), candidate countries and neighbouring countries are also covered at NUTS-3 level. Within this thesis, this dataset was used as a primary source in order to determine the required regional indicators for the three developed transport models.

7.2.5.3 ETISplus rail and road network models

The ETISplus rail and road network models covering Europe and neighbouring countries[20] are provided in the GIS vector format. They are used as a basis for the computation of the travel impedances for each pair of travel zones and the traffic loads at link level. Centroids are defined for each travel zone and the network models are linked to these centroids.

7.2.5.4 Corine land cover (CLC) database

The CLC database covering Europe provides information on land use types at the level of grid cells with an edge length of 100 m (EEA, 2007). It is provided in the GIS raster format. The CLC database covers 44 land use types that can be broadly assigned into five categories: artificial surfaces; agricultural areas; forest and semi-natural areas; wetlands; water bodies. Within this thesis, the CLC database is used to determine the spatial distributions of artificial areas in the NUTS-3 regions. This information is required for the downscaling of regional indicators from NUTS-3 level to LAU-2 level and for the definition of artificial city districts.

7.2.5.5 EuroBoundaryMap (EBM)

The EBM dataset is the European reference database of administrative units (BKG, 2013). It is provided in the GIS vector format and basically covers the 28 member states of the EU as well as the EFTA countries. The dataset consists of several layers encompassing the geometry of the administrative units of the European countries at six aggregation levels (four NUTS levels and two LAU levels),

[17] URL: `http://ec.europa.eu/eurostat/web/gisco` (Retrieved January 9, 2018).
[18] The coverage of this dataset is described in section 2.3.3.
[19] The attribute table provides regional statistics, such as region name, population and the number of workplaces.
[20] The geographic coverage and the provided link attributes of both network models are discussed in section 2.3.2.

their borderlines as well as the official names and unique codes according to the NUTS classification maintained by Eurostat. The dataset is therefore superior to the official GISCO dataset that only includes the four NUTS levels. Given that the EBM dataset is provided in a seamless database (i. e. there are no gaps or overlaps between the polygons representing the administrative units), spatial analyses can be carried out on this dataset, e. g. the EBM layers can be merged with the CLC database and the corresponding LAU- and NUTS regions can be linked. Within this thesis, the EBM v8.1 dataset is used as a primary source for the definition of the travel zones at LAU-1 and LAU-2 levels in the HIPAT model. It should be noted, however, that the EBM dataset refers only to the existing European administrative units. As units at LAU-1 and LAU-2 levels are not defined for all European countries, artificial travel zones were added at these levels.

7.2.5.6 European GEOSTAT 2011 grid dataset

The GEOSTAT dataset reports the population at the level of grid cells with an edge length of 1 km (Haldorson et al., 2017). It is provided by the European Environment Agency (EEA) in the GIS vector format and covers the 28 member states of the EU as well as the EFTA countries. The dataset was compiled from 28 national grid datasets, based on information from the 2011 Population and Housing census. Gaps were filled with modelled information. It should be mentioned that, according to the initial vision of the GEOSTAT project (Bloch et al., 2012), further indicators besides the population could be provided at grid level in future versions of this dataset. Within this thesis, the GEOSTAT 2011 grid dataset is used as a primary source for the estimation of the population for the LAU-2 regions. In addition, it is combined with the CLC database in order to derive the spatial distribution of the population in cities and to define artificial, homogeneous city districts.

7.2.5.7 OpenStreetMap (OSM)

The OSM dataset provides comprehensive georeferenced data, particularly about network models and land use types. It was used for validation purposes within this thesis. The dataset is continuously extended through a collaborative project by the voluntary contribution of many users (Ramm and Topf, 2008). It can be downloaded in a compressed format but covering Europe only, its size is about 20 gigabytes[21] and specific tools like Osmosis have to be applied in order to extract the required topologies (i. e. georeferenced indicators). Although it was initially planned to complement the ETISplus road network model based on OSM data, in the end this project was not implemented due to the difficulty of processing the OSM dataset and matching it to the ETISplus road network model.

[21] URL: `http://download.geofabrik.de/europe.html` (Retrieved January 12, 2018).

7.3 Generating the input data at LAU-2 level

A well-known challenge in European transport modelling is related to the availability of indicators at a regional level, which are required for transport modelling. These indicators include, for instance, population, GDP and the number of workplaces. Although regional indicators are not always available at LAU-2 level, it is possible to downscale these indicators from NUTS-3 to LAU-2 level. In addition, as LAU-2 regions are not defined for many European cities, artificial city districts are generated. The chapter is therefore structured as follows: section 7.3.3 describes the procedure for downscaling regional indicators based on land use data. Section 7.3.2 outlines the procedure for generating artificial city districts. The last section discusses limitations of the applied downscaling procedures and provides and outlook on how the European datasets could be improved.

7.3.1 Downscaling regional indicators

The basic idea behind downscaling is the disaggregation of available, space-related indicators to smaller regions by proration factors. These proration factors can be derived on the basis of the size of a disaggregated region in relation to the aggregated region or on the basis of available indicators like population and land use data for both regions. A common approach is to exploit the CLC land cover database (EEA, 2007) that provides the dominant land use type for grid cells at a spatial resolution of 100 m edge length for the whole of Europe. By assigning specific load factors to each land use type, regional indicators can be disaggregated from NUTS-3 level to the grid and then aggregated to LAU-2 level. It can be assumed, for instance, that the population density is higher for grid cells of the type "Continuous urban fabric" than for cells of the type "Industrial or commercial units". Furthermore, it can be assumed that the population density is close to zero for those land use types referring to non-artificial surfaces, like agricultural areas, forests or water bodies.

According to Gallego (2007), load factors $l(t)$ for each land use type t can be estimated by regression analysis.[22] Having determined $l(t)$, the population P of region R can then be downscaled to the corresponding grid cells c_i ($\cup c_i = R$) as follows:

$$p_{i|R} = P \cdot \omega_{i|R} \qquad \text{with} \qquad \omega_{i|R} = \frac{s_{i|R}\, l_i(t)}{\sum_i s_{i|R}\, l_i(t)} \tag{7.1}$$

where $p_{i|R}$ identifies the population that is assigned to the part $s_{i|R} \in [0, ..1]$ of the grid cell c_i (since some grid cells are not entirely located in one region) and $\omega_{i|R}$ identifies the applied proration factor. In a second step, the disaggregated indicators $p_{i|R}$ can be aggregated from the grid cells to nearly any spatial level, e. g. to the level of LAU-2 regions.

Besides population, other regional indicators can also be disaggregated following (7.1). Figure 7.8 outlines the procedure of downscaling the indicator "Number of workplaces" from NUTS-3 level to LAU-2 level, based on the CLC database and load factors from Table 7.2. The left-hand image drafts the boundaries of the aggregated region, providing the indicator number of workplaces. The image in the centre shows the subdivision of the aggregated region in different clusters of equal land use types.

[22] For the regression analysis, georeferenced data is required which provides the spatial distribution of the population at a very disaggregated level, e. g. household and census data.

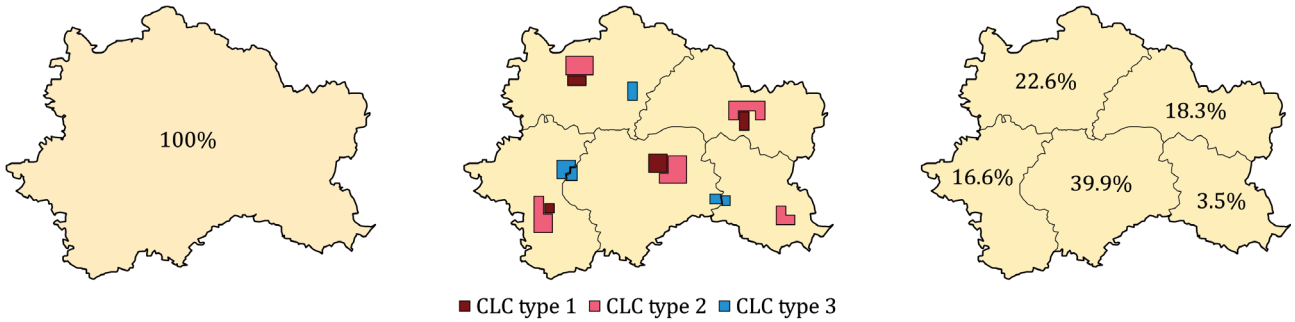

Figure 7.8: Downscaling regional indicators based on land cover data and load factors

Border-crossing clusters are subdivided, i. e. each cluster can be clearly assigned to one disaggregated region. The clusters were computed based on the grid cells (that are not shown). The right-hand image shows the proration factors determined for each disaggregated region that are applied for downscaling the regional indicator.

Within this thesis, the disaggregation of the required regional indicators is carried out under application of specific load factors that are distinguished by land use type. These load factors are taken from Schoch (2004) and summarised in Table 7.2. In total, the CLC database covers 44 land use types, but only the first nine types are relevant for the current disaggregation, i. e. the load factors for the land use types related to vegetated and non-artificial areas are 0.
Although the load factors originally refer to the indicator "Sozialversicherungspflichtige am Arbeitsort" (employees subject to social insurance contribution at their workplaces), it is assumed within this thesis that they are also valid for downscaling the regional indicator "Number of workplaces", as there is a high correlation between both indicators.

Table 7.2: Load factors by land use type - workplaces

Land use type	Type id(s)	Load factor(s)
Continuous urban fabric	1	150.9
Discontinuous urban fabric	2	10.2
Industrial or commercial units	3	41.9
Road and rail networks and associated land	4	3.7
Port areas and airports	5, 6	61.5
Mine, dump and construction sites	7, 8, 9	1.6
Vegetated and non-artificial areas	10, .., 44	0

7.3.2 Generating artificial city districts

The NUTS zoning system relies on the definition of administrative units of the member countries of the European Union. It introduces a hierarchy of four NUTS levels and two LAU levels. However, the hierarchy of the six levels is not consistently applied for all regions. For instance, the German capital Berlin is captured by an administrative unit at NUTS-1 level while other German cities are given by administrative units at NUTS-3 level. Those regions which are not further disaggregated to the level of LAU-2 regions have to be subdivided in order to generate artificial city districts (i. e. zone

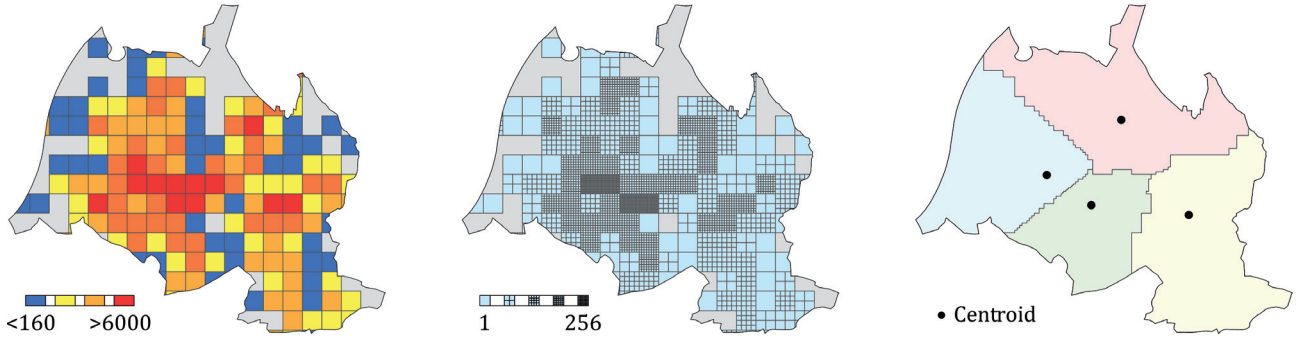

Figure 7.9: Generating artificial city districts at LAU-1 level for Karlsruhe (DE122)

districts) and to complement the hierarchy of six levels. These zone districts should be consistent with the LAU-2 regions in terms of size and weight in order to enable the modelling of regional trips. Zone districts can be constructed bottom-up on the basis of the European GEOSTAT dataset (Haldorson et al., 2017) providing population data at the level of grid cells with an edge length of 1 km.

In order to fulfil the criteria discussed in section 7.1.1, zone districts are derived in three steps in this thesis by use of the "Grouping Analysis" tool of the proprietary GIS software ArcGIS (GI, 2015). In the first step, the grid cells of the GEOSTAT dataset providing georeferenced population data are assigned to the aggregated travel zones, e. g. to the NUTS-3 regions. Grid cells referring to more than one travel zone are split. In the second step, the cells are homogenised by sequentially subdividing those cells with more than 160 inhabitants, i. e. cells of different size but equal weight are computed. In the last step, four compact zone districts with equal weight (i. e. equal numbers of cells) are generated by grouping adjacent cells.

The centroid C of a zone district is calculated based on the spatial distribution of the population in this zone according to (7.2) in which c_n identifies the centroid of cell n located in the zone district and p_n identifies the population of this cell. Although not explicitly stated, all centroids are given in latitude/ longitude coordinates, i. e. $C \in \mathbb{R}^2$ and $c_n \in \mathbb{R}^2$.

$$C = \frac{1}{P} \sum_{n=1}^{N} p_n \, c_n \qquad \text{with} \qquad P = \sum_{n=1}^{N} p_n \tag{7.2}$$

The centroids of the zone districts are then connected to the network models. The number of inhabitants P of these artificially generated zone districts is computed bottom-up.

Within this thesis, the outlined algorithm is applied twice for splitting NUTS-3 regions referring to an aggregated city region into 16 districts. Four artificial city districts are generated at LAU-1 level which are then split into 16 districts at LAU-2 level.

Figure 7.9 illustrates the approach for generating four artificial LAU-1 regions for the German city of Karlsruhe (DE122). The left-hand image shows the boundaries of the NUTS-3 region and the grid cells from the GEOSTAT dataset that are located within the boundaries of the city of Karlsruhe. The cells are coloured according to a colour ramp where blue identifies the lowest category of less than 160 inhabitants and red the highest one of more than 6,000. In the second image, all cells refer to less

than 160 inhabitants. The disaggregated cells were generated based on the first image by sequentially subdividing those cells with more than 160 inhabitants into four parts. Hence, those grid cells referring to the lowest category in the first image (i. e. less than 160 inhabitants) did not have to be subdivided while those grid cells referring to higher categories had to be subdivided into 4, 16, 64 or even 256 disaggregated cells. The right-hand image shows the four artificial LAU-1 regions that are generated by the cluster algorithm and their derived centroids.

7.3.3 Limitations and outlook

The non-availability of a consistent database covering socio-economic and demographic indicators below NUTS-3 level is a well-known problem in European transport modelling. This obstacle can be overcome by downscaling regional indicators to LAU-2 level and by constructing artificial travel zones at LAU-2 level. However, this immediately raises questions concerning the reliability of the disaggregated data source. It is quite evident that the application of an oversimplified downscaling method, e. g. based on the region size, would be worthless. If calculated based on the region size, the reported indicators for an aggregated region would be equally assigned to the disaggregated regions even though the spatial distribution of the population and the workplaces inside a NUTS-3 region is commonly very heterogeneous. On the other hand, if the downscaling method relies on indicators which are available at a very disaggregated level, like data on land use types and population data at grid level, the reliability of the disaggregated data source can be ensured to some extent.

Within this thesis, the CLC database, which covers several artificial land use types like settlements and industrial areas, is considered for the disaggregation of regional indicators and for the generation of artificial travel zones. Moreover, the GEOSTAT dataset is used that even reports the population at the level of grid cells. Hence, the population of a LAU-2 region is aggregated bottom-up and the quality of this indicator can be expected to be very good.

A minor flaw is related to the non-availability of an official GIS layer at European scale providing information on the boundaries of city districts. To overcome this gap, artificial LAU regions must be generated. It is important to stress that the outlined methods in the last sections for the generation of artificial LAU-2 regions and for the downscaling of regional indicators to LAU-2 level are applied exclusively in this thesis to facilitate a more accurate modelling of regional trips below 100 km. Thus, the full comparability to other European transport models operating on different sets of travel zones is not given even if the model output is only analysed at an aggregated level.

In order to improve transport modelling at a European scale and to avoid a potential MAUP when disaggregating regional indicators between different sets of travel zones (e. g. from NUTS-3 to LAU-2 level), an extension of the EBM dataset with regard to the provision of official city district boundaries would be desirable. In addition, the availability of other important transport demand indicators at the level of grid cells, like GDP and the number of workplaces, should be ensured through an official dataset comparable to the GEOSTAT dataset providing information on the population at the level of grid cells.

Chapter 8

Hierarchical modelling

The main bottleneck of transport models following the four-step approach is the computation of the O/D trip matrix by the trip distribution model. Given the fact that the complexity of the trip matrix increases quadratically with the number of travel zones, recent European transport models are limited to operate at NUTS-3 level. The NUTS-3-based travel zones have an average diameter of over 50 km and are therefore not particularly suited for the modelling of regional trips that are shorter than 100 km. In order to facilitate a more accurate modelling of these trips it is necessary to significantly decrease the diameter of the travel zones. However, halving the diameter of the travel zones increases the number of O/D relations by a factor of 16. Moreover, if a transport model is applied at LAU-2 level, encompassing 100 thousand travel zones, 80 gigabytes[1] of computer memory are required for storing the computed trip estimates for all all O/D relations, i. e. the full trip matrix. In contrast, in order to store all trip estimates for one thousand travel zones and, accordingly, one million O/D relations, only 8 megabytes are required. In both cases, additional memory is required in order to store the information on the origin and the destination for each O/D relation.

Taking into account that trip matrices are commonly distinguished by several demand segments, e. g. by purpose and transport mode, a multiple of the 80 gigabytes would be required if complete matrices encompassing all O/D relations were produced. It is therefore obviously not possible to increase the number of travel zones without simultaneously reducing the complexity of the trip matrix and the number of modelled O/D relations by the same extent.

A possible way out of this dilemma is the development of an innovative, hierarchical trip distribution model. Such a model would provide the flexibility needed to compute trip estimates between travel zones at different aggregation levels, for instance, long-distance trips at a more aggregated level and short-distance trips at a more disaggregated level. Modelling of O/D relations at different hierarchical levels, however, leads to several follow-up problems. First of all, it has to be decided which relations should be modelled at which level, although it is not entirely clear how the accessibility indicators can be computed. The accessibility indicator underlies the improved trip distribution model (5.22) described in chapter 4.4. In addition, the computation of aggregated output indicators like pkm demand generated by NUTS-3 travel zones is not obvious, given that parts of the demand are also modelled at more aggregated levels, e. g. NUTS-2.

[1] It is assumed that the so-called double-precision floating-point format (binary64) is used where eight bytes (i. e. 64 bits) are allocated for storing a number (cf. IEEE, 2008, Tab. 3.5).

The intended objective of this chapter is to introduce a hierarchical modelling approach capable of competing with the approved methodology of the IPAT model discussed in chapter 6. This means in particular that any methodological revision implemented for the IPAT model shall also be implemented for the hierarchical modelling approach. Thus, the hierarchical modelling approach affects all stages of the four-step model, including trip generation, trip distribution, modal split, and network assignment.

This hierarchical approach is developed in the following sections, and this chapter is structured as follows: section 8.1 describes the basic idea behind a hierarchical modelling concept and quantifies the savings potential, based on a theoretical example relying on quadratic travel zones. In section 8.2, the theoretical example is revised to take account of the differences between quadratic travel zones and NUTS regions that can have almost any shape. Furthermore, the requirements that need to be met by a hierarchical trip distribution model, including a consistent data exchange between the different hierarchical levels, are investigated. Section 8.3 introduces the hierarchical accessibility indicator problem and elaborates an approach for the computation of a competition-weighted accessibility indicator at different aggregation levels by recursive algorithms. This is a precondition for the application of an accessibility-based trip distribution model. The main findings of this chapter are summarised in section 8.4.

8.1 Agglomerative hierarchical approach using quad-trees

Agglomerative hierarchical approaches are, inter alia, applied in informatics for storing image data at different aggregation levels in a compact data format. These so-called image pyramids facilitate the computation of indicators at different aggregation levels and therefore speed up the processing of image data. For instance, if global indicators like the average grey value or contrast information are required, the image does not have to be processed at the most detailed aggregation level (cf., for example, Gonzalez and Woods, 2008). Moreover, if detailed indicators are required but only for a specific subsection of an image, this subsection can be processed at a more detailed aggregation level, while the overhead of having to process the whole image at a detailed level can be avoided.

If a similar concept were applied in European transport modelling with regard to the computation of the O/D trip matrix, the runtime complexity could be significantly reduced. It is quite obvious that for the computation of long-distance trips over 1000 km at NUTS-3 level rather than at NUTS-2 level the value added can be expected to be negligible. Even the accuracy of modelling long-distance trips at NUTS-1 level is satisfactory in such cases. It can be assumed, for instance, that an aggregated trip distribution model operating at country level and a disaggregated model operating at NUTS-3 level, both computing the number of flights between Belgium and Greece, provide very similar results. However, the estimated number of flights computed by the disaggregated model relies on more than four thousand O/D relations[2] while the aggregated model only considers two O/D relations.

The computation of the trip matrix by a hierarchical approach would save many O/D relations and therefore significantly reduce the complexity of the trip distribution model and, accordingly, the transport model. This could be an important way forward in European transport modelling given the need to model regional trips with a higher accuracy than is currently provided by European transport models operating at NUTS-3 level. Moreover, modelling trip estimates between the travel zones at different hierarchical levels would break the polynomial dependency between the number of travel zones and the number of O/D relations down to an almost linear level. This is shown by a theoretical example (Tab. 8.1) relying on quadratic travel zones, which are organised according to a quad-tree structure.

8.1.1 Brief description of quad-trees

Quad-trees define an efficient hierarchical data structure. They are frequently used in computer science to store two-dimensional data, like the XY coordinates of randomly distributed points (Janser et al., 1996). The underlying concept of quad-trees is to successively divide a square into four sub-squares until the desired resolution is achieved so that each square stores exactly one data element.

Figure 8.1 shows a quad-tree structure which consists of three hierarchical levels as well as the successively divided squares. The quad-tree structure provides significant advantages compared to a two dimensional data matrix in the storing of random coordinates, as empty squares or rather useless subtrees containing no information are not stored. Moreover, the exact number of coordinates stored in a quad-tree data structure does not have to be known in advance since the tree is successively extended when inserting new coordinates by introducing new hierarchical levels.

[2] According to the NUTS-2006 system (Eurostat, 2007), Belgium consists of 44 NUTS-3 regions and Greece of 51.

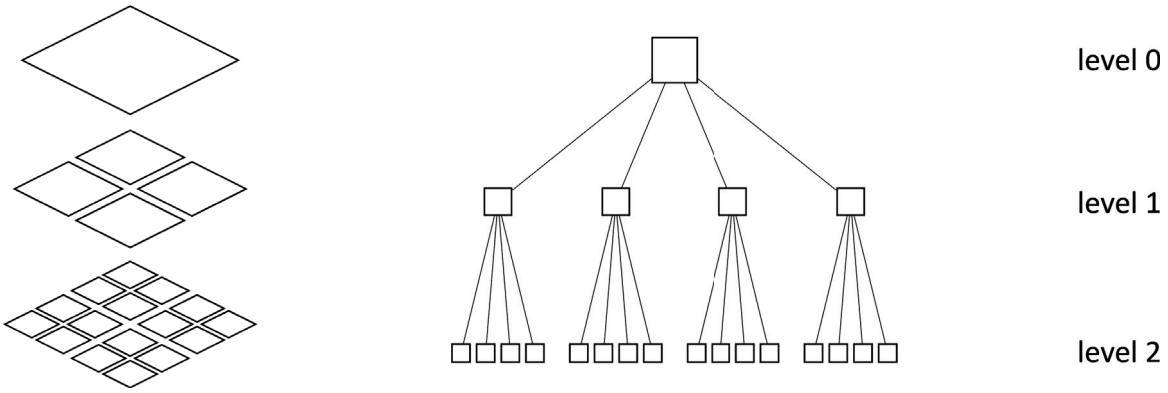

Figure 8.1: Structure of the quad-tree

However, the quad-tree data structure has some disadvantages compared to a data matrix that are related to the insertion, retrieval and deletion of a data element. According to Finkel and Bentley (1974), in a quad-tree structure the complexity of an average insert and an average retrieval of a coordinate is $O(\log_4 n)$ where n is the number of coordinates to be stored in the tree. Hence, the complexity for constructing the whole tree is $O(n \log_4 n)$. As a general sorting of the two-dimensional coordinates is not provided, deleting an element is more complex (Samet, 1980).

Taking into account the costs related to the construction of the quad-tree structure and the costs related to retrievals and updates of stored coordinates (i. e. transport demand and supply indicators), the quad-tree structure should be generated prior to the first model run. The generated structure can then be stored, provided that the overall settings (e. g. zoning system) of the transport model are kept constant. Given the large discrepancies between the structures of the NUTS classification and the quad-tree, it is obvious that an adequate agglomerative hierarchical approach still needs to be developed for transport modelling.

8.1.2 Savings potential of quad-trees for trip matrix computation

In order to quantify the savings potential when computing a trip matrix at different hierarchical levels following the outlined quad-tree approach in comparison to the standard computation of a full trip matrix at the most detailed level, a hypothetical example is discussed below. In this example, the travel zones are defined according to a quad-tree image pyramid, i. e. each travel zone is sequentially sub-divided into four sub-squares. The savings potentials are then computed by comparing the complexity of the full matrix with the complexity of two hierarchical trip matrices that are generated following two different rules H_1 and H_2.

Table 8.1 provides an overview of the hypothetical example. The first two columns show the IDs of the hierarchical levels of the quad-tree and the numbers of grid cells (i. e. travel zones) that are generated at a specific level. The third column shows the number of O/D relations that are produced if a standard model is applied, computing a full trip matrix. The fourth and fifth columns list the numbers of O/D relations that are produced by the hierarchical modelling according to two different "neighbourhood criteria", H_1 and H_2, to be explained shortly. The last two columns list the savings potentials, S_1 and S_2, of the hierarchical modelling. For example, if we set eight hierarchical levels,

Table 8.1: Savings (S_1, S_2) of hierarchical model (H_1, H_2) compared to O/D matrix

Level	Cells	O/D Matrix	H_1 Relations	H_2 Relations	S_1	S_2
0	1	1	1	1	0.0 %	0.0 %
1	4	16	16	16	0.0 %	0.0 %
2	16	256	256	256	0.0 %	0.0 %
3	64	4,096	1,756	3,196	57.1 %	22.0 %
4	256	65,536	9,016	20,536	86.2 %	68.7 %
5	1,024	1,048,576	40,756	102,676	96.1 %	90.2 %
6	4,096	16,777,216	173,296	458,416	99.0 %	97.3 %
7	16,384	268,435,456	714,796	1,937,356	99.7 %	99.3 %
8	65,536	4,294,967,296	2,903,656	7,966,696	99.9 %	99.8 %

i. e. if we subdivide the root travel zones eight times, we generate 65.5 thousand travel zones. Thus, a full trip matrix covering all O/D relations consists of 4.3 thousand million relations while the two hierarchical matrices only consist of 2.9 or 8.0 million relations. The corresponding savings potentials are therefore 99.9% or 99.8%. Hence, it can be concluded that the number of O/D relations is reduced to less than 0.1% in comparison to the full trip matrix by hierarchical modelling according to the H_1 criterion.

This tremendous reduction of complexity can be achieved as the trip demand between two travel zones is always computed at the highest aggregation level possible. This is the case if the two travel zones are not adjacent (i. e. neighbours), which will be called "neighbourhood criterion" in this thesis. On the other hand, if the two travel zones of an O/D relation are adjacent, both have to be sub-divided into four sub-squares resulting in 16 disaggregated O/D relations. For each of these O/D relations, it is checked again whether the two respective travel zones are adjacent or not (unless we have reached the highest level of disaggregation, i. e. level eight, where the process is halted). Accordingly, some relations can now be modelled at the current level while others have to be further disaggregated.

This pattern for disaggregating the O/D relations is outlined by Figure 8.2 in which the origin is located in the left corner and the sixteen destinations (including the origin itself) are drawn in different colours according to the distance between the origin and the destination. The example refers to the H_1 criterion which determines whether origin and destination of an O/D relation are neighbours:

H_1: only directly adjacent squares (i. e. travel zones) are neighbours, i. e. each origin can have up to nine neighbouring travel zones that are located around the origin (including the origin itself).

According to Figure 8.2 the H_1 criterion reveals that four of the sixteen O/D pairs are defined by neighbouring travel zones, namely the intra-zonal relation and the three relations to the adjacent orange-coloured cells. These four O/D relations are therefore disaggregated and transformed into 64 relations. Hence, 76 O/D relations have to be computed in total, taking into account the twelve relations that could already be computed at the more aggregated level. This means that almost three quarters of the calculations could be saved in comparison the computation of a complete trip matrix at the disaggregated level encompassing 64 travel zones. At this level, the origin is modelled by four cells and, accordingly, 4×64 O/D relations have to be computed for a complete trip matrix.

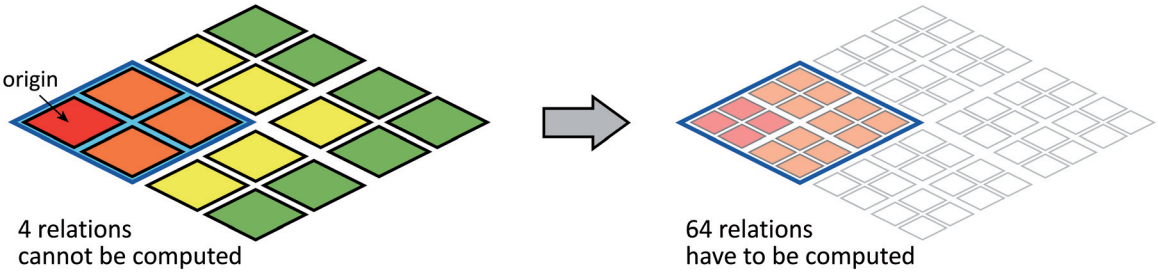

Figure 8.2: Disaggregation pattern for O/D relations (H_1 criterion)

Referring to the structure of the quad-tree outlined in Figure 8.1, the trip demand between any two cells cannot be computed at level 0 and level 1 without violating the neighbourhood criterion H_1. Hence, all O/D relations have to be disaggregated and modelled at a more disaggregated level (i. e. from level 0 to level 1 and subsequently from level 1 to level 2). Starting at level 2, however, several pairs of cells can be identified that do not violate the neighbourhood criterion H_1. These O/D relations are not further disaggregated but rather computed at level 2, even if cells could be broken down to level 8. The first savings can therefore be observed at level 3 in Table 8.1.

Besides the H_1 criterion that only treats adjacent cells as neighbours, the savings are also investigated for the stronger neighbourhood criterion H_2:

H_2: two squares (i. e. travel zones) are neighbours if both squares are directly adjacent to a common square, i. e. each origin can have up to 25 neighbouring travel zones.

Referring to the disaggregation pattern outlined by Figure 8.2, the H_2 criterion additionally requires disaggregation of the O/D relations between the origin and the five yellow-coloured squares. Hence, nine out of sixteen relations have to be disaggregated. In this case, the complexity savings are lower than for the H_1 criterion but the quadratic dependency of the complexity of the trip matrix to the number of travel zones is again reduced to a linear dependency at higher hierarchical levels. For instance, between levels 7 and 8, the number of travel zones increases by a factor of 4 while the number of O/D relations increases by a factor of 4.06 if the H_1 criterion is applied and by a factor of 4.11 if the H_2 criterion is applied.

It can therefore be concluded that the hierarchical approach overcomes the quadratic dependency of the complexity of the trip matrix on the number of travel zones, and reduces it to a linear dependency. This facilitates the attempt to increase the number of travel zones for modelling European transport at LAU-2 level by more than 100,000 travel zones.[3] Taking into account the hierarchical structure of the NUTS classification, savings in terms of model runtime and data complexity can already be expected for a European transport model operating at NUTS-2 level.[4]

[3]The NUTS-2013 zoning system comprises 118,504 LAU-2 regions (Eurostat, 2015b).
[4]The hierarchical model then consists of four aggregation levels which are: EU28, NUTS-0, NUTS-1, and NUTS-2.

8.2 Trip distribution modelling following the quad-tree approach

The NUTS zoning system and quad-trees refer to hierarchical structures. Thus, tremendous savings can be expected if sparse trip matrices at different aggregation levels are computed instead of a single trip matrix at the most detailed level (including every possible O/D relation). However, besides the savings arising from a hierarchical trip distribution model, specific challenges and obstacles emerge that have to be investigated.

A first challenge related to the NUTS zoning system is that the defined regions are not particularly suited for the construction of a quad-tree data structure. A second challenge is the definition of adequate neighbourhood criteria in order to decide which O/D relations have to be disaggregated and which do not. In addition, a major obstacle emerges with regard to the consistent adaptation of a standard trip distribution algorithm for application at different hierarchical levels, given that these algorithms were initially developed to solve the trip distribution problem at one single hierarchical level. This obstacle is mainly related to the challenge of defining adequate propagation rules in order to ensure the consistency between the trip estimates computed at different levels of the tree-based data structure. In addition, the hierarchical algorithm to be developed should also be efficient and benefit from parallel computing.

In the following section, three neighbourhood criteria are defined that can be applied for the NUTS regions. It is then investigated whether the trip distribution model can be applied consistently at different aggregation levels. Two basic propagation rules for aggregating and disaggregating region-based indicators in the tree structure are introduced. Finally it is investigated how the data consistency can be ensured in the tree when scheduling the trip distribution model in parallel.

8.2.1 Setting the neighbourhood criterion

The basic difference between quad-trees and the NUTS zoning system is that a NUTS region is commonly not divided into exactly four sub-regions. This obstacle can easily be overcome by applying a more general tree structure allowing a flexible number of lower-level nodes in order to reproduce the hierarchical structure of the NUTS zoning system. In this case, the root node of the tree refers to Europe as a whole and its children to the European countries.

The main challenge is, however, the definition of adequate neighbourhood criteria in order to determine whether two NUTS regions are adjacent and whether the trip demand between them should be computed at a more disaggregated level. Referring to the theoretical example (cf. sec. 8.1.2), the adjacency of two squares can easily be determined based on their size and their centroids. Given an almost arbitrary shape of the NUTS regions, the adjacency of two regions cannot be determined based on simple geometric rules, such as sharing a common borderline or the distance between the respective capitals. For example, Denmark and Sweden do not share a common land-based borderline and the travel distance between the two capitals, Copenhagen and Stockholm, is over 600 km. Taking into account the transnational European metropolitan region of Greater Copenhagen, formerly called the Øresund Region, it is obvious that the trip demand between Denmark and Sweden should not be modelled at NUTS-0 level. In contrast, the trip demand between Portugal and Finland can be adequately modelled at this high level.

Following the ideas behind the two neighbourhood criteria, H_1 and H_2, that were applied for the quad-tree example (cf. sec. 8.1.2) and taking into account that two neighbouring NUTS regions might not share a common land-based borderline, three reasonable neighbourhood criteria (N_1, N_2, N_3) could be defined as follows:

N_1: two regions are neighbours if the distance between their centroids is less than the sum of their radii;

N_2: two regions are neighbours if the distance between their centroids is less than the sum of their diameters;

N_3: two regions are neighbours if the minimum distance between the region's borderlines is less than a variable threshold depending on the distance between their centroids.

The criteria N_1 and N_2 rely on the simplifying assumption that all regions are compact, i. e. that the actual shapes of all regions are similar to the respective bounding circles. In this case, radii and diameters can be computed based on simplified geometric rules. In contrast, criterion N_3 requires the application of spatial algorithms in order to determine the minimum distance between the borderlines of two regions. Given that the neighbourhood criterion is applied in order to determine an optimal set of O/D relations (i. e. an optimal trade-off between the number of O/D relations and the model accuracy), it can be reasonable to define more sophisticated criteria and to combine several criteria.

It is important to stress that exceptions exist where all criteria fail. For instance, if two adjacent travel zones resemble ellipses where the eccentricity is close to 1, N_1 and N_2 fail. This is the case for the O/D relation between the French island Corsica and the Italian island Sardinia if the centroids are set to the city of Bastia located in the north of Corsica and the city of Cagliari located in the south of Sardinia. In this case, the distance between the two centroids is quite large while the radii and the diameters are quite small. Hence, N_1 and N_2 fail to disaggregate both islands. For this specific example, N_3 can be successfully applied if the threshold value is set high enough. However, other examples can be constructed where the third criterion fails as well. It is therefore necessary to validate the set of O/D relations that is generated by the neighbourhood criterion and to refine it by hand.

The a priori definition of the set of O/D relations is of great value in the implementation of hierarchical data structures in a programming language. If these structures had to be dynamically constructed during the runtime of the transport model, rather inefficient dynamic data structures like hash-maps or concatenated lists would have to be used. In contrast, if the number of O/D relations has been fixed in advance, very efficient static data structures like arrays can be used that provide a low complexity of $O(1)$ for retrieving and updating a value related to an O/D relation. At this point, the main difference between a standard trip distribution model computing a complete matrix and the hierarchical model computing a hierarchical matrix becomes obvious: while a data cell in a standard matrix is commonly accessed through the *column-ID* and the *row-ID*, i. e. through the origin and the destination, a data cell in the hierarchical matrix has to be accessed through the *relation-ID*. These differences between a standard matrix and the hierarchical matrix have significant implications, in particular, on the computation of the accessibility indicator underlying the trip distribution model. This issue will be discussed below in section 8.3.

8.2.2 Transferability of deterrence function and trip distribution model

The computation of trip estimates between two travel zones requires information on three indicators, namely: (i) trip generation, (ii) number of trip endings and (iii) generalised cost of travelling. If a hierarchical trip distribution model considering NUTS-1, NUTS-2 and NUTS-3 regions is applied, the trip estimates between the two German cities of Karlsruhe and Stuttgart are computed at NUTS-3 level while the trip estimates between Karlsruhe and Lyon are computed at NUTS-1 level as part of the O/D relation Baden-Württemberg to Centre-Est.

While the computation of the three above-mentioned indicators is straightforward, the question immediately arises whether it is permitted to apply the trip distribution model simultaneously at different NUTS levels. Given the scale dependency of the deterrence function, it is commonly not allowed to transfer the trip distribution model between different aggregation levels. However, it is important to stress at this point that the trip distribution model applied in the IPAT model relies on the improved deterrence model that has been outlined by (5.22) (cf. sec. 4.4). Bearing in mind the scaling function ρ_d (5.8) that is defined in relation to the average diameter of the travel zones, the improved deterrence model can be applied more consistently at different hierarchical levels.[5] This ensures the transferability of the trip distribution model between different hierarchical levels.

8.2.3 Basic propagation rules for hierarchical tree-based data structures

One further question arising almost instantly is related to the data exchange of regional indicators between different aggregation levels in the tree. Some indicators like trip generation and trip endings should be computed only at the most disaggregated level.[6] These indicators then have to be propagated bottom-up towards their parent nodes in the tree, i. e. from LAU-2 level to LAU-1 level, from LAU-1 level to NUTS-3 level, etc. In contrast, transport demand indicators like long-distance trips over 1000 km should initially be computed at a more aggregated level in order to reduce the number of O/D relations. These indicators then have to be propagated top-down towards their successor nodes in the tree by applying adequate disaggregation rules.

Figure 8.3 drafts the two directions for propagating data in a hierarchical tree-based structure, bottom-up and top-down propagation. The most basic rule for propagating indicators bottom-up is by adding their values while the most basic top-down propagation (i. e. disaggregation) is carried out on the basis of proration factors (p_k).[7] It should be emphasised that for specific purposes, like the computation of the hierarchical accessibility indicator that will be discussed below in section 8.3, more sophisticated propagation rules are required.

8.2.4 Parallel scheduling of the trip distribution model

The runtime of a transport model can be reduced by parallel scheduling, i. e. through the application of a high-performance computer. For instance, the latest high-performance computer that has been

[5] The scaling function computes different values if it is applied at NUTS-1 and NUTS-3 level. This adapts the deterrence model to the respective hierarchical level.

[6] Income modelled through "GDP per capita" is an important mobility determinant, but wealth is distributed unevenly among the population. If trip generation is computed at LAU-2 level, different income levels can be considered.

[7] Of course, the two outlined rules are consistent. It holds that $y_1 + y_2 = Y$.

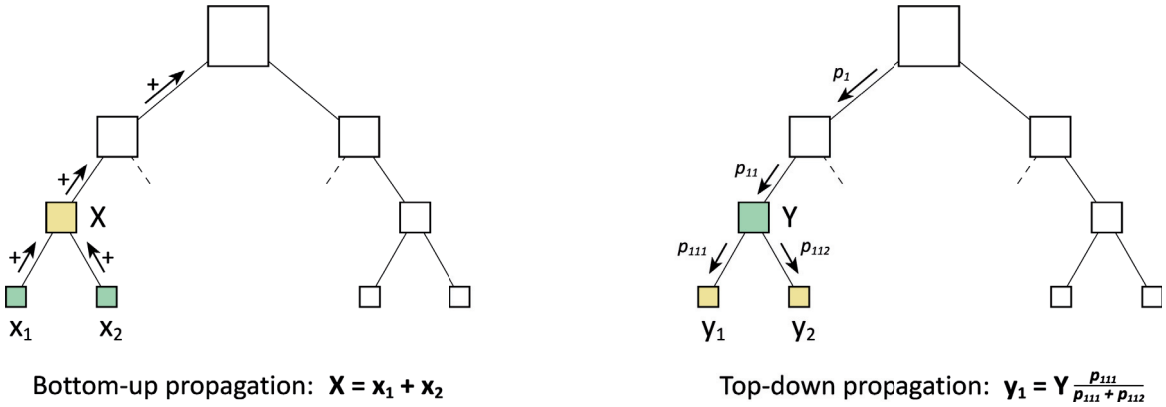

Figure 8.3: Basic rules for bottom-up and top-down propagation

available at the Karlsruhe Institute of Technology (KIT) since 2016 relies on 24,000 cores (KIT, 2016). Although it was not used for scheduling the new HIPAT model that was developed in this thesis and will be discussed in chapter 9, the application of a high-performance computer provides significant potential for further increasing the complexity of transport models operating at European scale.

With regard to the hierarchical data structure underlying the HIPAT model, and with regard to the propagation rules (cf. Fig. 8.3), it is important to stress that precisely these propagation rules open the door for possible dirty reads and dirty writes (i. e. inconsistent data access).[8] This can be an obstacle for scheduling the trip distribution model underlying the HIPAT model in parallel.

For a standard distribution model computing a full O/D trip matrix, this problem can be systematically avoided by implementing the model straightforwardly by means of two for-loops.[9] For the hierarchical trip distribution model, in contrast, inconsistent data access cannot be ruled out, given that each propagation rule changes several data elements in the tree. In this case, the data elements have to be protected by so-called locks, ensuring that only one thread can access an element at the same time.[10] It should be stressed that the strategy for implementing the required locks has to be well thought out in order to avoid deadlocks.[11]

[8]Dirty reads and dirty writes are a well-known problem in parallel computing. The problem can occur if two or more threads access and manipulate the same data element simultaneously.

[9]The for-loops can be scheduled in parallel without significant knowledge of parallel programming techniques by using, for instance, the OpenMP application programming interface (Chandra et al., 2001).

[10]If two threads want to access a protected data element, the second thread is blocked until the first thread holding the lock releases the lock (cf., for example, Goetz, 2008).

[11]A deadlock is an undesired state in parallel computing where two or more threads are blocking each other.

8.3 Hierarchical accessibility indicator problem

The hierarchical modelling approach has an important disadvantage. In comparison to a standard matrix encompassing all possible pairs of travel zones, the hierarchical matrix only encompasses some pairs due to the modelling of the O/D relations at different aggregation levels. Accordingly, it is not possible to determine the accessibility of a travel zone following the standard approach. In this dissertation, this is referred to as the "hierarchical accessibility indicator problem". In order to determine the accessibility indicator for Karlsruhe (DE122), for instance, all O/D relations that are somehow related to this travel zone have to be picked out of the hierarchical matrix and analysed.

The hierarchical accessibility indicator problem is solved by a recursive algorithm that is outlined in pseudo code at the end of this section. The algorithm, relying on fragmentary indicators and sophisticated propagation rules, is rather complex, and the sections before therefore provide more background information. It is first described how the standard accessibility indicator, computed at a single hierarchical level, can be computed in an efficient way. The problem of computing the hierarchical indicator is subsequently introduced, then the propagation rules are developed before the algorithm is described in detail.

8.3.1 Efficient computation of the standard accessibility indicator

The hierarchical accessibility indicator refers to the competition-enhanced accessibility indicator that was introduced in section 3.16 satisfying the following equations:

$$A_i^{(3)} = \sum_{j=1}^{n} \frac{e_j}{V_j^{(2)}} f_{ij} \qquad \text{with} \qquad V_j^{(2)} = \sum_{k=1}^{m} \frac{g_k}{J_k} f_{kj} \qquad \text{and} \qquad J_k = \sum_{l=1}^{n} e_l f_{kl} \tag{8.1}$$

where $A_i^{(3)}$ refers to the standard accessibility indicator computed for origin i, e_j to trip endings in destination j, f_{ij} to the deterrence indicator for the respective O/D relation, g_i to trip generation in origin i, and m or n to the number of travel zones ($m = n$).

If the accessibility indicator were computed for one travel zone following a straightforward implementation by means of three nested for-loops, the complexity would be $O(n^3)$.[12] This would already exceed the capacity of common computers if only n=1,000 zones were considered in the model. The more promising implementation for computing the accessibility indicator is therefore an iterative approach. It basically consists of three steps in which each for-loop is passed consecutively. This reduces the overall complexity for computing the competition-weighted accessibility indicator to approximately $O(3n)$.[13]

It should be stressed that the two outlined complexity values ($O(n^3)$, $O(3n)$) refer to the computation of the accessibility indicator for one origin by a standard transport modelling approach computing a full trip matrix. In this case, the accessibility indicator is computed at a single aggregation level encompassing all travel zones (i. e. all accessible destinations) and all O/D relations.

[12] The reason is that intermediate results like the indicator J_k in (8.1) are not computed only once but several times.

[13] In this case, the indicator J_k in (8.1) is computed only once and then stored temporarily. Likewise $V_i^{(2)}$.

8.3.2 Root of the problem

The problem of computing the hierarchical accessibility indicator originates from the fragmentary visibility of O/D relations at each aggregation level. In a hierarchical trip distribution model the O/D relation Karlsruhe/Stuttgart, for instance, is modelled at NUTS-3 level, whereas trips from Karlsruhe to Lyon, which are part of the aggregated O/D relation Baden-Württemberg/Centre-Est, are modelled at NUTS-1 level. Thus, the number of accessible destinations cannot be derived straightforwardly, neither for Karlsruhe nor for Baden-Württemberg nor for any other travel zone. It it obvious that only a fraction of these trips from Baden-Württemberg to Centre-Est originate in Karlsruhe. In consequence, the hierarchical accessibility indicator problem cannot be solved in a similar way to the above-mentioned standard algorithm for a standard model. It is therefore necessary to revise the standard algorithm in order to compute the hierarchical accessibility indicator in an efficient way. For instance, even if the overall transport model relies on more than one hundred thousand disaggregated travel zones, the hierarchical accessibility indicator is ideally built on only a few hundred O/D relations.

For each of these O/D relations a fragment of the overall accessibility indicator is computed. Moreover, each computed fragment is linked to a specific travel zone and has to be propagated bottom-up to every parent node in the hierarchical tree and top-down to every subsequent node. For instance, the fragment computed for the O/D relation Karlsruhe/Stuttgart is linked to the origin Karlsruhe. It can be directly applied when deriving the overall accessibilities for the city of Karlsruhe and for the respective city districts at LAU-2 level. However, it cannot be directly applied when deriving the overall accessibility for the federal state of Baden-Württemberg. For these reasons, the fragments computed for each O/D relation have to be processed in a specific way by propagation rules.

8.3.3 Solution concept for a simple example

Referring to the previous example where trip estimates between Karlsruhe and Stuttgart are computed at NUTS-3 level while trip estimates between Baden-Württemberg and Centre-Est are computed at NUTS-1 level, a reduced example is discussed in the following that refers to the four travel zones Karlsruhe (KA), Stuttgart (ST), Baden-Württemberg (BW) and Centre-Est (CE). For the purpose of simplification it is assumed that Baden-Württemberg only consists of Karlsruhe and Stuttgart, i. e. all travellers living in Baden-Württemberg either live in Karlsruhe or in Stuttgart. The reduced example consists of the following O/D relations:

- at NUTS-3 level: KA/KA, KA/ST, ST/KA, ST/ST; and
- at NUTS-1 level: BW/CE, CE/BW.

The starting point is the standard accessibility indicator $A_i^{(3)}$ given by (8.1) which refers to an origin i. For the purpose of simplification, for this example, we only discuss the first part of (8.1) assuming $V_j^{(2)} = 1 \ \ \forall j$. The accessibility for an origin i can then be computed according to (8.2) where $\tilde{A}(ij)$ is a fragmentary accessibility indicator computed for O/D relation ij.

$$A_i = \sum_{j=1}^{n} \tilde{A}(ij) \qquad \text{with} \qquad \tilde{A}(ij) = e_j f_{ij} \tag{8.2}$$

Consider travellers living in Karlsruhe. They can explore their own city, they can travel to Stuttgart and to Centre-Est. According to (8.2) the accessibility indicator for Karlsruhe A_{KA} is computed as follows:

$$A_{\mathrm{KA}} = \tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST}) + \tilde{A}(\mathrm{KA/CE}) \tag{8.3}$$

However, the issue in (8.3) is that the O/D relation KA/CE is not modelled in our example so that the fragmentary indicator $\tilde{A}(\mathrm{KA/CE})$ is not computed. For the computation of this missing element it is important to understand that trips from Karlsruhe to Centre-Est are modelled as part of the respective O/D relation at NUTS-1 level. In this specific case, the accessibility of Centre-Est is identical for Baden-Württemberg and for Karlsruhe. Hence, Equation (8.4) holds true and the accessibility indicator A_{KA} can be computed according to (8.5).

$$\tilde{A}(\mathrm{KA/CE}) = \tilde{A}(\mathrm{BW/CE}) \tag{8.4}$$

$$A_{\mathrm{KA}} = \tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST}) + \tilde{A}(\mathrm{BW/CE}) \tag{8.5}$$

Consider travellers living in Baden-Württemberg. They can explore their own region and they can travel to Centre-Est. According to (8.2) the accessibility indicator for Baden-Württemberg A_{BW} is computed as follows:

$$A_{\mathrm{BW}} = \tilde{A}(\mathrm{BW/BW}) + \tilde{A}(\mathrm{BW/CE}) \tag{8.6}$$

However, the issue in (8.6) is that the O/D relation BW/BW is not modelled in our example. With regard to the missing element $\tilde{A}(\mathrm{BW/BW})$ it is important to understand that the O/D relation BW/BW is modelled at NUTS-3 level by four relations: two relations referring to Karlsruhe and two to Stuttgart. Hence, four fragmentary indicators are computed: $\tilde{A}(\mathrm{KA/KA})$, $\tilde{A}(\mathrm{KA/ST})$, $\tilde{A}(\mathrm{ST/KA})$, $\tilde{A}(\mathrm{ST/ST})$. These indicators have to be combined according to the concept of weighted sum (8.7) in order to compute the missing indicator $\tilde{A}(\mathrm{BW/BW})$:

$$\tilde{A}(\mathrm{BW/BW}) = \omega_{\mathrm{KA}}\big(\tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST})\big) + \omega_{\mathrm{ST}}\big(\tilde{A}(\mathrm{ST/KA}) + \tilde{A}(\mathrm{ST/ST})\big) \tag{8.7}$$

$$\omega_{\mathrm{KA}} = g_{\mathrm{KA}}/g_{\mathrm{BW}} \quad \text{and} \quad \omega_{\mathrm{ST}} = g_{\mathrm{ST}}/g_{\mathrm{BW}} \quad \text{with} \quad g_{\mathrm{BW}} = g_{\mathrm{KA}} + g_{\mathrm{ST}} \tag{8.8}$$

where g_i is the generated trip demand in travel zone i so that $\omega_{\mathrm{KA}}, \omega_{\mathrm{ST}}$ are two weighting factors ($\omega_{\mathrm{KA}} + \omega_{\mathrm{ST}} = 1$, $\omega_i \geqslant 0$). The necessity to apply these weighting factors becomes clear when it is considered that the two fragmentary indicators $\tilde{A}(\mathrm{KA/ST})$ and $\tilde{A}(\mathrm{ST/ST})$ both assess the accessibility of the destination Stuttgart. If these two fragmentary indicators were simply added together, Stuttgart would be double-counted as a destination. Using the weighting factors ω_i, the accessibility indicator A_{BW} can be computed according to (8.9).

$$A_{\mathrm{BW}} = \omega_{\mathrm{KA}}\big(\tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST})\big) + \omega_{\mathrm{ST}}\big(\tilde{A}(\mathrm{ST/KA}) + \tilde{A}(\mathrm{ST/ST})\big) + \tilde{A}(\mathrm{BW/CE}) \tag{8.9}$$

In this example, two propagation rules have been introduced. Top-down propagation of a fragmentary accessibility indicator from NUTS-3 to NUTS-1 level is carried out based on a rather simple rule which is outlined in (8.4). In contrast, bottom-up propagation of a fragmentary accessibility indicator has to be carried out based on a weighting factor as outlined in (8.7).

8.3.3.1 Step by step calculation of the accessibility indicators

Table 8.2 outlines an algorithm for the step by step calculation of the two accessibility indicators A_{KA} and A_{BW}. The calculation starts with $A_{\mathrm{BW}} = 0$ and $A_{\mathrm{BW}} = 0$. The six O/D relations of the discussed example are passed consecutively. In each step a fragmentary accessibility indicator $\tilde{A}(ij)$ is computed which is linked to an origin i. It is added to A_i so that the values of A_{KA} and A_{BW} expand (until they reach their final values in step six). The computed fragmentary indicator is then propagated top-down to every subsequent node in the hierarchical tree and bottom-up to every parent node using the two propagation rules outlined by (8.4) and (8.7). While top-down propagation is rather simple, bottom-up propagation applies a weighting factor ω_i that is computed according to (8.8).

Table 8.2: Step by step calculation of accessibility indicators A_{KA} and A_{BW}

Step	$\tilde{A}(ij)$ [1]	Current values
1	$\tilde{A}(\mathrm{KA/KA})$	$A_{\mathrm{KA}} = \tilde{A}(\mathrm{KA/KA})$ $A_{\mathrm{BW}} = \omega_{\mathrm{KA}}\tilde{A}(\mathrm{KA/KA})$
2	$\tilde{A}(\mathrm{KA/ST})$	$A_{\mathrm{KA}} = \tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST})$ $A_{\mathrm{BW}} = \omega_{\mathrm{KA}}\big(\tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST})\big)$
3	$\tilde{A}(\mathrm{ST/KA})$	$A_{\mathrm{KA}} = \tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST})$ $A_{\mathrm{BW}} = \omega_{\mathrm{KA}}\big(\tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST})\big) + \omega_{\mathrm{ST}}\tilde{A}(\mathrm{ST/KA})$
4	$\tilde{A}(\mathrm{ST/ST})$	$A_{\mathrm{KA}} = \tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST})$ $A_{\mathrm{BW}} = \omega_{\mathrm{KA}}\big(\tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST})\big) + \omega_{\mathrm{ST}}\big(\tilde{A}(\mathrm{ST/KA}) + \tilde{A}(\mathrm{ST/ST})\big)$
5	$\tilde{A}(\mathrm{BW/CE})$	$A_{\mathrm{KA}} = \tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST}) + \tilde{A}(\mathrm{BW/CE})$ $A_{\mathrm{BW}} = \omega_{\mathrm{KA}}\big(\tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST})\big) + \omega_{\mathrm{ST}}\big(\tilde{A}(\mathrm{ST/KA}) + \tilde{A}(\mathrm{ST/ST})\big) + \tilde{A}(\mathrm{BW/CE})$
6	$\tilde{A}(\mathrm{CE/BW})$	$A_{\mathrm{KA}} = \tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST}) + \tilde{A}(\mathrm{BW/CE})$ $A_{\mathrm{BW}} = \omega_{\mathrm{KA}}\big(\tilde{A}(\mathrm{KA/KA}) + \tilde{A}(\mathrm{KA/ST})\big) + \omega_{\mathrm{ST}}\big(\tilde{A}(\mathrm{ST/KA}) + \tilde{A}(\mathrm{ST/ST})\big) + \tilde{A}(\mathrm{BW/CE})$

[1] Fragmentary accessibility indicator computed in this step.

In the first step, $\tilde{A}(\mathrm{KA/KA})$ is computed referring to the origin Karlsruhe. It is added to A_{KA} and propagated bottom-up to the NUTS-1 level following the concept of weighted sum. Thus, $\omega_{\mathrm{KA}} \cdot \tilde{A}(\mathrm{KA/KA})$ is added to A_{BW}. In the second step, $\tilde{A}(\mathrm{KA/ST})$ is computed and added to A_{KA}. Using the weighting factor ω_{KA} it is propagated bottom-up and added to A_{BW}. In steps 3 and 4, two fragmentary indicators referring to the origin Stuttgart are computed that are not relevant for A_{KA}. Using the weighting factor ω_{ST} they are propagated bottom-up and added to A_{BW}. In the fifth step, $\tilde{A}(\mathrm{BW/CE})$, referring to the origin Baden-Württemberg, is computed and added to A_{BW}. It is propagated top-down to the NUTS-3 level and added to A_{KA}. In the last step, $\tilde{A}(\mathrm{CE/BW})$ is computed, which is neither relevant for A_{BW} nor for A_{KA}.

It should be noted that all fragmentary accessibility indicators $\tilde{A}(ij)$ as well as all accessibility indicators A_i are stored temporarily as they will be needed for the calculation of the trip distribution.

8.3.4 Expansion of example and solution concept

Dropping the simplifying assumption $V_j^{(2)} = 1 \quad \forall j$ the computation of the hierarchical accessibility indicator is carried out in a similar way to (8.1), i. e. iteratively based on fragmentary indicators. Firstly, the indicator J_k is computed, then the indicator $V_j^{(2)}$, and subsequently the indicator $A_i^{(3)}$. The fragmentary indicators are computed for each O/D relation and their processing in the hierarchical tree-based data structure requires propagation rules. In the above example, the two rules for top-down propagation (8.4) and bottom-up propagation (8.7) that are applied for J_k and $A_i^{(3)}$ have been introduced.[14] For the propagation of the fragmentary indicators $\tilde{V}(ij)$ related to $V_j^{(2)}$ two different rules are required, as explained in the following example:

$$\tilde{V}(\text{BW/BW}) = 1 \cdot \left(\tilde{V}(\text{KA/KA}) + \tilde{V}(\text{KA/ST})\right) + 1 \cdot \left(\tilde{V}(\text{ST/KA}) + \tilde{V}(\text{ST/ST})\right) \tag{8.10}$$

$$\tilde{V}(\text{CE/KA}) = \tilde{V}(\text{CE/BW}) \cdot \left(e_{\text{KA}}/e_{\text{BW}}\right) \tag{8.11}$$

where e_j refers to estimated trip endings in destination j. In this case, the bottom-up propagation rule (8.10) is quite simple, as all trips ending in Karlsruhe and Stuttgart, of course, end in Baden-Württemberg as well. In contrast, the top-down propagation rule (8.11) relies on a split factor paying attention to the fact that trips ending in Baden-Württemberg either end in Karlsruhe or in Stuttgart.

8.3.5 Formal description of the algorithm

The hierarchical accessibility indicator problem can be solved efficiently by the subsequent computation of the three accessibility indicators J_k, V_j and A_i (8.1). Accordingly, six propagation rules have to be defined that are implemented by the Algorithms 2, 3 and 4 shown below in pseudocode.[15] The computation of the hierarchical accessibility indicator requires the two input datasets R and Z:

- Let $R = \{r_1, r_2, ..., r_k, ..., r_t\}$ be the pre-determined set of O/D relations. For each of these relations let $r_k = (i, j, f_{ij})$, where (i, j) are the origin and the destination belonging to this specific relation, and $f_{ij} \in \mathbb{R}^+$ is the deterrence indicator computed by the deterrence model.
- Let $Z = \{z_1, z_2, ..., z_i, ..., z_n\}$ be the pre-determined set of travel zones. For each of these travel zones let $z_i = (A_i, V_i, J_i, g_i, e_i)$, where $z_i \in \mathbb{R}^5$. A_i, V_i and J_i identify the three accessibility indicators, g_i refers to trip generation and e_i to trip endings in travel zone i.

[14] Given that both indicators are linked to the origins, the applied propagation rules are identical.

[15] Pseudocode is a standard concept used in computer science in order to briefly outline the operation principle of an algorithm. It is intended for human reading. In general, pseudocode follows the structural conventions of a programming language. It also uses elements of a natural language and mathematical notation to describe details of the algorithm. A standard syntax for pseudocode does not exist. For more information see, for example, Mehlhorn and Sanders (2008).

The starting point for the overall calculation of the hierarchical accessibility indicator is given by Algorithm 1. This algorithm first initialises the three sets of accessibility indicators J, V and A with zeros (lines 6 to 8), then executes the three Algorithms 2, 3 and 4 that compute J, V and A one after another (lines 3 to 5). Each of the three algorithms loops over the whole set of O/D relations. In the respective line 3, the values of i, j, f_{ij} are determined for the current relation and in line 4, a fragmentary accessibility indicator b is computed. The computed fragmentary indicator is then added in line 5 to the current value of the accessibility indicator for the respective travel zone in the hierarchical tree. In a subsequent step, two recursive procedures are executed (lines 6, 7) in order to propagate the computed fragmentary indicator bottom-up to parent nodes and top-down to successor nodes in the tree.

Algorithm 1: Main entry point for computing the hierarchical accessibility indicator

```
1 def main(R,Z):
2 |   initialise JVA(Z)
3 |   compute J(R,Z)
4 |   compute V(R,Z)
5 |   compute A(R,Z)

6 def initialise JVA(Z):
7 |   foreach zone z in Z do
8 |   |   (0, 0, 0, −, −) → z
```

Algorithm 2: Recursive propagation rules for computing the indicator J

```
1 def compute J(R,Z):
2 |   foreach relation r in R do
3 |   |   (i, j, f_ij) ← r
4 |   |   b  = f_ij * e_j
5 |   |   J_i = J_i + b
6 |   |   propagationRuleTopDown(i,b)
7 |   |   propagationRuleBottomUp(i,b)

8 def propagationRuleTopDown(i,b):
9 |   foreach successor node k of node i do
10|   |   J_k = J_k + b
11|   |   propagationRuleTopDown(k,b)

12 def propagationRuleBottomUp(i,b):
13|   foreach parent node k of node i do
14|   |   p  = b * g_k/g_i
15|   |   J_k = J_k + p
16|   |   propagationRuleBottomUp(k,p)
```

Algorithm 3: Recursive propagation rules for computing the indicator V

```
1  def compute V(R,Z):
2      foreach relation r in R do
3          (i, j, f_ij) ← r
4          b  = f_ij * g_i * 1/J_i
5          V_j = V_j + b
6          propagationRuleTopDown(j,b)
7          propagationRuleBottomUp(j,b)

8  def propagationRuleTopDown(j,b):
9      foreach successor node k of node j do
10         p  = b * e_k/e_j
11         V_k = V_k + p
12         propagationRuleTopDown(k,p)

13 def propagationRuleBottomUp(j,b):
14     foreach parent node k of node j do
15         V_k = V_k + b
16         propagationRuleBottomUp(k,b)
```

Algorithm 4: Recursive propagation rules for computing the indicator A

```
1  def compute A(R,Z):
2      foreach relation r in R do
3          (i, j, f_ij) ← r
4          b  = f_ij * e_j * 1/V_j
5          A_i = A_i + b
6          propagationRuleTopDown(i,b)
7          propagationRuleBottomUp(i,b)

8  def propagationRuleTopDown(i,b):
9      foreach successor node k of node i do
10         A_i = A_i + b
11         propagationRuleTopDown(k,b)

12 def propagationRuleBottomUp(i,b):
13     foreach parent node k of node i do
14         p  = b * g_k/g_i
15         J_k = J_k + p
16         propagationRuleBottomUp(k,p)
```

8.4 General conclusions

8.4.1 Summary

A well-known problem in transport modelling is that trip estimates between neighbouring regions cannot be computed with sufficient accuracy. Given an average diameter of over 50 km for the NUTS-3 regions, European transport models operating at this level are not particularly suited for modelling regional trips up to 100 km. Taking into account that the travel distance is computed based on the centroids, the travel distance for two neighbouring NUTS-3 regions can vary between a few and one hundred kilometres. An obvious solution enabling the modelling of regional trips with a higher accuracy is to apply more disaggregated travel zones, for instance, LAU-2 regions with an average diameter of only 5 km. However, the complexity of the trip matrix and, accordingly, the complexity of the trip distribution model increase quadratically with the number of travel zones. Thus, their number can only be increased if the complexity of the trip matrix can be reduced to the same extent.

Recent European transport models operating at NUTS-3 level, like the TRANS-TOOLS-, the VACLAV- and the PAT models, which follow a standard approach, only deal with about 1,500 travel zones. However, these models run into complexity problems as they model all trips at NUTS-3 level, thus leading to 2.25 million O/D relations in total. In contrast, hierarchical models relying on the idea of estimating long-distance and regional trips at different aggregation levels constitute a way forward in European transport modelling. If a hierarchical approach is followed, it seems to be feasible to consider even more than 100,000 travel zones, given that huge complexity savings can be expected according to Table 8.1. As over 90% of the trips carried out with passenger cars are shorter than 50 km[16], regional trips have to be modelled at LAU-2 level rather than at NUTS-3 level.

The implementation of a hierarchical trip distribution model is not straightforward. The first challenge to be faced is to determine the optimal set of O/D relations referring to different aggregation levels. Therefore an adequate neighbourhood criterion is required in order to decide which relations have to be disaggregated and which relations can be modelled at a higher aggregation level. It should also be stressed that the neighbourhood criterion generating the set of O/D relations should be applied in advance. This helps to validate the set and to improve the criterion. In addition, this allows the implementation of optimised and thread-safe data structures that are required to solve the hierarchical trip distribution problem efficiently by parallel computing.

The main challenge of the hierarchical trip distribution model is the computation of the hierarchical accessibility indicator, given that the O/D relations are modelled at different levels. For instance, some relations are modelled at NUTS-1 level, some at NUTS-3 level and others at LAU-2 level. In this context, it is important to understand that the overall accessibility indicator of a travel zone is a compound of several fragments. These fragments are initially computed for each O/D relation. They are either linked to the origin or to the destination and have to be propagated bottom-up to every parent travel zone according to the hierarchical tree structure and top-down to every subsequent travel zone. The propagation is carried out based on more or less sophisticated rules that are implemented by three recursively applied algorithms. Once the hierarchical accessibility indicators are computed, the

[16] E. g. 90.8% in Germany (infas and DLR, 2010a), 93.3% in Denmark (Christiansen, Hjalmar, 2011), 92.0% in the Netherlands (MVW, 2010), 94.7% in the United Kingdom (DfT, 2008), and 98.1% in Spain (Pérez Lou et al., 2007).

computation of the trip estimates and the computation of the modal split values for each O/D relation is straightforward.

8.4.2 Outlook and continuing challenges

Besides the important challenge of computing the hierarchical accessibility indicators, a minor challenge arises in the computation of traffic loads at the level of network links within the assignment model. For example, trip estimates related to the aggregated O/D relation Baden-Württemberg/Centre-Est are treated exclusively as trips from the centre of Stuttgart to the centre of Lyon given that these two cities or, to be more precise, their respective centroids represent the two aggregated NUTS-1 regions. The aggregated O/D relation, however, also includes trips from Karlsruhe to Clermont-Ferrand, for instance, that have a different routing. Aggregated O/D relations should therefore be disaggregated in order to improve the quality of the computed traffic loads. This task can be carried out following the approach described by Schoch (2004) by breaking down the trip estimates T_M referring to the aggregated O/D relation M to the respective sub-relations $s \in M$ as follows:

$$T_s = p_s T_M \qquad \text{with} \qquad p_s = \mathrm{GM}(s) \left(\sum_{s \in M} \mathrm{GM}(s) \right)^{-1} \tag{8.12}$$

where T_s identifies the trip estimates referring to a disaggregated sub-relation (s) and GM to an adequate gravity modelling approach, e. g. the zero-constrained gravity model satisfying (4.11) (cf. sec. 4.2).

It is important to emphasise that T_s are computed during the runtime of the assignment model when deriving the traffic loads at link level. The computation of the proration factors p_s according to (8.12), however, requires information on the travel costs referring to the shortest paths between the origins and the destinations of all disaggregated O/D relations. This information is not precomputed and must therefore also be determined during the runtime. However, shortest path retrieval for every individual O/D relation is extremely inefficient. A faster but simplified method could be to compute the proration factors p_s as follows:

$$p_s = f(d_s) \left(\sum_{s \in M} f(d_s) \right)^{-1} \tag{8.13}$$

where f is an adapted deterrence function and d_s the crow-flies distance between the two travel zones.

Another point relates to the computation of well-known output indicators, like mileage travelled by the population living in a NUTS-3 region, e. g. Karlsruhe. For a standard transport model computing a complete trip matrix and a complete pkm matrix referring to a single aggregation level, this indicator can be produced by adding up all cells in the respective row of the pkm matrix. For the hierarchical transport model, however, O/D relations are modelled at different levels. While those O/D relations that are modelled at a more disaggregated level, like trips between the city districts at LAU-2 level, can be added together, trips modelled at a more aggregated level, like trips from Baden-Württemberg to Centre-Est, have to be disaggregated.

It should be stressed in this context that the rules applied for deriving the indicator mileage travelled by the population living in a travel zone are very similar to the propagation rules that were applied to compute the hierarchical accessibility indicator. For instance, in order to disaggregate the O/D relation Baden-Württemberg/Centre-Est and to derive the number of trips originating in Karlsruhe and Stuttgart, the respective shares can be computed according to (8.8), based on the generated trip demand in Karlsruhe and in Stuttgart. It should be also stressed that the computation of all required output indicators should be carried out during the runtime of the transport model, given that the processing of the hierarchical matrices is complex and requires additional indicators, including information on the trip demand generated in each travel zone, information on the travel costs for all O/D relations and information on the centroids of each travel zone. The set of output indicators could include several trip matrices referring to different aggregation levels, e. g. a complete trip matrix at NUTS-3 level and several trip matrices at LAU-2 level with a limited geographic focus.[17]

[17]The computation of these additional output indicators is not discussed within this thesis.

Chapter 9

The HIPAT modelling approach

This chapter presents the hierarchical, integrated passenger transport (HIPAT) modelling approach that has been developed in this thesis. The HIPAT approach facilitates the modelling of European passenger transport at LAU-2 level encompassing more than 100,000 travel zones. It was successfully implemented for a prototype model with a limited scope. The chapter is structured as follows: section 9.1 provides an overview of the main objectives and challenges of the HIPAT approach. In the next two sections 9.2 and 9.3, the methodological improvements and the assumptions made to simplify the implementation of the prototype are discussed. Section 9.3.3 describes the process of generating the input database for the prototype. Section 9.4 discusses the model application and analyses the computed results. The last section summarises the main conclusions and provides an outlook towards the implementation of a more complete model for Europe.

9.1 Recapitulation of the main challenges and objectives

Current EU passenger transport models operating at NUTS-2 and NUTS-3 levels lack an accurate modelling of regional trips. This is due to the large size of the travel zones and accordingly to the large share of intra-zonal trips. The average diameter of a NUTS-2 region is over 125 km, and that of a NUTS-3 region is over 50 km. Besides the well-known problem of modelling intra-zonal trips, it has to be stressed that inter-zonal trips between neighbouring travel zones can also not be modelled with sufficient accuracy. The reason is that the trip demand between two travel zones depends on the travel distance that is computed on the basis of the two centroids. A centroid, however, can be located anywhere inside a travel zone. This means that the representative travel distance between two neighbouring, average-sized NUTS-3 regions can vary between a few and 100 km. Accordingly, the trip demand between neighbouring travel zones is very likely to be underestimated or overestimated.

The inaccurate modelling of regional trips within the trip distribution step has several disadvantages for the modelling of traffic loads at link level in the assignment step. For this reason, the IPAT model discussed in section 6.2 does not provide assignment maps. It replaces the assignment step by a lighter conversion step in order to translate the trip matrices into passenger-kilometres and vehicle-kilometres. For the same reason, Schoch (2004) revised the assignment step of the NUTS-3-based VACLAV model by disaggregating the trip matrix and simulating the distribution of intra-zonal trips inside the travel zones. This improves the quality of the computed assignment results but does not solve the general issue of modelling intra-zonal and regional trips within the trip distribution model.

It is therefore essential to reduce the share of intra-zonal trips to a minimum by reducing the size of the travel zones in order to improve existing transport models. However, increasing the number of travel zones from 300 NUTS-2 regions to more than 100,000 LAU-2 regions raises serious complexity problems as the number of O/D relations increases from 90,000 to over 10,000,000,000. It is quite obvious that a European transport model can only be implemented at LAU-2 level if the number of O/D relations can be reduced significantly. Within this thesis, this is facilitated by the developed HIPAT approach in which trip estimates between two travel zones are computed at different hierarchical levels depending on the travel distance. In most cases, for instance, long-distance trips of over 1000 km can be computed at NUTS-1 level with sufficient accuracy, while regional trips below 50 km should be generally calculated at LAU-2 level.

At this point, it is necessary to draw attention to a specific parameter in the equation of the revised deterrence model (5.21) that has been developed within this thesis in section 5.4, namely the scaling function $\rho_d(x)$ modelling the number of accessible destinations in relation to the travel distance x and the average diameter d of the travel zones. Considering a trip distance of 250 km, for instance, 40 destinations can be reached from the origin if the model is applied at NUTS-3 level. In contrast, only 16 destinations can be reached if the model is applied at NUTS-2 level.[1] It is therefore necessary that the scaling function adjusts the deterrence model to the relevant hierarchical level. This property is an essential prerequisite. It facilitates consistent application of the deterrence model and accordingly the HIPAT model at different hierarchical levels. In contrast, a transport model relying on a standard deterrence function cannot be transferred between different hierarchical levels.

The methodological development of the revised HIPAT modelling approach raises several questions. For its application, an input database is required that is similar to the input database of the IPAT model - but it is required at LAU-2 level. As the availability of regional indicators at LAU-2 level is generally not given at European scale, these indicators have to be estimated by downscaling. For this purpose, tailored downscaling approaches have to be developed. A method of achieving this has been described for the indicator "Number of workplaces" for the LAU-2 regions in section 7.3.1. A further challenge stems from the fact that the two local levels, LAU-1 and LAU-2, are not consistently defined in Europe according to the NUTS classification scheme. In order to close existing gaps and to improve the quality of the applied zoning system, artificial travel zones have to be generated to complement the official definition of LAU regions. This is essential in order to subdivide large city regions that are defined at NUTS-3 level into several artificial city districts corresponding to LAU-2 level. The developed approach discussed in section 7.3.2 may also be applied in a future version of the HIPAT model for disaggregating the existing LAU-2 regions towards more homogeneous grid cells.

The main objectives of this chapter are to provide a general proof of concept of the HIPAT modelling approach and to demonstrate the advantages of modelling transport at LAU-2 level. This is facilitated by implementing a prototype model relying on 33,191 travel zones. The prototype focuses on a single transport demand segment and an important European transport corridor. Regions inside this corridor are modelled at LAU-2 level, and regions outside at NUTS-2 and NUTS-0 levels.

[1] These numbers of accessible destinations $\nu_d(x)$ have been computed following Equation (5.8), assuming an average diameter of 50 km for the NUTS-3 regions and 125 km for the NUTS-2 regions. It holds that $\nu_d(x) = 8(x/d)$.

9.2 Methodological improvements

This section discusses the methodological improvements of the HIPAT model compared to its predecessor, the IPAT model. In principle, both transport modelling approaches are very similar with regard to the methodology of the trip generation model, the integrated trip distribution/ modal split model and the conversion model.[2] An essential improvement of the HIPAT approach is the hierarchically structured zoning system, in which the travel zones and accordingly the O/D relations are ordered in a specific way. In contrast to the IPAT model, the HIPAT approach is complemented by a fully operating network assignment model. Another major improvement of the HIPAT approach is the partial integration of the trip distribution, the modal split and the assignment models by applying these three submodels in several incremental steps. This facilitates modelling of the impacts of road congestion already at the level of the trip distribution model.[3] The three submodels could be almost fully integrated if a large number of incremental steps were chosen. It would then be possible to model effects with the HIPAT model that are usually reserved for microscopic models, like the "first come, first served" principle with regard to the booking of hotels for vacation trips (e. g. tourists' decisions depend on the decisions of other tourists).

It is important to emphasise that the hierarchically structured zoning system has to be established in advance. Within this step, artificial LAU-1 and LAU-2 regions are generated and regional indicators are disaggregated from NUTS-3 to LAU-2 level. The neighbourhood criterion is then applied in order to determine the optimal set of O/D relations. Besides these two input datasets, network models are required as well as modelling and calibration coefficients. In the following section, the structure of the HIPAT model and the methodological improvements are discussed in more detail.

9.2.1 Structure of the HIPAT model

Figure 9.1 outlines the overall structure of the HIPAT model, including the four sub-models, trip generation, indicators as inherited from the IPAT model (cf. sec. 6.2.2), distribution/ modal split, and assignment/ conversion. In addition, the main data flows are shown, drawing specific attention to the three indicators of generated trip demand, number of available destination opportunities, and current travel impedances. These indicators are continuously updated if the HIPAT model is applied in incremental steps. Furthermore, the most important input- and output datasets are shown.

In the first step, the HIPAT model generates the origin trip demand and estimates the number of destination opportunities for all travel zones at the most disaggregated level (i. e. at LAU-2 level). These demand and supply indicators are then added up for all aggregated travel zones following the most basic rule for bottom-up propagation (cf. Fig. 8.3, p. 166). The generated trip demand can then be subdivided into small portions in order to compute the trip matrices and the network loads successively in incremental steps. Prior to each incremental step, the number of destination opportunities still available is updated for all travel zones. In addition, the set of travel impedances is updated, taking into account changing traffic loads in the networks and congestion effects.

[2] The conversion model replaces the assignment model. It translates the trip matrices into pkm and vkm on the basis of precomputed trip lengths (without route choice modelling). Furthermore, it computes aggregated output indicators, such as pkm per country.

[3] Following a standard four-step transport modelling approach, road congestion would only have an impact on the assignment model.

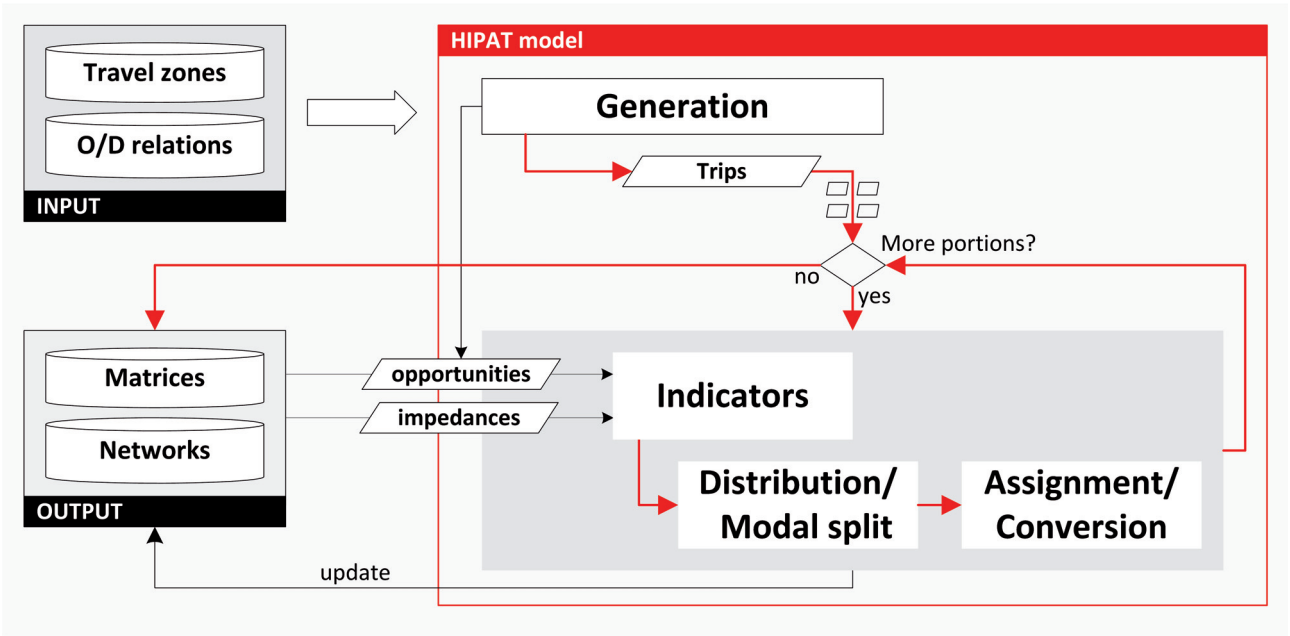

Figure 9.1: Structure of the HIPAT modelling approach

In the second step, the four indicators[4], GTT, EMC, deterrence value and EMC accessibility are recalculated considering changes in available destination opportunities and travel impedances. These indicators feed into the integrated trip distribution/ modal split model and into the assignment/ conversion model that follow in the third and in the fourth step.

The HIPAT model produces O/D-based output indicators in terms of trips, pkm and vkm as well as assignment results at the level of network links. In addition, the conversion model computes several aggregates of the O/D-based output indicators, such as pkm demand by mode at country level, that are particularly important in order to calibrate and to validate the HIPAT model.

9.2.2 Hierarchically structured zoning system

One of the most important decisions with regard to the implementation of any transport model concerns the definition of the travel zones. Accurate model results can only be expected if the travel zones are chosen in line with the homogeneity criteria discussed in section 7.1, and if the modeller is aware of the modifiable areal unit problem. This is a major challenge of the HIPAT modelling approach, which deals simultaneously with travel zones at different hierarchical levels. As these levels are not consistently defined for all European countries, artificial travel zones have to be generated to close existing data gaps. Following this approach, for instance, artificial city districts at LAU-1 and LAU-2 are generated. Having closed all data gaps, the hierarchical structure is established by linking each travel zone to its direct parent and children zones.[5] This is important for the propagation of indicators between the different hierarchical levels of travel zones. In addition, the hierarchically structured zoning system is an essential precondition in order to apply the neighbourhood criterion and to generate the set of O/D relations referring to different hierarchical levels.

[4]The four indicators are not shown by Figure 9.1. They are discussed in more detail in section 6.2.2 (Fig. 6.4, p. 102).

[5]For instance, the NUTS-1 region of Baden-Württemberg (DE1) is linked to its parent region at NUTS-0 level, Deutschland (DE), and to the NUTS-2 regions DE11, DE12, DE13 and DE14.

9.2.3 Neighbourhood criterion

The neighbourhood criterion is one of the key elements of the HIPAT modelling approach. It is applied in order to decide whether two travel zones are adjacent or not. In the current version of the HIPAT prototype model, the following criterion is applied:

- two regions are neighbours if the distance between their centroids is less than the sum of their diameters.

If this criterion suggests that two travel zones are not adjacent, trip estimates between them are directly computed. In the other case, both travel zones are disaggregated and the neighbourhood criterion is applied for the newly created O/D relations. This procedure is repeated until the lowest hierarchical level is reached (i. e. LAU-2 level). At this level of detail, all remaining O/D relations have to be computed, including intra-zonal trips as well as trips between neighbouring travel zones. Hence, the recursive application of the neighbourhood criterion produces all relevant O/D relations referring to different hierarchical levels (e. g. NUTS-0 and LAU-2 level). It is important to stress that the set of O/D relations is precomputed in order to enable the implementation of more efficient, static data structures in `Java`.[6]

The applied neighbourhood criterion (see above) corresponds to criterion N_2 defined in section 8.2.1 on page 164. It is certainly stronger than criterion N_1, which promises larger savings according to the outcomes of the theoretical example (cf. Tab. 8.1, sec. 8.1.2). However, inspection of the results produced by the two criteria showed several shortcomings. It emerged, for instance, that N_1 produces poor results for real travel zones with regard to the number of "false negatives" (i. e. two regions that should be classified as neighbours are not classified as neighbours). In consequence, such O/D relations are not disaggregated but rather modelled at an aggregated level and the derived trip estimates can be very inaccurate. The weak performance of N_1 in deciding on the basis of the two radii and the distance between the two centroids can be explained by the almost arbitrary shape of real administrative units and the almost arbitrary location of a centroid inside a region.[7] On the other hand, the stronger criterion N_2, which decides on the basis of the two diameters, produces many "false positives" (i. e. two regions that should not be classified as neighbours are classified as neighbours). Such O/D relations are then disaggregated as well, leading to an excess of detail and reducing the complexity savings of the hierarchical approach.

An open question of the HIPAT approach is the coverage of countries outside Europe that are currently modelled only at NUTS-0 level, like Russia. Given the large size of Russia, this leads to many false positives. For instance, all O/D relations from the Baltic countries Estonia, Latvia and Lithuania to Russia are modelled at LAU-2/NUTS-0 level. Thus, the elaboration of a more sophisticated neighbourhood criterion in order to reduce the number of false negatives and false positives could be the subject of further investigations. Another solution could be to cover countries outside Europe at a more detailed level.

[6] For instance, static data structures like arrays with a fixed size can be used in order to store large data tables, given that they commonly provide a better performance and less overhead than dynamic structures like hash maps.

[7] For instance, the two NUTS-2 regions DE13 and DE22 are not classified as neighbours by the N_1 criterion due to the large distance between their capitals Tübingen and Augsburg.

9.2.4 Hierarchical accessibility indicator

Modelling trip estimates between travel zones at different hierarchical levels disables the straightforward computation of the accessibility indicator for a travel zone following the standard approach. This leads to the formulation of the hierarchical accessibility indicator problem and its solution by three recursively applied algorithms that are discussed in section 8.2. The basic idea followed by the algorithm is to compute fragmentary accessibility indicators for each O/D relation and to propagate them bottom-up to parent nodes and top-down to successor nodes in the tree by the application of specific propagation rules.[8] It should be noted that propagation of the fragmentary indicators in the tree requires numerous and multiple data updates of the tree elements (i. e. read/write data access). It should therefore be carried out in parallel by so-called "worker threads".[9] Having computed the hierarchical accessibility indicator, the application of the trip distribution and the modal split model is straightforward.

9.2.5 Network assignment/ conversion model

A major improvement from the IPAT to the HIPAT model is the implementation of a network assignment model. It computes the travel impedances for each O/D relation and the traffic loads at the level of network links. Its development was carried out on the basis of the hypernet model that is part of the specific implementation of the IPAT model for the HIGH-TOOL model, called the PAD module (Van Grol et al., 2016). The implemented assignment model mainly relies on the free `Java` graph library `jgrapht` version 0.9.1.[10] It can be outlined by the following procedure, which is performed several times if the HIPAT model is applied in incremental steps (cf. Fig. 9.1):

1. For every link of the network modelling graph, impedance attributes and cost indicators (e. g. travel time and GTT) are computed based on the current congestion level and accordingly the current travelling speed.

2. For every origin node in the graph, the full topology of shortest paths to all destination nodes is generated.

3. For every O/D relation, the entire path is passed through in the backward direction. In this process, link-based traffic load attributes are updated along the trajectory as well as O/D-based travel impedance and GTT values.[11]

To speed up the assignment algorithm, all origin nodes are determined in advance and the full topology of shortest paths is only generated once for these nodes.[12]

The assignment model is complemented by a conversion model in order to translate the trip matrices into passenger-kilometres and vehicle-kilometres as well as to compute several aggregates of these O/D-based indicators. It should be stressed that the computation of aggregated output indicators is not

[8] The fragmentary indicators are either related to the origin or the destination of the corresponding O/D relation.

[9] To avoid dirty reads and dirty writes, data access has to be synchronised, i. e. only one thread is allowed to access a data element at the same time (cf. sec. 8.2).

[10] Online documentation available, URL: `http://jgrapht.org/` (Retrieved April 10, 2017).

[11] GTT and impedance values are required for translating the trip matrices into pkm, vkm, and expected travel costs.

[12] Given the small size of the LAU-2-based travel zones, two or more origins can be connected to the same feeding node in the network modelling graph.

straightforward, given that the O/D relations are modelled at different levels. For instance, in order to derive the pkm demand generated by the travel zone Karlsruhe (DE122), the pkm demand computed for the O/D relation Baden-Württemberg/Centre-Est modelled at NUTS-1 level is disaggregated to NUTS-3 level based on the generated trip demand for Karlsruhe in relation to the generated trip demand for Baden-Württemberg.

9.2.6 Incremental application

A major improvement of the HIPAT modelling approach compared to existing macroscopic transport models is the integration of the trip distribution, the modal split and the assignment model. For this purpose, the generated trip demand is broken down into small portions and the trip matrices and network loads are computed in several incremental steps. This facilitates modelling of impact of congestion already at the level of the trip distribution model, provided that the O/D-based travel costs are updated between each incremental step and the trip distribution model is sensitive to these costs.[13] The more incremental steps are considered, the better the accuracy of the results to be expected. In contrast, following a standard approach, O/D-based travel costs are only computed once, based on preloaded networks including congested parts, and the entire trip matrix is computed based on these costs.

The incremental approach can be outlined as follows. In the first step, unloaded networks are used for the computation of the travel impedance and the GTT values.[14] Based on this information, the first portion of the generated trip demand is assigned to the trip matrices and to the networks. The network impedance and the GTT values are then updated, taking into account the impact of rising traffic loads on congestion and on travel times. In addition, the number of destination opportunities still available is reduced for each travel zone by the number of arriving trips.[15] Subsequently, the second portion of the generated trip demand passes through the trip distribution, the modal split and the network assignment model. Once again, these trips are assigned to the trip matrices and to the networks, before the impedances and the GTT values are updated together with the number of destination opportunities still available. This procedure is repeated until all portions are processed.

In general, the size of the portions should be chosen variably. In addition, the first portion should be the largest and the last should be the smallest. Following this rule, the number of incremental steps can be reduced while the accuracy of the model is still ensured. It can be assumed, for instance, that the capacity of most EU motorways is sufficient for half of the traffic volumes. In this case, congestion effects become more and more relevant when traffic volumes are reaching 100%. Hence, it makes almost no difference if the first 50% are modelled in one step or in several. Depending on the purpose of application, the number of incremental steps as well as the size of the fractions should be chosen variably. Table 9.1 outlines four settings that are appropriate for different use cases in which the HIPAT model is applied in incremental steps.

[13] An incremental approach is also applied in the VACLAV model, but only for the computation of the network assignment (cf. sec. 2.2.1). Hence, the impact of congestion on the destination and mode choices cannot be modelled.

[14] Although not explicitly mentioned, all transport demand segments should be taken into account, i. e. all transport modes and all trip purposes.

[15] For instance, if the model computes that 100 tourists travel to Paris, the number of available accommodation facilities is reduced accordingly. This is particularly relevant with regard to the accessibility-based trip distribution model.

Table 9.1: Predefined settings for different use cases

	Allocation of the fractions [%]									
Description	1	2	3	4	5	6	7	8	9	10
All-or-nothing	100	-	-	-	-	-	-	-	-	-
Moderate runtime	60	30	10	-	-	-	-	-	-	-
Trade-off	45	25	15	10	5	-	-	-	-	-
High fidelity	25	20	15	15	10	5	5	2	2	1

The first setting, "all-or-nothing", is particularly important in order to calibrate and to test the HIPAT model, notably with regard to the integrated trip distribution and modal split sub-models. As the generated trip demand is processed in one step, feedback loops can be avoided and the effects of the initial calibration parameters on the computation of the trip matrices and the network assignment can be observed directly.[16] This facilitates continuous improvement of the set of calibration parameters by comparing modelled output indicators with observed transport performance indicators, such as generated pkm per year by country.

It should be stressed that accurate assignment results cannot be expected if all traffic loads are assigned to the shortest paths between two travel zones. Once the set of calibration parameters is determined, the HIPAT model should be applied in several incremental steps in order to validate the final model output, e. g. by comparing modelled traffic loads with observed AADT values. If the discrepancies are too large and if internal modelling coefficients of the assignment model have to be changed, the whole calibration process has to be repeated.

In contrast to the all-or-nothing setting, the "moderate runtime" setting subdivides the generated trip demand into three portions, allowing specific tests to be carried out, e. g. checking the robustness of the implemented model, the correctness of the data flows in the model, and the consistency of the feedback loops. In order to produce accurate output indicators, more than three incremental steps should be used, e. g. five according to the "trade-off" setting or rather ten according to the "high fidelity" setting. It is also possible to define a user-specific setting encompassing more than ten steps.

It should be noted that the incremental approach of the HIPAT model is quite well suited to modelling the decision process of travellers with regard to trip planning. Nowadays, real time information is often used, including Internet-based systems for booking hotels, flights and rail journeys as well as real-time information on traffic jams on motorways. However, if the HIPAT model is not properly calibrated, it can nevertheless produce results that appear realistic due to the corrective effects introduced by the feedback loops. It is therefore very important to apply the all-or-nothing setting in order to check the calibration of the HIPAT model before producing accurate results.

[16] Feedback loops imply implicit constraints or, to be more precise, correction factors that are applied in each incremental step. Assuming a very poor set of calibration parameters, for instance, the HIPAT model could assign the whole demand for vacation trips to Paris in the first incremental step. By means of the feedback loops, the number of available accommodation facilities in Paris is then reduced to zero and no further vacation trips to Paris will be computed. Thus, the final model output might appear reasonable even though the applied calibration parameters are not reasonable.

9.3 Implementation of a prototype

This section describes the implementation of a prototype following the HIPAT modelling approach. The prototype started as a copy of the IPAT model, and the new features were implemented subsequently. In order to reduce the complexity and the runtime of the prototype and to facilitate its implementation, several simplifying assumptions were made. The prototype covers the whole of Europe and its neighbouring countries, but only a part of Europe is modelled at LAU-2 level. In addition, only one transport demand segment is considered (i. e. private road commuting trips). In this context, it should be stressed that the primary objectives of implementing a prototype were to gather experience and to demonstrate the advantages of modelling transport at LAU-2 level rather than producing absolutely accurate results. Moreover, at the beginning of this exercise it was not clear how the runtime of the prototype would effectively increase. Given that the prototype encompasses 33,191 travel zones (i. e. almost a third of the intended objective, which was to apply more than 100,000 travel zones), it was expected that the prototype can provide a proof of concept for the HIPAT modelling approach with sufficient reliability. It was also expected that useful lessons can be learned from the prototype facilitating the implementation of the HIPAT modelling approach for the whole of Europe.

9.3.1 Overview and key model features

The implementation of the prototype had to overcome two fundamental obstacles: the non-availability of regional indicators at LAU-2 level and the non-availability of respective LAU-2 regions for many cities. Both issues had to be solved with significant effort following the methods outlined in section 7.3, by downscaling regional indicators from NUTS-3 level and by generating artificial city districts. The required input data at LAU-2 level was generated for regions located along an important European transport corridor between Paris and Budapest which is called the “Magistrale”. Regions outside the Magistrale are modelled at NUTS-2 and at NUTS-0 levels. The decision to focus the prototype on this specific transport corridor was made with reference to an earlier study carried out at KIT for the Magistrale (IWW et al., 2001). Even though that study investigated rail transport, the prototype focuses on private road commuting trips, as only the ETISplus road network model can provide an acceptable basis for modelling regional trips below 50 km.[17] The geographic scope of the prototype, i. e. the Magistrale transport corridor situated in central Europe, is outlined in Figure 9.2. The key model features are summarised in Table 9.2.

The implemented prototype follows the incremental four-step approach in transport modelling outlined in Figure 9.1. The trip demand is generated in the first step and can be subdivided into smaller fractions that are then processed in several incremental steps by the trip distribution-, modal split- and the assignment models. Given that the prototype was developed on the basis of the IPAT model, both models share basic features, e. g. the overall geographic scope and the implementation in `Java`. The main extension of the prototype is the modelling of the travel zones located along the Magistrale transport corridor at LAU-2 level, and its main simplification is the consideration of only one transport demand segment (private road commuting trips between 0 and 280 km).[18] In particular this specific

[17] The density of the ETISplus rail network model is rather low. Even at NUTS-3 level, some regions have to be connected to long-distance train stations in neighbouring regions by virtual feeding links (cf. Szimba et al., 2012, p. 4).

[18] It should be emphasised that the modelling of long-distance trips with the prototype would not be a problem.

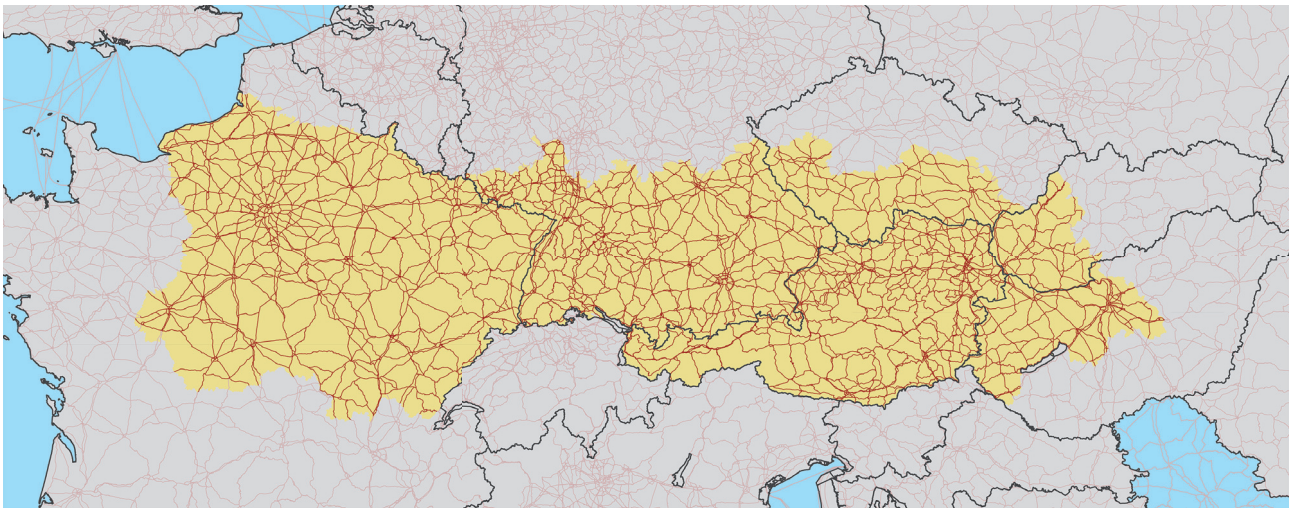

Figure 9.2: Geographic scope of the investigated corridor and density of the road network model

Table 9.2: Key features of the HIPAT prototype model

Feature	Description
Methodology	Incremental four-step approach
Scope & zoning	Magistrale corridor, European and neighbouring countries; hierarchical zoning system (NUTS-0, ..., LAU-2 level)
Complexity	33,191 travel zones; 10,914,221 O/D relations
Demand segment	Private road commuting trips ranging from 0 to 280 km
Remarks	Pre-calibrated `Java` model for testing purpose
Model runtime	About 2 minutes including data exchange with the file system

demand segment, which contributes heavily to capacity overloads (congestion) on the trans-European road networks, cannot be modelled with sufficient accuracy by current European transport models operating at NUTS-3 level.

9.3.2 Simplifying assumptions and model calibration

The HIPAT prototype considers only private road commuting trips. For this reason, the model indicators are not distinguished by trip purpose and by transport mode. This enables the simplification of several modelling steps underlying the prototype. These simplifications are discussed in the following section.

9.3.2.1 Travel zones

The travel zones underlying the HIPAT prototype are defined at seven hierarchical levels, including four NUTS levels, two LAU levels and a "root" level referring to the whole geographic scope (i. e. all of Europe and its neighbouring countries). However, only those regions located along the Magistrale corridor are disaggregated up to LAU-2 level while other European regions are only disaggregated up to NUTS-2 level. Neighbouring countries are modelled at NUTS-0 level.

Another simplification is the definition of artificial LAU-1 zones. In order to close existing gaps, some LAU-2 zones have been used directly as LAU-1 zones, without any merging; e. g. the travel zone Eisenstadt located in Austria (LAU-2 code: 10101) was also defined at LAU-1 level.[19] This method certainly closes existing data gaps but impairs the efficiency of the hierarchical modelling approach. According to Table 9.4 (see sec. 9.3.3) the number of travel zones increases by a factor of 100 between the levels NUTS-3 and LAU-2 and, accordingly, O/D relations that are directly disaggregated from NUTS-3 to LAU-2 level have to be modelled by 10,000 relations. In contrast, if artificial LAU-1 zones were constructed by merging ten LAU-2 zones, the same O/D relations would be disaggregated into 100 relations at LAU-1 level. Taking into account that the majority of these relations can be modelled at LAU-1 level, only a few relations have to be disaggregated into 100 relations each at LAU-2 level. The concept of defining artificial LAU-1 zones by merging LAU-2 zones can therefore save many O/D relations.

9.3.2.2 Generation model

The number of private road commuting trips T_i per year is generated for all travel zones i at the most disaggregated level (i. e. at LAU-2 level for regions inside the Magistrale, at NUTS-2 level for other EU regions, and at NUTS-0 level for neighbouring countries). It is computed by a simplified approach, based on the working population and on country-specific (ci) trip rate factors R:

$$T_i = R_{ci} \cdot \text{labour}_i \tag{9.1}$$

where the working population (labour force) is not distinguished by age and gender. Other indicators, like income and education level, are also not considered. For aggregated travel zones, e. g. LAU-1 zones, the trip demand is added up.

9.3.2.3 GTT indicator

For each O/D relation, generalised travel costs c_{ij} in terms of generalised travel times (GTT) are computed based on a simplified approach:

$$c_{ij} = d_{ij} \tag{9.2}$$

where d refers to the impedance indicator travel distance from origin i to destination j. The impedance indicators initially precomputed can be updated during the assignment step taking congestion effects into account. Though (9.2) seems to be an oversimplification of the HIPAT prototype with regard to the rather complex cost function of the IPAT model (cf. sec. 6.2.2.2), this is acceptable given a high correlation between road travel time costs and road travel distance costs, and given that only private road commuting trips are considered in the prototype. In addition, the GTT indicator defined in a virtual utility scale has to be transformed back into the ordinary distance scale before it can be applied in the deterrence model.[20] It should be emphasised, however, that the application of a more complex cost function would also have been possible for the prototype.

[19] In order to generate artificial LAU-1 zones by merging LAU-2 zones, several methods were tested within this thesis based on the "Grouping Analysis" tool of ArcGIS. However, none of the tested methods produced satisfactory results and this idea was finally given up. It should be mentioned that the Grouping Analysis tool is not optimally suited for this task. Better results could be expected if a tailored grouping algorithm were developed, e. g. via the `ArcPython` interface.

[20] The respective scale transformation is defined by (6.9) in section 6.2.2.2.

9.3.2.4 EMC indicator

Expected minimum cost refer to the expected value of the different mode-specific GTT costs. If only one transport mode is considered, the EMCs are identical to the GTT costs for this transport mode.[21]

9.3.2.5 Deterrence model and deterrence function

The deterrence model applied in the prototype basically follows the methodology outlined in (5.21) in section 4.4. For the purpose of simplification, the two factors related to country barriers, economic and other differences between origin and destination, and the correction factor δ relevant for intra-zonal trips, are not considered. The simplified deterrence model follows:

$$M_d(c_{ij}) = \rho_d(c_{ij}) f(c_{ij}), \quad (i \in I,\, j \in J) \tag{9.3}$$

$$f(c_{ij}) = \omega_S c_{ij}^{\alpha_S} e^{-\beta_S c_{ij}} + \omega_M e^{-\beta_M c_{ij}} + \omega_L c_{ij}^{-\alpha_L} \tag{9.4}$$

where f is the composite deterrence function (4.24) introduced in section 4.3.3 and ρ_d is the scaling function. As ρ_d is defined in relation to the average diameter d of the travel zones, it enables the simultaneous application of the deterrence model at different hierarchical levels. c refers to the GTT indicator or, to be more precise, to the travel distance in the HIPAT prototype. ω_i are weighting factors and α_i, β_i are decay coefficients. As the deterrence model is applied for road commuting trips, the deterrence value M_{ij} is only computed if the crow-flies distance between origin and destination is less than 200 km. The longest commuting trips are therefore about 280 km. If the crow-flies distance is larger than 200 km, $M_{ij} = 0$ is assumed. In this case, no commuting trips are computed for the respective O/D relations.

9.3.2.6 Integrated trip distribution, modal split and assignment step

In the prototype only one transport mode is considered and the modal split step is skipped. In addition, the all-or-nothing setting is used (cf. Tab. 9.1) in order to avoid counterbalancing effects. In this case, all generated trips are distributed and assigned in one single step. This enables a comparison of the model results of different scenarios and, in particular, the identification of advantages when applying the prototype at LAU-2 level instead of NUTS-2 level.[22]

Given the limited density of the ETISplus road network model, particularly for urban roads, short-distance trips should only be proportionately assigned to the networks links. With increasing travel distance, the share of the trips assigned to the networks can be increased.[23] Table 9.3 summarises the setting of these share factors in the prototype, taking into account that short-distance trips are more likely to use regional roads that are not covered by the ETISplus road network model while, in contrast, long-distance trips are likely to use motorways and federal roads that are covered by the ETISplus road network model. For the generation of more accurate assignment maps, the outlined

[21] Equation (3.18) discussed in section 3.4.1 outlines the computation of the EMU indicator. The EMC indicator is computed analogously.

[22] Only travel zones located along the Magistrale are disaggregated to LAU-2 level.

[23] If only a small portion of all intra-city trips were assigned to the ETISplus road network model, the respective urban motorways would be highly congested. This would affect the computation of the travel costs for long-distance trips based on the network model and on shortest paths.

Table 9.3: Share of trips assigned to the networks by distance band

Distance band [km]	< 2	[2; 5)	[5; 10)	[10; 20)	[20; 50)	> 50
Percentage [%]	0	10	20	40	80	100

share factors should be distinguished by country or even at a regional level in relation to the density of the applied network model.

It should be noted that these share factors were not applied for the analysis of the prototype that was carried out based on different model runs. These model runs refer to different hierarchical levels. Their results can only be compared if the whole content of the trip matrix is visualised (i. e. assigned to the networks).

9.3.2.7 Calibration

The prototype is calibrated at a European scale by adjusting the coefficients of the composite deterrence function in order to meet an average trip length of 19.6 km.[24] The derived calibration coefficients are not refined for each country in a second step as is the case with the IPAT model. Hence, the deterrence function is not distinguished by country in the prototype. This decision was made with regard to the fact that only regions along the Magistrale are modelled at LAU-2 level. For this reason, the derived calibration coefficients could not have been refined for the majority of the countries.

Though this simplifies the calibration, the overall process is still very challenging due to the distinct modelling of short-distance trips at LAU-2 level. In this case, all three components of the composite deterrence function have to be refined in an iterative process. However, the three components interfere. If the combined function is improved for short-distance trips, for instance, the composite deterrence function changes as a whole, having an impact on the distribution for medium-distance and long-distance trips as well. In contrast to the HIPAT prototype, the calibration of the IPAT model operating at NUTS-2 level was relatively simple, as the deterrence function basically had to be calibrated only for long-distance trips. If O/D trips are modelled at NUTS-2 level, the majority of short-distance trips are carried out intra-zonally.

Besides the challenge of calibrating the deterrence function to all trip lengths ranging from 1 to 280 km, a major obstacle was the correct processing of the hierarchical trip matrix. For the IPAT model, the computation of a complete trip matrix was comparably simple. For instance, the total number of arriving trips for a travel zone could be computed by adding all values in the respective column of the O/D trip matrix. This approach fails for the HIPAT prototype, given that the O/D relations are modelled at different hierarchical levels. For instance, trips arriving in the city of Karlsruhe (DE122) can be modelled at LAU-2, LAU-1, NUTS-3, NUTS-2, NUTS-1 and NUTS-0 level. In order to derive the total number of arriving trips, two cases have been distinguished:

1. O/D trips that are modelled at LAU-2, LAU-1 and NUTS-3 level; and

2. O/D trips that are modelled at NUTS-2, NUTS-1 and NUTS-0 level.

[24] For the calibration only those trips generated for travel zones located along the Magistrale are considered, as almost 100% of the trips generated for neighbouring countries (modelled at NUTS-0 level) are carried out intra-zonally.

In the first case, all trips definitively end in the city of Karlsruhe. In the second case, only some of the trips end in the city of Karlsruhe, e. g. for the NUTS-2 relation DE27/DE12 (Schwaben/Karlsruhe). In this case, the number of arriving trips was determined based on the number of workplaces in the city of Karlsruhe (DE122) in relation to the region of Karlsruhe (DE12).

It has to be noted that slightly simplified disaggregation rules are applied in the prototype in order to break down relation-based output indicators modelled at higher hierarchical levels to lower hierarchical levels. These disaggregation rules are not fully consistent with the propagation rules applied for the hierarchical accessibility indicator in order to compute the O/D trip matrix. For this reason, small discrepancies can be observed between the generated trip demand for a travel zone and the number of originating trips (i. e. $T_i \neq \sum_j T_{ij}$).[25] In total, the discrepancy between the number of generated trips and the number of O/D trips in the matrix is 0.3% and is therefore acceptable. The reason not to define more consistent disaggregation rules is the complexity of the propagation rules underlying the hierarchical accessibility indicator, that are implemented by the three recursively applied Algorithms 2, 3 and 4 outlined in section 8.3. Intermediate results computed by these algorithms are currently not stored, but rather only the final hierarchical accessibility indicator. If all intermediate results are stored, more consistent disaggregation rules can be be defined. Given that the produced text file for the trip matrix is about 360 megabytes, and given the complexity of processing the hierarchical matrix, no further effort was invested at this point to eliminate the current model error of 0.3% that results from the disaggregation of O/D-based output indicators.

9.3.2.8 Summary

This section discussed several simplifying assumptions that are used in the prototype. Besides the general limitation to one demand segment (private road commuting trips) and to the Magistrale corridor, several modelling steps of the prototype follow simplified procedures. These simplifications are acceptable given that the prototype was implemented, first of all, in order to gather experience and to provide a proof of concept for the HIPAT modelling approach. For those purposes, the complexity of the prototype encompassing 33,191 travel zones is expected to be large enough. It is furthermore expected that the lessons learned from the prototype will help to streamline the implementation of a complete HIPAT model for the whole of Europe, encompassing more than 100,000 travel zones and all demand segments. This assessment is also supported by the fact that the current implementation of the prototype has not been fully optimised. Its runtime can be further reduced, for instance, by parallel computing and by the optimisation of the zoning system.

A key question related to the implementation of the HIPAT model for the whole of Europe will be how the calibration process can be simplified. Given that the trip matrix produced by the prototype is about 360 megabytes, it is expected that the trip matrix produced by the HIPAT model covering all transport modes and trip purposes will increase to several gigabytes. After this key question has been solved, the complexity of the current prototype version can then be progressively increased, e. g. by applying the more complex cost functions underlying the IPAT model for the computation of the GTT indicator.

[25] O/D-based originating trips are disaggregated based on the generated trip demand in the respective travel zones.

9.3.3 Generation of the input database

This section describes the process of generating the required input data sets for the HIPAT prototype, including:

1. the travel zones and their hierarchical ordering;
2. the regional indicators for each travel zone;
3. the centroids and feeding links for each travel zone;
4. the set of O/D relations referring to different hierarchical levels; and
5. the travel impedances for each O/D relation.

9.3.3.1 Travel zones

The prototype relies on 33,191 travel zones that are defined at seven different hierarchical levels based on the administrative units provided by the EBM database (BKG, 2013) and the zoning system of the IPAT model. The top level encompasses the "root" travel zone referring to the whole geographic scope (Europe and neighbouring countries). The second level refers to the NUTS-0 regions (i. e. the countries) and the next levels to the NUTS-1, NUTS-2, NUTS-3, LAU-1 and LAU-2 regions. Up to NUTS-2 level, the zoning system applied in the HIPAT prototype was derived based on the zoning system applied in the IPAT model. The difference between the two systems is that in the HIPAT prototype these regions located along the Magistrale corridor are further disaggregated up to LAU-2 level. Hence, the two zoning systems are identical for regions outside the Magistrale.

The majority of the applied travel zones are defined in accordance with the NUTS classification. Existing data gaps at LAU-1 level were closed by adding LAU-2 regions. In addition, artificial city districts at LAU-1 and LAU-2 level were added for 39 cities located along the Magistrale.[26] Table 9.4

Table 9.4: Structure of the zoning system underlying the prototype

Id	Level	Zones	Diameter	Comment
0	EU	1	3650.7 km	Root travel zone
1	NUTS-0	40	440.2 km	Defined for the whole scope
2	NUTS-1	99	13.8 km	Only defined for Europe
3	NUTS-2	284	127.7 km	Only defined for Europe
4	NUTS-3	229	42.8 km	Only defined for Magistrale
5	LAU-1[1]	5,647	8.5 km	Only defined for Magistrale
6	LAU-2[2]	26,891	4.2 km	Only defined for Magistrale

[1] The figure includes 156 artificial city districts and 3,435 LAU-2 regions.
[2] The figure includes 624 artificial city districts.

briefly outlines the basic parameters of the applied zoning system. The hierarchical order of the travel zones (i. e. the hierarchical tree structure) was established bottom-up by linking each LAU-2

[26] For Paris (FR101), the EBM database uses the same region at NUTS-3, LAU-1 and LAU-2 level. In this case, the LAU-1 and the LAU-2 regions were replaced by four and 16 artificial city districts respectively.

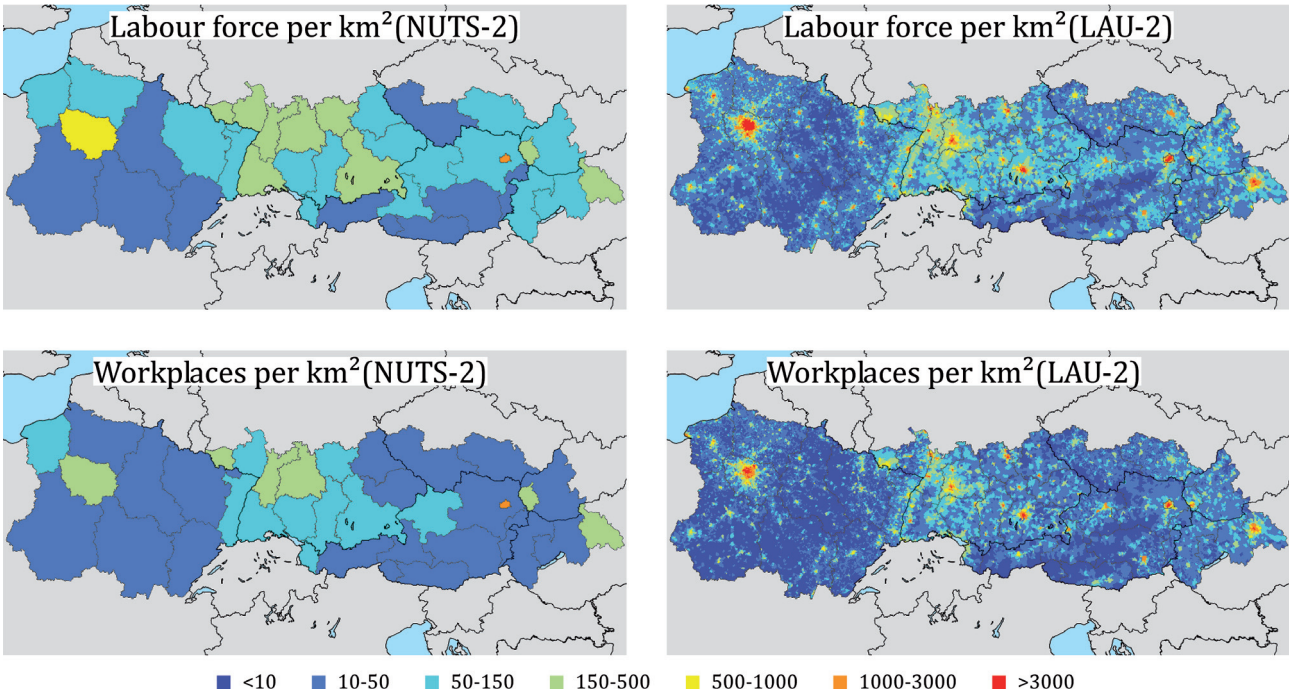

Figure 9.3: Evaluation of density-based indicators at NUTS-2 and LAU-2 levels

zone to its parent LAU-1 zone and simultaneously linking the LAU-1 zones to their children zones. This procedure was repeated for all travel zones at each hierarchical level, in order to determine all child/parent relations and accordingly all parent/child relations.[27]

9.3.3.2 Regional indicators

The prototype only considers private road commuting trips. Thus, information on the working population and on the number of workplaces is required for all travel zones. These indicators were taken from the ETISplus socio-economic dataset discussed in section 2.3 and refer to the base year 2010. Provided at NUTS-3 level, they were disaggregated to LAU-2 level in the first step on the basis of proration factors according to the methodology discussed in section 7.3.1. The factors for the indicator working population were derived based on the European GEOSTAT dataset, which provides population data at the level of grid cells with an edge length of 1 km. The factors for the indicator number of workplaces were derived based on land use data provided by the Corine land cover database at the level of grid cells with an edge length of 100 m and specific load factors for each land use type (cf. Tab. 7.2). In the second step, these two regional indicators were derived for all aggregated travel zones by adding the indicators from LAU-2 level.

Figure 9.3 illustrates the spatial distribution of the working population (labour force) and the workplaces by seven density categories along the Magistrale corridor. While the two figures produced at NUTS-2 level are dominated by large, uniform areas generally representing lower density categories, the LAU-2 figures resemble a patchwork of heterogeneous areas encompassing all density categories. Given the strong correlation between the district type and the mobility behaviour of its residents, the LAU-2 level provides a much better basis for transport modelling.[28]

[27] Establishing the hierarchical tree structure is required for the propagation of indicators between different levels.
[28] Residents of cities tend to travel less than residents of rural regions (cf. Fig. 2.1, sec. 2.1.2).

9.3.3.3 Centroids and feeding nodes

For each travel zone, further indicators were derived including the centroid, the feeding node in the road network model and the crow-flies distance between the centroids. In the first step, the centroids were determined. For all NUTS-3 zones, the centroids were taken from ETISplus and for all NUTS-2 zones from HIGH-TOOL. At NUTS-1 level the most relevant NUTS-2 centroid was selected, and at NUTS-0 level the capital of each country, and Brussels for Europe. At LAU-2 level, the centroids refer to the geographic location of the indicator "ResidenceOfAuthority" provided by the EBM dataset. If not available, the centroid was computed based on the GEOSTAT dataset, which provides information on the spatial distribution of the population in the zone. For unsettled zones, the geographic centroid was chosen. For the travel zones at LAU-1 level, the centroid of the most populated LAU-2 zone was chosen.

In the second step, the feeding nodes were determined. For each travel zone, the node of the road network model located closest to the centroid was selected. The crow-flies distance between both locations was then computed. Due to the limited density of the ETISplus road network model, several feeding nodes are used more than once.[29] It should be noted that the crow-flies distance between the centroid of a LAU-2 region and its respective feeding node is 7.3 km on average. For this reason, it is expected that the density of the ETISplus road network model is currently not sufficiently high in order to derive reliable travel impedance indicators for the majority of the O/D relations that are modelled at LAU-2 level.

9.3.3.4 O/D relations

The set of O/D relations referring to different hierarchical levels was generated by the recursive application of the neighbourhood criterion on the root relation EU/EU and then on each O/D relation that is generated. In the very initial step, the criterion determined that the two travel zones EU and EU are neighbours. In consequence, the number of trips between these two travel zones (i. e. all trips carried out in Europe) cannot be modelled at the root level. Hence, the root O/D relation was disaggregated into 1,600 O/D relations at NUTS-0 level, given that 40 countries are considered in the HIPAT prototype. Some of these relations could be accepted (e. g. DK/BE) while others had to be further disaggregated to NUTS-1 level (e. g. DK/DE or DK/DK). However, some O/D relations (e. g. UA/BY) could not be further disaggregated, as both neighbouring countries are only modelled at NUTS-0 level in the prototype. Outside the Magistrale, European O/D relations could be disaggregated up to NUTS-2 level at maximum, and along the Magistrale up to LAU-2 level. Hence, long-distance trips within the Magistrale can be computed very efficiently at NUTS-2 level while short-distance trips can be modelled very accurately at LAU-2 level.

Following the outlined approach and applying the neighbourhood criterion recursively generated the complete set of O/D relations. In total, the prototype encompasses 10,914,221 relations referring to different aggregation levels that are summarised in Table 9.5. As expected the hierarchical structure of the O/D relations is symmetrical, i. e. there are as many NUTS-0/NUTS-1 relations as there are NUTS-1/NUTS-0 relations. In addition, the number of O/D relations increases from NUTS-0 to LAU-2 in which LAU-2/LAU-2 relations have the largest share with 72.2 % followed by the share of LAU-1/LAU-1 relations (26.5 %). This means that almost all O/D relations (98.7 %) underlying the prototype are used for the modelling of short-distance trips.

[29] Only 5,261 network nodes are located along the Magistrale corridor compared to 26,891 travel zones at LAU-2 level.

Table 9.5: Hierarchical structure of the set of O/D relations

O\D	EU	NUTS-0	NUTS-1	NUTS-2	NUTS-3	LAU-1	LAU-2
EU	-	-	-	-	-	-	-
NUTS-0	-	1,190	146	366	18	161	13,729
NUTS-1	-	146	3,736	-	-	-	-
NUTS-2	-	366	-	9,864	858	6,577	33,398
NUTS-3	-	18	-	858	17,206	-	-
LAU-1	-	161	-	6,577	-	2,887,788	-
LAU-2	-	13,729	-	33,398	-	-	7,883,931

However, Table 9.5 also reveals two shortcomings of the prototype with regard to the computed set of O/D relations, the large number of LAU-1/LAU-1 relations and the existence of very heterogeneous relations such as LAU-2/NUTS-0. The large number of LAU-1/LAU-1 relations can be explained by the fact that 3,435 out of 5,647 LAU-1 zones are defined based on the respective LAU-2 zones (cf. Tab. 9.4). Their number could be significantly reduced if artificial LAU-1 zones were defined by grouping about ten LAU-2 zones.[30] This would avoid many LAU-1/LAU-1 relations and, accordingly, reduce the runtime complexity of the HIPAT prototype to almost the same extent.

The existence of very heterogeneous relations, such as LAU-2/NUTS-0, is not optimal for two reasons. Given that origin and destination are modelled at very different levels, the two travel zones are not alike. This is highly critical for the robustness of the trip distribution model.[31] The second reason concerns the application of the neighbourhood criterion generating the hierarchical set of O/D relations. If the criterion suggests, for instance, that the trip demand between the two neighbouring countries the Slovak Republic[32] and the Ukraine cannot be modelled at NUTS-0 level, it is very likely that the criterion will demand the disaggregation of the Slovak Republic up to LAU-2 level due to the large size of the Ukraine that is modelled at NUTS-0 level. This leads to an overhead of O/D relations modelled at LAU-2/NUTS-0 level. These relations, referring rather to long-distance trips, are not equivalent to LAU-2/LAU-2 relations, which refer exclusively to short-distance trips.

With regard to the above findings, it would generally make sense to improve the set of O/D relations by revising the zoning system. First of all, the number of O/D relations could be significantly reduced if artificial LAU-1 zones were generated by grouping about ten LAU-2 zones. Secondly, too heterogeneous O/D relations could be avoided if neighbouring countries were modelled at NUTS-1 level and the regions surrounding the Magistrale corridor at NUTS-3 level. It would then be possible to limit the differences between the origin and the destination to one hierarchical level, i. e. NUTS-2/NUTS-1 relations would be used for trips between European and neighbouring countries and NUTS-3/NUTS-2 relations for trips between regions located along the Magistrale and regions outside the Magistrale. These measures should be taken into account when implementing a complete HIPAT model covering the whole of Europe at LAU-2 level.

[30] Between the NUTS-3 and LAU-2 levels the number of zones increases by a factor of about one hundred (cf. Tab. 9.4).

[31] For instance, if a small error is introduced for a LAU-2/NUTS-0 relation, the negative implications for competing LAU-2/LAU-2 relations will be much larger.

[32] In the prototype, only the western part of the Slovak Republic, as part of the defined Magistrale corridor, is modelled at LAU-2 level.

9.3.3.5 Travel impedances

The current version of the prototype only requires information on the travel distance between origin and destination. This information feeds into the computation of the GTT indicator (9.2) or straight into the deterrence model (9.3). For each O/D relation the travel distance was precomputed by an exogenous model based on the crow-flies distance between the two centroids and a detour factor of $\sqrt{2}$. The travel impedances are not updated during the iterative calibration process of the prototype, given that the whole generated trip demand is processed in a single step according to the all-or-nothing setting (cf. Tab. 9.1).

The decision to compute the travel impedances following a simplified approach was made in particular with respect to the limited density of the applied ETISplus road network model. Given that only 5,261 network nodes are located along the Magistrale corridor compared to 26,891 travel zones at LAU-2 level, about five LAU-2 zones had to be connected on average to each node of the network model. In addition, the crow-flies distance of 7.3 km between the centroid of a LAU-2 zone and the corresponding feeding node is much larger than the average radius of a LAU-2 zone, i. e. about 2.1 km. For this reason, for too many O/D relations modelled at LAU-2 level, the travel impedances could not be derived with high confidence based on the network model.

It should be noted that for an early version of the HIPAT prototype that was only applied at NUTS-3 level, the travel impedance indicators (travel time and travel distance) were computed based on the network model. In this version, the "moderate runtime" setting was implemented as well (cf. Tab. 9.1) in which the travel impedance indicators were continuously updated between each incremental step.[33] From a technical point of view, the extension of the current version of the prototype by the computation of the travel impedance indicators based on the network model is therefore not an obstacle. However, starting from the ETISplus network, the density of the network model has to be increased considerably in order to compute the travel impedance indicators with high confidence for relations modelled at LAU-2 level. The optimal enlargement of the network model could be subject of further studies.

9.3.3.6 Summary

The process of generating the input database for the prototype was complex, time-consuming and far from optimal. It consisted of several intermediate steps that had to be revised several times. Amongst other things, artificial city districts had to be generated and the availability of regional indicators at LAU-2 level had to be ensured by downscaling. In addition, several methods were tested in order to generate artificial LAU-1 zones by merging LAU-2 zones. However, the artificially generated LAU-1 zones were not in line with the homogeneity criteria discussed in section 7.1.1. In particular the two criteria of inter-homogeneity of shape and size and compactness were violated. Hence, the generation of artificial LAU-1 zones was finally rejected.

The generated input datasets were validated and the results already indicate the potential of the LAU-2-based HIPAT approach for transport modelling compared to a standard approach at NUTS-2

[33] In this case, the network assignment model was applied at the beginning of each incremental step in order to update the travel impedance indicators, taking into account the current congestion level in the transport networks. The information on all shortest paths was temporarily stored in the computer's main memory. It could then be used again in order to update the traffic loads at link level after the trip distribution/ modal split step (cf. Fig. 9.1).

level. While at NUTS-2 level, for instance, differences between cities and rural areas are not reflected due to the large size of the travel zones, the analysed Magistrale corridor resembles a patchwork of different areas when modelled at LAU-2 level (cf. Fig. 9.3). Besides the regional indicators, the set of O/D relations was also evaluated and the hierarchical structure is summarised in Table 9.5. Given the comparatively large share of LAU-1/LAU-1 relations it has to be concluded that the application of LAU-2 zones to close data gaps at LAU-1 level is not optimal. For this reason, the approach for generating artificial travel zones at LAU-1 level by merging several LAU-2 zones should be further investigated. With regard to the HIPAT prototype, this would save over 50% of the current travel zones at LAU-1 level and accordingly about 75% of the O/D relations.

In general, the hierarchical zoning system should be revised. In particular, any two adjacent travel zones forming an O/D relation should not differ by more than one hierarchical level. With regard to the HIPAT prototype, this would require the modelling of neighbouring countries outside Europe at NUTS-1 level and regions near the Magistrale at NUTS-3 level. Then, too heterogeneous O/D relations between adjacent travel zones such as LAU-2/NUTS-0 could be avoided. In addition, it would make sense to replace the LAU-1 level by two artificial levels, given that the number of travel zones increases by a factor of about 120 between NUTS-3 and LAU-2 level (cf. Tab. 9.4). Currently, an O/D relation that cannot be modelled at NUTS-3 level is disaggregated into 120 LAU-1/LAU-1 relations instead of only 25 relations if two artificial levels are defined.[34] Hence, the definition of two artificial levels replacing the LAU-1 level would reduce the overall number of O/D relations in the model and therefore its runtime.

It should be stressed that the current availability of European reference datasets providing regional indicators at a disaggregated level is far from optimal. With regard to the European GEOSTAT database providing information on the population at the level of grid cells, only one important transport demand indicator is available. Other indicators (e. g. the number of workplaces) have to be disaggregated from NUTS-3 to LAU-2 level under application of land use data and specific load factors. Given that the LAU regions are not defined consistently for all European countries by the NUTS classification, artificial travel zones have to be generated at LAU-1 and LAU-2 level in order to close data gaps. This is not optimal. For this reason, it is necessary to establish a European reference database providing the required regional indicators for transport modelling at LAU-2 level as well as the travel zones including city districts. This European reference database should also include the European reference network models at the correct level of detail and the travel zones should already be connected to the network models.

[34] The number of zones increases by a factor of about eleven between NUTS-3 and LAU-1 as well as between LAU-1 and LAU-2. If three hierarchical levels were considered between NUTS-3 and LAU-2, the travel zones would only increase by a factor of about five between each level.

9.4 Model application and results

This chapter deals with the application of the HIPAT prototype and the discussion of the model output. The prototype was implemented for two main reasons: to prove that about 33,000 travel zones can be handled without causing runtime problems and to demonstrate the advantages of disaggregated travel zones for transport modelling at a European scale.

9.4.1 Overview of the test scenarios

In order to investigate the advantages of modelling European passenger transport at LAU-2 level rather than at a more aggregated level, three test scenarios are defined. These scenarios differ concerning the most detailed aggregation level for the travel zones that are located along the Magistrale. For the first scenario, the respective travel zones are only disaggregated up to NUTS-2 level, for the second up to NUTS-3 level and for the third up to LAU-2 level. Table 9.6 outlines the specific properties of

Table 9.6: Overview on the test scenarios

Scenario	Level	Origins[1]	Destinations[2]	Relations	Savings
Scenario 1	NUTS-2	36	$36 + 248 + 10$	3,068	71.0%
Scenario 2	NUTS-3	229	$229 + 248 + 10$	24,273	78.2%
Scenario 3	LAU-2	26,891	$26{,}891 + 248 + 10$	10,846,130	98.5%

[1] The figure includes only travel zones along the Magistrale.

[2] The figure includes also destinations outside the Magistrale.

each scenario that are related to the Magistrale corridor.[35] The conditions outside the Magistrale are identical for all scenarios: EU28, Norway and Switzerland are modelled at NUTS-2 level, neighbouring countries at NUTS-0 level. Accordingly, the number of travel zones and the number of O/D relations is different for the three test scenarios.

The prototype is then applied for each scenario computing trip estimates for the whole of Europe. However, for the calibration and investigation of each scenario, only trips originating within the travel zones located along the Magistrale are considered. The destinations of these trips can be located anywhere in Europe including regions outside the Magistrale (248 NUTS-2 regions and 10 countries). Hence, for the first test scenario the trip demand between 36 origins and 294 destination is considered (cf. Tab. 9.6). Given that the respective O/D relations are modelled at different hierarchical levels, the HIPAT prototype saves 71.0% of the O/D relations compared to a standard modelling approach computing a complete trip matrix.[36] For the second scenario, the HIPAT prototype saves 78.2% of the O/D relations and for the third scenario 98.5%.

It has to be stressed that the three test scenarios are calibrated by adjusting the composite deterrence function (9.4) underlying the deterrence model in order to meet an average trip length of 19.6 km. Though the average trip length of private road commuting trips is different for the countries located along the Magistrale, this simplification does not limit the analyses carried out based on the produced

[35] A specific test scenario related to the LAU-1 level was not applied, as the LAU-1 layer is not completely defined in Europe (e. g. in Austria).

[36] A complete trip matrix consists of 10,584 O/D relations while the hierarchical trip matrix computed by the HIPAT prototype only consists of 3,068 relations.

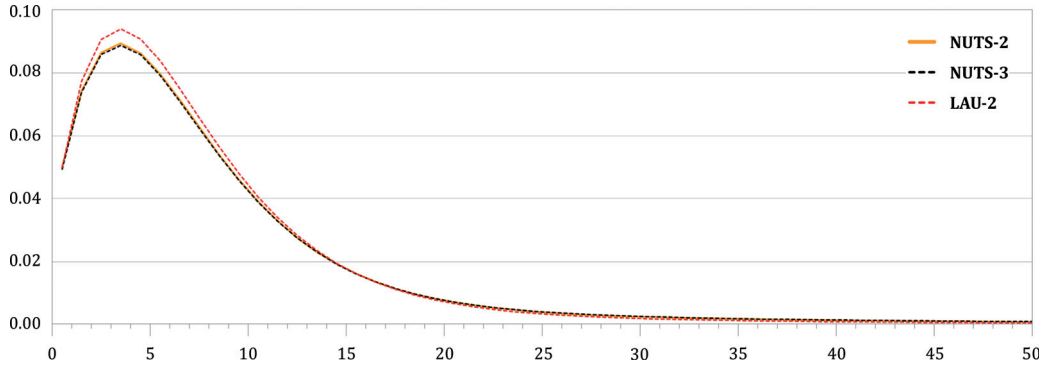

Figure 9.4: Comparison of the deterrence functions underlying the test scenarios

trip matrices. The applied modelling coefficients for adjusting the composite deterrence function are summarised in Table 9.7. While the modelling coefficients applied for the NUTS-2 and the NUTS-3 scenarios are almost identical, the coefficients applied for the LAU-2 scenario are slightly different.

Table 9.7: Applied coefficients for the deterrence function

Scenario	Level	ω_S	ω_M	ω_L	α_S	β_S	β_M	α_L
Scenario 1	NUTS-2	0.80	0.15	0.05	1.2	0.3	0.052	0.8
Scenario 2	NUTS-3	0.80	0.15	0.05	1.2	0.3	0.050	0.8
Scenario 3	LAU-2	0.84	0.15	0.01	1.2	0.3	0.600	1.3

Figure 9.4 outlines the shape of the three deterrence functions $f(x)$ in which the x-axis refers to the travel distance given in kilometres and the y-axis to the deterrence value that is computed as follows:

$$y(x) = f(x) \cdot Y^{-1} \qquad \text{with} \qquad Y = \int_0^\infty f(x)\,dx \tag{9.5}$$

At first glance it can be observed that the two functions underlying the NUTS-3 and the NUTS-2 scenarios are almost identical while the function underlying the LAU-2 scenario is slightly different. Hence, it can be concluded that the transferability of the deterrence function between the NUTS-2 and the NUTS-3 level could be established as expected by the application of the scaling function ρ_d in the deterrence model (9.3). On the other hand, the transferability between these two levels and the LAU-2 level is not entirely given. For the LAU-2 scenario, the share of long-distance trips was initially overestimated and the composite deterrence function or, to be more precise, the power function that is particularly related to long-distance trips, had to be adjusted. This was mainly implemented by changing the weighting factor ω_L from 0.05 to 0.01 and the decay factor α_L from 0.8 to 1.3. It must therefore be expected that longer trips are underrepresented in the LAU-2 scenario in comparison to the NUTS-2 and the NUTS-3 scenarios.

One of the reasons for the initial overestimation of long-distance trips in the LAU-2 scenario might lie in the computation of trip estimates between travel zones that are too heterogeneous, e. g. between a LAU-2-based origin and a NUTS-2-based destination (cf. Tab. 9.5). Possible solutions to overcome this issue could be to adjust the scaling function ρ_d for heterogeneous O/D relations and to improve

the zoning-system in order to avoid relations that are too heterogeneous. For the analysis that is carried out in the following section it is assumed that the differences between the applied deterrence functions are within an acceptable range and do not impair the comparability of the results of the three test scenarios.

9.4.2 Analysis of scenario output

This section discusses the model output of the three test scenarios. Based on the outcomes it then assesses the suitability of each scenario and, accordingly, of the three aggregation levels for the modelling of commuting trips. The trip length distributions (TLDs) and the passenger load values at link level are therefore derived and evaluated for each scenario. The load values are computed following an all-or-nothing assignment method for the three scenarios. In the first scenario, only the load values along the Magistrale are investigated, in the second load values for the whole of Europe and in the third the load values between two neighbouring NUTS-2 zones.[37] In order to demonstrate the impact of disaggregating the travel zones from NUTS-2 to LAU-2 level, a specific case is also investigated in which only the transport demand occurring within two NUTS-2 regions is considered for the analysis.

9.4.2.1 Frequency of trip occurrences by distance band

For this analysis, only trips originating along the Magistrale area are considered (cf. Tab. 9.6) in order to derive the TLDs based on the computed trip matrices. Table 9.8 summarises the basic properties of the trip matrices with regard to the share of intra-zonal (T_{ii}) and inter-zonal trips (T_{ij}) as well as the related demands in terms of pkm. In addition, for the inter-zonal trip demand, a distinction is made between trips that are shorter than 50 km and those that are longer. A quick analysis of Table 9.8 reveals that the NUTS-2 scenario differs significantly from the other two scenarios, given that the majority of the trip demand is intra-zonal in this scenario. This leads to the presumption that NUTS-2-based travel zones are not suited for the modelling of commuting trips.

Table 9.8: Share of intra-zonal and inter-zonal trip demand for the test scenarios

Scenario	Level	T_{ii}	$T_{ij} \leqslant 50$	$T_{ij} > 50$	pkm_{ii}	$\text{pkm}_{ij} \leqslant 50$	$\text{pkm}_{ij} > 50$
Scenario 1	NUTS-2	96.6%	0.9%	2.5%	83.1%	2.0%	14.9%
Scenario 2	NUTS-3	73.2%	21.4%	5.4%	45.0%	23.6%	31.4%
Scenario 3	LAU-2	27.1%	62.5%	10.4%	4.1%	45.8%	50.1%

Figure 9.5 compares the applied deterrence functions and the derived distributions of trip lengths in which intra-zonal trips are drafted in grey and inter-zonal trips in red. The y-axis is drawn in different levels for the three graphs due to different peak values for the three TLDs. The x-axis is limited to 50 km and the distribution is not shown for trips longer than 50 km.

At first sight, a large discrepancy between the TLD and the deterrence function can be observed for the NUTS-2 scenario. This can be explained by the large size of the applied NUTS-2 zones, particularly

[37] Although only the trip demand generated within the Magistrale is considered for the calibration of the prototype, trip estimates are computed for all O/D relations, including for instance trips from Denmark to Belgium that are modelled at NUTS-0 level.

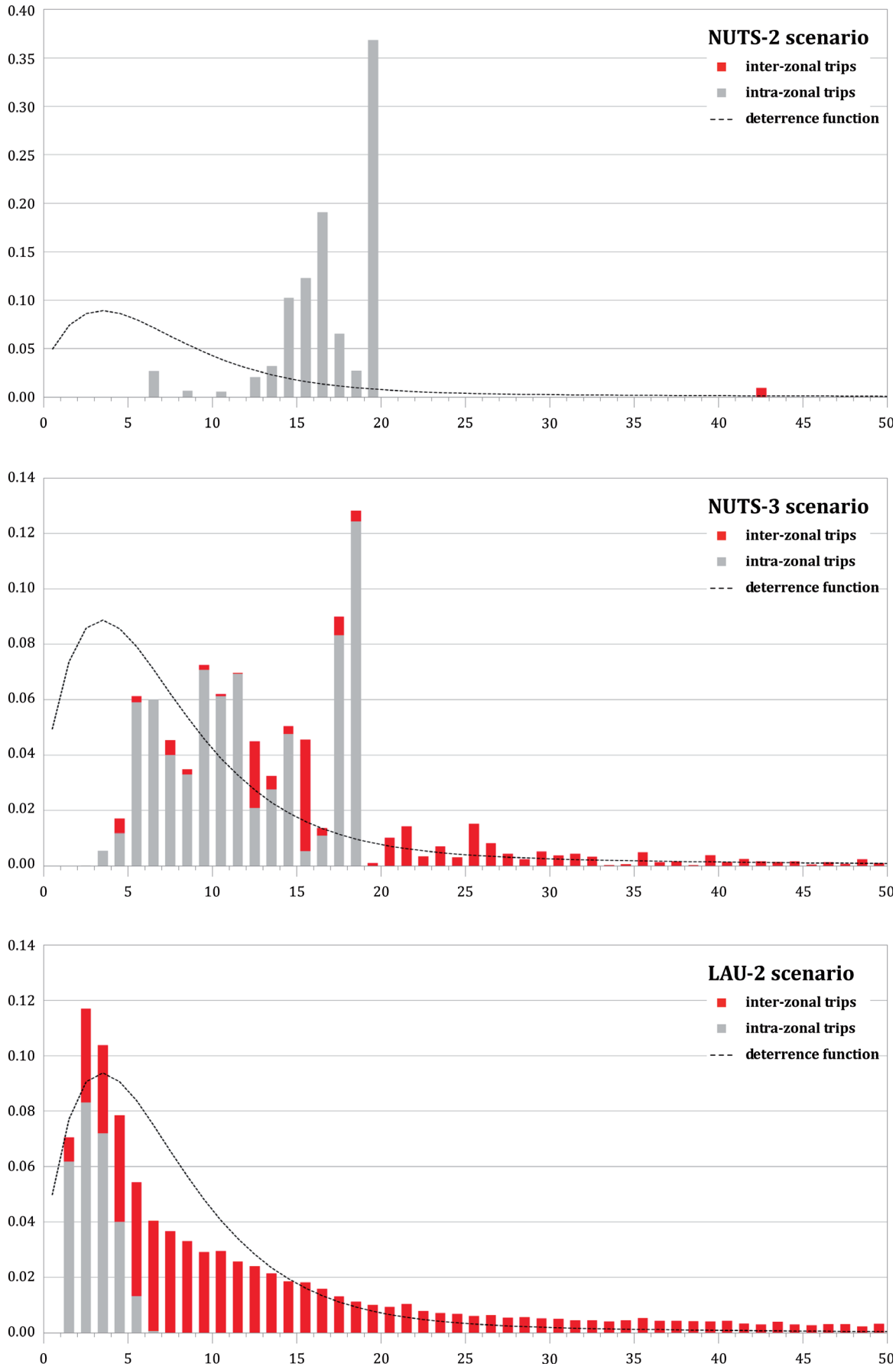

Figure 9.5: Distribution of trip lengths for the test scenarios

for France. Hence, 96.6% of the generated trips are carried out intra-zonally. The travel distance for these intra-zonal trips varies between the distance bands of 6-7 km and 19-20 km depending on the size of the zones. The high peak value of the TLD for the distance band of 19-20 km can be explained by the large size of the 36 selected NUTS-2 zones located along the Magistrale. Given an average diameter of about 120 km for the NUTS-2 zones, intra-zonal trips are longer than 19 km for many of these zones.[38] Besides the small share of inter-zonal trips and the high peak value of the TLD, a further inconvenience arises from the fact that the first inter-zonal O/D relation only appears in the distance band of 42-43 km or, to be more precise, O/D relations between 20 and 42 km do not exist in the model, given the large size of the NUTS-2 regions. Hence, it has to be concluded that the NUTS-2 scenario does not provide any basis for analysing the distribution of trip lengths based on the computed trip matrix.

For the NUTS-3 scenario, the discrepancy between the TLD and the deterrence function is not as large as for the NUTS-2 scenario due to the smaller size of the NUTS-3 regions. In this case, 73.2% of the trips are carried out intra-zonally. The peak value appearing in the distance band of 18-19 km can be explained by the comparably large size of the NUTS-3 regions in France. Given that the TLD only follows a reasonable pattern for inter-zonal trips starting from the distance band of 20-21 km, it has to be concluded that NUTS-3-based travel zones are not suited for analysing the distribution of trip lengths based on the computed trip matrix.

For the LAU-2 scenario, the distortion of the TLD is within an acceptable range. In this case, the TLD resembles a continuous distribution and its pattern is very similar to that of the applied deterrence function. In particular, it must be pointed out that even the distribution of trip lengths for intra-zonal trips follows a reasonable pattern. In addition, only 27.1% of the trips are carried out intra-zonally. It can therefore be concluded that LAU-2-based travel zones provide a sound basis for analysing the distribution of trip lengths based on the computed trip matrix and for modelling the trip distribution. Better results can be expected if different deterrence functions are applied for different countries and if the HIPAT prototype is applied in incremental steps.

A comparison of the TLDs produced for the three scenarios shows that the trip matrix computed at LAU-2 level provides the best basis for analysing the distribution of trip lengths. For the NUTS-3 scenario, it only makes sense to analyse the distribution of trip lengths for inter-zonal trips starting from the distance band of 20-21 km. Given that 55.0% of the generated demand in terms of pkm is inter-zonal, under certain conditions NUTS-3-based trip matrices are suitable for the modelling of traffic loads at link level by the network assignment model.[39] In contrast, NUTS-2-based trip matrices, such as those provided by the European transport models HIGH-TOOL and ASTRA-EC, are not suited for the modelling of traffic loads at link level.

A detailed analysis of the findings shows that the TLD provided by the LAU-2 scenario also differs slightly from the applied deterrence function, e. g. for trips within the distance bands of 5-10 km and for trips >20 km. These deviations might be reduced if the deterrence function were slightly modified in order to distinguish between different mobility patterns, e. g. for city residents carrying out many intra-

[38] The prototype is calibrated to an average trip length of 19.6 km. Hence, the length of intra-zonal trips is always below 19.6 km.

[39] Commonly, only inter-zonal trips are used for the modelling of traffic loads at link level.

city trips and residents living in rural regions carrying out rather longer trips (cf. Fig. 2.1, sec. 2.1.2). Currently, the applied deterrence function is calibrated to an average European level, and therefore cannot exactly reproduce the specific mobility patterns of city residents and residents of rural regions. However, it must be clearly pointed out that this high-scrutiny analysis of the distribution of intra-city trips is only possible for the LAU-2 scenario due to the application of artificial city districts.

9.4.2.2 Passenger loads at link level

For this analysis, the whole trip demand originating along the Magistrale area that is summarised by the trip matrix was visualised. To this end, passenger loads at link level were computed by the network assignment model following an all-or-nothing method. This ensures comparability of the three scenarios. Figure 9.6 outlines the computed load values along the Magistrale for the three scenarios. Seven categories are distinguished: the highest is shown in bright red, the lowest in light blue and unloaded links are shown in light grey.

As might be expected from the previous analysis, the NUTS-2 assignment looks rather empty. As 96.6% of the trip demand is for intra-zonal trips (cf. Tab. 9.8), the majority of the trip demand summarised by the trip matrix cannot be assigned to the network links. Nevertheless, the NUTS-2 assignment shows some network links in the highest load category, for instance the network links between Paris located in the NUTS-2 region Île de France (FR10) and Amiens located in the region Picardie (FR22). In this case, the trip distribution model assigns 7.5% of the trip demand generated for the region FR22 to the O/D relation FR22/FR10. This pattern can be explained by the very large number of workplaces provided by the region Île de France, which attract commuters from the region Picardie. However, for the majority of the network links (72.2%) no passenger loads can be computed. Hence, it has to be concluded that the computed trip matrix at NUTS-2 level is not suited for modelling the commuter passenger loads at link level by the network assignment model.

For the NUTS-3 scenario, the derived passenger load values look much better. In this case, the share of intra-zonal trips is only 73.2% and load values can be computed for about two thirds of the network links (61.0%). In addition, the distribution of the load categories follows a reasonable pattern in which the highest are computed for those network links located around big cities. For instance, Paris and Munich can easily be recognised. However, besides the high load values computed for links around big cities, the NUTS-3 assignment still has significant gaps (i. e. unloaded links shown in light grey). The largest gaps can be observed for France due to the comparably large size of the French NUTS-3 regions. It has to be concluded that the NUTS-3 level is not optimal for modelling demand segments like commuting trips with a rather regional character.

The best assignment results can be achieved for the LAU-2 scenario in which load values are computed for 95.5% of the network links. Particularly for France a significant improvement from NUTS-3 to LAU-2 level can be observed. For example, the “star-shaped” pattern indicating high load values around Paris is more apparent. The improved quality can be clearly attributed to the smaller size of the LAU-2 regions. 72.9% of the trip demand is for inter-zonal trips and contributes to the computation of the passenger load values at link level. It can therefore be concluded that LAU-2 travel zones provide a sound basis for modelling the distribution of trips with a rather regional character, such as commuting trips.

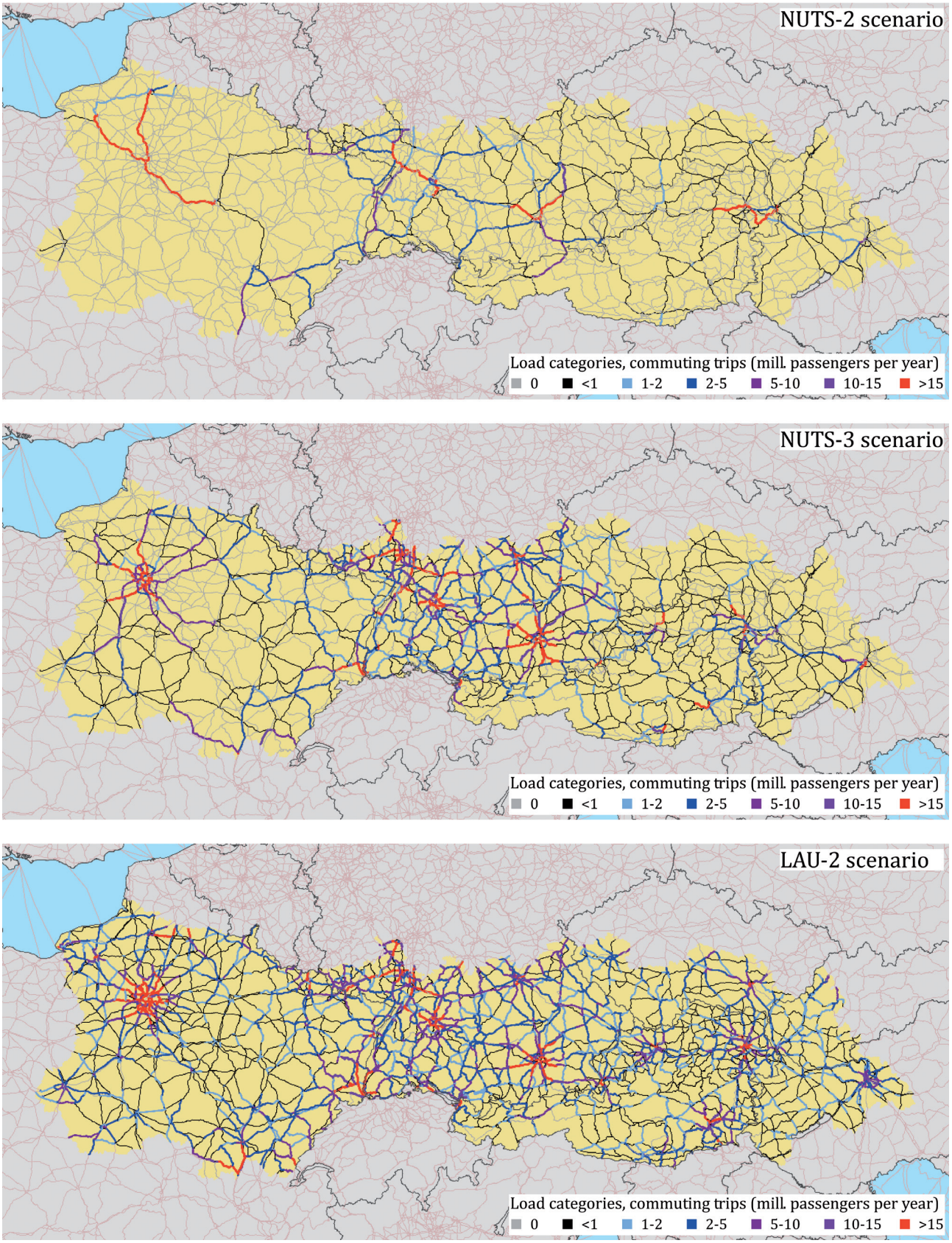

Figure 9.6: Private road commuting trip flows

A comparison of the three assignment results clearly indicates the potential of LAU-2 travel zones for transport modelling at European scale, given that the majority of the trips are carried out inter-zonally and are therefore considered for the network assignment. The quality of the NUTS-3 assignment seems to be acceptable, but only for countries like Germany which are covered by rather small NUTS-3 zones. In comparison to the LAU-2 map, however, many links in the NUTS-3 scenario are unloaded, e. g. in France and peripheral regions in the northeastern part of the Magistrale that are modelled by comparably large travel zones. The NUTS-2 assignment, in contrast, is almost worthless.

9.4.2.3 Investigating passenger loads outside the Magistrale area

In addition to the previous analysis, assignment maps for the whole of Europe were produced at NUTS-2 and LAU-2 level by considering the whole set of O/D relations (cf. Tab. 9.5) and not the reduced set that was used for the previous analysis. As regions outside the Magistrale are exclusively modelled at NUTS-2 level, no commuter passenger load values are computed for most of the network links in these regions as illustrated by Figure 9.7.

Both assignment maps seem to be comparable for regions outside the Magistrale. However, the map produced for the NUTS-2 scenarios shows some of those links in a higher load category than the LAU-2 scenario, for instance the two links around the Spanish capital Madrid, several links in Italy, in south-eastern European countries, in Great Britain and in the Netherlands.

Taking into account that more similar results could have been expected for regions outside the Magistrale, these discrepancies are inconvenient. They originate from the application of different deterrence functions for the three scenarios. In particular, the composite deterrence function (9.3) underlying the LAU-2 scenario was adjusted in order to overcome an overestimation of long-distance trips. This was mainly carried out by reducing the weighting factor ω_L related to the power function (f_p) from 0.05 to 0.01 and by changing the decay factor α_L from 0.8 to 1.3 (cf. Tab. 9.7), where:

$$f_p^{(\text{NUTS-2})} := 0.5\, c^{-0.8} \quad \gg \quad f_p^{(\text{LAU-2})} := 0.1\, c^{-1.3} \tag{9.6}$$

Thus, long-distance trips are overestimated in the NUTS-2 scenario compared to the LAU-2 scenario. This explains the higher load categories for regions outside the Magistrale.

The differences between the load categories outside the Magistrale area highlight an important issue with regard to the interpretation of model results: if we apply the same transport model at different hierarchical levels, the output indicators are not entirely comparable. The differences between the computed load categories for areas located along the Magistrale and areas located outside demonstrate another important issue: if we apply heterogeneous travel zones of different sizes for transport modelling, the output indicators differ significantly depending on the sizes of the travel zones. This again emphasises the need to apply homogeneous travel zones of equal size according to the homogeneity criteria discussed in section 7.1.1. Though LAU-2-based travel zones are not completely homogeneous, their size is small enough to analyse the distribution of commuting trips from a European perspective.

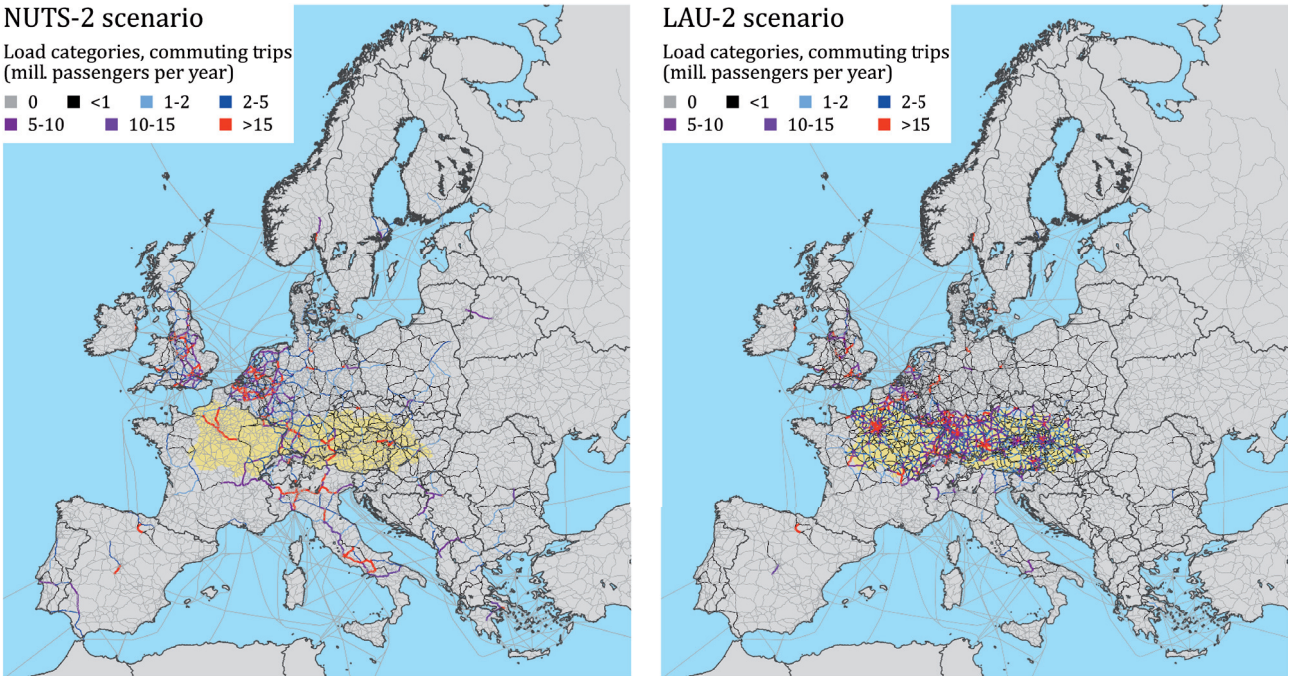

Figure 9.7: Comparison of load categories computed at NUTS-2 and LAU-2 levels

9.4.2.4 Investigating traffic between two neighbouring NUTS-2 regions

For the following analysis, only trips starting and ending within the Karlsruhe-Stuttgart region are considered. The respective O/D relations were extracted from the much larger data sets produced by the three main model runs of the prototype (cf. Tab. 9.5). The region is defined at NUTS-2 level by the two travel zones DE12 (Karlsruhe) and DE11 (Stuttgart). Accordingly, trips within this region are modelled by four O/D relations (two intra-zonal and two inter-zonal relations) for the first scenario in which NUTS-2 is the most detailed level.

Table 9.9: Overview of scenarios related to the Karlsruhe-Stuttgart region

Scenario	Level	Zones[1]	Diameter	Relations	Routings[2]	mill. trips[3]	mill. pkms[3]
Scenario 1	NUTS-2	2	104.9 km	4	1	1,943.0	33,534.7
Scenario 2	NUTS-3	25	28.0 km	625	300	1,943.0	32,938.9
Scenario 3	LAU-2	658	5.3 km	65,756	8,794	1,943.0	32,446.4

[1] Overall travel zones: 658 LAU-2, 286 LAU-1, 25 NUTS-3, 2 NUTS-2 regions.
[2] Number of routings considered for the assignment (excluding double counts).
[3] Total demand generated for this region (including trips ending outside this region).

Table 9.9 shows how the number of zones in this region increases from NUTS-2 to NUTS-3 and to LAU-2 level. In consequence, the numbers of O/D relations and of routings increase as well. For the LAU-2 scenario the number of routings seems to be comparably low. This is can be explained by the limited density of the applied network model. In consequence, several LAU-2 regions were often connected to the same feeding node and the computed routings are then identical for these regions. The three investigated scenarios are identical with regard to the overall trip demand that is generated within the Karlsruhe-Stuttgart region. However, they are slightly different with regard to the generated demand in terms of pkm. These differences are not a limitation for this analysis.

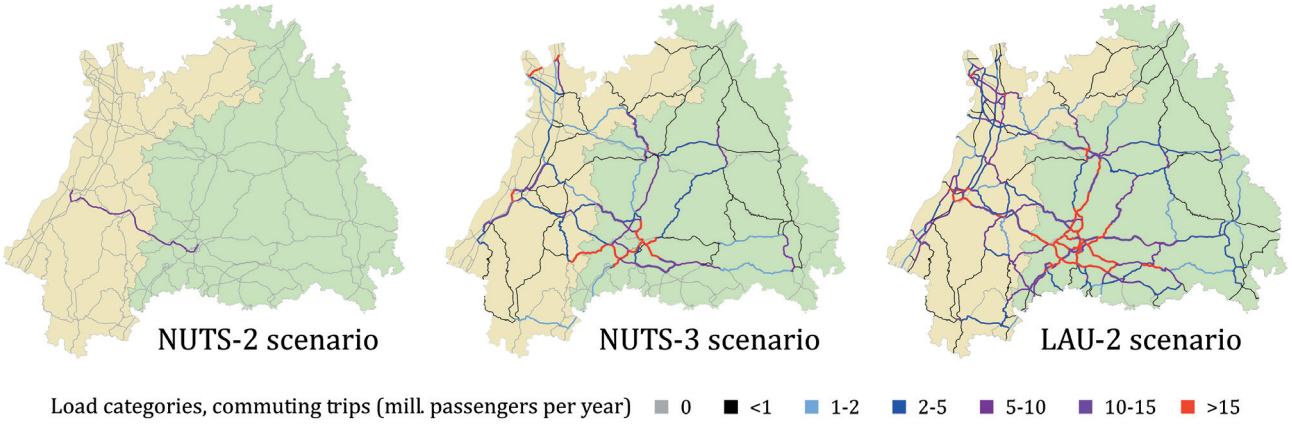

Figure 9.8: Modelled passenger load values for the Karlsruhe-Stuttgart region

The respective O/D relations, i. e. the trip matrix referring to the Karlsruhe-Stuttgart region, were processed by the network assignment model in order to visualise the computed trip demand. The passenger load values for the network links were computed by an all-or-nothing assignment. Figure 9.8 shows the computed load values for the three scenarios. Seven load categories are distinguished in which the highest is indicated in bright red, the lowest in light blue, and unloaded links are indicated in light grey.

At first sight, we immediately recognise enormous variations in quality. While the NUTS-2 assignment looks almost empty, the quality increases towards the LAU-2 assignment. The reason for this pattern is obvious: at NUTS-2 level, only one routing, i. e. the shortest path between Karlsruhe and Stuttgart, is considered for the assignment.[40] In contrast, over 8,794 different routings are considered for the assignment at LAU-2 level (cf. Tab. 9.9).

It should be emphasised that the computed load values also change from generally lower categories in the NUTS-2 assignment to higher ones in the NUTS-3 and in the LAU-2 assignments.[41] This pattern can be explained by the much larger size of the NUTS-2 zones compared to the LAU-2 zones and a large travel distance of over 70 km between Karlsruhe and Stuttgart. As a result, the majority of the trip demand is intra-zonal if the Karlsruhe-Stuttgart region is modelled at NUTS-2 level, while only one quarter is intra-zonal if the region is modelled at LAU-2 level.

Table 9.10: Share of intra-zonal and inter-zonal trips in the Karlsruhe-Stuttgart region

Scenario	Level	T_{ii}[1]	T_{ij}[1]	T_{ik}[1]	pkm_{ii}[1]	pkm_{ij}[1]	pkm_{ij}[1]
Scenario 1	NUTS-2	98.2%	0.3%	1.5%	92.8%	1.7%	5.5%
Scenario 2	NUTS-3	63.9%	29.1%	7.0%	32.7%	38.1%	29.2%
Scenario 3	LAU-2	25.9%	62.2%	11.9%	4.8%	49.7%	45.6%

[1] Travel zones i, j are located within the region, and travel zone k is located outside.

Table 9.10 shows how the share of intra-zonal trips in this region decreases with decreasing size of the travel zones. Taking into account the low share of intra-zonal trips for the LAU-2 scenario it can be

[40] Intra-zonal O/D relations exist as well, but they are not relevant for the computation of traffic loads at link level.
[41] For some links in the NUTS-3 scenario higher load values are computed than for the LAU-2 scenario, because at LAU-2 level more routings are considered for the network assignment.

concluded that LAU-2-based travel zones are the most appropriate for transport modelling. Given that only about one third of the demand in terms of pkm is intra-zonal for the NUTS-3 scenario, NUTS-3-based travel zones might also be appropriate for transport modelling if the network assignment is computed in incremental steps.[42]

9.4.3 Summary

This section discussed the application of the HIPAT prototype for three test scenarios, in which the Magistrale is modelled at NUTS-2, NUTS-3 and LAU-2 level. Each test scenario is calibrated for an average trip length of 19.6 km by adjusting the deterrence function (cf. Tab. 9.7). Given that slightly different modelling coefficients are applied, the output indicators produced by the three test scenarios are not entirely comparable. For instance, the deterrence function had to be adjusted for the LAU-2 scenario in order to reduce the occurrence of long-distance trips. In consequence, the LAU-2 scenario tends to compute lower load categories for regions outside the Magistrale than the NUTS-2 scenario even though these regions are modelled in both scenarios at NUTS-2 level (cf. Fig. 9.7).

The three test scenarios were investigated and evaluated based on four analyses demonstrating the necessity to apply the prototype at LAU-2 level rather than at NUTS-2 level. In the first step, the produced TLDs were compared to the applied deterrence functions, revealing a good compliance only for the LAU-2 scenario, while the discrepancy between the TLD and the deterrence functions was significant for the NUTS-2 scenario. In the second step, passenger loads were computed following an all-or-nothing method for the Magistrale area. The analysis revealed very good results for the LAU-2 scenario while the assignment computed at NUTS-2 level was almost worthless. The third and fourth analyses were also carried out based on network assignments and confirmed the finding that LAU-2 zones are the most appropriate for transport modelling.

[42] The incremental application of the HIPAT model also makes it possible to consider different routings at NUTS-3 level when assigning the trip demand to the network links.

9.5 General conclusions

This chapter discussed the HIPAT modelling approach operating at LAU-2 level that was developed based on the IPAT model operating at NUTS-2 level. As the number of travel zones increases from 300 to over 100,000 between the two models, a prototype with a reduced scope was implemented in order to test the model concept and to gain experience. The implemented prototype focuses on private road commuting trips up to a maximum length of 280 km. It covers the whole of Europe, but only an area around the Magistrale corridor is modelled at LAU-2 level by 26,891 travel zones. Outside the Magistrale, travel zones are modelled at NUTS-2 level and neighbouring countries at NUTS-0 level. In total, the prototype relies on 33,191 travel zones defined at different hierarchical levels.

9.5.1 Overview

The HIPAT prototype was implemented in `Java` based on the existing code of the IPAT model. Its runtime is about two minutes, including the four modelling steps of trip generation, trip distribution, mode choice and network assignment. Taking into account that the current prototype has not been optimised concerning parallel computing[43], and that the entire ETISplus road network model covering the whole of Europe is already used, it is expected that the HIPAT modelling approach can be implemented at LAU-2 level for the whole of Europe. Furthermore, a runtime of less than one hour seems to be feasible on a standard computer even if the HIPAT model is executed in incremental steps.

The major evolutionary step between the IPAT and the HIPAT models is the development of the hierarchical zoning system, which enables O/D relations to be modelled at different hierarchical levels. This computes long-distance trips, for instance, very efficiently at NUTS-1 level and medium-distance trips efficiently at NUTS-3 level. Following this hierarchical concept, only short-distance trips are computed at LAU-2 level.

The optimal level of detail is determined by the so-called neighbourhood criterion that was developed within this thesis. Based on the locations of the origin and the destination, the criterion determines whether origin and destination are far enough away from each other. If this is the case, the respective O/D relation can be modelled at the considered level of detail. Otherwise, the trip estimates between the origin and the destination have to be modelled at a more disaggregated level. Following this concept and applying the neighbourhood criterion recursively, trip estimates between travel zones are always computed at the most aggregated level possible. For the HIPAT prototype, this reduces the number of O/D relations to 0.1% in comparison to a complete trip matrix.[44]

9.5.2 General findings

The advantages of applying LAU-2-based travel zones can be clearly identified when analysing the produced output indicators for the three investigated test scenarios in which the Magistrale was modelled at NUTS-2, NUTS-3 and LAU-2 levels. Given the large size of the NUTS-2 regions underlying the first scenario, the share of inter-zonal trips is only 3.4% of all trips (cf. Tab. 9.8) and the computed

[43] In the current version of the prototype, the assignment step is scheduled in parallel on eight cores.

[44] The hierarchical trip matrix consists of only 10,914,221 O/D relations (Tab. 9.5). In contrast, a complete trip matrix would consist of 737,068,201 O/D relations (cf. Tab. 9.6).

network assignment looks almost empty (cf. Fig. 9.6). However, the produced assignment map outlines some network links in the highest load categories. For this reason, the NUTS-2 assignment is difficult to interpret and, in fact, of little use, either for the validation of the HIPAT prototype or for the presentation of the model results. Moreover, the TLD derived at NUTS-2 level shows no similarity at all to the applied deterrence function (cf. Fig. 9.5). It can therefore be concluded that NUTS-2-based travel zones do not provide any basis for the modelling of commuting trips.

In contrast to the NUTS-2 scenario, the share of the inter-zonal demand in terms of pkm is 95.9% for the LAU-2 scenario. Consequently, the computed network assignment looks reasonable. As it provides detailed insights into the distribution of passenger load values, it can be used for the validation of the prototype and for the presentation of the model results. In addition, the TLD derived at LAU-2 level resembles a continuous distribution for all trips and even for intra-zonal trips (cf. Fig. 9.5). It can therefore be concluded that LAU-2-based travel zones provide a sound basis for the modelling of commuting trips.

Besides the NUTS-2 and the LAU-2 scenarios, a further scenario was investigated in which the travel zones located along the Magistrale are modelled at the intermediate NUTS-3 level. This scenario outperforms the NUTS-2 scenario, but the quality of the model output is clearly below that of the LAU-2 scenario. It can therefore be concluded that NUTS-3-based travel zones are an option for the modelling of vacation trips, for instance, given that these trips are generally longer than commuting trips.

9.5.3 Further findings

The HIPAT model requires input data at LAU-2 level. Given the lack of respective European datasets, these input datasets had to be generated. Regional indicators were computed by downscaling based on regional data available at NUTS-3 level, gridded population data and land cover data. Artificial city districts were generated taking into account the distribution of the population in the cities. In addition, LAU-2 regions were copied to the LAU-1 layer to close existing data gaps in this layer, e. g. for Austria. The feeding nodes were subsequently determined for each travel zone, the set of O/D relations was generated and the travel impedances were computed. It can therefore be concluded that the non-availability of disaggregated European datasets covering regional indicators and city districts is not a major obstacle for the development of a European transport model operating at LAU-2 level.

However, it has to be stressed that the applied process to generate the required input data for the prototype was very complex. Within this thesis, the input data was produced following a chain of tailored scripts. These scripts were implemented in different programming languages and software environments including, for instance, `Python`, `Java`, `Postgres`, `Excel` and `ArcMap`. In order to ensure the consistency of the input data produced, each intermediate dataset produced along the chain was checked manually. Several times, individual scripts or the whole chain of scripts had to be revised. Besides this very time-consuming and complex process, it was found that applying the neighbourhood criterion to determine the set of O/D relations generated many false positives. Hence, some trip estimates between two travel zones are currently computed at a more detailed level than necessary.

A further shortcoming of the current input datasets is the low density of the ETISplus road network model that was used in the prototype. Its density is sufficient for transport modelling at NUTS-3 level,

but at LAU-2 level about five travel zones had to be connected on average to each node of the network model. If two travel zones are connected to the same node, the respective travel impedances cannot be derived based on the shortest paths but have to be estimated based on the crow-flies distance between the centroids, and the diameter of the two travel zones.

Besides the challenge of providing the required input datasets for the prototype, its calibration was also very challenging and time-consuming, for several reasons. First of all, the produced text file storing the hierarchical trip matrix is about 360 megabytes. Given that several output indicators had to be joined in order to calibrate and to validate the prototype, the produced text files could not be processed by standard tools such as `Excel`, so tailored scripts had to be developed, e. g. in `Python`. In addition, aggregated output indicators such as "originating pkm" cannot be derived straightforwardly for a travel zone, given that the O/D relations are modelled at different hierarchical levels according to the HIPAT modelling approach. In this case, rather complex disaggregation rules have to be applied in order to process and to disaggregate the hierarchical trip matrix.[45]

It should be stressed that the specific indicator of originating pkm is required in order to ensure that the transport model also produces reliable results for each individual travel zone rather than just reliable results for each country on average. For the prototype an extensive analysis of this indicator for each LAU-2 zone was not judged practicable, given that the deterrence function underlying the prototype was calibrated to an average European mobility level. This average mobility level is neither representative for city residents carrying out many intra-city trips nor for residents living in rural regions carrying out rather longer trips. For this reason, the results of the current version of the prototype can only be analysed at a more aggregated level.

9.5.4 Outlook

For the implementation of the complete HIPAT model covering the whole of Europe as well as all transport modes and trip purposes, several prerequisites will be required. First of all, the complex process related to the generation of the input data has to be simplified. This could be achieved by migrating the whole chain of scripts into a powerful GIS software environment. It should be noted that within this thesis `ArcMap` was applied for several tasks but the software was often pushed beyond its limits and crashed.

In addition, the zoning system should be improved. Neighbouring countries should be modelled at a more disaggregated level in order to avoid O/D relations between adjacent travel zones that are too heterogeneous and therefore not consistent, e. g. between a LAU-2 and a NUTS-0 region. In such cases, the deterrence model cannot be properly adjusted by the scaling function that computes a scaling factor based on the average diameter of the travel zones. The scaling factor can only be computed with high accuracy if the origin and the destination are modelled at the same hierarchical level and if the diameter of all travel zones related to this level is almost identical. This is currently not the case, as travel zones are defined according to the NUTS classification. For instance, the French NUTS-3 regions are much larger than the German NUTS-3 regions.

[45] For instance, trips originating in Karlsruhe (DE122) and ending in Augsburg (DE271) are modelled as part of the aggregated O/D relation DE12/DE27 (Karlsruhe/Schwaben) at NUTS-2 level.

This non-comparability raises the question of how the current zoning system underlying the HIPAT prototype can be improved. One possible solution is to subdivide very large NUTS-3 zones and, in consequence, to extend the NUTS-3 layer. Furthermore, a spatial cluster algorithm should be developed in order to compute more homogeneous travel zones at LAU-1 level by merging the respective LAU-2 zones. However, the concept of defining the travel zones based on administrative units should be carefully reconsidered, given that the definition of homogeneous grid cells arranged in a quadtree data structure promises greater improvements. While the prototype uses about 27,000 LAU-2 zones (cf. Tab. 9.4), the quadtree example uses about 66,000 grid cells at the most detailed level (cf. Tab. 8.1, sec. 8.1.2). However, about 11,000,000 O/D relations have to be modelled in the prototype while only about 3,000,000 O/D relations have to be modelled in the quadtree example.

Besides reducing the number of O/D relations, false positives that are currently produced by the neighbourhood criterion could be systematically avoided if the travel zones were defined based on homogeneous grid cells following the quadtree example. However, the concept of grid cells cannot easily be combined with the NUTS classification. Hence, the trade-off between the concepts should be a subject of further research, along with a general revision of the hierarchical zoning system.

Though the implementation of the prototype highlighted many pending issues, the primary goals could be reached. First of all, the prototype was successfully implemented and a runtime of about two minutes was achieved. This is a significant improvement compared to the TRANS-TOOLS model, which has a runtime of over two days (cf. Ibañez-Rivas, 2010). In addition, the advantages of LAU-2-based travel zones for transport modelling could be demonstrated. Taking into account the significant discrepancies between the applied deterrence functions and the TLDs produced at NUTS-2 and NUTS-3 level outlined by Figure 9.5, it has to be concluded that only LAU-2-based travel zones provide a sound basis for transport modelling. However, the path towards the implementation of the HIPAT model for the whole of Europe will be very challenging, and additional research is required in order to solve some of the pending issues.

Part IV

Résumé

Chapter 10

Summary and Outlook

10.1 Summary

In this thesis, a series of transport models and modelling methods have been developed, culminating in the HIPAT modelling approach. The aim was to model European passenger transport at LAU-2 level with more than 100,000 travel zones with an estimated runtime of about one hour on a standard computer. Due to the comparably small size of the LAU-2-based travel zones, regional trip flows can be modelled, maybe for the first time, with high accuracy within one consistent European transport model. This is of particular importance with regard to private road commuting trips, which contribute heavily to capacity overloads in the trans-European networks. In contrast to current European transport models, which rely on much larger travel zones, the HIPAT modelling approach enables European policy makers to investigate and to assess the effectiveness of potential transport policy measures, including the impacts on regional trip flows. This will improve the basis for decision-making.

The starting point of the HIPAT approach is the standard macroscopic four-step transport modelling approach. The major difference between existing European transport models and the HIPAT approach is the number of applied travel zones and their size. While current European transport models operating at NUTS-3 level rely on about 1,500 travel zones with an average diameter of about 50 km (and a large variation of diameters), the LAU-2 level encompasses more than 100,000 travel zones with an average diameter of about 5 km. Moreover, some LAU-2-equivalent (or even smaller) zones are created if this is recommended, e. g. cities are subdivided into artificial districts. The disaggregation of travel zones from NUTS-3 to LAU-2 level addresses a well-known critical issue in transport modelling, namely the computation of intra-zonal trips and trips between neighbouring travel zones.

Prima facie, such disaggregation has a serious drawback: the complexity of the trip matrix depends quadratically on the number of travel zones. The HIPAT approach tackles this issue by modelling trip estimates between travel zones at different hierarchical levels, so only short-distance trips are modelled at LAU-2 level while long-distance trips are modelled at higher levels, up to the NUTS-0 level. However, the key target to develop a European passenger transport model operating simultaneously at different hierarchical levels raised several follow-up problems, including (i) transferability of the deterrence function and logit model between different levels, (ii) computation of the accessibility indicator underlying the trip distribution model, (iii) data availability of regional indicators at LAU-2 level and (iv) generation of artificial city districts. The first two issues were solved by methodological

revisions of the trip distribution and the modal split models in order to ensure the transferability of the HIPAT approach between different levels. The last two issues concern well-known data availability problems that were tackled by exploiting European GIS datasets providing georeferenced indicators up to a spatial resolution of one hectare that are available, free of charge, to the public.

The development of the HIPAT transport modelling approach was a long process carried out in three stages. In the first stage, a fairly consistent European data basis was compiled for the year 2010 at NUTS-3 level, covering regional indicators, transport networks and mobility coefficients. The PAT model was then developed in accordance with existing transport modelling approaches, in order to compute the trip matrices for updating the European Transport Policy Information System (ETIS). Besides the limitation to only inter-zonal rail and road passenger trips, the development of the PAT model revealed several methodological issues pertaining to received European transport modelling. The most crucial issues were scale dependencies of the deterrence function and logit model.[1] Within the PAT model these problems were circumvented by, for example, introducing three additional sets of calibration parameters related to the trip generation, the trip distribution and the modal split models.

In the second stage, these methodological issues were solved and the IPAT model, covering four modes of transport, four trip purposes, intra-zonal and inter-zonal trips, was developed. The main revisions were carried out for the trip distribution and the modal split models in order to overcome the scale dependencies of the deterrence function and the logit model, and to enable the transferability of the IPAT model without the need to apply additional calibration parameters. First of all, the deterrence function underlying the trip distribution model was replaced by a more complex deterrence model relying, inter alia, on a scaling function that is based on the diameter of travel zones. In addition, an often overlooked issue related to the multinomial logit model was tackled by making the scaling parameter, or rather the error term, dependent on the travel distance.

It should be stressed that both methodological improvements were developed from scratch within this thesis and are therefore unique features of the IPAT model. However, given that both revisions are crucial for the consistent application of a transport model at European scale and for different policy scenarios, the question arises as to which second-best solutions might currently be implemented in other European transport models.

In the third stage, the HIPAT modelling approach was developed, starting from the existing NUTS-3 data basis and from the consistent methodology of the IPAT model. The main improvements are (i) the development of the hierarchical data structures underlying the HIPAT model, (ii) the development of the neighbourhood criterion determining the set of O/D relations that are modelled at different hierarchical levels and (iii) the development of the hierarchical accessibility indicator ensuring the overall model consistency. The hierarchical procedure reduces the number of O/D relations by about 99%, enabling the HIPAT model to be applied at LAU-2 level.

As already mentioned, a crucial step towards the formulation of the HIPAT modelling approach was the development of the IPAT model during the HIGH-TOOL research project that was recently awarded

[1] Scale dependencies limit the transferability of transport models. If the deterrence function is calibrated at NUTS-3 level for travel zones with a diameter of 50 km, it cannot be applied at another level, e. g. LAU-2 and NUTS-2. If the modelling coefficients of the logit model are determined for a certain level of travel costs (e. g. for regional trips from 50 to 100 km), the logit model cannot be applied for other cost levels (i. e. for short-distance and long-distance trips).

the German Mobility Prize 2017. In this project, the HIGH-TOOL model operating at NUTS-2 level was developed for the European Commission to assess the economic, social and environmental impacts of transport policies in 5-year steps from 2010 to 2050. The HIGH-TOOL model follows a modular approach in which the passenger transport demand was computed by the PAD module using a slightly extended version of the IPAT model (but at NUTS-2 level only). In the context of the publication of the HIGH-TOOL model, the methodology of the IPAT model was intensively tested and validated by a broader audience through a series of case studies and consistency checks. Though it was found that the sensitivity of the IPAT model to changes in the cost variables seems to be rather low, all model results and reactions are plausible. It should be stressed in this context that the limited sensitivity of the IPAT model operating at NUTS-2 level can be clearly attributed to the very high proportion of intra-zonal trips.

The development of the HIPAT model started as a copy of the source code of the IPAT model. Bearing in mind the large increase in the number of travel zones between these two models, only a prototype implementation was realised within this thesis in order to investigate the performance of the HIPAT modelling approach and to provide a proof of concept. The prototype focuses (i) on private road commuting trips with a rather regional character and (ii) on a particular region of central Europe called the "Magistrale". This region is an important European transport corridor between Paris and Budapest. Only the Magistrale region has been modelled at the detailed LAU-2 level, while the remainder of Europe has been modelled at NUTS-2 level as inherited from the IPAT model. This implementation of the HIPAT model is built on about 33,000 travel zones and has a runtime of about two minutes. It can therefore be expected that the HIPAT approach can be implemented successfully for the whole of Europe, i. e. for more than 100,000 travel zones, for four transport modes and four trip purposes. Moreover, a runtime of less than one hour seems to be possible on a standard computer, even if the HIPAT model is executed in incremental steps in order to further improve its accuracy.

In addition to this general proof of concept regarding implementability and tractability, the need to apply LAU-2-based travel zones for transport modelling could be clearly demonstrated. This was achieved by comparing two output indicators of three test scenarios in which the HIPAT prototype was applied at NUTS-2, NUTS-3 and LAU-2 levels. Both indicators were computed based on the corresponding trip matrices.[2]

Taking into account the huge discrepancy for the NUTS-2 scenario between the derived trip length distribution (TLD) and the applied deterrence function as well as the weak assignment results, it can be concluded that NUTS-2-based travel zones do not actually provide an acceptable basis for the modelling of regional trips. If calculated at NUTS-3-level, the assignment results improve significantly but the discrepancy between the TLD and the applied deterrence function is still very large. It can therefore be concluded that NUTS-3-based travel zones are also not really suitable for the modelling of regional trips. The best results were achieved when the HIPAT prototype was applied at LAU-2 level. In this case, the TLD resembles a continuous distribution and its pattern is very similar to that of the applied deterrence function. Moreover, the assignment results improve further from NUTS-3 to LAU-2 level. It can therefore be concluded that only LAU-2-based travel zones provide a sound basis for the modelling of regional trips.

[2] See Figures 9.5 and 9.6 in section 9.4.2 and the subsequent discussion.

10.2 Outlook

In this thesis, the HIPAT modelling approach was introduced and successfully implemented for a prototype model with a limited scope. Taking into account the significant progress achieved between the results produced at NUTS-3 and LAU-2 levels by the prototype, a worthwhile future task will be to complete the implementation of a HIPAT model for the whole of Europe, encompassing more than 100,000 travel zones, four transport modes and four trip purposes. In addition, the HIPAT modelling approach might also be considered for national transport models facilitating, for instance, the modelling of urban trips at a very disaggregated level and inter-city trips at NUTS-3 level.

The step-by-step procedure encompassing the implementation of the PAT model, the IPAT model and the HIPAT prototype revealed the complexity of revising the well-known four-step approach in transport modelling. For instance, methodological issues related to the scale dependencies of the deterrence function and the logit model that were identified during the development of the PAT model had been intensively studied and solved within the development of the IPAT model, before the main target of this thesis could finally be achieved by the development of the HIPAT modelling approach. The findings for the logit model might also be relevant for other researchers, besides transport modellers, given the popularity of this model.

Taking into account the findings of this thesis, the HIPAT modelling approach promises considerable improvements for transport models operating at large scale, and therefore also for the simulation and assessment of European transport policies. However, the intensive work with all relevant components forming a transport model within this thesis, reveals several challenges of proceeding from the HIPAT prototype towards a fully sufficient European HIPAT model.

Although the revision of the programming code seems to be merely a formality, a crucial question arises with regard to the calibration and the validation of the final model. As the trip matrix produced by the prototype is about 360 megabytes, it can be expected that the trip matrix produced by a full-scale European HIPAT model will increase to several gigabytes. For the processing of this matrix tailored tools must be developed, as the O/D relations refer to different hierarchical levels. Furthermore, the prototype was only calibrated for one transport demand segment, while the calibration of a European model will encompass at least 16 transport demand segments in 40 countries.

Another crucial question arises with regard to the compilation of the required input data basis. First of all, the hierarchical zoning system should be revised by introducing more artificial travel zones in order to improve their homogeneity. In the second step, the density of the network models should be increased, facilitating the connection of all travel zones to different feeding nodes. In addition, the applied method for downscaling regional indicators from NUTS-3 level to LAU-2 level then has to be extended to enable the computations of all required indicators besides population and workplaces. Besides these crucial questions, several other parts of the methodology could be further investigated in the future, for instance the hierarchical accessibility indicator and the neighbourhood criterion. The continued implementation of the HIPAT modelling approach therefore provides plenty of opportunities to carry out additional investigations and to improve European transport modelling.

Bibliography

Ahrens, G.-A., Ließke, F., Wittwer, R., and Hubrich, S. (2009). *Endbericht zur Verkehrserhebung "Mobilität in Städten – SrV 2008" und Auswertungen zum SrV-Städtepegel*. Institut für Verkehrsplanung und Straßenverkehr, Technische Universität Dresden.

Annoni, A., Luzet, C., Gubler, E., and Ihde, J. (2001). *Map Projections for Europe*. European Commission, Joint Research Centre, Institute for Environment and Sustainability.

Anselin, L., and Rey, S. J. (2010). *Perspectives on Spatial Data Analysis*. Advances in Spatial Science. Springer-Verlag Berlin Heidelberg.

Bal, H. E., and Haines, M. (1998). *Approaches for integrating task and data parallelism*. IEEE Concurrency, *6*(3), 74–84.

Batz, G. V., and Sanders, P. (2012). *Time-Dependent Route Planning with Generalized Objective Functions*. In *Algorithms – ESA 2012: 20th Annual European Symposium, Ljubljana, Slovenia, September 10-12, 2012. Proceedings*, (pp. 169–180). Berlin, Heidelberg: Springer Verlag.

Ben-Akiva, M., and Lerman, S. (1985). *Discrete choice analysis: theory and application to travel demand*. MIT Press, Cambridge, MA.

Berglund, S., and Algers, S. (2016). *Final report on the passenger demand model in TT3*. Department of Management Engineering, Lyngby, Denmark. Deliverable 8.2 of the TRANS-TOOLS 3 project funded by the European Commission.

Bickel, P., Friedrich, R., Burgess, A., Fagiani, P., Hunt, A., De Jong, G., Laird, J., Lieb, C., Lindberg, G., Mackie, P., Navrud, S., Odgaard, T., Ricci, A., Shires, J., and Tavasszy, L. (2005). *Deliverable 5, Proposal for Harmonised Guidelines*. HEATCO (Developing Harmonised European Approaches for Transport Costing and Project Assessment). Funded by the European Commission under the 6th Framework Programme.

Bloch, V. V. H., Gundersen, G. I., Thorsdalen, B., Backer, L. H., van Leeuwen, N. F., Lipatz, J., Tammilehto-Luode, M., Tammisto, R., Makarenko-Piirsalu, D., Marsik, K., Jablonski, R., Santos, A. M., Kuzma, I., and Kaminger, I. (2012). *GEOSTAT 1A – Representing Census data in a European population grid, Final Report*. European Forum for Geography and Statistics. Data set.

Bonnafous, A., Crozet, Y., Mercier, A., Ovtracht, N., and Thiebaut, V. (2009). *MOSART (MOdélisation et Simulation de l'Accessibilité aux Réseaux et aux Territoires): un prototype d'outil d'aide à la décision, individuelle et collective pour une mobilité durable*. Rapport Final: Laboratoire d'Economie des Transports (LET), Lyon.

Bonnafous, A., Le Nir, M., Raux, C., Routhier, J.-L., and Tabourin, E. (1989). *Les effets frontières*. Archive ouverte, HAL id: halshs-00817850.

Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Springer-Lehrbuch. Heidelberg: Springer.

Box, G., and Cox, D. (1964). *An Analysis of Transformations*. Journal of the Royal Statistical Society, *26*, 211–243.

Bundesamt für Kartographie und Geodäsie (BKG) (2013). *EuroBoundaryMap v8.1*. EuroGeographics, BKG. Data set.

Bundesamt für Statistik (BFS) (2017). *Grenzgängerstatistik im 4. Quartal 2016 - Zunahme der Anzahl Grenzgängerinnen und Grenzgänger*. Statistik der Schweiz. Medienmitteilung.

Burgess, A., Chen, T., Snelder, M., Schneekloth, N., Korzhenevych, A., Szimba, E., Kraft, M., Krail, M., Nielsen, O., Hansen, C., Martino, A., Fiorello, D., and Christidis, R. (2008). *Final Report TRANS-TOOLS (TOOLS for TRansport forecasting ANd Scenario testing), Deliverable 6, funded by the European Commission under the 6th Framework Programme*. TNO Inro, Delft, Netherlands.

Cascetta, E. (2009). *Transportation Systems Analysis : Models and Applications*. Springer Verlag.

Cascetta, E., and Coppola, P. (2014). *High Speed Rail (HSR) induced demand models*. Procedia - Social and Behavioral Sciences, *111*, 147–156.

Cascetta, E., Nuzzolo, A., Russo, F., and Vitetta, A. (1996). *A modified logit route choice model overcoming path overlapping problems. Specification and some calibration results for interurban networks*. In *Proceedings of the 13th International Symposium on the Theory of Road Traffic Flow (Lyon, France)*. Pergamon.

Cascetta, E., Pagliara, F., and Papolla, A. (2007). *Alternative approaches to trip distribution modelling: A retrospective review and suggestions for combining different approaches*. Papers in Regional Science, *86*(4), 597–620.

Chandra, R., Dagum, L., Kohr, D., Maydan, D., McDonald, J., and Menon, R. (2001). *Parallel programming in OpenMP*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Christiansen, Hjalmar (2011). *The Danish National Travel Survey - declaration of variables*. DTU Transport. Data set, TU 2006-10, version 1.

Christidis, P., and Ibañez-Rivas, J. N. (2012). *Measuring road congestion*. European Commission, Joint Research Centre; Institute for Prospective Technological Studies, Luxembourg: Publications Office of the European Union. Technical note.

Communication from the Commission to the Council and the European Parliament (COM) (2002). *Towards an integrated European railway area*. Commission of the European Communities, Brussels. COM(2002)18 final.

COWI, IWW, NEA, OBET, ETC and others (COWI et al.) (2006). *Feasibility study on Rail Baltica Railways, Final Report*. project on behalf of the European Commission, DG Regional Policy.

Crozet, Y. (2012). *The Three Stages of Accessibility: The Coming Challenge of Urban Mobility*. In *Sustainable Transport for Chinese Cities*, vol. 3 of *Transport and Sustainability*, (pp. 79–97). Emerald Group Publishing Limited.

Crozet, Y., Mercier, A., and Ovtracht, N. (2012). *Évaluer les impacts sociaux des politiques de mobilité urbaine : de l'accessibilité spatiale à l'accessibilité sociale*. Cahiers de géographie du Québec, *56*(158), 381–403.

De Ceuster, G., van Herbruggen, B., Ivanova, O., Carlier, K., Martino, A., and Fiorello, D. (2007). *TREMOVE service contract for the further development and application of the transport and environmental TREMOVE model Lot 1 (Improvement of the data set and model structure)*. European Commission Directorate General Environment, Brussels. Final Report.

de Grange, L., Fernández, E., and de Cea, J. (2010). *A consolidated model of trip distribution*. Transportation Research Part E: Logistics and Transportation Review, *46*(1), 61–75.

de Jong, G., Daly, A., Pieters, M., and van der Hoorn, T. (2007). *The logsum as an evaluation measure: Review of the literature and new results*. Transportation Research Part A: Policy and Practice, *41*(9 SPEC. ISS.), 874–889.

de Jong, G., Pieters, M., Daly, A., Graafland, I., Kroes, E., and Koopmans, C. (2005). *Using the Logsum as an Evaluation Measure: Literature and Case Study*. Working Paper WR-275-AVV, RAND Europe, Prepared for AVV Transport Research Centre.

Dechoux, C. (2013). *Quelle place pour la Lorraine dans la Grande Région et dans l'Union européenne? Etude des dynamiques économiques d'une région au coeur de différents espaces européens*. Institut d'études politiques de Lyon, Université Lumière Lyon 2, Lyon. Mémoire de Séminaire.

Delling, D., Dibbelt, J., Pajor, T., Wagner, D., and Werneck, R. F. (2013). *Computing Multimodal Journeys in Practice*. In *Experimental Algorithms: 12th International Symposium, SEA 2013, Rome, Italy, June 5-7, 2013. Proceedings*, (pp. 260–271). Berlin, Heidelberg: Springer Verlag.

Demuth, N. (2004). *Der ICE als Pendler- und Vorortzug? Die ICE-Bahnhöfe Montabaur und Limburg – Impulse für Wohnstandortwahl, Wohnsiedlungsentwicklung und berufliche Mobilität*. Master's thesis, Fachbereich VI – Geographie/Geowissenschaften, Universität Trier, Trier.

Department for Transport (DfT) (2008). *National Travel Survey: 2008*. National Statistics, London, United Kingdom. Data set.

DGB Abteilung Arbeitsmarktpolitik (DGB) (2016). *Mobilität in der Arbeitswelt: Immer mehr Pendler, immer größere Distanzen*. DGB publication "Arbeitsmarkt aktuell".

Dijkstra, E. W. (1959). *A note on two problems in connexion with graphs*. Numerische Mathematik, *1*(1), 269–271.

Ding, C. (1998). *The GIS-based Human-Interactive TAZ Design Algorithm: Examining the Impacts of Data Aggregation on Transportation-Planning Analysis*. *25*, 601–616.

Eberhard, C., Heinitz, F., Last, J., Mandel, B., Rothengatter, W., and Schoch, M. (1998). *An Integrated Framework for the STEMM Passenger Model, STEMM WP1: Passenger Modelling Methodology*. Report of the STEMM project funded by the European Commission, Karlsruhe/Strasbourg: IWW, MKm/ BETA.

Ellinghaus, A., Haupt, S., Wiggenhauser, V., Pfeifer, J., and Pellinghoff, B. (2017). *Planfeststellungsbeschluss - B10/18, Bau einer zweiten Rheinbrücke zwischen Karlsruhe und Wörth im Zuge der B10*. Regierungspräsidium Karlsruhe, Baden-Württemberg, Karlsruhe.

European Commission (EC) (2008). *Nomenclature of Territorial Units for Statistics (NUTS) 2006 - Statistical Units*. European Commission, Eurostat (ESTAT), GISCO. Data set.

European Commission (EC) (2009). *EU energy and transport in figures: statistical pocketbook*. Publications Office of the European Union, Luxembourg.

European Commission (EC) (2013). *EU energy, transport and GHG emissions: trends to 2030 : Reference Scenario 2013*. Luxembourg: Publication Office of the European Union.

European Commission (EC) (2015). *EU transport in figures: statistical pocketbook*. Publications Office of the European Union, Luxembourg.

European Commission (EC) (2016). *EU transport in figures: statistical pocketbook*. Publications Office of the European Union, Luxembourg.

European Environment Agency (EEA) (2007). *CLC-2006 technical guidelines*. Office for Official Publications of the European Communities, Luxembourg.

Eurostat (2004). *Regions in the European Union - Nomenclature of territorial units for statistics, NUTS-2003/EU-25*. Office for Official Publications of the European Communities, Luxembourg.

Eurostat (2007). *Regions in the European Union - Nomenclature of territorial units for statistics, NUTS-2006/EU-27*. Office for Official Publications of the European Communities, Luxembourg.

Eurostat (2011). *Regions in the European Union - Nomenclature of territorial units for statistics, NUTS-2010/EU-27*. Office for Official Publications of the European Communities, Luxembourg.

Eurostat (2015a). *People in the European Union: who are we and how do we live?*. Office for Official Publications of the European Communities, Luxembourg.

Eurostat (2015b). *Regions in the European Union - Nomenclature of territorial units for statistics, NUTS-2013/EU-28*. Office for Official Publications of the European Communities, Luxembourg.

Everitt, B., and Hothorn, T. (2011). *An introduction to applied multivariate analysis with R*. New York: Springer.

Finkel, R. A., and Bentley, J. L. (1974). *Quad trees - a data structure for retrieval on composite keys*. Acta Informatica, *4*(1), 1–9.

Fiorello, D., De Stasio, C., Köhler, J., Kraft, M., Newton, S., Purwanto, J., Schade, B., Schade, W., and E., S. (2009). *The iTREN-2030 reference scenario until 2030. Deliverable 4 of iTREN-2030 (Integrated transport and energy baseline until 2030)*. Milan, Italy: Project co-funded by European Commission 6th Framework Programme.

Flacke, W. (Ed.) (2010). *Koordinatensysteme in ArcGIS : Praxis der Transformationen und Projektionen*. Norden: Points Verlag, 2nd ed.

Forschungsgesellschaft für Straßen- und Verkehrswesen, Arbeitsgruppe Verkehrsplanung (FSV) (1997). *Empfehlungen für Wirtschaftlichkeitsuntersuchungen an Straßen (EWS), Aktualisierung der RAS-W 86*. Forschungsgesellschaft für Strassen- und Verkehrswesen, Köln.

Furness, K. P. A. (1965). *Time function iteration*. Traffic Engineering and Control, (7), 458–460.

Gallego, F. (2007). *Downscaling population density in the European Union with a land cover map and a point survey*. European Environment Agency (EEA).

Gaudry, M. (1993). *Asymmetric Shape and Variable Tail Thickness in Multinomial Probabilistic Responses to Significant Transport Service Level Changes*. Transport Demand and Mode Choice: The Impact of Functional Form.

Gaudry, M., Mandel, B., and Rothengatter, W. (1994a). *Entwicklung eines gekoppelten Verkehrserzeugungs- und -verteilungsmodells für den Personenfernverkehr*. Universität Montreal, Kanada. Forschungsauftrag des Bundesministers für Verkehr, FENr.: 60307/92.

Gaudry, M., Mandel, B., and Rothengatter, W. (1994b). *Introducing Spatial Competition through an Autoregressive Contiguous Distributed (AR-C-D) Process in Intercity Generation-Distribution Models within a Quasi-Direct Format (QDF)*. Centre de recherche sur les transports, Université de Montréal.

Geoinformatik GmbH (GI) (2015). *ArcGIS 10.3 : das deutschsprachige Handbuch für "ArcGIS for Desktop Basic & Standard"; mit Funktionen von ArcGIS Online für Desktopanwender*. Berlin: Wichmann.

Goetz, B. (Ed.) (2008). *Java concurrency in practice*. Upper Saddle River, NJ: Addison-Wesley, 6th print. ed.

Gonzalez, R. C., and Woods, R. E. (2008). *Digital image processing*. Upper Saddle River, NJ: Pearson Prentice Hall, 3rd ed.

Guo, J., Trinidad, G., and Smith, N. (2000). *MOZART: A Multi-Objective Zoning and AggRegation Tool*. In *Proceedings of the Philippine Computing Science Congress 2000*.

Haldorson, M., Moström, J., Hedeklint, K., Kaminger, I., Tammisto, R., Loonis, V., Nordbeck, O., Engelien, E., Wardzińska-Sharif, A., and Santos, A. (2017). *A Point-based Foundation for Statistics, Final Report from the GEOSTAT 2 project*. European Forum for Geography and Statistics (EFGS) and Eurostat. Data set.

Hansen, W. (1959). *Accessibility and Residential Growth*. Massachusetts Institute of Technology, Department of City and Regional Planning.

Heining, J., and Möller, S. (2009). *Grenzpendler in Deutschland - Wer sie sind, woher sie kommen, wohin sie gehen*. IAB Kurzbericht.

Hennermann, K., and Woltering, M. (2014). *Kartographie und GIS: eine Einführung*. Darmstadt: Wissenschaftliche Buchgesellschaft (WBG), 2nd ed.

Horni, A., Nagel, K., and Axhausen, K. W. (2016). *The Multi-Agent Transport Simulation MATSim*. London: Ubiquity Press. License: CC-BY 4.0.

Ibañez-Rivas, J. N. (2010). *Peer review of the TRANS-TOOLS reference transport model*. European Commission, Joint Research Centre; Institute for Prospective Technological Studies, Luxembourg: Publications Office of the European Union. Technical note.

IEEE Computer Society (IEEE) (2008). *IEEE Standard for Floating-Point Arithmetic*. IEEE Std 754-2008.

Ihrig, J. (2009). *Automatische Ermittlung der optimalen Beleuchtung für unbekannte Farbaufdrucke auf transparenten Hohlkörpern*. Master's thesis, Fraunhofer Institut für Informations- und Datenverarbeitung (IITB), Universität Karlsruhe (TH).

Ihrig, J. (2012). *Passenger Modelling Methodology*. In *D6 ETISplus Database - Content and Methodology*, chap. 38. Zoetermeer. Report of the ETISplus project funded by the European Commission.

Ihrig, J. (2014). *Methodology of the IPAT model: equations for generation, distribution, modal split and conversion*. Chair of Network Economics, Karlsruhe Institute of Technology (KIT), Karlsruhe. Model documentation, v2.5.

Institut für angewandte Sozialwissenschaft GmbH, Deutsches Zentrum für Luft- und Raumfahrt e.V. – Institut für Verkehrsforschung (infas and DLR) (2010a). *Mobilität in Deutschland 2008*. Im Auftrag des Bundesministerium für Verkehr, Bau und Stadtentwicklung. Data set.

Institut für angewandte Sozialwissenschaft GmbH, Deutsches Zentrum für Luft- und Raumfahrt e.V. – Institut für Verkehrsforschung (infas and DLR) (2010b). *Mobilität in Deutschland 2008, Ergebnisbericht, Struktur – Aufkommen – Emissionen – Trends*. Im Auftrag des Bundesministerium für Verkehr, Bau und Stadtentwicklung.

International Association of Oil & Gas Producers (OGP) (2012). *Using the EPSG Geodetic Parameter Dataset*. OGP Publication 373-7-1 – Geomatics Guidance Note number 7, part 1. Data set.

IWW, SMA, and TU Wien/ sfr (IWW et al.) (2001). *Magistrale für Europa, Schlussbericht*. Karlsruhe. Co-funded by the European Commission under the INTERREG II C program.

Janser, A., Luther, W., and Otten, W. (1996). *Computergraphik und Bildverarbeitung*. Vieweg Lehrbuch : Mathematik/Informatik. Wiesbaden [u.a.]: Vieweg.

Karasmaa, N. (2003). *The transferability of travel demand models - An analysis of transfer methods, data quality and model estimation*. Ph.D. thesis, Laboratory of Transportation Engineering, Helsinki University of Technology, Helsinki.

Karlsruher Institut für Technologie (KIT) (2012). *Ausgezeichnet: Routenplaner der nächsten Generation, IT am KIT: Dorothea Wagner und Peter Sanders erhalten den "Google Focused Research Award"*. Press release 17. URL: www.kit.edu/kit/pi_2012_8851.php (Retrieved October 4, 2017).

Karlsruher Institut für Technologie (KIT) (2016). *Superhirn im Dienst der Spitzenforschung*. Press release 33. URL: www.kit.edu/kit/pi_2016_033_superhirn_im_dienst_der_spitzenforschung.php (Retrieved April 2, 2018).

Karlsruher Institut für Technologie (KIT) (2017). *Deutscher Mobilitätspreis geht an Projekt des KIT*. Press release 85. URL: www.kit.edu/kit/pi_2017_085_deutscher-mobilitatspreis-geht-an-projekt-des-kit.php (Retrieved October 4, 2017).

Kato, H., Kaneko, Y., and Inoue, M. (2003). *Measurement of transport investment benefit: Empirical comparisons between OD-based approach and route-based approach*. Journal of the Eastern Asia Society for Transportation Studies, *5*, 2962–2971.

Kiel, J., Laparidou, K., Smith, R., Li, T., Larrea, E., Szimba, E., Kraft, M., Ihrig, J., Mandel, B., Purwanto, J., Corthout, R., van Grol, R., Székely, A., and Berki, Z. (2016a). *Validating the HIGH-TOOL model: Results of checks and implemented Case Studies, HIGH-TOOL Deliverable D8.2*. Karlsruhe. Project co-funded by the European Commission under the 7th Framework Programme.

Kiel, J., Smith, R., and Laparidou, K. (2016b). *Updated Input Database for the HIGH-TOOL Model - version: 368*. Project co-funded by the European Commission under the 7th Framework Programme, Zoetermeer/Karlsruhe. PostgreSQL Database.

Knudsen, M. A., and J., R. (2013). *Ex post socio-economic assessment of the Oresund Bridge*. Transport Policy, *27*(Supplement C), 53–65.

Kohlstock, P. (2014). *Kartographie : eine Einführung*. UTB; 2568: Geowissenschaften. Paderborn: Schöningh, 3rd ed.

Koppelman, F., and Bhat, C. (2006). *A Self Instructing Course in Mode Choice Modelling: Multinomial and Nested Logit Models*.

Kraft, M. (2017). *Integration von Netz-basierten und sozio-ökonomischen Verkehrsmodellen*. Institut für Volkswirtschaftslehre, Karlsruher Institut für Technologie (KIT), Karlsruhe. Laufende Doktorarbeit, Version 4.1.

Krail, M., and Schade, W. (2014). *ASSIST Final Report – Summary of the Project Approach and Findings*. Fraunhofer-ISI, Karlsruhe, Germany. Deliverable D8.3 of ASSIST (Assessing the social and economic impacts of past and future sustainable transport policy in Europe).

Larrea, E. (2016). *Final Version of the HIGH-TOOL Model: Documentation, HIGH-TOOL Deliverable D5.3*. Barcelona/Karlsruhe. Project co-funded by the European Commission under the 7th Framework Programme.

Larrea, E., and Szimba, E. (2016). *HIGH-TOOL User Guide*. Barcelona/Karlsruhe. Project co-funded by the European Commission under the 7th Framework Programme.

Lenormand, M., Huet, S., and Gargiulo, F. (2012a). *Generating French virtual commuting network at municipality level*. Arxiv preprint.

Lenormand, M., Huet, S., Gargiulo, F., and Deffuant, G. (2012b). *A Universal Model of Commuting Networks*. Arxiv preprint.

Lohninger, H. (2012). *Grundlagen der Statistik*. URL: www.statistics4u.info/fundstat_germ/ (Retrieved October 4, 2017).

Lohse, D., Schiller, C., Teichert, H., Vrtic, M., Fröhlich, P., Schüssler, N., and Axhausen, K. W. (2006). *Ein zweiseitig gekoppeltes Modell zur simultanen Berechnung der Verkehrserzeugung, Verkehrsverteilung und Verkehrsaufteilung*. Verkehrsforschung Online, *3*.

Lohse, D., Teichert, H., Dugge, B., and Bachner, G. (1997). *Ermittlung von Verkehrsströmen mit n-linearen Gleichungssystemen unter Beachtung von Nebenbedingungen einschließlich Parameterschätzung (Verkehrsnachfragemodellierung: Erzeugung, Verteilung, Aufteilung)*. In *Schriftenreihe des Instituts für Verkehrsplanung und Straßenverkehr*, vol. 5/1997. Fakultät Verkehrswissenschaften, "Friedrich List", Technische Universität Dresden.

Maier, G., and Weiss, P. (1990). *Modelle diskreter Entscheidungen : Theorie und Anwendung in den Sozial- und Wirtschaftswissenschaften*. Springer Verlag.

Mandel, B., Gaudry, M., and Rothengatter, W. (1997). *A disaggregate Box-Cox Logit mode choice model of intercity passenger travel in Germany and its implications for high-speed rail demand forecasts*, vol. 31. Springer Verlag, Heidelberg.

Mandel, B., Kraft, M., Schnell, O., Klar, R., Ihrig, J., Szimba, E., Smith, R., Laparidou, K., Chahim, M., Corthout, R., and Purwanto, J. (2016). *Final Structure of the HIGH-TOOL Model, Deliverable D2.2 of the HIGH-TOOL project*. Karlsruhe. Project co-funded by the European Commission under the 7th Framework Programme.

Martis, K. C. (2008). *The original gerrymander*. Political Geography, *27*, 833–839.

Mehlhorn, K., and Sanders, P. (2008). *Algorithms and Data Structures: The Basic Toolbox*. Springer Verlag.

Ministerie van Verkeer en Waterstaat; Rijkswaterstaat (MVW) (2010). *Mobiliteitsonderzoek Nederland 2007*. Dienst Verkeer en Scheepvaart. Data set.

MKmetric (2012). *Air Passenger Matrix*. In *D6 ETISplus Database - Content and Methodology*, chap. 29. Zoetermeer. Report of the ETISplus project funded by the European Commission.

Monzón, A., Ndiaye, A. B., Pfaffenbichler, P. C., and Wegener, M. (2010). *Evaluation of TRANS-TOOLS Version 2*. Report for the European Commission Joint Research Centre (JRC) Institute for Prospective Technological Studies (IPTS) Sevilla.

Nauwelaers, C., Maguire, K., and Ajmone Marsan, G. (2013). *The case of Oresund (Denmark-Sweden) – Regions and Innovation: Collaborating Across Borders*. OECD Publishing. OECD Regional Development Working Papers.

NEA, IWW, COWI, NESTEAR, PWC Italy, TINA, IVT, HERRY, and MKmetric (NEA et al.) (2004). *Final Report of the TEN-STAC project*. Funded by the European Commission (DG TREN), Rijswijk.

NEA, NESTEAR, ISIS, Universität Karlsruhe (TH), MDS Transmodal, TNS, MKmetric, VTT, and ETH Zürich (NEA et al.) (2005). *Final technical report v1, ETIS-BASE Deliverable D9*. Core Database Development for the European Transport Policy Information System (ETIS). Project funded by the European Commission.

NEA, OSC, IWW, and MKmetric (NEA et al.) (2009). *Intermodal Impedances and Links to TRANSTOOLS*. Zoetermeer. Deliverable D9 of the WORLDNET project funded by the European Commission under the 6th Framework Programme.

Newton, S., Enei, R., de Stasio, C., Szimba, E., Laugesen, M. S., Carvalho, D., and Chakarova, K. (2013). *Final Report, ETISplus Deliverable D8*. Zoetermeer. Project funded by the European Commission.

Newton, S. E. (2010). *D1 ETISplus Inception Report*. Zoetermeer. Report of the ETISplus project funded by the European Commission.

Nielsen, O. A. (2009). *Trans-Tools overview*. Technical University of Denmark (DTU), Department of Transport. Lecture notes, TRANS-TOOLS training course April 27-29, 2009.

Obe, R. O., and Hsu, L. S. (2015). *PostGIS in action*. Shelter Island, NY: Manning, 2nd ed.

Openshaw, S. (1984). *The Modifiable Areal Unit Problem*. CATMOG, (2). Geo Books, Norwich.

Openshaw, S., and Taylor, P. J. (1979). *A million or so correlation coefficients: three experiments on the modifiable areal unit problem*. Statistical Applications in the Spatial Sciences, (pp. 127–144). Pion, London.

Ortúzar, J. d. D., and Willumsen, L. G. (2001). *Modelling Transport*. Chichester: Wiley, 3rd ed.

OSCE/ Office for Demographic Institutions and Human Rights (OSCE) (2014). *Hungary, Parliamentary elections 6 April 2014*. OSCE/ODIHR Limited Election Observation Mission, Final Report, Warsaw.

Ottmann, T., and Widmayer, P. (2002). *Algorithmen und Datenstrukturen*. Spektrum Akademischer Verlag, 4th ed.

Papadimitriou, G., Ntziachristos, L., Wüthrich, P., Notter, B., Keller, M., Fridell, E., Winnes, H., Styhre, L., and Sjödin, R. (2013). *Transport data collection supporting the quantitative analysis of measures relating to transport and climate change, Final Report of the TRACCS project*. European Commission, Directorate-General for Climate Action (DG CLIMA). Data set.

Petersen, M. S., Bröcker, J., Enei, R., Gohkale, R., Granberg, T., Hansen, C. O., Hansen, H. K., Jovanovic, R., Korchenevych, A., Larrea, E., Leder, P., Merten, T., Pearman, A., Rich, J., Shires, J., and Ulied, A. (2009). *Report on Scenario, Traffic Forecast and Analysis of Traffic on the TEN-T, taking into Consideration the External Dimension of the Union – Final Report*. Copenhagen, Denmark. Funded by DG TREN.

PTV GROUP (2015). *PTV VISUM - MODULE*. PTV GROUP, Karlsruhe.

Purwanto, J., De Ceuster, G., and Vanherle, K. (2017). *Mobility, Vehicle fleet, Energy use and Emissions forecast Tool (MOVEET)*. In *Transportation Research Procedia*, vol. 25, (pp. 3425–3438).

Pérez Lou, J. I., Ropero Ortega, J. M., Cuesta Rilo, L., Torrijos Prieto, J. A., Sánchez Núñez, J. A., Gómez Merchán, J., Domínguez Azuara, C., Yuste Martín, P., and Povedano Povedano, P. (2007). *Mobility survey of persons resident in Spain (Movilia 2006/2007)*. Ministry of Development - Government of Spain. Data set.

Ramm, F., and Topf, J. (2008). *OpenStreetMap : die freie Weltkarte nutzen und mitgestalten*. Berlin: Lehmanns Media.

Samet, H. (1980). *Region Representation: Quadtrees from Boundary Codes*. Commun. ACM, *23*(3), 163–170.

Samuelson, P. A., and Nordhaus, W. D. (2007). *Volkswirtschaftslehre : das internationale Standardwerk der Makro- und Mikroökonomie*. Landsberg am Lech: mi-Fachverlag, 3rd ed.

Schade, W. (2005). *Strategic Sustainability Analysis: Concept and application for the assessment of European Transport Policy*. Ph.D. thesis, Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW), Universität Karlsruhe (TH), Baden-Baden: Nomos Verlagsgesellschaft.

Schick, N., and Wieczorek, T. (2008). *Verkehrsuntersuchung Zweite Rheinbrücke mit Nordtangente*. PTV AG, Karlsruhe. Studie im Auftrag der Stadt Karlsruhe, Stadtplanungsamt.

Schimke, A., Stoller, A., Siegele, J., and Ihrig, J. (2012). *Socio-economic data, base year 2010*. In *D6 ETISplus Database - Content and Methodology*, chap. 4. Karlsruhe. Report of the ETISplus project funded by the European Commission.

Schmedding, D. (2006). *Erfassung und Bewertung von Strassenverkehrslärm auf der Basis von geographischen Informationssystemen*. Ph.D. thesis, Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW), Universität Karlsruhe (TH), Baden-Baden: Nomos Verlagsgesellschaft.

Schneider, J. (2008). *Geschäftsreisende 2008 - Strukturen, Einstellungen, Verhalten*. Bonn: Internationale Fachhochschule Bad Honnef.

Schoch, M. (2004). *Verwendung feinräumiger geographischer Informationen in aggregierten Verkehrsprognosen*. Ph.D. thesis, Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW), Universität Karlsruhe (TH), Baden-Baden: Nomos Verlagsgesellschaft.

Schulz, C. (2006). *Modellierung des Personenverkehrs in der Türkei*. Master's thesis, Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW), Universität Karlsruhe (TH), Karlsruhe.

Schulz, C. (2012). *The Identification of Critical Road Infrastructures - The Case of Baden-Wuerttemberg*. KIT Scientific Publ.

Schulze, C., and Waßmuth, V. (2009). *Verkehrsuntersuchung Karlsruhe-Nord, Untersuchung von Planfällen*. PTV AG, Karlsruhe. Studie im Auftrag der Stadt Karlsruhe, Stadtplanungsamt.

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.

Siegele, J., Schimke, A., and Ihrig, J. (2012). *Annex report A: Socio-economic data, base year 2005*. In *D6 ETISplus Database - Content and Methodology*. Karlsruhe. Report of the ETISplus project funded by the European Commission.

Spiekermann, K., and Wegener, M. (1993). *Zeitkarten für die Raumplanung*. Institut für Raumplanung, Fakultät Raumplanung, Universität Dortmund. Working paper.

Statistisches Bundesamt (Destatis) (2007). *Statistik lokal - Daten für die Gemeinden, kreisfreie Städte und Kreise Deutschlands*. Status 31.12.2005, CD-ROM. Statistische Landesämter des Bundes und der Länder.

Stelzl, I. (1982). *Fehler und Fallen der Statistik : für Psychologen, Pädagogen, Sozialwissenschaftler*. Huber.

Stewart, G. (1973). *Introduction to Matrix Computations*. Computer Science and Applied Mathematics. Elsevier Science & Technology.

Szimba, E. (2008). *Interdependence between transport infrastructure projects : an analytical framework applied to priority transport infrastructure projects of the European Union*. Ph.D. thesis, Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW), Universität Karlsruhe (TH), Baden-Baden: Nomos Verlagsgesellschaft.

Szimba, E., and Ihrig, J. (2016). *High-Level Strategic Transport Model (HIGH-TOOL)*. Poster session at the TRA Marketplace.

Szimba, E., Ihrig, J., Kraft, M., Mitusch, K., Chen, M., Chahim, M., van Meijeren, J., Kiel, J., Mandel, B., Ulied, A., Larrea, E., De Ceuster, G., Van Grol, R., Berki, Z., Székely, A., and Smith, R. (2018). *HIGH-TOOL – a strategic assessment tool for evaluating EU transport policies*. Journal of Shipping and Trade, *3*(Supplement C), 53–65.

Szimba, E., Ihrig, J., Kraft, M., Ulied, A., Larrea, E., Biosca, O., Martínez, C. L., Chen, M., van Meijeren, J., Chahim, M., Mandel, B., Kiel, J., Smith, R., Laparidou, K., Purwanto, J., Corthout, R., van Grol, R., van Eck, G., Berki, Z., and Székely, A. (2016). *Final Report, HIGH-TOOL Deliverable D10.5*. Karlsruhe. Project co-funded by the European Commission under the 7th Framework Programme.

Szimba, E., and Kraft, M. (2011). *The Strategic Passenger Model VACLAV: Methodology and Application*. Institute for Economic Policy Research (IWW), Karlsruhe Institute of Technology (KIT), Karlsruhe. Model documentation, v3.1.

Szimba, E., and Kraft, M. (2012a). *Rail network, base year 2010*. In *D6 ETISplus Database - Content and Methodology*, chap. 6. Karlsruhe. Report of the ETISplus project funded by the European Commission.

Szimba, E., and Kraft, M. (2012b). *Road network, base year 2010*. In *D6 ETISplus Database - Content and Methodology*, chap. 5. Karlsruhe. Report of the ETISplus project funded by the European Commission.

Szimba, E., and Kraft, M. (2012c). *Road passenger impedances, base year 2010*. In *D6 ETISplus Database - Content and Methodology*, chap. 15. Karlsruhe. Report of the ETISplus project funded by the European Commission.

Szimba, E., Kraft, M., and Adams, M. (2012). *Rail passenger impedances, base year 2010*. In *D6 ETISplus Database - Content and Methodology*, chap. 16. Karlsruhe. Report of the ETISplus project funded by the European Commission.

Szimba, E., Mandel, B., Kraft, M., and Ihrig, J. (2017). *A Decision Support Tool for the Strategic Assessment of Transport Policies – Structure of the Tool and Key Features*. Transportation Research Procedia, *25*(Supplement C), 2843 – 2860. World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016.

Szimba, E., Zeitler, M., Mandel, B., and Ulied, A. (2015). *Development of a high-level strategic assessment model for the evaluation of transport policies by the European Commission – the HIGH-TOOL project*. HIGH-TOOL session at the European Transport Conference (ETC), Frankfurt, Germany.

TINA Transport Strategies and Dorsch Consult and IWW and NEA and NESTEAR and Su-Yapi and Geomatic (TINA et al.) (2007). *Technical Assistance to Transport Infrastructure Needs Assessment for Turkey, Final Report*. Project on behalf of the European Commission and Central Finance and Contracts Unit (CFCU), Ankara.

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

TRT (2012). *Tram and Metro Passenger Matrix & Cycling and Walking Matrix*. In *D6 ETISplus Database - Content and Methodology*, chap. 33-34. Zoetermeer. Report of the ETISplus project funded by the European Commission.

Ullenboom, C. (2014). *Java ist auch eine Insel - Programmieren mit der Java Platform, Standard Edition 8*. Galileo Computing. Bonn: Galileo Press, 11th ed.

Van Grol, R., De Bok, M., De Jong, G., Van Eck, G., Ihrig, J., Kraft, M., Szimba, E., Mandel, B., Ivanova, O., Boonman, H., Chahim, M., Corthout, R., Purwanto, J., Smith, R., Laparidou, K., Helder, E., Grebe, S., and Székely, A. (2016). *Elasticities and Equations of the HIGH-TOOL Model (Final Version), HIGH-TOOL Deliverable D4.3*. Karlsruhe. Project co-funded by the European Commission under the 7th Framework Programme.

van Meijeren, J., Davydenko, I., Chahim, M., Szimba, E., Kraft, M., Ihrig, J., Smith, R., Laparidou, T., Purwanto, J., and Corthout, R. (2016). *Validation by Coherence Checks, HIGH-TOOL Deliverable D8.1*. Karlsruhe. Project co-funded by the European Commission under the 7th Framework Programme.

Vickerman, R. (1996). *High-speed rail in Europe: experience and issues for future development*. The Annals of Regional Science, (31), 21–38.

Walter, P., Bläuer Herrmann, A., Cangemi, V., Murier, T., Perrenoud, S., Reutter, R., Saucy, F., Schmassmann, S., and Zimmermann, M. (2016). *Arbeitsmarktindikatoren 2016*. Neuchâtel: Bundesamt für Statistik (BFS). Statistik der Schweiz.

Wang, F. (2001). *Explaining intraurban variations of commuting by job proximity and workers' characteristics*. Environment and Planning B: Planning and Design, *28*(2), 169–182.

Washington, S. P., Karlaftis, M. G., and Mannering, F. L. (2004). *Statistical and econometric methods for transportation data analysis*. Boca Raton, Florida: CRC Press LCC.

Weiß, C., Chlond, B., von Behren, S., Hilgert, T., and Vortisch, P. (2016). *Deutsches Mobilitätspanel (MOP) – Wissenschaftliche Begleitung und Auswertungen Bericht 2015/2016: Alltagsmobilität und Fahrleistung*. Im Auftrag des Bundesministerium für Verkehr und digitale Infrastruktur Referat G34 "Prognosen, Statistik und Sondererhebungen". Data set.

Wildermuth, B. R., Delaney, D. J., and Thompson, K. E. (1972). *Effect of Zone Size on Traffic Assignment and Trip Distribution*. Highway Research Record, *392*, 58–75.

Zumkeller, D. (2002). *Verkehrsplanung - Modellierung von Verkehrsverhalten, Bearbeitungsstand 8/2002*. Institut für Verkehrswesen (IfV), Universität Karlsruhe (TH), Karlsruhe. Materialsammlung zur Vorlesung.

Zumkeller, D., Chlond, B., Ottmann, P., Kagerbauer, M., and Kuhnimhof, T. (2007). *Panelauswertung 2006 – Datenaufbereitung, Plausibilisierung, erste Auswertungen zu den Erhebungen zur Alltagsmobilität 2004/06 sowie zu Fahrleistungen und Treibstoffverbräuchen 2005/07 für das Mobilitätspanel*. Im Auftrag des Bundesministerium für Verkehr, Bau und Stadtentwicklung. Data set.