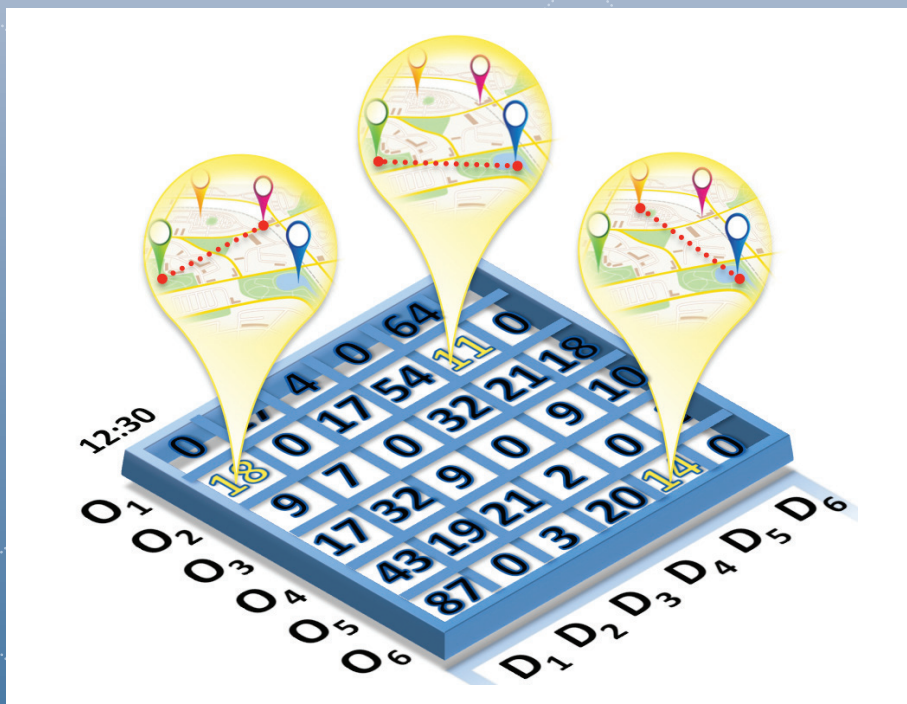




TRAIL THESIS SERIES

Tamara Djukic



Dynamic OD Demand Estimation and Prediction for Dynamic Traffic Management

Dynamic OD Demand Estimation and Prediction for Dynamic Traffic Management

Tamara Djukic

Delft University of Technology

This thesis is the result of a project funded by ITS Edulab (a cooperation between Rijkswaterstaat and Delft University of Technology) and the Netherlands Research School for Transport, Infrastructure and Logistics TRAIL.



Cover illustration: Tamara Djukic

Dynamic OD Demand Estimation and Prediction for Dynamic Traffic Management

Proefschrift
ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op dinsdag 18 November 2014 om 12.30 uur
door

Tamara DJUKIC

Master of Science in Civil Engineering
University of Belgrade, Serbia
geboren te Cacak, Servië

Dit proefschrift is goedgekeurd door de promotoren:

Prof. dr. ir. S.P. Hoogendoorn

Prof. dr. ir. J.W.C van Lint

Samenstelling promotiecommissie :

Rector Magnificus

voorzitter

Prof. dr. ir. S.P. Hoogendoorn

Technische Universiteit Delft, promotor

Prof. dr. ir. J.W.C van Lint

Technische Universiteit Delft, promotor

Prof. dr. M.E. Ben-Akiva

Massachusetts Institute of Technology

Prof. dr. M. Bierlaire

Swiss Federal Institute of Technology in Lausanne

Prof. dr. ir. E. de Romph

Technische Universiteit Delft

Prof. dr. ir. C. Witteveen

Technische Universiteit Delft

Prof. dr. S. Vukanovic

University of Belgrade

TRAIL Thesis Series no. T2014/9, the Netherlands Research School TRAIL

TRAIL

P.O. Box 5017

2600 GA Delft

The Netherlands

Phone: +31 (0) 15 278 6046

E-mail: info@rsTRAIL.nl

ISBN: 978-90-5584-179-0

Copyright © 2014 by Tamara DJUKIC

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the author.

Printed in the Netherlands

"There is nothing either good or bad, but thinking makes it so."
- William Shakespeare, Hamlet

Acknowledgments

When somewhere in the near or distant future I decide to revisit the pages of this thesis, I will read this page to relive, even for a moment, a time spent in research with all of you. I would like to take this opportunity to thank you for your contribution and to record my dearest memories.

I would like to express my special appreciation and thanks to my promoters Serge and Hans, you have been a tremendous mentors for me. As I am finishing this thesis I am more aware than ever of how working with you has changed me. Your maddening attention to plan (everything) drove me to finally learn how to deal with planning process. Both of you have contributed a lot in my research by making valuable suggestions, asking right questions and challenging my ideas. I am specially grateful for the times when your enthusiasm and believe in my research were all that kept me going.

I extend special thanks to Smiljan Vukanovic and Svetlana Vukanovic without whom I would never have considered the possibility to pursue a PhD degree. I will be always grateful to you for enlightening me with the first glance of research and for giving me a great deal of freedom to explore the career path that I want.

I am especially grateful to members of MULTITUDE project that have been fully supporting the part of work described in this thesis. In particular, special thanks to Gunnar Flötteröd that has hosted my visit at KTH, Sweden, and whose valuable comments contributed in joint publication presented in Chapter 5 of this thesis. Also, I would like to thank Vincenzo, Costas, Jaume, Lidia, Christine, Tomer, Ernesto, Marialisa, Vittorio, Jordi, Mark and Biagio for being a great company through hours spent together in the meetings and for many fruitful discussions about dynamic OD matrices that kept me on the good course.

My sincere thanks go to all my committee members Prof. Moshe Ben-Akiva, Prof. Michel Bierlaire, Prof. Cees Wittewein, Prof. Erik de Romph and Prof. Smiljan Vukanovic for their time in evaluating thesis manuscript. Their critical comments and constructive suggestions have led to significant improvement of my thesis.

The work presented here wouldn't have been possible without support of my research group members. Thank you for sharing your exceptional knowledge, for supporting me during my research and for having a great outings after work. I really enjoyed working with you, showing me diverse cultures from all over the world. Also, I want to thank my roommates Goof, Erica, Giselle, Mario, Jeroen and Gerdien for contributing to pages in Dutch of this thesis and for making a coffee break enjoyable moment fulfilled with many interesting discussions. Special thanks to Olga and Giselle for our adventures

in New York and so many great moments we have spend together last five years. I am happy that two of you accepted to be my paranympths and support me during the defense. Finally, I am very grateful to have shared my days in department with Pavle and Nikola, a great friends, patient listeners, always ready to make a short break from work and have inspiring discussions about survival in the Netherlands.

Surviving in this male-dominated world of traffic research wouldn't be possible without all my dear girlfriends on the outside: Natalija, my fashion and style guru; Ksenija with whom I can always forget everything about work and just relax and enjoy; Kica whose positive energy, open heart and amazing hospitality keeps my visits in Belgrade and our trips unforgettable. Special thanks to Gaga and Darko, for being true friends and for our precious times together. No matter where you are, all the cheerful time that we spend together would never fade away.

To my dear Agata, my housemate and soul-mate, that knows how I fill before I say it. Thank you for making my life so enjoyable in Delft, for supporting me in the moments when there was no one to answer my queries, being good listener and simply making me happy. Luigi thank you for your patience and continued support of Agata and me, for great dinners and accommodation.

*In the end, a special thanks to my mom and dad, Milunka and Slobodan, that have always given me freedom to find my own way, but making sure there is a loving home to come back to. Words cannot express how grateful I am to my brother Nikola, his wife Aleksandra, and my sweet little niece Ema, the cutest kid on the planet. No matter how serious the things could get, we have always had our way to laugh and make it right all together. I miss you more than anything...,
...and few more words,...*

A boy, eyes sparkling with great dreams, gathered a few twigs and leaves. He tied them in his familiar way and placed the creation between two stones. The stream's cascade propelled the leaves, turning the twigs. The small Serbian town of Smiljan was witness to the first water turbine. Who would have thought? A few decades later the boy made his dream come true. In a distant land, at Niagara Falls, he designed the world's first hydroelectric power plant.

The boy grew up into a man, the man became a legend - Nikola Tesla.

Sometimes beginnings seem ordinary, inept, even funny. However, it is these first steps, these twigs and leaves, that lead us to tremendous advances in science, life, and our understanding of consciousness, the universe and God. I hope that my beginning carries the spark of something big. I hope that this work contributes new insights and research directions, something great and significant, to make everybody who helped and supported me proud.

Tamara Djukic, November 2014

Summary

A desirable feature of dynamic traffic management systems (DTMS) is the ability to estimate network states and predict their short-term evolution. Unreliability and lack of knowledge about past and prevailing traffic conditions may well lead to poor predictions that would, for example, render computed Intelligent Transportation System (ITS) measures irrelevant or outdated by the time they take effect. One of the key input traffic variables required by this system is traffic demand represented by *dynamic origin destination (OD) matrices*. Each cell of this matrix represents the number of vehicles departing from origin zone during one time interval to destination zone.

The inability to provide high quality dynamic OD demand estimates makes prediction with advanced simulation models simply impossible, irrespective of how well these models have been calibrated. In this respect, the estimation and prediction of a dynamic OD demand with sufficient data and sufficient granularity are critical in establishing the credibility of simulation tools for real-time purposes. Driven by the aforementioned issues and requirements for improvement, this thesis addresses several problems pertaining to the provision of efficient and reliable dynamic OD demand information for DTMS applications.

The literature review is largely based on a newly developed categorization of the dynamic OD demand estimation and prediction methods. A rich variety of methods developed so far and in use today are classified based on the *modeling-steps* with which the OD demand estimation and prediction is described; the types of input data, the way in which their relationship with OD flows is modeled, and the solution approaches for the estimation and prediction of dynamic OD demand. This approach shows better how various challenges within each modeling-step have been solved and how different methods relate to each other.

Dynamic OD demand estimation methods differ in many aspects, such as mapping methodology of traffic data and OD flows, measures of error, solution approaches and type of networks; adding to the difficulty of creating generic assessment of OD estimation methods. In this thesis, a benchmarking methodology for the qualitative assessment of dynamic OD demand estimation methods is developed. The methodology presented here is generic, in the sense that various OD estimation approaches can be tested under numerous diverse circumstances related to, for example, data availability and quality, and network lay out. The objective of the benchmark methodology is

not to conclude that one approach is the "best", but to provide support for comparison in a variety of settings and conditions. With this benchmark methodology one can, for example, determine the particular situations and conditions under which one approach might behave more favorably than another. One can also use the methodology to perform sensitivity analyses on single or multiple dynamic OD estimation methods.

The dynamic OD demand estimation and prediction problem is computationally intensive because solution methods have to deal with high-dimensional structures of OD matrices and computational complexity of these methods. One possible solution approach to solve the issue of high-dimensionality is to approximate prior OD demand dataset into the lower-dimensional space without significant loss of accuracy. The new reduced set of variables is defined instead of the OD flows. As a result, the dimensionality of the state is reduced substantially and the complexity of the estimation and prediction problem is likewise reduced. For real-time application, at the end of each interval, the observed traffic counts would be used to sequentially update OD flows for the current time interval. The problem is formulated as state-space model solved by colored Kalman filter. In a case study, the proposed method demonstrates that by significantly reducing the dimensionality of the OD data while, preserving the structural patterns, the computational costs can be dramatically reduced.

The spatial correlation between OD pairs carries important information about the structures in OD matrices. An emerging alternative performance indicator, structural similarity (SSIM) index, has been presented to quantify these correlations. For example, under the assumption that any prior OD matrix contains the best possible pattern information, the SSIM index can be viewed as an indication of the quality of the estimated OD matrix compared to the prior OD matrix. This is important in applications where it is necessary to know whether a particular OD demand estimation method can reproduce actual OD demand. This quality metric has been shown to have several advantages over existing statistical measures that measure pointwise deviations between two OD matrices. Therefore, the proposed measure can be applied as additional performance indicator for benchmarking tasks of dynamic OD estimation methods.

This thesis gives new insights in real-time dynamic OD demand estimation and prediction for large-scale networks and provides efficient methodologies to assess the performance of existing methods. The presented methods are ready to use in practice and can be compared with existing methods.

Samenvatting

Een aantrekkelijk kenmerk van dynamische verkeersmanagementsystemen (DVMS) is het vermogen de toestand in een netwerk te schatten en hun ontwikkeling op de korte termijn te voorspellen. Onbetrouwbaarheid en gebrek aan kennis over historische en huidige verkeersomstandigheden kan mogelijk leiden tot slechte voorspellingen waardoor, bijvoorbeeld, gecomputeriseerde Intelligente Transport Systemen (ITS) maatregelen niet meer relevant of gedateerd zijn op het moment dat ze worden ingezet. Een van de cruciale input verkeersvariabelen die dit systeem nodig heeft, is dat de verkeersvraag weergegeven wordt in *dynamische herkomst-bestemming (HB) matrixen*. Elke cel van deze matrix geeft het aantal voertuigen weer die tijdens n tijdsinterval uit de herkomstzone vertrekken naar de bestemmingszone.

Het niet in staat zijn om hoogwaardige schattingen van de dynamische HB-vraag te leveren zorgt ervoor dat voorspellingen met geavanceerde simulatiemodellen eenvoudigweg onmogelijk zijn, ongeacht hoe goed deze modellen ook gekalibreerd zijn. In dit opzicht zijn het schatten en voorspellen van een dynamische HB-vraag met voldoende data en voldoende granulariteit cruciaal in het vaststellen van de geloofwaardigheid van simulatie-instrumenten voor directe doeleinden. Gedreven door de hiervoor genoemde kwesties en behoeften voor verbetering, stelt dit proefschrift diverse problemen aan de orde die betrekking hebben op het leveren van efficiënte en betrouwbare informatie over de dynamische HB-vraag voor DVMS-applicaties.

De literatuurstudie is grotendeels gebaseerd op een nieuw ontwikkelde categorisering van de schattings- en voorspelmethodes voor de dynamische HB-vraag. De rijke variëteit aan methoden die tot nu toe zijn ontwikkeld en toegepast, zijn geclassificeerd op basis van de modelleer stappen waarmee de schatting en/of voorspelling van de HB-vraag wordt beschreven. Voorbeelden zijn het type data-invoer, de manier waarop hun relatie met HB-verkeersstromen is gemodelleerd, en de oplossingsstrategieën voor het schatten en voorspellen van dynamische HB-vraag. Deze aanpak laat beter zien op welke wijze de uitdagingen binnen iedere modelleer stap zijn opgelost en hoe de verschillende methoden zich tot elkaar verhouden.

Dynamische HB-vraag schattings- en voorspelmethoden verschillen op veel eigenschappen, zoals de indelingsmethodiek van verkeersdata en de herkomsten en bestemmingen van verkeersstromen, de meting van de fout, de oplossingsaanpakken en de type verkeersnetwerken; Deze diversiteit draagt bij aan de moeilijkheid van het creëren

van een algemene beoordelingsmethodiek voor HB-schattingmethoden. In dit proefschrift wordt een methodiek voor de kwalitatieve beoordeling van dynamische HB-vraag schattingsmethode ontwikkeld. De methodologie die wordt besproken is generiek in de zin dat verschillende HB-schattingmethoden getest kunnen worden onder vele verschillende omstandigheden. Deze zijn bijvoorbeeld gerelateerd aan de beschikbaarheid van data, de kwaliteit van de data en de verkeersnetwerk lay-out. Het doel van de standaard beoordelingsmethodiek is niet om te beoordelen wat de beste aanpak is, maar om de verschillende aanpakken voor verschillende instellingen en omstandigheden onderling te vergelijken. Met behulp van deze standaard beoordelingsmethodiek kan bijvoorbeeld worden bepaald onder welke omstandigheden de ene aanpak de voorkeur heeft boven de andere. De beoordelingsmethodiek kan ook worden gebruikt om gevoeligheidsanalyses met n or meerdere dynamische HB-schattingmethoden uit te voeren.

Het dynamische HB-vraag schatting- en voorspellingsprobleem is rekenkundig intensief omdat oplossingsmethoden om moeten gaan met de hoogdimensionele structuren van HB-matrices en de rekenkundige complexiteit van deze methodes. Een mogelijke aanpak om het probleem van de hoge dimensionaliteit op te lossen is om het voorgaande HB-vraagdataset te benaderen in de lager-dimensionele ruimte zonder significant verlies van de nauwkeurigheid. Het nieuwe gereduceerde set van variabelen wordt gedefinieerd in plaats van de HB-stromen. Dit resulteert in een substantiele afname van de dimensionaliteit van de toestand en daarmee een afname van de complexiteit van het schatting- en voorspellingsprobleem. Voor realtime toepassing worden, aan het eind van elk interval, de geobserveerde verkeersstellingen gebruikt voor de opeenvolgende update van de HB-stromen van het huidige tijdsinterval. Het probleem is geformuleerd als state-space model welke wordt opgelost met het gekleurde Kalman filter.

The ruimtelijke correlatie tussen HB-paren bevat belangrijke informatie over de structuren van de HB-matrices. Een alternatieve prestatie-indicator, de structurele gelijkenis (SSIM) index, is gepresenteerd om deze correlaties te kwantificeren. De SSIM index kan bijvoorbeeld, onder de aanname dat een voorgaande HB-matrix de best mogelijke patrooninformatie bevat, worden gezien als een indicatie van de kwaliteit van de geschatte HB-matrix vergeleken met de voorgaande matrix. Dit is belangrijk voor toepassingen waarbij het nodig is om te weten of een specifieke HB -vraagschattingmethode de werkelijke HB-vraag kan reproduceren. Deze kwaliteitsmeettenheid heeft laten zien dat deze verschillende voordelen heeft boven bestaande statistische maten welke puntsgewijs afwijkingen meten tussen twee HB-matrices. Daarom kan de voorgestelde maat worden toegepast als additionele prestatie-indicator voor de benchmarking taken van dynamische HB-schattingmethoden.

Dit proefschrift geeft nieuwe inzichten in realtime dynamische HB-vraag schatting en voorspelling voor grootschalige netwerken en levert efficiënte methodologieën om de prestatie van bestaande methoden te schatten. De gepresenteerde methodes zijn gereed om in de praktijk te gebruiken en kunnen worden vergeleken met bestaande methodes.

Contents

Preface	i
Summary	iii
Samenvatting	v
List of Figures	xiii
List of Tables	xv
1 Introduction	1
1.1 Research motivation	2
1.2 Research objectives and scope	4
1.2.1 Research objectives	4
1.2.2 Research scope	5
1.3 Thesis Contributions	6
1.3.1 Scientific Contributions	7
1.3.2 Practical Contributions	8
1.4 Thesis Outline	9
2 State-of-the-art dynamic OD demand estimation and prediction	11
2.1 Generic formulation of the dynamic OD estimation and prediction problem	12
2.1.1 Generic formulation of the dynamic OD estimation problem .	12
2.1.2 Generic formulation of the dynamic OD prediction problem .	15
2.2 Categorization of dynamic OD demand estimation methods	16

2.3	Dynamic OD demand estimation: State-of-the-art	17
2.3.1	Types of input data used in dynamic OD demand estimation and prediction	17
	OD flow data	19
	Route flow data	22
	Traffic condition data	22
	Discussion	23
2.3.2	Mapping of OD flows to input data	23
	Analytic DTA models	25
	Simulation DTA models	26
	Discussion	27
2.3.3	Objective functions and solution frameworks	28
	Generalized least squares (GLS) formulation	28
	State space formulation	33
	Maximum entropy / minimum information formulation	37
	Discussion	38
2.4	Dynamic OD demand prediction: State-of-the-art	38
	Discussion	40
2.5	Conclusions	40
3	Methodology for benchmarking dynamic OD demand estimation methods	43
3.1	Introduction	44
3.2	The concept of benchmarking framework based on LHC	46
3.2.1	The Latin hypercube sampling	47
3.3	Components in the benchmarking framework	48
3.3.1	Computational platform design and implementation	49
3.3.2	OD demand and traffic data generator	51
	Prior OD demand simulation scenarios	51
	Traffic data simulation scenarios	52
3.3.3	Output data processor	54
3.4	Case study	55

3.4.1	Network topology and method selection	55
3.4.2	Considered scenarios and performance indicators	56
3.4.3	Results and discussion	56
	Total demand analysis	56
	OD pair analysis	59
	Computational efficiency	62
3.5	Conclusions	62
4	Dimensionality reduction methods in OD demand estimation	65
4.1	Introduction	66
4.2	OD demand representation in low-dimensional space	68
4.3	Exploring the temporal variability of dynamic OD matrices	70
4.3.1	Setting up OD matrices database	70
4.3.2	Reducing the dimensionality of dynamic OD matrices	71
	Criteria for selecting a limited number of principal components	71
	Projecting the data onto a limited set of eigenvectors	72
4.3.3	Finding temporal patterns in time series of OD matrices	72
	Contribution of principal demand components to OD pairs	73
	Temporal variability patterns in time series of OD matrices	74
	Decomposition of OD flows	76
4.4	Conclusions	78
5	Methodology for real time OD demand estimation and prediction	79
5.1	Overview of methodology	80
5.2	A Reduced state space OD estimation and prediction model formulation	82
5.2.1	State Augmentation	84
5.3	Solution approach	85
5.3.1	Temporal correlation between observations introduced by dimensionality reduction	85
5.3.2	Colored noise Kalman filter solution algorithm	86
5.4	Numerical experiments	89

5.4.1	Network topology	89
5.4.2	Simulating historical daily OD demand	90
5.4.3	State vector reduction	91
5.4.4	Method performance	92
5.5	Conclusions	94
6	Measures of performance in OD demand estimation	95
6.1	The concept of structural similarity	96
6.2	Theoretical background of the SSIM	99
6.3	The sensitivity of the MSE error and SSIM index	102
6.4	The sensitivity of the SSIM index on OD pairs ordering	103
6.5	Towards a new research directions	106
6.5.1	SSIM index as a performance indicator	107
6.5.2	SSIM index as the error measure	107
6.6	Discussion	109
7	Conclusions and future work directions	111
7.1	Conclusions	112
7.1.1	Contributions achieved for dynamic OD demand modeling . .	112
7.1.2	Implications for dynamic traffic management	113
7.2	Future research directions	114
	Appendices	117
A	The Latin hypercube sampling method	119
	Notation	121
	Bibliography	123
	About the author	133
	TRAIL Thesis Series	137

List of Figures

1.1	Outline of the content of this thesis	10
2.1	Network description: origins, destinations and intermediate nodes, connected by direct links	12
2.2	The overview of dynamic OD matrix estimation (and prediction) framework	15
2.3	Types of input data used in dynamic OD demand estimation and prediction	19
2.4	Examples of some traffic data and their sources	20
2.5	Illustration of full and partial information on OD flows from OD flow data sources: a) full information on OD flows and b) examples of partial information on OD flows with respect to full information on particular OD pairs (dark blue color) or their sample (light blue color) . .	21
2.6	Dynamic traffic assignment framework	25
2.7	Mapping in congested traffic conditions	27
3.1	Example of LHC method for two variables with a normal distribution (Budiman & Alex (2006)).	48
3.2	Framework for comparison of dynamic OD estimation methods. . . .	49
3.3	Flowchart with the main elements of the computational platform (Antonioni et al. (2014))	50
3.4	Urban corridor network, Kruithuisweg street, the Netherlands.	56
3.5	Average total demand estimates for different scenarios	57
3.6	Average total demand for two origins with all links covered with detectors	58
3.7	Average total demand for two origins with 60% links covered with detectors	59
3.8	Total demand with all and 20% links covered with detectors	59

3.9	Total origin demand with all and 20% links covered with detectors . . .	60
3.10	Simulated OD flow distribution for several OD pairs	60
3.11	a) Mean error and b) Absolute mean error per path length	61
3.12	a) Mean error and b) Absolute mean error per OD demand volume . .	62
4.1	Freeway network, A12, Gouda Utrecht, the Netherlands.	70
4.2	Scree plot of eigenvalues from centered OD matrix X	71
4.3	Cumulative percentage of total variation explained by eigenvalues. . .	72
4.4	Original and approximated OD demand of one OD pair with five eigen- vectors.	73
4.5	CDF of number of significant principal demand components that con- stitute each OD pair.	74
4.6	Correlation of OD flows volume and their principal demand components.	75
4.8	Decomposition of OD pairs into the temporal trends.	76
4.7	Three different temporal trends captured by temporal vectors on prin- cipal demand components.	77
5.1	Overview of common OD estimation (and prediction) methodology .	80
5.2	Overview of proposed OD estimation (and prediction) methodology .	81
5.3	The Vitoria network, Basque Country, Spain	90
5.4	The set of decisions and alternatives for traveler	91
5.5	Cumulative percentage of total variation explained by eigenvalues for time interval k_2	92
6.1	Illustration of spatial interaction between OD pairs	96
6.2	Comparison of patterns in real and estimated OD matrices: a) real OD matrix; b) and c): estimated OD matrices that have the same MSE with respect to the real OD matrix, but different structural patterns.	97
6.3	Computation of local SSIM index per sliding $N \times N$ square box	99
6.4	Computation of local SSIM index per sliding $N \times N$ square box in OD matrices	100
6.5	Comparison of patterns in reference and generated OD matrices: a) ground truth OD matrix; b and c): generated OD matrices that have the same MSE with respect to the reference OD matrix, but different structural patterns	102

6.6	Comparison of patterns in reference and estimated OD matrices: a) "ground truth" OD matrix; b and c): estimated OD matrices that have the same SSIM with respect to the reference OD matrix, but different MSE values.	104
6.7	Sensitivity of SSIM index on OD pairs ordering: a) "ground truth" OD matrix; b) reordered "ground truth" OD matrix; c) prior OD matrix and d) reordered prior OD matrix	105

List of Tables

2.1	Overview of dynamic OD demand estimation methods	18
5.1	The number of state variables that capture 95% of variance per time interval k	91
5.2	Error in the solution given by number of state variables in the state vector	93
5.3	CPU time computations in seconds	94
6.1	Sensitivity of SSIM index on OD pairs ordering in OD matrix	106

Chapter 1

Introduction

Dynamic traffic management system (DTMS) aims to identify traffic problems in a network before they appear, and are used to find a better match between demand and supply to meet the desired network performance. The goal of this dynamic control and influence on the transport system is to anticipate and prevent the occurrence of unfavorable traffic conditions, and to optimize the efficiency, costs and safety of the transport system. Dynamic OD demand estimation and prediction methods have been a major input for a DTMS for many years. These methods are subject to continuous improvement, expanding their capabilities and prediction power. In this thesis dynamic OD demand estimation and prediction methods for real-time applications are developed and analyzed.

The outline of this introductory chapter is as follows. Section 1.1 describes the context and the background of this research, where the dynamic OD demand estimation and prediction problem is introduced. In section 1.2 the objectives and scope of the thesis are discussed. The scientific and practical contributions of this research are listed in section 1.3. Section 1.4 gives an overview of the chapters in this thesis.

1.1 Research motivation

Given the exponential growth of vehicle mobility in the past decades and the increased complexity of traffic and travel patterns over heterogeneous road networks both within and between large metropolitan areas, transport and traffic simulation models have become critically important tools providing road agencies and city planners with support for their decisions in both operations and longer-term planning. An exciting and growing field of application is the use of simulation tools for real-time intelligent transportation systems (ITS) and dynamic traffic management systems (DTMS). DTMS aim to identify traffic problems in a network before they appear, and are used to find a better match between demand and supply to meet the desired network performance. With advanced real-time simulation tools that are able to integrate data from many different sensors, both traffic operators and travelers can be provided with up-to-date and even projected travel and traffic information. The goal of this dynamic control and influence on the transport system is to anticipate and prevent the occurrence of unfavorable traffic conditions, and to optimize the efficiency, costs and safety of the transport system.

A desirable feature of such systems is the ability to estimate network states and predict their short-term evolution. Unreliability and lack of knowledge about past and prevailing traffic conditions (the supply side of road transport operations) may well lead to poor predictions that would, for example, render computed ITS measures irrelevant or outdated by the time they take effect. The traffic assignment tool is a central component of many (more advanced) DTMS measures, and steers decisions regarding the activation and the intensity of these measures. Regardless of either the simulation tools or the application, a necessary input into the operation process is the underlying *traffic demand* imposed on a transport network. This traffic demand is often expressed in matrix form, known as *dynamic or time-varying origin destination (OD) matrix*, where each cell represents the number of vehicle trips departing from origin zone in particular time interval and traveling to destination zone on transportation network. In contrast to static OD matrices (aggregated traffic demand over relatively long period), dynamic OD matrices reflect temporal variation of vehicle departure times over the analysis period on network. They are important input for DTMS applications that capture, for example, dynamic and spatial effects of congestion, dynamic link and path flows or changes in travel times. The inability to provide high quality dynamic OD matrix estimates makes prediction with advanced simulation models simply impossible, irrespective of how well these models have been calibrated. In this respect, the estimation and prediction of an OD matrix with sufficient data and sufficient granularity are critical in establishing the credibility of simulation tools for real-time purposes.

Direct observation of dynamic OD demand is extremely difficult and rare. Full knowledge of OD demand would require the tracking of vehicles on a network, and the extraction of trips' characteristics, i.e. origin of trip, destination of trip, departure time, and mode of transportation. For example, there are emerging data collection systems today that can deduce (sample of) the OD flows using GPS devices, Automatic Vehicle

Identification (AVI), and cell phones. Concerns about these surveillance systems have concentrated on privacy fears, small and not representative sample rates. For example, GPS navigation systems are typically less frequently used for daily recurring trips. As a result of the limitations of surveillance systems, in most networks it is still not possible to observe the OD flows directly.

Since the OD demand cannot be observed, another approach is to turn to demand models that provide an estimate of the OD demand. There are two types of demand models: trip-based and activity-based demand models. These demand models provide reliable estimate of the structure and the order of magnitude of the OD demand for an average day. However, considering the sensitivity of advanced simulation models for small changes in OD demand, the resulting OD demand seldom obtains the required accuracy that is necessary when used as input for these systems. Therefore, the usual procedure is to estimate dynamic OD demand indirectly from the observed traffic data and network conditions they induce on the links and routes of the network. The latter can be obtained using surveillance equipment such as loop detectors, GPS, Bluetooth and WiFi scanners, transit smart cards, and cameras. The estimation and prediction procedure also includes any available prior information on OD demand, which typically comes from results of previous estimations. This research focuses on estimation and prediction of dynamic OD demand for dynamic traffic management (DTMS) support. Because of real-time requirements of DTMS, current and future OD flows must be estimated and predicted at any point in time, on the basis of the most up-to-date observed traffic data. Then, as time proceeds and more traffic data become available, the estimates and predictions must be updated to reflect the evolution of the OD demand and network conditions.

Most real-time applications, such as dynamic traffic control or route guidance generation, require a model to provide output in nearly real-time fashion. Moreover, the stochastic nature of simulation models (e.g., simulation-based dynamic traffic assignment (DTA)) implies the need to make multiple runs to generate statistically robust results. This adds more time constraints on the computational efficiency of dynamic OD demand estimation and prediction model. Speeding up the solution process is desirable even for off-line planning applications, where the evaluation of a single plan may require many simulation runs, adding an excessive cost of time to the evaluation of potential plans. In general, there are three factors that increase the computational effort: a) the size of the state vector, b) the complexity of model components (e.g., assignment matrix, covariance matrices), and c) the number of traffic observations to be processed.

Unfortunately, in contrast to the richness of literature on the topic of dynamic OD estimation and prediction methods, few are tested and proved to be successful for real-world networks. Despite their impact on traffic state estimation, control and management, our understanding of the strengths and weaknesses of OD demand estimation methods for real-time applications has been limited by the lack of tools to assess the performance of existing methods in a generic way. These tools are necessary, because

methodologies developed for small networks with high quality traffic data may not work effectively in practice. Benchmark methodologies are therefore needed to provide a support comparison in a variety of input data settings and network scales. The benefits of determining the particular situations and conditions on the network under which one dynamic OD estimation and prediction method might behave more favorably than another can be significant for practitioners and researchers.

1.2 Research objectives and scope

1.2.1 Research objectives

Driven by the aforementioned issues and requirements for improvement, this thesis addresses several problems pertaining to the provision of efficient and reliable dynamic OD demand information for dynamic traffic management operational applications. The focus of the thesis are dynamic OD demand estimation and prediction, and qualitative analysis in the context of large-scale, real-world networks with various traffic data sources. There are two subproblems presented in this aim listed as follows:

1. *Formulate and develop a real-time dynamic OD demand estimation and prediction model that satisfies a real-time computation constraints for large-scale networks.*

The first objective is important to the successful deployment of dynamic traffic management systems, where dynamic OD demand information serves as an essential input for traffic simulation tools. The dynamic OD demand estimation and prediction problem is computationally intensive because solution methods have to deal with high-dimensional structures of OD matrices and computational complexity of these methods. One possible solution approach to solve the issue of high-dimensionality is to approximate prior OD demand dataset into the lower-dimensional space without significant loss of accuracy. The quantitative methods that explore the information and structure in the (estimated, predicted or realized) OD flows themselves, where the major concern is to reduce dimensionality of the OD matrices, are required. As a result, the dimensionality of the state can be reduced and the complexity of the estimation and prediction problem is likewise reduced. The importance and originality of this approach lies in the possibility of capturing the most important structural information in OD demand without loss of accuracy and considerably decreasing the model dimensionality and computational complexity for real-time applications.

2. *Develop an efficient and generic benchmarking framework to assess the performance of dynamic OD demand estimation methods.*

One of the key traffic variables required for both ex-post and ex-ante evaluation of traffic management and policy measures are OD demand matrices. This

implies that the effect of traffic management measures highly depends on the quality of OD demand estimates and predictions. Thus, the second objective is to handle a common issue in the choice and performance assessment of dynamic OD demand estimation and prediction methods in practice. The assessment methodology needs to be generic, in the sense that various OD estimation approaches can be tested under numerous diverse circumstances related to, for example, data availability and quality, and network lay out. In addition, there is no undisputed performance indicator for assessment of estimated and predicted OD matrices. Instead, there are many candidate statistical metrics that are limited to evaluate temporal and spatial patterns in estimates. The consequence of this sensitive, wide range of statistical metrics is that researchers and practitioners may use different metrics according to their needs, rather than as objective assessment criteria. A structured and generic methodology for benchmarking dynamic OD estimation methods under different circumstances would enable researchers to pinpoint the strengths and weaknesses of various OD estimation methods.

1.2.2 Research scope

This thesis concerns real-time estimation and prediction of dynamic OD demand through the use of dynamic traffic assignment (DTA) models and the observed traffic data. The observations of link traffic counts have been obtained by loop detectors in a real network. Though continuing advances in wireless technologies have made it possible to track each suitably-equipped or electronically-tagged vehicle, widespread adoption of these technologies is not likely in the near future. Companies specializing in emerging data collection (e.g., TomTom, Vodafone, Inrix, and Google) have built their own tools and methods employing their practical expertise in collecting, pre-processing and securing these data. However, this practice often seems diverging and has not yet been formalized in a comprehensive way to supply suitable input to models of complex transport systems, and as a consequence, advanced traffic data are not generally available. However, the problem formulations and solutions presented in this thesis are generic, meaning that they are flexible with respect to the type of traffic data and corresponding modeling assumptions they use. This implies that additional traffic data, such as speeds and density measurements, or even a sample of direct OD flow observations, can be included in problem formulation.

In this thesis, the focus is on the aggregate OD demand modeling approach to estimating and predicting dynamic OD demand in the context of DTA. This approach views trips between every origin to every destination as the unit of analysis, that is, the observed link counts per time interval are used to directly estimate the OD trips over time intervals. Demand correlations across subsequent time intervals are represented in an aggregated way, e.g. by auto-regressive process. It should be noted that this approach differs from the disaggregate demand modeling approach. The disaggregate

demand modeling approach views individual's behavior and activities as the unit of analysis, using the observed link counts to calibrate the parameters of the behavioral model to estimate OD trips. The behavioral models can range from a simple departure-time choice model (as in Lindveld et al. (2003)) to a finer activity-based model (as in Flötteröd et al. (2011)). A shortcoming of the current practice is that despite the disaggregate nature of activity-based travel demand models, dynamic traffic simulation models used for dynamic traffic management are still based on aggregate OD demand. Thus, this thesis focuses on estimation and prediction of aggregate OD demand.

Furthermore, the scope of this thesis is limited to the unimodal transport mode of motorized traffic (e.g., cars and trucks). Other transport modes, such as bicycles, trains and scheduled buses, are not considered. Another limitation is that only a single user class has been considered, as opposed to multiple user classes each group among multiple groups of drivers exhibits a different behavior (e.g. route choice behavior). In addition to these limitations pertaining to the problem formulations and solution algorithms proposed and applied in this thesis, they are also generally applicable to all dynamic OD demand estimation and prediction approaches used in practice (e.g., transport management centers, traffic simulation software).

In this thesis, a real-time dynamic OD demand estimation and prediction method will be developed for DTMS applications. The term real-time has been defined to avoid misinterpretation with the term on-line, as these terms are often used interchangeably in literature. The term real-time in this thesis refers to the ability of the method to handle traffic data continuously and automatically acquired from traffic surveillance equipment. This means that current and future OD flows must be estimated and/or predicted at any point in time, in as short time as possible, based on the most up-to-date traffic data. Then, as time proceeds, the method uses new sets of observed traffic data to update the solution and reflect the evolution of transport demand. In fact, many authors refer to their studies as on-line without explicitly quantifying their response time or any time performance bound, while referring only to the on-line availability of traffic data. Although this research concerns methodological developments for dynamic OD demand estimation and prediction, it aims to be oriented to application. When appropriate, the goal is to use real networks and datasets to illustrate proposed models and algorithms.

1.3 Thesis Contributions

The main contribution of this research has been the presentation of various approaches to the identification of correlation structures (patterns) in OD flows and the methodology to solve the dynamic OD demand estimation and prediction problem efficiently for real-time applications. More specifically, this has been achieved through following scientific and practical contributions.

1.3.1 Scientific Contributions

In summarizing this research's main scientific contributions, this section combines information from previous sections with details from the rest of the thesis, in order to make the scientific contributions concrete.

Literature review classified based on modeling-steps

A rich variety of methods developed so far and in use today are classified based on the *modeling-steps* with which the OD demand estimation and prediction is described; the types of input data, the way in which their relationship with OD flows is modeled, and the solution approaches for the estimation and prediction of dynamic OD demand. This approach shows better how various challenges within each modeling-step have been tackled and how different methods are related to each other.

Dimensionality reduction methods in OD demand estimation

A method for dimensionality reduction and approximation of OD demand has been proposed based on the feature extraction technique, that is, principal component analysis (PCA). OD flows, obtained for several previous days or months, subsume various kinds of information about trip making patterns and their spatial and temporal variations. The research shows how results of the PCA method can be used to reveal structures in the underlying temporal variability patterns in dynamic OD matrices. The results indicate that three main patterns in dynamic OD matrices can be distinguished: structural, structural deviation and stochastic trends. Insight is presented into how each OD pair contributes to these trends and how this information can be used further in predicting dynamic OD matrices on the basis of a set of dynamic OD matrices obtained from real data. By applying PCA, we find that the dimensionality of dynamic OD demand can indeed be significantly reduced. The research provides illustration of how PCA can be applied to linearly transform the high-dimensional OD matrices into the low-dimensional space without significant loss of accuracy.

Methodology for real-time dynamic OD demand estimation and prediction

A methodology to solve the high-dimensionality problem in real-time OD demand estimation has been developed, and the research shows the efficiency of the resulting approximation for large-scale networks. With the acquired knowledge of the potential to linearly transform the high-dimensional OD matrices into low-dimensional space without a significant loss of accuracy, a new problem formulation and solution approach are developed. A new, transformed set of variables (demand principal components) is defined and used to represent the dynamic OD demand in low-dimensional space. These new variables are defined as state variables in a novel reduced-state space model for real-time estimation and prediction of OD demand. The enhanced quality of dynamic OD demand estimates is demonstrated using this new formulation and a so-called colored Kalman filter approach for dynamic OD demand estimation and prediction, in which correlated observation noise is taken into account. Moreover, this demonstrates that by significantly reducing the dimensionality of the dynamic OD demand while, preserving the structural patterns, the computational costs can be dramatically reduced.

Methodology for benchmarking dynamic OD demand estimation

In this research, a methodology for the qualitative assessment of dynamic OD demand estimation methods is developed. Over the last three decades, many dynamic OD demand estimation methods using various traffic data have been developed. These methods differ in many aspects, such as time dimension, mapping methodology of traffic data and OD flows, measures of error, solution approaches and type of networks; adding to the difficulty of creating generalized assessment of OD estimation methods. The methodology discussed here is generic, in the sense that various OD estimation approaches can be tested under numerous diverse circumstances related to, for example, data availability and quality, and network lay out. One of the central components of the methodology presented is an efficient Monte Carlo sampling method, the so-called Latin hypercube (LHC) method. With the results of the benchmark study, it is easier to decide which methods should be subject to further improvement and which can be neglected because of their qualitatively undesirable features. The objective of the benchmark methodology is not to conclude that one approach is the "best", but to provide support for comparison in a variety of settings and conditions. One can use the methodology to perform sensitivity analysis on single or multiple dynamic OD estimation methods.

Measures of performance in OD demand estimation

The structural relationships in OD demand can generally be explained by two types of patterns in OD demand, i.e., temporal and spatial pattern. In this thesis, a new performance indicator, Structural SIMilarity (SSIM) index, that quantifies the dependencies between OD pairs is proposed. To illustrate, under the assumption that any prior OD matrix or available true OD matrix contains the best possible pattern information, the SSIM index can be viewed as an indication of the quality of the estimated OD matrix compared to the prior OD matrix or true OD matrix, respectively. This is important in applications where it is necessary to know whether a particular OD demand estimation method can reproduce actual OD demand. Therefore, the performance indicator can be applied as additional quality measure for benchmarking task of dynamic OD estimation methods.

1.3.2 Practical Contributions

The proposed dynamic OD demand estimation method is developed for specific applications on large-scale networks. When compared to other dynamic OD demand estimation methods, the dimension of the state vector is drastically reduced. Because of this dimensionality reduction the development of a solution approach that allows for efficient computations of dynamic OD demand using traffic data in real time was possible. This is well-suited for applications that require a methodology using observed up-to-date data from the traffic network and gives fast, accurate estimation and prediction results. The practical relevance of the methodology presented here emerges in the following applications: dynamic traffic assignment, real-time traffic state estimation

and prediction, dynamic traffic management in uncongested and congested networks, or in case of unforeseen events. These applications have a great impact on the development of tools and information- generation tailored to influencing individual behavior. For example, city and transport authorities will get more efficient transport networks at a within-day level through better individualized travel advice. On the other hand, the users of those networks will have more reliable journey times, either through reduced journey times or perhaps higher quality experiences.

Insights into the performance of dynamic OD demand estimation methods also greatly benefit practitioners for whom a benchmarking tool would provide a way to assess the quality of the estimated dynamic OD matrices and construct confidence bounds around them. This would, in turn, facilitate the calibration and validation of simulation models using OD demand estimations. With this benchmark methodology one can, for example, determine the particular situations and conditions under which one approach might behave more favorably than another. In the case of transport modelers and ITS operators, these advanced tools will allow them to update their applications to meet the latest modeling advances and evaluate the transport benefits of different management and operation scenarios. Road agencies and city planners will also profit from the tools to make reliable decisions regarding investment in data-collection technologies.

1.4 Thesis Outline

This thesis is structured on four papers and the chapters are given in their chronological order. The content of each chapter and how they relate are schematically outlined in Figure 1.1.

The outline of the thesis is presented in the following and for each chapter reference to the publication is given. **Chapter 2** gives an extensive overview and discussion of the methods previously proposed in literature to estimate and predict dynamic OD demand. The discussion points out relevant challenges and research gaps, laying the foundation for the rest of this thesis.

Chapter 3 presents the efficient benchmark methodology to assess the performance of dynamic OD demand estimation methods. Special attention is dedicated to the design of a generic approach and efficient generation of scenarios and simulations. Each chapter consists of a methodological part and an illustrative case study for real networks. This chapter has been published as: *Djukic, T., J. W. van Lint, S. Hoogendoorn, An Efficient Methodology For Benchmarking Dynamic OD Demand Estimation Methods*. Transportation Research Record: Journal of the Transportation Research Board 2263(1): 35-44, 2011.

In **Chapter 4** the concept of dimensionality reduction and approximation of OD demand is presented as a solution for the high dimensionality of the OD demand estimation and prediction problem. This chapter has been published as: *Djukic, T., J. W. van*

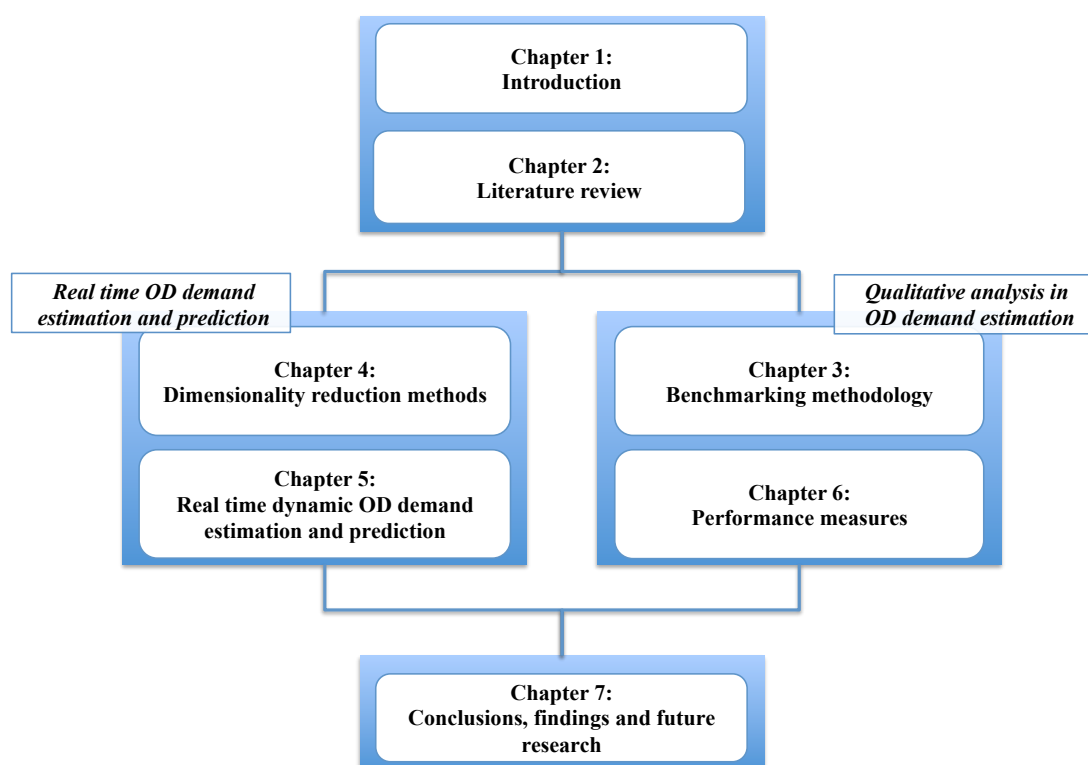


Figure 1.1: Outline of the content of this thesis

Lint, S. Hoogendoorn, Application of Principal Component Analysis to Predict Dynamic Origin-Destination Matrices. Transportation Research Record: Journal of the Transportation Research Board 2283(1): 81-89, 2012.

In **Chapter 5**, dynamic OD estimation and prediction methodology is adapted to improve computational efficiency and provide reliable OD demand estimates for real-time applications. In this way, the dimensionality of the state vector is reduced, while ensuring that most of the structural information about demand is preserved. This chapter has been published as: *Djukic, T., G. Flötteröd, H. van Lint, S. Hoogendoorn*, Efficient real time OD matrix estimation based on Principal Component Analysis. Proceedings of the Intelligent Transportation Systems Conference, Anchorage, Alaska, 2012.

The focus of **Chapter 6** is placed on the choice and evaluation of performance indicators in dynamic OD demand estimation. Potential drawbacks in standard measures are identified and a new performance indicator is proposed to evaluate the patterns in OD matrices. This chapter has been published as: *Djukic, T., S. Hoogendoorn, H. van Lint*, Reliability assessment of dynamic OD estimation methods based on structural similarity index. Proceedings of the of the Transportation Research Board: 13p, 2013.

Finally, **Chapter 7** summarizes the main conclusions and implications of this thesis and gives recommendations for future work on this topic.

Chapter 2

State-of-the-art dynamic OD demand estimation and prediction

This chapter presents a review of dynamic OD matrix estimation and prediction methods. The review follows the evolution of the development of dynamic OD demand estimation and prediction methods since they were first introduced in the 1980's. It started with the use of link traffic counts at intersections (Cremer & Keller (1987)), which were translated into a practical generic optimization problem by Cascetta et al. (1993). This OD demand problem formulation has received a lot of attention in the literature and has been continuously improved and extended. Increasing effort has been put into making dynamic OD estimation and prediction methods produce empirically more realistic outcomes and improving their computational efficiency, resulting in a variety of approaches to the dynamic OD estimation and prediction problem.

The chapter starts with a generic description and formulation of the dynamic OD estimation and prediction problem in Section 2.1. A brief overview of the categorization based on the modeling-steps with which the OD demand estimation and prediction is described is given in Section 2.2. Further, special attention is paid to each of these modeling-steps and detailed literature review is provided in Section 2.3. How the models evolved to predict OD flows is described in Section 2.4. Section 2.5 connects the discussions of all chapters and gives an overview of the main findings.

2.1 Generic formulation of the dynamic OD estimation and prediction problem

The purpose of this section is to provide a generic formulation of the dynamic OD demand estimation and prediction methods that would accommodate all methods currently proposed in the literature.

2.1.1 Generic formulation of the dynamic OD estimation problem

In transportation modeling, the study area is divided into zones, and each zone's characteristics, like attraction (e.g., vicinity to employment) and production (e.g., population) are determined. The transportation network, which is illustrated in Figure 2.1, is represented as a directed graph $G(U, L)$.

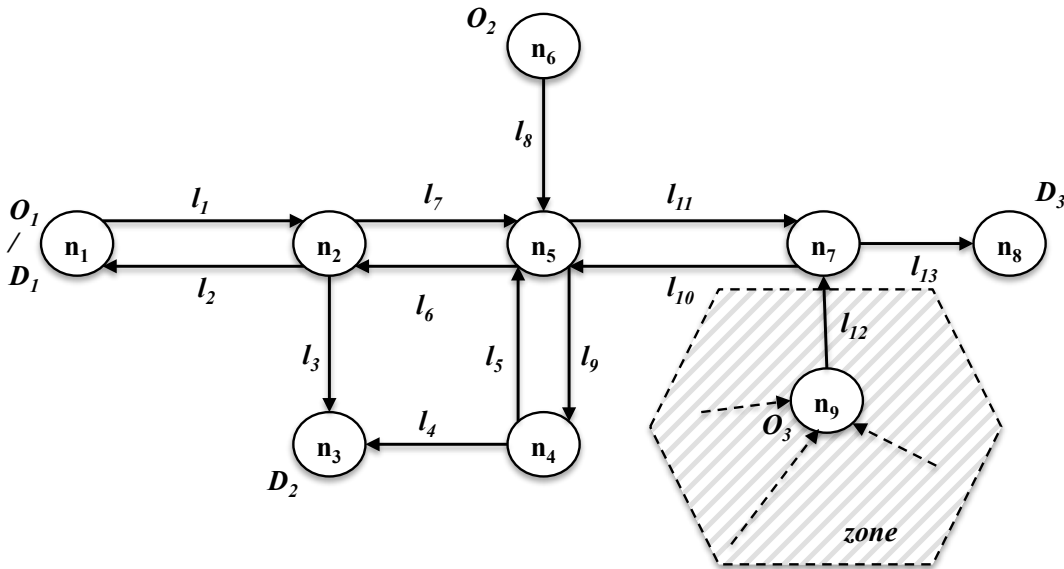


Figure 2.1: Network description: origins, destinations and intermediate nodes, connected by direct links

This *network* consists of directed *links* $l \in L$, where L is the set of links which are connected by *nodes* $u \in U$, where U is set of nodes. Three node types exist: *origins* $o \in O$, where O is the set of origin nodes, i.e. the locations from where the trip starts; *destinations* $d \in D$, where D is the set of destination nodes, i.e. the locations at which trip ends; and *intermediate nodes* $z \in Z$, where Z is the set of intermediate nodes, i.e., intersections on network. Each zone is connected to the network via centroids, the centroid corresponds to origin or destination nodes in the zone. Let $\Omega \subseteq U \times U$ be set of all n OD pairs in the network, and $\hat{L} \subseteq L$ be the set of r links where traffic data observations are available. The time horizon under consideration is discretized into K time intervals of equal duration, indexed by $k = 1, 2, \dots, K$. If $\mathbf{x} \in \mathbb{R}^n$ represents the

OD demand for each OD pair in Ω , the \mathbf{x}_k represents the OD demand at departure time interval k , $i = 1, \dots, K$. In this chapter the dynamic OD demand is represented by a vector, rather than a matrix. It is also important to define κ , the maximum number of time intervals needed to travel between any OD pair in the network. For instance, in dynamic context, depending on the size of the network and its complexity (travel times and distance from the origin o to the destination d), some vehicles could need more than one time interval to reach their destination d or pass traffic sensor at link l . The vector $\mathbf{y}_{k,\hat{L}} = A(\mathbf{x}_h) \in \mathbb{R}^r$, for time interval $h = k, k-1, \dots, k-\kappa$, represents the observed link traffic data at time interval k (e.g. link traffic counts) for each link in \hat{L} .

Before defining the dynamic OD demand estimation problem, it is necessary to express the relationship between the vector of observed link traffic data and the OD flows, given by assignment function $A(\mathbf{x}_h)$. Clearly, the assignment function $A(\mathbf{x}_h)$ plays an important role in estimation process of dynamic OD demand. To explain the traffic assignment model we assume that vector \mathbf{y}_k represents link traffic counts observed in time interval k . The assignment process can be decomposed as follows, independently of the nature of the traffic assignment model.

For all $i \in \Omega$, let P_i be the set of p_i feasible paths linking OD pair i . The total number of paths in the network is given by $p = \sum_{i \in \Omega} p_i$. The formal dependence between link and route flows is given by *link-route proportion matrix*, $\mathbf{R}_k^h \in \mathbb{R}^{r \times p}$, whose elements denote the proportion of route flow i departing in time interval h contributing to link flow l in time interval k . These proportions depend on how link flows are defined, when each route flow reaches link l , and how flows move on links. Commonly, path flows are modeled as space-discrete packets, which means that for this approach the elements of \mathbf{R}_k^h are either 0 or 1, depending on whether packet $[p, h]$ crosses the detector on link l during time interval k .

The formal dependence between OD flows and route flows is defined by *demand-route proportion matrix*, $\mathbf{B}_h \in \mathbb{R}^{p \times n}$. This matrix express the proportion of OD flow i choosing a route p given the departure interval h . Clearly, in uncongested networks, the matrix \mathbf{B}_h vary moderately as a function of the OD flows. However, in congested networks, the dependence of the matrix \mathbf{B}_h to the OD flows becomes more pronounced and significantly complicates dynamic OD estimation problem (more elaborate discussion will be provided in Section 2.3.2).

By combining link-route and demand-route proportion matrix, the *traffic assignment matrix* $\mathbf{A}_k^h \in \mathbb{R}^{r \times n}$ is defined as

$$\mathbf{A}_k^h = \sum_{h=k-\kappa}^k \mathbf{R}_k^h \mathbf{B}_h \quad (2.1)$$

and assignment model is given by

$$A(\mathbf{x}_{k-\kappa}, \dots, \mathbf{x}_k) = \sum_{h=k-\kappa}^k \mathbf{A}_k^h \mathbf{x}_h \quad (2.2)$$

Thus, the relationship between observed link traffic counts and OD demand can be expressed as:

$$\hat{\mathbf{y}}_k = \sum_{h=k-\kappa}^k \mathbf{A}_k^h \mathbf{x}_h \quad (2.3)$$

Now, dynamic OD demand estimation problem can be defined. Given a vector of observed traffic data at time interval k , $\mathbf{y}_k \in \mathbb{R}^r$, the dynamic OD estimation problem consists of finding an OD demand for departure time k , \mathbf{x}_k , such that $\hat{\mathbf{y}}_{k,\hat{\mathbf{x}}_k}(\mathbf{x}_h)$ is as close as possible to observed values \mathbf{y}_k . Therefore, the dynamic OD estimation problem is formulated as:

$$\hat{\mathbf{x}}_k = \underset{\mathbf{x} \geq 0}{\operatorname{argmin}} f\left(\sum_{h=k-\kappa}^k \mathbf{A}_k^h \mathbf{x}_h, \mathbf{y}_k\right) \quad (2.4)$$

Function f is given in the form of functions measuring the deviation between estimated and observed traffic data, as will be discussed in section 2.3.3.

Usually, the information on dynamic OD flows contained in link traffic counts, represented by the system of stochastic equations depicted in (2.3), is insufficient to estimate the dynamic OD flows. Indeed, even if we assume linear system of equations, the number of r independent equations is usually much less than the number of unknown OD flows n to be estimated. Thus, for most practical applications, the dynamic OD estimation problem is underdetermined. That is, there is an infinite number of valid OD matrices that, when assigned on the network, exactly reproduce the link traffic counts observed on the links.

In summary, the information contained in link traffic counts must be combined with that from other sources to estimate the unknown OD demand flows. To overcome that problem, it is common to use a historical OD demand for time interval k , $\tilde{\mathbf{x}}_k$, referred here as a prior OD demand, and to select among the infinite number of potential candidates the one that is closest to the prior OD demand to reach a unique solution. The prior OD demand is usually obtained from a transportation studies, travel surveys, or estimations of OD flows from previous days. Further techniques, based on additional data to identify a structure of the OD demand, have also been proposed as an attempt to address underdetermination (see, for example, Bierlaire and Toint, (1995)). Section 2.3.1 provides a detailed overview of different types of traffic data that provide additional information on OD flows.

Finally, the dynamic OD estimation problem can be rewritten as follows:

$$\hat{\mathbf{x}}_k = \underset{\mathbf{x} \geq 0}{\operatorname{argmin}} [\alpha f(\mathbf{x}_k, \tilde{\mathbf{x}}_k) + (1 - \alpha) f\left(\sum_{h=k-\kappa}^k \hat{\mathbf{A}}_k^h \mathbf{x}_h, \mathbf{y}_k\right)] \quad (2.5)$$

Regardless of the function f used, the purpose is to obtain an OD matrix that yields OD flows and traffic data as closely as possible to their observed values. Note that the weighted formulation can be adopted to combine the two sets of deviations, with respective weights α and $(1 - \alpha)$ for the first and second function. The weights could

be interpreted as the decision maker's relative preferences or importance belief for the different objectives. For example, if provided prior OD demand information is not reliable a small value of α is used, and vice versa.

2.1.2 Generic formulation of the dynamic OD prediction problem

Let $t \in K$ be a prediction horizon, $t > 0$, and x_{k+t} is a vector of predicted dynamic OD demand for time interval $[k, k+t]$. Let $\tau, k > \tau \geq 0$, denotes the number of past time intervals that capture the effect of previous state estimates on the state in current time interval k . Given a vector of prior OD demand $\tilde{\mathbf{x}}$, and vector of real-time OD demand estimates $\hat{\mathbf{x}}$, the dynamic OD prediction problem can be formulated as:

$$\mathbf{x}_{k+t} = g(\hat{\mathbf{x}}_k, \hat{\mathbf{x}}_{k-1}, \dots, \hat{\mathbf{x}}_{k-\tau}, \tilde{\mathbf{x}}_{k+1}, \tilde{\mathbf{x}}_k, \dots, \tilde{\mathbf{x}}_{k-\tau}) \quad (2.6)$$

where g is function that represents the spatial and temporal OD flow relations between intervals $k - \tau$ and $k + 1$. For example, function g can be formulated as a random walk model, linear trend model, or autoregressive model. Section 2.4 provides more elaborate discussion on various examples of dynamic OD demand prediction methods.

The dynamic OD matrix estimation and prediction framework is depicted in figure 2.2.

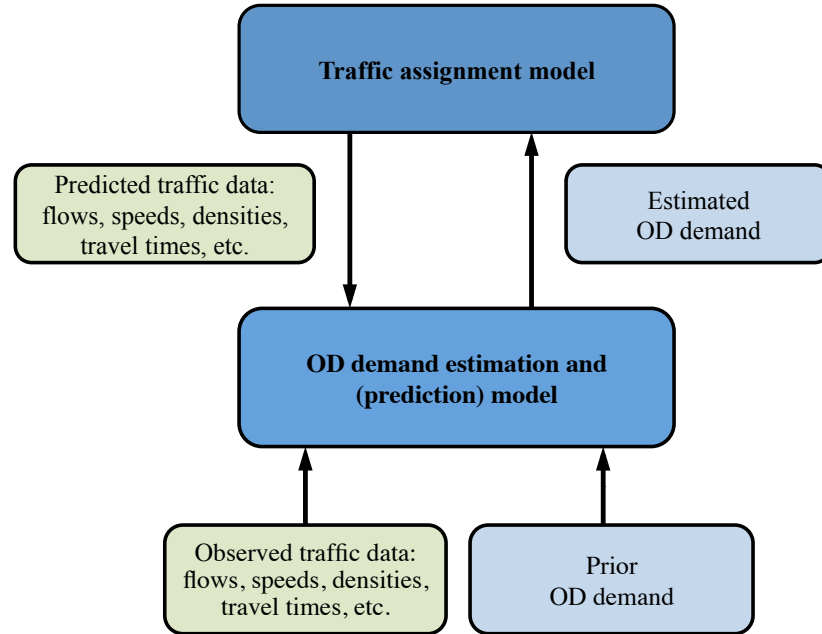


Figure 2.2: The overview of dynamic OD matrix estimation (and prediction) framework

2.2 Categorization of dynamic OD demand estimation methods

In this section, categorization of dynamic OD demand estimation methods proposed in literature is presented. It serves as reference for the overall picture in the detailed discussions of the problem formulations and solution approaches in Section 2.3.

The focus in literature in the dynamic OD estimation problem has been quite diverse. To present the different methods in organized fashion, a classification is therefore necessary. Categorization of dynamic OD demand estimation and prediction methods has been done so far according to different criteria such as whether the DTA model is needed for generating link flow proportions - traffic assignment-based or non-traffic assignment-based (Zhou (2004)), recursive vs. non-recursive approaches (Nie (2006)), the scale of the application - closed networks vs. open networks (Ashok (1996)), off-line vs. on-line (Antoniou et al. (2004), Zhou (2004), Peterson (2007)), traffic conditions - uncongested vs. congested (Bert (2010)). These classifications are closely related since they have in common that they refer to the application domain of the dynamic OD estimation methods.

The overview given in this chapter takes different approach and gives a literature overview through the most important steps involved in dynamic OD demand estimation and prediction modeling. A rich variety of methods developed so far and in use today are classified based on the *modeling-steps* with which the OD demand estimation and prediction is described; the types of input data y , \bar{x} , the way in which their relationship with OD flows is modeled $A(x)$, and the solution approaches f for the estimation and prediction of dynamic OD demand. This approach shows better how various challenges within each modeling-step have been tackled and how different methods are related to each other. This review will mainly consider theoretical issues of each modeling-step derivations and characteristics. However, some practical issues, such as application domain, are discussed as well. Furthermore, it forms the basis for further research directions within each modeling step. Static OD demand estimation methods are omitted in literature review. This would make review much more extended without adding much insight since all dynamic OD estimation methods has been formulated as extensions of static methods.

To this end, the discussed dynamic OD estimation methods are classified according to the following modeling steps:

- **Input data:** The input data can represent link flows, OD flows, travel times, traffic densities, route paths, network design, etc. This distinction is important because different types of input data provide different information on OD flows and result in different problem formations and assumptions. In this respect, OD flow data, route flow data and link condition data will be distinguished.

- **Mapping of OD flows to input data:** From a modeling point of view, the most distinguishing difference between the OD demand estimation approaches, is how the relationship between state variables (e.g., OD flows, OD proportions) and any available traffic data (e.g., link traffic counts) is defined, calculated and recalculated through out the estimation process. An accurate description of this relationship leads to an accurate description of traffic state reality, but to more complexity as well. With respect to the operationalization criterion, this relationship can be operationalized either as analytical solutions of sets of equations or as a simulation model.
- **Objective function and solution approach:** The definition and/or choice of objective functions depends on whether explicit assumptions are made on the probability distribution of the random residuals of OD flows and traffic observations. In this respect, existing approaches are distinguished by applied objective functions, i.e., least squares, state-space and entropy maximization. The solution framework has to take into account the formulated objective function and constraints imposed in the previous two steps. The properties of a solution framework and algorithm determine the application domain of dynamic OD demand estimation model, i.e., for off-line or real-time applications.

Table 2.1 presents an overview of some well known dynamic OD estimation methods based on proposed criteria. While not being exhaustive, the table provides insights into dynamic OD demand estimation efforts during the last three decades of related research.

In the reminder of this chapter, the dynamic OD estimation methods are discussed in more detail, in the order of the *modeling – steps* classification (input data, mapping of OD demand to input data and solution approaches). The aim of the discussion is to provide some insight into main challenges and goals of each modeling-step and resulting solution approaches.

2.3 Dynamic OD demand estimation: State-of-the-art

2.3.1 Types of input data used in dynamic OD demand estimation and prediction

OD flows are difficult to observe directly, because this would require continuous access to the trips and tracking of vehicles on network. Since it is not often possible to directly observe OD matrices, they must be estimated from available traffic data. In the last decades, the amount of empirical traffic data becoming available for both on-line and off-line use has increased, particularly in terms of the wide range of sensor technologies developed and applied to collect traffic data. Traffic sensors may range

Table 2.1: Overview of dynamic OD demand estimation methods

Author	Input data			Mapping		Objective function		
	ODF	RF	LC	A-DTA	S-DTA	GLS	SS	ME
Bell (1991)	+		+	+		+		
Cascetta et al. (1993)	+		+	+	+			
Chang & Wu (1994)	+		+		+		+	
Chang & Tao (1996)	+	+	+	+			+	
Wu (1997)	+		+		+			+
Van Der Zijpp & De Romph (1997)	+	+	+	+			+	
Tavana (2001)	+		+		+	+		
Sherali & Park (2001)	+		+		+	+		
Dixon & Rilett (2002)	+		+	+		+		
Ashok & Ben-Akiva (2002)	+	+	+	+			+	
Mishalani et al. (2002)	+		+		+		+	
Tsekeris (2003)	+		+		+			+
Lindveld et al. (2003)	+		+		+	+		
Bierlaire & Crittin (2004)	+		+		+	+		
Kwon & Varaiya (2005)	+	+	+		+		+	
Zhou & Mahmassani (2006)	+		+		+	+		
Antoniou et al. (2006)	+	+	+	+			+	
Balakrishna (2006)	+		+		+	+		
Zhou & Mahmassani (2007)	+	+	+		+		+	
Barcelo (2010)	+	+	+	+			+	
Cipriani et al. (2011)	+		+		+	+		
Frederix et al. (2011)	+		+		+	+		

Input data: ODF, OD flow data; RF, route flow data; LC, link condition data; Mapping of OD flows to input data: A-DTA, analytical based dynamic traffic assignment; S-DTA, simulation based dynamic traffic assignment; Objective functions: GLS, generalized least squares; SS, state-space; EM, entropy maximization.

from inductive loop detectors, radar, ultrasonic sensors, and electronic toll collection devices to infrared cameras (automatic vehicle identification (AVI) data) and in-vehicle GPS/GSM receivers/transmitters (automatic vehicle location (AVL) data). Nowadays, AVL data (or floating car data (FCD)) can also be obtained with mobile phones and Bluetooth devices, among others. Sensors typically measure traffic characteristics, which are the result of not only OD demand, but also route choice and traffic operations.

Although data from traffic sensors come in many forms and qualities, they can essentially be subdivided into three categories depending on the source of information on OD flows and traffic operations. The types of input data that are used in literature for dynamic OD estimation and prediction are subdivided as depicted in Figure 2.3: (1) *OD flow data*; (2) *route flow data*; and (3) *traffic condition data*. The first type of input data, OD flow data, represent direct observations of OD flows obtained from surveys or probe vehicles. The second type of input data, route flow data, are determined by the travel behaviour process. This process describes travel choices: when to depart, which mode to use, which route to choose. The third category of input data, traffic condition data over network, is determined by traffic operations. These data describe traffic state on a network: link flows, travel speeds, travel times, densities, etc. The sources of input data and their application in dynamic OD demand estimation and prediction process are discussed below.

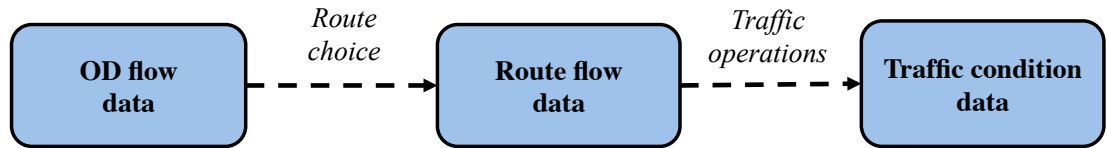


Figure 2.3: Types of input data used in dynamic OD demand estimation and prediction

To illustrate on a very simple network example, suppose there are four OD flows going from A to C, A to D, B to C and B to D (see Figure 2.4). Figure 2.4 gives some examples of different traffic sensors and the spatio-temporal semantics of the traffic variables that can be observed with those sensors. This figure will be used as a reference to explain the most important sources of OD data, and their features for dynamic OD demand estimation and prediction.

OD flow data

OD flow data may be collected from traffic data collection points and surveys that provide observation of OD flows on a network. These data, that depend on the degree of information on OD flows available, encompass *full information* (i.e., OD flows over all OD pairs are known) or *partial information* on OD flows (i.e., a fraction of OD flows over few or all OD pairs are known).

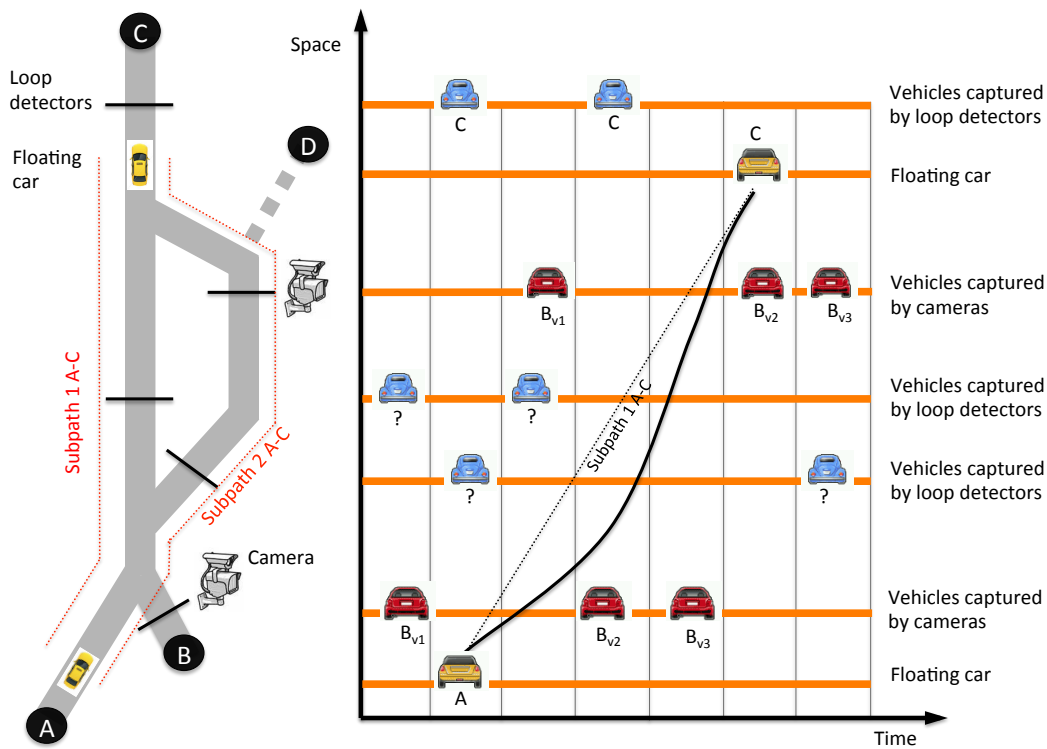


Figure 2.4: Examples of some traffic data and their sources

Observations of OD flow data are rare. Data on OD flows collected from preference surveys among a sample of travelers (derived from home- and road-side interviews) may provide a detailed picture of trip patterns and travelers' destination choices. These surveys collect sample data related to the OD flows and spatial distribution of OD flows. However, this information on OD demand represents an aggregate value over fixed time, i.e. static OD demand and does not provide temporal distribution of demand over time. Different proposals have been elaborated for pre-processing static OD demand to obtain dynamic OD flows. One approach is a time slicing of the static OD matrix based on available traffic counts per time interval (TSS (2006)). Accordingly, each OD pair flow for one time interval is modified by the same percentage, resulting in unreliable input information for the OD demand estimation and prediction problem. The practical and theoretical limitations of survey OD generation techniques have led to an exploration of how such data could be derived from equipped vehicles with in-vehicle traffic sensors that act as probes by transmitting their origin and intended trip destination when they initiate a trip.

Automatic vehicle location (AVL) data are receiving attention for their potential to provide a large sample of OD flow data. The observation of OD flows from in-vehicle traffic sensors (e.g., GPS and GSM) allows the detection of vehicles in multiple locations as they traverse the network. This feature makes the re-identification and tracking of these probe vehicles possible, which in turn may (under certain conditions) provide

information on particular OD pairs (e.g., OD pair AC in Figure 2.4). In an ideal case, if data on OD flows are collected from all vehicles equipped with in-vehicle traffic sensors, full information on OD flows over all OD pairs can be extracted. Today, probe vehicles constitute only a fraction of the total number of vehicles in a network. Thus, in-vehicle data provide partial information on total OD flows as depicted in Figure 2.5(b). Several models have been developed for the estimation of OD flows using AVL data (N. Caceres (2007), Ashok & Ben-Akiva (2000), Van Aerde et al. (1993)). Ashok & Ben-Akiva (2000) introduced the notion of direct measurements for the incorporation of AVL data into the solution of the OD estimation and prediction problem.

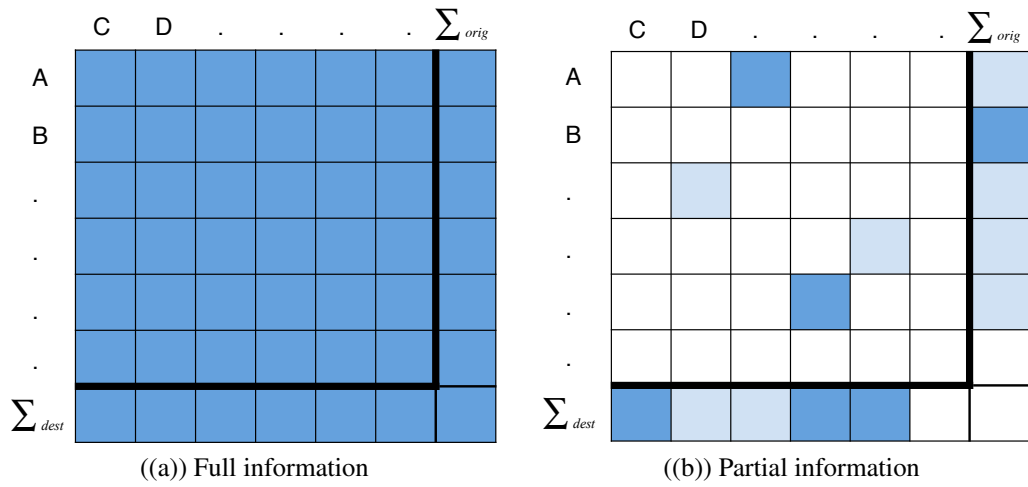


Figure 2.5: Illustration of full and partial information on OD flows from OD flow data sources: a) full information on OD flows and b) examples of partial information on OD flows with respect to full information on particular OD pairs (dark blue color) or their sample (light blue color)

Automatic vehicle identification (AVI) data represent another OD flow data source of growing importance for estimating dynamic OD demand flows. The observation of OD flows from AVI sensors (e.g., electronic-toll collection devices, infrared cameras, Bluetooth, WiFi, etc.) depends on: a) the location of these traffic sensors on a network, as depicted in Figure 2.4 and b) the sample of tagged vehicles. In an ideal case, if cameras are located on links connected to origin and/or destination nodes on a network, they can provide under some assumptions total demand that departs from origin B or arrives at destination C, as represented in Figure 2.5(b). If only a subset of vehicles is equipped with transponder tags or only a subset of vehicles is correctly identified by the AVI readers, then these OD flow data need to be explicitly considered in order to infer OD flows over all OD pairs. Several models have been developed for the estimation of OD flows using AVI data (Asakura et al. (2000), Dixon & Rilett (2002), Antoniou et al. (2006)). In brief, these models require estimating the sample rate (either market penetration rates or identification rates) so as to relate the AVI samples to the OD demand. The estimation of sample rates, however, is a difficult problem in its own right,

as these rates are essentially time-dependent and location-dependent random variables. Moreover, the inclusion of sample rates in the OD demand estimation problem could dramatically increase the number of unknown variables and impact the reliability of OD demand estimates. To circumvent primary difficulties associated with estimating sample rates, Zhou & Mahmassani (2007) developed an OD demand estimation model using partially observed AVI data.

Route flow data

New technologies for probe vehicle re-identification and tracking (e.g. AVI systems and AVL systems) might provide route data, such as partial point-to-point travel times, route choice fractions, vehicle paths, and turning fractions. The data may come from cameras that capture and compare vehicle plates or from floating car data which may report the vehicle's location at certain intervals to construct trajectories, as is depicted in Figure 2.4. The difficulty for the OD demand model formulation is to define the relationship between traffic flow data and OD flows. Thus, the identification of trajectories or route travel times can help to identify or estimate route flows. Therefore, they provide constraints on the traffic conditions resulting from assigning the OD flows to the network. Estimating OD matrices only from link flow data can be rather challenging given the indeterminate relation between link flow observations and route flows (Parry & Hazelton (2012)). Hence, few researchers have tried to integrate route flow data into the dynamic OD demand estimation problem. Examples include: turning fractions (e.g., (Van Der Zijpp & De Romph (1997), Mishalani et al. (2002))) and route flows (e.g., Sun & Feng (2011), Antoniou et al. (2006)).

Traffic condition data

Traffic link flow data collected from loop detectors at specific locations on a network are the most common type of input data used in dynamic OD demand estimation and prediction. The traffic link flow data could either be collected in the middle of a roadway segment, at entry or exit ramps on highways, or across a screen-line in an urban area. The number and position of loop detectors on an urban or highway network plays an important role, since traffic link flow data from these detectors can provide different information on OD flows. In an ideal case, if link flow data are collected on road segments belonging exclusively to routes used to serve one particular OD pair, they can provide information on OD volume for that particular OD pair. In addition, if loop detectors are located on links connected to origin or destination nodes on a network, they can provide under some assumptions total demand that departs from origin B or arrives at destination C, as represented in Figure 2.5(b). The traffic link flow data observed by loop detectors located on links between nodes 1 and 2 in Figure 2.4 are comprised of contributions from several OD flows (i.e., OD pairs: AC, AD, BC, BD). Thus, such link flow data require adequate specification of relation and mapping with OD flows. This procedure is discussed in detail in Section 2.3.2.

Apart from traffic link flow data, loop detectors are able to detect speeds and occupancy at links in the network. The available speed or derived density measurements can help to identify whether traffic link flow data represents a congested or uncongested traffic state on a network. As such, they can facilitate correct interpretation of traffic link flow data, and identification of OD flows that need to be adjusted, and in which direction. The simplest approach to including this type of input data is to include speed or density measurements in the goal function of the dynamic OD estimation problem given by equation 2.5 (e.g., Balakrishna (2006), Cipriani et al. (2011), Frederix et al. (2011)); or travel times (e.g., Barcelo (2010), Cipriani et al. (2011)).

Discussion

The above summary of the traffic data used to estimate the OD demand follows the historical lines of the development of dynamic OD matrix estimation methods since they were first introduced in the 1980's. The application of link flow data for estimation of dynamic OD flows has largely remained present in relevant literature since the pioneering work of Willumsen (1978), Van Zuylen (1978) and Nguyen (1977). However, today, the development of new traffic surveillance technologies has turned researchers' attention and attempts to applying observations of OD flows and traffic condition data in OD demand estimation and prediction.

The traffic data may be unreliable, inaccurate or have limited time-space coverage. Concerns about emerging data collection systems have concentrated on privacy fears, high error rates, misidentification and high investments. Therefore, in this thesis, link flow data and OD flow data will be taken as input data, as they provide the minimum data needed to complete an OD estimation, and in a majority of cases, they are the only available data for the concerned study area (including other type of input data is an interesting area for further research, see Chapter 7.2).

The next sections will show how the existing OD demand estimation methods use the above-mentioned input data to estimate dynamic OD flows.

2.3.2 Mapping of OD flows to input data

This section describes the most critical issue in OD matrix estimation, whether static or dynamic, the relationship (mapping) of the observed link flow data and traffic condition data with unobserved OD flows. From a modeling point of view, the most distinguishing difference between the OD demand estimation approaches presented in literature, is how the relationship between OD flows and any available traffic data (e.g. link traffic counts, speeds, densities, etc.) is defined, calculated and re-calculated through out the estimation process. This relationship is accomplished by means of an assignment matrix introduced in Section 2.1.1. In the dynamic problem, the assignment matrix depends on link and path travel times and traveler route choice fractions - all of which

are time-varying, and the result of dynamic network loading models and route choice models. These dynamics are reflected in travel times between each origin and destination trips on a network, influenced by traffic link flow. While a vast body of literature has been developed in this area over the past three decades, this section focuses on some of the efforts that highlight the basic problem dimensions.

The assignment function that provides mapping between the unknown OD flows and link traffic counts is a complex function of the OD demand, driver route choice decisions, departure time choice and resulting network travel times. Here we recall assignment function defined in section 2.1, where the relationship between observed link traffic counts and OD demand can be expressed as:

$$\hat{\mathbf{y}}_k = \sum_{h=k-\kappa}^k \mathbf{A}_k^h \mathbf{x}_h + \boldsymbol{\varepsilon}_k \quad (2.7)$$

where random vector $\boldsymbol{\varepsilon}_k$ is the sum of two (independent) modeling and observation errors.

The elements of assignment matrix are generally a function of the OD flows, and they represent both the propagation of the OD flows, the departure time and the route choice decisions related to an OD flow. The assignment matrix can be seen as the product of two terms (see Cascetta et al. (1993)):

$$\mathbf{A}_k^h = \sum_{h=k-\kappa}^k \mathbf{R}_k^h \mathbf{B}_h \quad (2.8)$$

Equation (2.8) shows that the actual computation of a dynamic assignment matrix depends on the modeling of users' route choice behavior (i.e., route choice model) and the modeling of route flow movement in the network (i.e., dynamic network loading model). These two models form part of the Dynamic Traffic Assignment (DTA) framework depicted in Figure 2.6.

In the equation (2.8), the elements $b_p^{i,h}$ of the *demand-route proportion matrix*, \mathbf{B}_h , express the proportion of OD flow i departed in time interval h choosing a route p . In practice, the demand-route proportion matrix can be obtained through use of a *Route Choice* (RC) model. The RC model distributes trips in the dynamic OD matrix over the available routes for each departure time and each OD pair. The route flows are transferred to a *Dynamic Network Loading* (DNL) model that simulates spatio-temporal propagation of route flows through the network and computes the dynamic link travel times and dynamic link flows. The elements, $r_{l,k}^{i,h}$, of the *link-route proportion matrix*, \mathbf{R}_k^h , denotes the proportion of route flow i departing in time interval h contributing to link flow l at time interval k . These proportions depend on how link flows are defined, when each route flow reaches link l , and how flows move on links. Route travel times are transferred back into the RC model and users may adapt their chosen route according to the new traffic conditions. The DTA is thus an iterative procedure, converging to dynamic traffic equilibrium. In addition to choosing a route, travelers can also decide

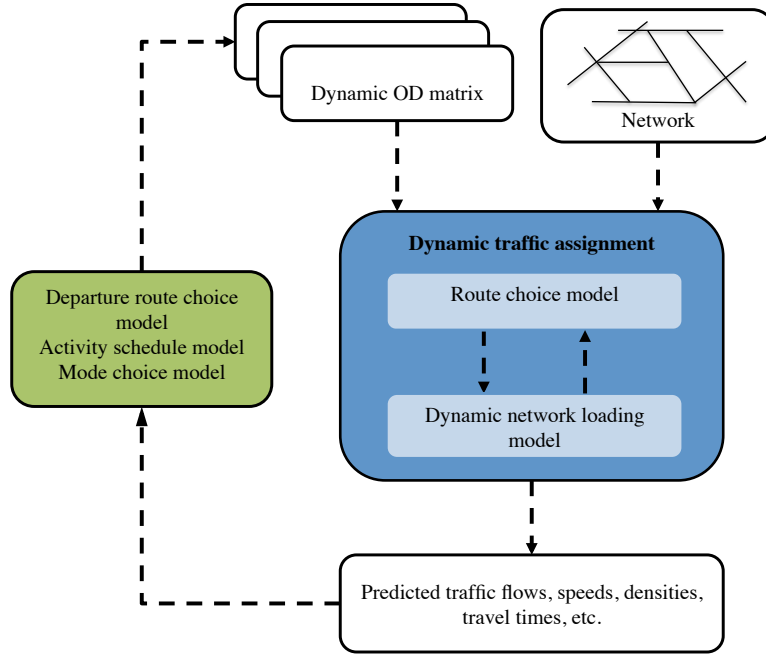


Figure 2.6: Dynamic traffic assignment framework

to change their departure time. The departure time choice model uses the OD travel times for each departure time to determine the optimal departure time for all travelers and produces departure time rates. The departure time choice model is an optional component of DTA models. To be as realistic and flexible as possible, here analytic and simulation based DTA approaches will be discussed briefly, including which traffic data (route flows and traffic condition data) are important for these approaches.

Analytic DTA models

When route flows or traffic condition data are observable (travel times in the network or route-choice fractions can be estimated), the analytical expressions can be derived to compute the traffic assignment matrix. The elements of assignment matrix can be calculated easily if travel times on all links are known under free-flow conditions on a network. The *analytic DTA models* describe the average behavior of traffic with macroscopic traffic flow variables such as inflow rates and travel times. In addition, to accommodate considerations of erroneous travel times or route-choice fractions, a stochastic assignment matrix based model might be preferred (Ashok & Ben-Akiva (2002), Van Der Zijpp (1997)). Although these models have potential as tools for deriving theoretical insights, well-behaved mathematical formulations are currently unavailable, which remains the primary difficulty in ensuring the first-in-first-out (FIFO) property (Peeta & Ziliaskopoulos (2001)). These negative properties severely restrict the function that maps route flows on travel times, i.e., the DNL model. There may be situations, however, where the available surveillance system only allows for measurement of link counts. In such a scenario, and to avoid mathematical intractability, the assignment matrix would be obtained through the iterative application of simulation-

based DTA models and OD estimation models, to describe certain traffic phenomena more accurately, even at higher computational costs.

Simulation DTA models

Simulation DTA models use a traffic simulator to replicate the complex traffic flow dynamics, critical for developing meaningful operational strategies for real-time deployment (Peeta & Ziliaskopoulos (2001)). Indeed, simulation DTA models keep track of individual vehicles, or vehicle packets, at each time step, and describe more accurately certain traffic flow phenomena than analytical models. The models use sophisticated algorithms and detailed macroscopic, microscopic, and mesoscopic simulation techniques to estimate current network performance, predict future conditions and generate traffic guidance. Both in academia and in practice, many different DTA simulation models are used to compute assignment matrices in the context of dynamic OD demand estimation. Examples include DYNASMART (Zhou & Mahmassani (2006), Tavana (2001)), Dynamaq (Cipriani et al. (2010)), MitSimLab (Balakrishna et al. (2007)), DynaMIT (Gupta (2003)), and Aimsun (Barcelo et al. (2010)).

These simulation models use different DNL models and RC models, which have different effects on dynamic OD demand estimation. The choice of DNL model and RC model within a DTA framework significantly influences the accuracy of dynamic OD demand estimation. The effect of DNL models on DTA based OD demand estimation has been studied by Frederix et al. (2010). Their results indicated that applying a DNL model that represents queuing properly is a necessary condition, but not sufficient to guarantee an accurate OD estimation. The authors note that convergence is not guaranteed when the historical OD matrix is not close to real OD demand. Including traffic condition data about the traffic state in the OD estimation problem, in the form of speeds or travel times can be an effective solution to this problem (see e.g., Balakrishna et al. (2007), Ashok & Ben-Akiva (2002), Tavana (2001)). In general, there are two recognized approaches to including traffic condition data in OD estimation process. The first approach uses speed measurements and travel times as an empirical alternative to an analytical DNL model to derive the assignment matrix as discussed above. The second approach uses speed measurements and densities as additional information within an objective function (equation 2.5). While the effect of DNL models has been studied in literature, the effect of RC models still remains open for research. Peeta & Ziliaskopoulos (2001) provided a state-of-the-art review and detailed discussion of formulation approaches, solution methodologies and traffic flow modeling strategies on the DTA problem.

Recently, research efforts have emphasized the importance of accurate modeling of the relationship between unknown OD flows and link flows under the congested traffic state to improve dynamic OD demand estimation. To determine OD flows, it becomes necessary to understand which information traffic observations provide in congested networks (Figure 2.7).

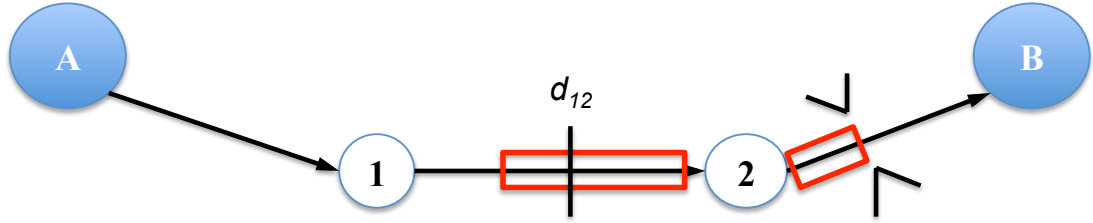


Figure 2.7: Mapping in congested traffic conditions

Suppose that there is a bottleneck between node 1, n_1 , and destination node B, n_B , with a capacity lower than the demand between A and B. Then, the queuing occurs upstream of this bottleneck (see Figure 2.7). As long as demand exceeds capacity, this queue will grow, and it will eventually spill back onto detector d_{12} . From that moment, link traffic counts will no longer provide information on OD demand, but will rather measure information on the supply, i.e., capacity of the bottleneck. Thus, the assumption that the traffic assigned to a path or a link is proportional to the demand values is not valid in congested traffic conditions. As can be deduced from the assignment relationship in equation (2.8), there are two main sources of non-linearity leading to the non-convexity of the dynamic OD estimation problem: the dependence of the link-route proportion matrix, D , on OD flows, and the dependence of the demand-route proportion matrix, R , on OD flows. The influence of non-linearity of the link-route proportion matrix on OD estimation has been studied by several authors (e.g., Flötteröd & Bierlaire (2009), Frederix et al. (2011)). In Frederix et al. (2011), the method of Marginal Computation is used to efficiently calculate the sensitivity of the link flows to the OD flows, which is not captured by the linear approximation of the relationship between OD flows and link traffic counts given by equation (2.7). The difficulty of deriving an exact calculation of the non-linear relationship between traffic condition data and OD flows has resulted in different solution frameworks that will be discussed in detail in following sections.

The following section will provide an overview of how this complex relationship between traffic data and OD flows was incorporated within solution frameworks.

Discussion

Current efforts, while continuing to address the fundamental issues, are increasingly focusing on real-time operational issues such as robustness of the DTA, given the inherent system randomness. An important research direction is to improve the representation capability of DTA models to adequately describe traffic dynamics and behavioral processes in a network. It is vital to have an accurate DTA models that would realistically describe network flow distributions, to capture the effects of congestion as well as route choices. If DTA models are not realistic, they will lead to biased OD demand estimates.

The work presented in this thesis focuses on simulation DTA approaches concerning traffic assignment. Indeed, simulation tools are more realistic and flexible in terms

of utilization, customization and validation in benchmark task with other models proposed by the scientific community. Moreover, the models deal better with the complex dynamic situations and high level of network detail than analytical models. Nevertheless, in this thesis it is assumed that a DTA model should always ensure that the evolution of congestion patterns and travel times are well-captured.

2.3.3 Objective functions and solution frameworks

This section discusses the error measures f , given in equation (2.5) as part of the dynamic OD demand estimation problem. The error measures are distinguished based on probability assumptions regarding OD flows, i.e., depending on whether explicit assumptions are made on the probability distribution of the random residuals of OD flows and traffic observations. This overview on error measures is complementary to the one given by Cascetta & Postorino (2001) for the static OD demand estimation problem. In this section, the focus is on a review of error measures for the dynamic OD demand estimation problem. Unlike the dominant studies on how the relationship between dynamic OD flows and traffic observations is made, the dynamic OD matrix estimation problem has been formulated incorporating multiple OD flow distribution rationales, such as maximum entropy, least squares, and Bayesian inference. Furthermore, some new relationships between the objective functions and approaches to dealing with non-linear constraints are discussed. The remainder of this section will consist of an overview of several prospects for the error measures, namely generalized least squares, maximum entropy, and the state space model, followed by a review of the solution frameworks and algorithms for specific problems.

Generalized least squares (GLS) formulation

The Generalized Least Squares (GLS) method is concerned with determining the most probable value of OD demand as the value that minimizes the sum of the squares of the residuals. The residuals of traffic condition data and prior OD demand are modeled as

$$\varepsilon_x = \hat{\mathbf{x}}_k - \tilde{\mathbf{x}}_k \quad \varepsilon_y = \hat{\mathbf{y}}_k - \mathbf{y}_k \quad (2.9)$$

where the residuals ε_x and ε_y have symmetrical distributions (i.e., normal or uniform) and $E(\varepsilon_x) = E(\varepsilon_y) = 0$ is assumed. The dispersion matrices of the ε 's must be known; they are formally defined as $Var(\varepsilon_x) = V_t$ and $Var(\varepsilon_y) = W_t$. If measure of error f given in equation (2.5) is defined as GLS, then objective function can be formulated as:

$$[\hat{\mathbf{x}}_1 \dots \hat{\mathbf{x}}_K] = \underset{[x_1 \dots x_K] \geq 0}{argmin} \sum_{k=1}^K [(\mathbf{x}_k - \tilde{\mathbf{x}}_k)^T V_k^{-1} (\mathbf{x}_k - \tilde{\mathbf{x}}_k)] + \sum_{k=1}^K [(\hat{\mathbf{y}}_k - \mathbf{y}_k)^T W_k^{-1} (\hat{\mathbf{y}}_k - \mathbf{y}_k)] \quad (2.10)$$

subject to

$$\hat{\mathbf{y}}_k = \sum_{h=k-\kappa}^k \mathbf{A}_k^h \mathbf{x}_h \quad (2.11)$$

This formulation of objective function provides estimation of OD demands for all departure time intervals ($[\hat{\mathbf{x}}_1 \dots \hat{\mathbf{x}}_K]$) in a single step by processing observed traffic condition data over all study period; and is referred in literature as *simultaneous approach*. The sequential formulation of GLS objective function, where OD demand is estimated for one departure time interval, can be formulated as

$$\hat{\mathbf{x}}_k = \underset{\mathbf{x}_k \geq 0}{\operatorname{argmin}} [(\mathbf{x}_k - \tilde{\mathbf{x}}_k)^T V_k^{-1} (\mathbf{x}_k - \tilde{\mathbf{x}}_k)] + [(\hat{\mathbf{y}}_k - \mathbf{y}_k)^T W_k^{-1} (\hat{\mathbf{y}}_k - \mathbf{y}_k)] \quad (2.12)$$

subject to

$$\hat{\mathbf{y}}_k = \sum_{h=k-\kappa}^{k-1} \mathbf{A}_k^h \hat{\mathbf{x}}_h + \mathbf{A}_k^k \mathbf{x}_k \quad (2.13)$$

Although there are some remarks in literature that discuss how to derive dispersion matrices V_t and W_t , it is often assumed that the matrices are diagonal. This implies that there is no covariance between error components and the diagonal elements of dispersion matrices are the variances. i.e., $\operatorname{Var}(\epsilon_x)$ and $\operatorname{Var}(\epsilon_y)$. Hence, the GLS estimator has the form of a weighted Euclidean norm, where the elements of the dispersion matrices V_t and W_t are weights selected to indicate the degree of confidence in the available data (that is, in prior OD matrix and traffic condition data, respectively). Note that the weighted GLS formulation can be adopted to combine the two sets of deviations, with respective weights α and $(1 - \alpha)$ for the first and second objectives as discussed in Section 2.1.1.

Having a closer look at the GLS estimation given in equations (2.10) and (2.12), the optimal solution is dominated by the representatives of the prior OD demand matrix. For example, the structure of the dynamic OD demand may change substantially compared to the prior OD demand (e.g, in case of irregular events, bad weather conditions). Difficulties arise when the prior OD demand matrix produces a traffic state different from the actual traffic state. Using only link traffic counts in vector \mathbf{y} and prior OD demand is not sufficient to determine optimal OD demand, specially in congested networks, as we discussed in Section 2.3.2. It follows that using traffic condition data, such as speed or density observations, can improve estimation accuracy. Thus, the simple modification of GLS estimator in equations (2.10) and (2.12) requires normalization of objective function.

In this section the solution frameworks for simultaneous GLS problem formulation for dynamic OD demand estimation are discussed. In case a given solution approach is proposed for sequential problem formulation this will be explicitly stated. Note that most of sequential GLS problem formulations for estimating dynamic OD demand in literature have resulted in equivalent state-space formulations; these are presented in the state space model formulation section.

Solution frameworks for GLS objective function

Different solution frameworks and algorithms have been developed to solve the dynamic OD demand estimation problem defined by the GLS objective function in equations (2.10) and (2.12). The issue of modeling the relationship between OD demand and observed traffic data (e.g., link traffic counts), given by equations (2.11) and (2.13), and its use in the estimation process, has led to development of multiple solution directions. First, the section gives an overview of a solution framework where assignment matrices have been determined exogenously, followed by solution methods that deal with endogenously-derived assignment matrices. In addition, alternative solution frameworks are presented, that do not rely on derivation of assignment matrices at all.

Solution framework when link costs are known

The GLS objective function in equation (2.10) for uncongested networks or congested networks for which link costs are known, is solved as a *quadratic optimization problem* using standard gradient methods, Cremer & Keller (1987), Bell (1991), Cascetta et al. (1993). These solution algorithms assume linear formulation between OD flows and link traffic counts and elements of assignment matrix, A_k^h , are exogenously computed using a path choice model (either deterministic or stochastic). The most sophisticated solution algorithm for least square problems, the LSQR algorithm (see Paige & Saunders (1982)), has been proposed by Bierlaire & Crittin (2004). The LSQR algorithm is analytically equivalent to a conjugate gradient method, but exhibits better numerical properties, especially when A is ill-conditioned. Since assignment matrices are large and sparse, a key property of the LSQR algorithm is that the matrix A is used only to compute products of the form Ax or $A^T y$, which is attractive for large sparse problems. Based on transponder-tag data collected from a freeway corridor in Houston, Dixon & Rilett (2002) applied the framework developed by Cascetta et al. (1993) to exogenously calculate the link flow proportions based on the observed travel time from AVI counts. It should be noted that AVI data only provide OD demand distribution information, so OD demand volume information from link traffic counts and historical OD matrices must be added to identify a unique solution.

Despite the authors' significant contributions to the estimation of dynamic OD demand through GLS linear model formulation, there is still a serious problem with linear formulation between OD flows and link traffic counts. In other words, in their formulation, like many other previously published works, the dependence of link-flow proportions on the OD demand is not explicitly included in the solution procedure. This dependency and non-linearity in link-flow proportions can be significant particularly in congested networks.

Bi-level solution framework

In the literature, the GLS formulated OD demand estimation problem for congested networks has typically been formulated as a *bi-level optimization problem* (Tavana (2001), Van der Zijpp & Lindveld (2001), Zhou et al. (2003), Lindveld et al. (2003),

Zhou & Mahmassani (2006)). Tavana (2001) were the first to endogenously formulate a non-linear relationship between dynamic OD flows and link traffic counts under congested conditions. The upper level represents an optimization step, in which an auxiliary OD demand solution is obtained through the optimization of a GLS objective function (equation 2.10). The lower level represents an assignment step, which uses OD demand estimated in upper level, $\hat{\mathbf{x}}$, to generate dynamic assignment matrix or traffic variables is given as:

$$\mathbf{y}_k(x_k) = \underset{l \in \hat{L}}{\operatorname{argmin}} \sum \int_0^{y_l} s_{l,k}(x) dx \quad (2.14)$$

subject to

$$\mathbf{y}_k = \sum_{h=k-\kappa}^k \mathbf{A}_k^h \mathbf{x}_h \quad (2.15)$$

where $s_{l,k}(\cdot)$ is a cost function which defines the delay depending on the flow for the link $l \in L$. If time is discretized the integration over time can be approximated by summation over observation intervals k . The function $s_{l,k}(\cdot)$ expresses the general objective function that can be formulated as dynamic user equilibrium (DUE) or stochastic user equilibrium (SUE). The solution algorithm iterates between the upper and lower levels. The new OD matrix is obtained by moving from the current solution in a search direction, which is defined by the difference between the auxiliary solution and that of the previous iteration step. The step size in this search direction may be pre-defined by exact or approximate line search methods, to minimize the value of the performance function along the descent direction.

Tavana (2001) proposed heuristic solution approach for solving the general bi-level framework, which uses a sensitivity analysis-based (SAB) algorithm to solve a generalized least-squares problem in the upper level. Based on the SAB algorithm, the changes of link flow proportions due to the adjustments in dynamic OD flows are explicitly considered, and numerical derivatives of link flow proportions with respect to OD flows are obtained from a mesoscopic DTA simulation program. However, the SAB algorithm has revealed a critical shortcoming when there is a significant difference between the prior OD matrix and the true OD matrix. This problem stems from the heavy dependence of the SAB algorithm on historical OD information. Such dependence may lead to a state in which the OD estimation cannot produce a correct solution, especially when travel patterns are dramatically changed. In addition, the SAB algorithm needs to approximate the derivatives through simulation for each OD pair and each time interval in every iteration, which is computationally intensive, especially for large-scale networks.

Single-level solution framework

To avoid the bi-level formulation of the dynamic OD demand estimation problem, Sherali & Park (2001) proposed a constrained least squares model with time-dependent path flows as decision variables (i.e., flows on the path connecting OD pair) instead

of OD flows based on observed link traffic counts. In literature, these methods are referred to as Path flow estimators (PFE). Final OD flow estimates can be obtained by summing up path flows for each OD pair. As such, a bi-level formulation is avoided in which the equilibrium conditions would implicitly be accounted for by the constraints. This is done by adding a term that penalizes the path flows that experience a high cost, leading to a single-level formulation where the equilibrium conditions are added as soft constraints. The developed algorithm uses a decomposition scheme that employs a simultaneous least square model, along with dynamic shortest path sub-problems, in order to generate additional path information as needed to solve the problem. The constrained least squares model is solved using a projected conjugate gradient method, while the sub-problem is a dynamic shortest path problem on an expanded time-space network (STEN). Unfortunately, convergence cannot be guaranteed.

Along the same lines, Nie et al. (2005) formulated the OD demand estimation problem as a variational inequality (VI) to take the response of travelers into account in the objective function in the case of deterministic user equilibrium (DUE). This formulation is achieved by balancing the path cost and the path deviation (which measures the deviation between measured and simulated traffic flows), weighted by a dispersion parameter. The dispersion parameter determines to which extent DUE is satisfied. Such methods become quite complex if the incorporated DNL model needs to capture congestion feedback, complex node intersecting, path set generation, etc. The latest work of Nie et al. (2013), utilize the mesoscopic DNL model (i.e., Newell's simplified KW model) to derive analytical, local gradients of different traffic observation, such as link flow, density and travel time, with respect to path flow. By dualizing the difficult DUE constraint into the GLS objective function, the authors have proposed an effective Lagrangian relaxation-based solution framework. The proposed PFE framework is highly dependent on the effect of multiple path flow solutions corresponding to the same equilibrium link flow pattern. Further tests and case studies using larger and real-size networks are required with these PFE based methods for better assessment.

Joint solution framework

In contrast to the bi-level and single-level solution frameworks of a GLS objective function, Balakrishna et al. (2007) proposed solution framework that does not rely on the calculation of assignment matrices. Rather, the complex relationship between the OD flows and traffic observations is captured directly, by treating the assignment model (e.g., any DTA model) as a black box. The proposed approach allows estimation of dynamic OD demand solving the GLS objective function given by equation (2.10), that is, using an additional set of traffic observations. This framework represents a non-linear, non-analytical optimization problem, due to the use of a sophisticated simulator to obtain fitted traffic condition data observations. Instead of assuming a specific analytic relationship between OD flows and for example, speeds, simulation is used to determine this relationship. For example, the speeds at sensors are outputs of running the DTA simulator with a given set of OD inputs, and can be compared against real-world speed measurements. Therefore, there are a number of advantages to this

methodology: while the assignment matrix is only a linear approximation of the interrelation between OD flows and input data (i.e., link traffic counts and traffic condition data), a DTA model captures these relationships directly. Thus, the proposed approach does not require the analytical form of these relationships to be known. The most studied solution algorithm of this problem in literature (Balakrishna (2006), Balakrishna & Koutsopoulos (2008), Cipriani et al. (2011), Frederix et al. (2011)) is the Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm (see Spall & Member (1992)). In the SPSA algorithm, the gradient is found by perturbing all elements in \mathbf{x} simultaneously. Thus, it requires just two performance function evaluations per iteration, irrespective of the size of the estimation problem (i.e., the number of unknown OD flows, d). In contrast, traditional gradient based approaches must evaluate $2 * d$ objective functions per iteration, since the gradient is approximated by perturbing each element in \mathbf{x} . There are different methods for defining a step size (see Cipriani et al. (2011)). While the algorithm's simplicity is appealing, it still requires extensive computation time, even for the medium size network, for example, 15 hours to estimate an OD demand with four departure time intervals for the Calgary network (734 links, 221 nodes and 77 centroids, see Cipriani et al. (2011)).

Solution framework based on meta-heuristic approaches

In addition, several authors have proposed meta-heuristic solution approaches that do not rely on the assignment matrix at all. These methods include advanced evolutionary algorithms (EA) (Kattan & Abdulhai (2006), Park & Schneeberger (2003)) and simulated annealing (Stathopoulos & Tsekeris (2004)). Solving the GLS objective function (equation 2.10) with an EA based framework eliminates the need for the iterative process (Kattan & Abdulhai (2006)). In other words, the EA approach computes the exact equilibrium flow corresponding to each candidate OD demand, simultaneously satisfying both levels of the bi-level and single-level formulation. An important advantage of these methods is that they are generally able to find global and not only local optima. However, they usually require a large number of performance function evaluations, which, in the context of dynamic OD demand estimation for real-time applications, can be computationally expensive.

State space formulation

The most widely utilized sequential GLS problem formulation applied in real-time dynamic OD demand estimation problems is the state space model and its Kalman Filter (KF) solution algorithm (and/or its variations). In this section, it is shown that the sequential GLS objective function presented in equation (2.12) can be viewed, under most circumstances, as a state-space based formulation. Here, the connection between the Kalman Filter and least square estimation theory is pointed out. More precisely, the application of results obtained by classical generalized least squares (GLS) to discrete stochastic linear processes leads to the Kalman Filter (Genin (1970)). The main advantage of Kalman Filters that makes them popular for real-time applications is that

the computational and memory cost can be lower compared to simultaneous solution approaches.

Okutani & Stephanedes (1984) were one of the first to formulated the dynamic OD estimation problem in state space form in such a way that it can be solved by a Kalman filter:

$$\mathbf{x}_{k+1} = \sum_{q=k-\tau}^k \Phi_q \mathbf{x}_q + \mathbf{w}_k \quad (2.16)$$

$$\mathbf{y}_k = \sum_{h=k-\kappa}^k \hat{\mathbf{A}}_h^k \mathbf{x}_h + \mathbf{v}_k \quad (2.17)$$

The process equation (2.16) represents an autoregressive process of order τ , with coefficients $f_{d,q}$. Note that if the noise \mathbf{w}_k is identical to zero and if Φ_q is the identity matrix, then the state is a constant for all k and one has returned basically to the least square method. The observation equation (2.17) represents relation between traffic observations and OD flows. Equations (2.16) and (2.17) constitute a linear state space system of equations that can be solved using the Kalman filter algorithm (Kalman (1960)). Although this formulation encapsulates the dynamics, there are still a number of serious problems with this linear formulation.

Solution frameworks for state space problem formulation

Solution framework regarding assumptions on OD flows

The first problem relates to the typical Kalman filter assumptions on the noise terms \mathbf{w}_k and \mathbf{v}_k in (2.16) and (2.17). Arrival processes (vehicles entering the network) are typical Poisson processes, with exponentially distributed headways. Moreover, both demand and link flows are by definition positive. Consequently, Gaussian assumptions on these noise terms make no sense, because this would allow negative OD flows and link flows, particularly for smaller average values. Ashok et al. (1993) have revised Okutani's work where the state vector is defined as the difference between the OD flow \mathbf{x}_k and average prior OD flow $\tilde{\mathbf{x}}_k$. Instead of predicting the actual number of trips between origins and destinations as in process equation (2.16), the proposed idea is to predict how much these trips deviate from a certain prior OD pattern. There are two reasons for doing this. First, by defining the state variables as deviations from certain historical dynamic OD demand, the state estimation errors are now more amenable to approximation by normal distribution (property of the Kalman filter). Secondly, this adaptation leads to a more valid process model. For example, series of prior OD matrices $\tilde{\mathbf{x}}_k$ have been estimated from historical data for several previous days or months. Hence, Ashok et al. (1993) formulated deviations of OD demand from historical averages as an autoregressive process. In general, historical averages reflect regular OD demand patterns, but prior OD demand subsume a wealth of information about the structural variations over both space and time (think of daily peak periods and the differences between Tuesdays and Saturdays). If the prevailing OD demand is structurally

different from the regular (average) demand pattern, demand deviations will not satisfy the fundamental stationary assumptions for an AR process and could degrade the prediction performance given by process equation (2.16). Therefore, by considering demand deviations from the a priori estimate of the regular pattern as a dynamic process with smooth trend, a polynomial trend filter in process equation (2.16) is proposed by Zhou & Mahmassani (2007) to capture structural deviations in prevailing OD demand. Furthermore, they proposed an adaptive day-to-day updating formulation based on the Kalman filter framework to update the a priori estimate, using the new demand estimate and new observations. The updating procedure aims to adaptively recognize and capture the systematic day-to-day evolution, and also maintain robustness under disruptions. Since the definition of process equation (2.16) refers to dynamic OD demand prediction, this issue will be revisited later in Section 2.4.

Solution framework regarding assumptions in congested networks

The second problem relates to the observation equation (2.17). Particularly in congested networks the assumption of linearity is highly debatable, due to traffic flow dynamics and queuing, and to the effect of these dynamics on route choice as discussed in Section 2.3.2. The consequence is that traffic condition data (speeds, travel times) and associated observation equations are needed to solve the problem for congested networks. Also, such a non-linear dynamic system has to be solved with modified Kalman filter solution approach. The most straight forward extension is the Extended Kalman Filter (EKF), which involves a first order Taylor linearization of the observation equation (2.17) about the best available estimate of the state vector. The EKF solution approach requires a computationally intensive linearization step and is sensitive to the quality of the initial estimates of the state or the process model. An improvement to the extended Kalman filter led to the development of the Unscented Kalman filter (UKF) and Iterated Extended Kalman Filter (IEKF). In the UKF, the probability density is approximated by a non-linear transformation of a random variable, leading to more accurate results than the first-order Taylor expansion of the non-linear functions in the EKF. The Iterated EKF method attempts to improve upon EKF, by using the current estimate of the state vector to linearize the measurement equation in an iterative mode. Details on KF solution algorithms and their variations can be found in many textbooks (e.g., Grewal & Andrews (2008), Anderson & Moore (1979)).

Chang & Wu (1994) were first to attempt to formulate non-linear relationship between OD distributions and link traffic counts under congested conditions. Their proposed model employs information from link traffic counts collected on on-ramps, off-ramps and mainlines, and macroscopic traffic characteristics to construct a set of non-linear observation equations. The set of non-linear equations, where the state variables are assumed to follow a random walk process in time, is solved by the EKF. Chang & Tao (1996) have extended the work of Chang & Wu (1994) to generalized networks. They introduced the concept of additional link traffic counts at cordon lines to increase the state observability. A cordon line is defined as a closed curve that intersects with a set of links, and subdivides the network into inside and outside parts. An additional set of

observation equations is formulated and the state space model was solved by EKF.

Solution framework regarding assumptions on traffic condition data

A weakness of these two models is that observation equations 2.17 with assignment fractions derived from stochastic assignment are specified only for the initial time interval, but not for subsequent intervals. To encapsulate the dynamic evolution of the assignment matrix and its uncertainty, Ashok & Ben-Akiva (2002) extend their previous framework (Ashok & Ben-Akiva (2000)) with additional process and observation equations. The state vector is augmented by the travel times and route choice fractions to explicitly model and estimate the stochastic assignment matrix, $\hat{\mathbf{A}}_h^k$. The non-linear equation system is solved by EKF algorithm. Along these lines, Hu & Chen (2004) augment the state vector of unknown OD flows with link travel times, and the observation equation with observed link speeds. An important difference with respect to the approach of Ashok & Ben-Akiva (2002) is the use of link speeds without route choice fractions. Good knowledge of the underlying process that describes the temporal evolution of speeds provides easier definition and calibration of the process equation.

Dixon & Rilett (2002) estimated OD flows by using AVI tag counts and OD split proportions derived from collected AVI data. In this method, the AVI penetration rate is assumed to be randomly distributed, with the AVI data collection points acting as the origins and the destinations, and AVI tag data provide OD split proportions that are used in addition to traffic counts as measurements in observation equation (2.17) in addition to traffic counts. The solution algorithm is a constrained Kalman filter on the assumption of a random walk process equation (2.16), where the non-negativity constraints are defined by Lagrangian multipliers (Bell (1991)) and added to state vector. This model, however, is limiting in terms of its applicability of the model. Antoniou et al. (2006) estimated dynamic OD demand by incorporating direct measurements of OD flows from AVI systems or probe vehicles, in addition to traffic counts. Since the direct observation of OD flows requires the location of AVI stations close to origin and destination points, Antoniou et al. (2006) relax this assumption to formulate indirect observation equation by extracting "sub-paths" from several connected AVI detector locations. They proposed several solution approaches, ranging from EKF to UKF. Further, Sohn & Kim (2008) used cell-based trajectories as probe phones to estimate the average path choice proportions. Unlike a previous direct approaches that used sample OD flows extracted from the cell-based location data as additional observations, they proposed an indirect approach wherein the assignment matrix is derived from passing time at observation locations, and the path choice proportions. The estimation of dynamic OD flows by cell-phone data detected by Bluetooth system was also studied by Barcelo et al. (2010) and Barcelo (2010). The tracking of vehicles is assumed by processing Bluetooth and WiFi signals, and the distribution of travel times between OD pairs or between each on-ramp and sensor locations are directly observable and are no longer state variables but measurements in the observation equation. A set of linear state and observation equations is solved by Kalman filter algorithm.

Maximum entropy / minimum information formulation

The theory of maximum entropy (EM) for dynamic OD matrix estimation follows the rational that the a priori probability function $p(\mathbf{x}_k|\tilde{\mathbf{x}}_k)$ is assumed to be the Poisson distribution function with parameters $\tilde{x}_{i,k}$ for each OD pair i , $i \in \Omega$. Then, the simultaneous EM generic dynamic OD demand formulation given in equation (2.4) becomes:

$$\lambda \sum_{k \in K} \mathbf{x}_k \left(\log \frac{\mathbf{x}_k}{\tilde{\mathbf{x}}_k} - 1 \right) \quad (2.18)$$

subject to

$$\hat{\mathbf{y}}_k = \sum_{h=k-\kappa}^k \mathbf{A}_k^h \mathbf{x}_h \quad (2.19)$$

where the vector λ , denotes relevant weights that allow for some prior OD matrix elements to be assumed more reliable than others.

Solution frameworks for maximum entropy objective function

The solution of an EM objective function can be derived with the following structure (Wu (1997)):

$$\hat{\mathbf{x}}_k = \tilde{\mathbf{x}}_k \prod_{i=1}^n [\lambda_i^{\mathbf{A}_k^i}] \quad (2.20)$$

where λ_i , $i \in \Omega$, are positive Lagrangian multiplier parameters. Equations (2.18) and (2.20) constitute an equation system that captures the optimal solution of the EM model.

Despite to the popularity of the GLS objective function and state space formulation, the formulation of dynamic OD demand estimation problem as an EM objective function and its solution have been studied by several authors (Wu (1997), Tsekeris (2003), Yamamoto et al. (2009)). The common assumption in proposed solution approaches is that the assignment matrices are derived exogenously, assuming a linear relationship between link traffic counts and OD flows. In addition, Yamamoto et al. (2009) used observed link travel speeds from probe vehicles to estimate the assignment matrices and link flows. Wu (1997) proposed two iterative algorithms to solve the EM objective function, namely the Multiplicative Algebraic Reconstruction Technique (MART) algorithm and Revised MART (RMART) algorithm. Although the RMART algorithm is designed to accelerate the convergence speed, it is not proved for conservation of the convergence. Therefore, Tsekeris (2003) proposed the double iterative matrix adjustment procedure (DIMAP), in the sense that the multiproportionally-adjusted matrix in each iteration is one that has already been corrected on the basis of the MART algorithm. Although algorithm convergence is ensured, its on-line application is very limited due to the extensive computational requirements.

Discussion

The problem formulations are distinguished based on probability assumptions on OD flows and whether traffic condition data are included in addition to link flow data. The problem formulation can have a large influence on the problem complexity. An example of this is that in the case of simultaneous problem formulations, the problem contains far more unknown variables to be estimated. When traffic condition data are included in the problem formulation, the problem complexity additionally increases since more unknowns and constraints are included in formulation.

The main difference between the solution frameworks presented above is how the relationship between OD demand and observed traffic data (e.g., link flow data, traffic condition data) is modeled and used in the problem formulation. Thus, the solution frameworks differ in whether the relationship between OD demand and observed traffic data is given by linear model or non-linear model. Once a solution framework is defined, different solution algorithms have been proposed that result in a global optimum or an approximate solution.

Many authors give attention to the performance of solution algorithms in terms of real-time applications and computational efficiency. However, the presented results are case-study-dependent and do not contain sufficient information to enable a fair comparison, presenting a limit to the overview included here.

2.4 Dynamic OD demand prediction: State-of-the-art

Unlike the substantial research efforts devoted to dynamic OD demand estimation, dynamic OD demand prediction problems have been formulated by few authors. Existing OD demand prediction methods can be categorized according to the underlying assumptions in representing dynamic OD demand process.

In general, prediction OD demand methods rely on time series prediction models, so-called ARIMA (AutoRegressive Integrated Moving Average) models and their derivatives. The main building blocks of these models are Autoregressive models of order q (AR(q)) and Moving Average of order s (MA(s)) models, while integrated c times (I(c)) holds for integrated application of previous models to remove a non-stationary in data.

The most straightforward short-term dynamic OD demand prediction model for incorporating the state estimate, $\hat{\mathbf{x}}_k$, for the previous time interval can be expressed as a random walk model corresponding to an AR(1) model with autocorrelation coefficient of 1. Cremer & Keller (1987) and Chang & Wu (1994) applied the random walk model, given as:

$$\mathbf{x}_{k+1|k} = \hat{\mathbf{x}}_{k|k} + w_{k+1} \quad (2.21)$$

to predict dynamic OD flows, by directly extending the latest state estimates as the future predictions. Equation (2.21) corresponds to the sequential generic formulation of dynamic OD demand estimation method given in equation (2.12). This simple prediction model does not explore the information in prior OD demand. In order to capture the effect of state estimates over lagged time intervals, Okutani & Stephanedes (1984) proposed an autoregressive process of order q , $AR(q)$ model, expressed as

$$\mathbf{x}_{k+1|k} = \sum_{q=k-\tau}^k F_k^q \hat{\mathbf{x}}_{k|q} + w_{k+1} \quad (2.22)$$

to predict dynamic OD flows at time interval $k+1$ by its previous q estimates: $\hat{\mathbf{x}}_k, \dots, \hat{\mathbf{x}}_{k-\tau}$. The matrix F_k^q denotes the effects of state estimates $\hat{\mathbf{x}}_{k|q}$ on future OD demand state $\mathbf{x}_{k+1|k}$. The elements of matrix F_k^q can be derived from the optimization process such as ordinary least square (OLS) procedure on prior OD demand information.

The autoregressive processes given in equations (2.21) and (2.22) can only capture temporal interdependencies among OD flows. However, the OD demand is a function of spatial and temporal distribution of activities and characteristics of a transportation system. Such information is present in prior OD demand information obtained from the off-line OD estimation results or day-to-day demand, updating using previous real-time demand estimates. One simple way then of incorporating structural relationships is to include all the prior OD demand information into the OD demand prediction model.

Ashok & Ben-Akiva (2002) formulated an OD demand prediction model using the deviations of OD demand from historical averages, $\bar{\mathbf{x}}_{k|k}$, instead of the estimated OD flows from previous intervals. Thus, the proposed OD demand prediction model represents the temporal evolution among deviations in OD flows by autoregressive process (AR), in which a fourth-order autoregressive model based on calibration results from several historical OD data sets is adopted. The OD demand prediction model then can be expressed by the following equation:

$$\mathbf{x}_{k+1|k} = \sum_{q=k-\tau+1}^k F_k^q (\hat{\mathbf{x}}_{k|q} - \bar{\mathbf{x}}_{k|q}) + w_{k+1} \quad (2.23)$$

where elements of matrix F_k^q are predetermined autoregression coefficients, w_{k+1} is system evolution noise, and τ is the order of autoregressive process.

An autoregressive model is suitable to describe a stationary random process with constant mean and variance. However, if prevailing OD demand were structurally different from the historical OD demand pattern, overall prediction of OD demand would be seriously degraded. Also, autoregressive models with higher order terms require extensive off-line calibration efforts to find the autocorrelation coefficients of matrix F_k^q . This will affect the computation time for on-line applications, especially for large-scale networks. In order to overcome the drawbacks of an autoregressive model, Zhou

& Mahmassani (2007) proposed the third-order polynomial model to represent temporal evolution of structural deviations. To obtain the future OD demand with prediction horizon $k + 1$, the prediction model is given as:

$$\mathbf{x}_{k+1|k} = \bar{\mathbf{x}}_{k+1|k} + M_k(\hat{\mathbf{x}}_{k|k} - \bar{\mathbf{x}}_{k|k}) + w_k \quad (2.24)$$

where one needs to first predict the demand deviation at time $k + 1$ based on estimated derivatives at the current time interval k , and then substitute the predicted demand deviation, $M_k(\hat{\mathbf{x}}_{k|k} - \bar{\mathbf{x}}_{k|k})$, and the historical OD demand for prediction horizon $k + 1$ into equation (2.24). Refer to the work of Zhou & Mahmassani (2007) for a more details on the derivation of the third-order polynomial filter.

Discussion

Previously presented dynamic OD prediction models emphasized the importance of incorporating the historical OD demand information, in terms of estimation of coefficients for autoregressive process or in terms of regular pattern, to ensure prediction accuracy and robustness. However, the dimensionality of the state vector imposes limitations on higher-order terms in an autoregression model and polynomial function. The reduction of the model order substantially decreases the size of the state vector, and therefore improves computational efficiency (Zhou & Mahmassani (2007)).

2.5 Conclusions

The importance of considering traffic as a typical dynamic process has become clear and during the last three decades, during which a variety of dynamic OD demand estimation and prediction models have been developed, as presented in this chapter. In the specific context of Intelligent Transportation Systems (ITS), the dynamic nature of the OD demand problem and the real-time requirements make the estimation and prediction of OD demand even more intricate. Thus, the estimation and prediction of OD demand has become an important element of Dynamic Traffic Management Systems (DTMS). These systems impose several criteria that must be captured:

1. The dynamic nature of the process must be captured in the modeling framework.
2. Input data are (easily) observable and available in real-time.
3. Relevant phenomena, especially in congested networks are reproduced and predicted by the model
4. The model estimates and predicts OD flows; as time proceeds and more data becomes available, the solution must be updated to reflect the evolution of the OD demand and network conditions.

5. The model allows efficient computation for different scales of network.

This chapter has shown that the dynamic OD demand estimation and prediction problem is very complex and formulations proposed in literature differ in multiple ways, both in functionality and problem complexity. For example, the relationship between OD flows and traffic data observations can be described by flow-independent functions, or detailed traffic flow models can be used. Differences also relate to the type of input data used in problem formulation. For example, link traffic counts, or traffic condition data in combination with link traffic counts can be used. The final difference is in the application possibilities. For example, sequential formulations of an OD estimation problem are better suited for real-time applications than the corresponding simultaneous formulations, due to computational and time constraints.

Studies often mentioned the importance of DTMS requirements to be satisfied, yet they are hardly ever incorporated or demonstrated in the proposed formulations. The question of trade-off between accuracy and efficient solution approaches is often pointed out. However, the importance of accuracy of OD demand estimates, especially in congested networks, has received a lot of attention despite adding more complexity in proposed methods. The overview has shown the limited attention to OD demand prediction methods that are very important input for DTMS. Therefore, the efficient solution approaches on dynamic OD demand estimation and prediction for dynamic traffic management will be dealt with in the following chapters.

Chapter 3

Methodology for benchmarking dynamic OD demand estimation methods

One of the key traffic variables required for both ex-post and ex-ante evaluation of traffic management and policy measures are dynamic OD demand matrices. However, a major contradiction is that quite frequently these evaluations have as main input very rough and low quality information on the dynamic OD demand. Consequently, errors in OD demand have an unintended consequences on subsequent stages of traffic state analysis. This state has fostered the interest in the assessment and comparison of various dynamic OD demand estimation methods under different circumstances (e.g. different network structures and different sets of data available in different qualities).

In this chapter such an assessment methodology is proposed, based on the Latin Hypercube (LHC) method, which is particularly suited for high-dimensional estimation problems. The objective of the benchmark methodology is not to conclude that one approach is the "best", but to provide support for comparison in a variety of settings and conditions. With this benchmark methodology one can, for example, perform sensitivity analysis on single or multiple OD estimation methods. The methodology is demonstrated on a real urban corridor network for a well-known OD estimation method (the entropy maximization estimation method) to illustrate which results can be obtained and how these can be used to benchmark different OD estimation methods.

This chapter is an edited version of the article:

Djukic, T., J. W. van Lint, S. Hoogendoorn, An Efficient Methodology For Benchmarking Dynamic OD Demand Estimation Methods. Transportation Research Record: Journal of the Transportation Research Board 2263(1): 35-44, 2011.

3.1 Introduction

The problem of estimating dynamic OD demand from various traffic data observations has been widely studied in the literature as presented in Chapter 2. Dynamic OD demand is important input to many models applied within DTMS for predicting traffic state on network. However, a major contradiction is that quite frequently these sophisticated models have as main input very rough and low quality information on the temporal variability of trip patterns as described by the dynamic OD demand. Consequently, errors in OD demand have an unintended consequences on subsequent stages of traffic state analysis. This results in situations in which it is hard for the analyst to identify whether flaws in the intended model are due to modeling mistakes, an improperly calibrated model or an unsuitable specification of the dynamic OD demand. This state has fostered the interest in the evaluation and cross-comparison of various dynamic OD demand estimation methods under different circumstances (e.g. different network structures and different sets of data available in different qualities). From the literature, however, three key difficulties in the evaluation and cross-comparison of various dynamic OD matrix estimation methods can be identified.

The first difficulty is that there is no undisputed performance indicator for estimated and predicted dynamic OD matrices. Instead, there are many candidate statistical metrics (e.g., RMSE, MSE, NMSE, MAE), which evaluate quality of estimated OD matrices with respect to available ground truth OD matrices (e.g., Marzano et al. (2009)). The consequence of this wide range of statistical metrics is that researchers and practitioners may use different metrics according to their needs, rather than as objective assessment criteria. Few studies have focused on evaluation of uncertainty and quality of the estimated dynamic OD matrices in absence of the ground truth OD matrices (Bierlaire (2002) and Yang et al. (1991)). For example, Bierlaire has proposed total demand scale measure, that evaluates the level of arbitrariness introduced by the prior OD matrix in estimation process. This measure is not depended on the prior OD matrix, neither on the traffic data observations. However, this performance indicator should be used in addition to, not instead of, standard statistical metrics.

A second reason why generalization of assessment of OD estimation methods is difficult is the many aspects in which these methods differ, as discussed in Chapter 2. Since these methods differ in so many aspects, it is difficult to make a priori statements as to which dynamic OD demand estimation method is most suitable for a particular traffic network problem or model employed by DTMS. Even when a proposed approach is compared with alternative approaches, it can be expected that due to various reason, such as familiarity with alternatives and selection of suboptimal parameter values, the comparison might be not completely fair and informative.

A third problem in comparison of dynamic OD matrix estimation methods is the lack of consistence in the experimental design and presented results. Each researcher or developer tests their algorithms and methods under different assumptions, with different networks and traffic conditions, using different traffic data and performance indicators.

Usually, demand and traffic condition data characteristics are generated by random and structural perturbations of ground truth data to reflect various traffic situations on network. In most performance evaluations, for each scenario, one or multiple runs are conducted. However, the focus is on the average response of a dynamic OD estimation method on the basis of simulated OD demand and traffic data characteristics, rather than an OD demand distribution. This leads to the main subject of this chapter.

A structured and generic methodology for benchmarking different OD estimation methods under different circumstances (e.g. network lay-out, data availability and quality of prior OD) would enable researchers to pinpoint the strengths and weaknesses of different OD estimation methods and their applicability and validity under different circumstances. For example, there may be clear limits to the type and nature of errors in available data for which a particular OD estimation method would still provide adequate estimation results. Likewise, some OD demand estimation methods may be fundamentally inappropriate for reliably estimating OD demands for certain network topologies. For practitioners, such a benchmarking tool would provide a way to assess the quality of the estimated OD matrix, and construct confidence bounds around these. This will be very helpful in calibrating and validating simulation models using OD demand estimations.

This chapter presents and demonstrates such a generic framework for benchmarking dynamic OD estimation methods. It is generic in the sense that a wide range of OD estimation approaches can be tested under a wide range of different circumstances, related to, for example, data availability and quality, and network lay out. One of the central components of framework is an efficient Monte Carlo sampling method, the so called Latin Hypercube (LHC) method, used to generate series of inputs for cross comparison. The framework considers ground truth dynamic OD demand as a benchmark against which quality of OD demand estimates from various dynamic OD estimation methods can be evaluated. In addition, proposed framework provides opportunity to generate wide range of performance indicators to get better grasp and insight on performance of considered methods. The objective of the benchmark methodology is not to conclude that one approach is the "best", but to provide support for comparison in a variety of settings and conditions. With this benchmark methodology one can, for example, perform sensitivity analysis on single or multiple OD estimation methods.

The chapter introduces the framework for comparison of dynamic OD demand estimation methods and highlights its main features. These include the methods for generating structural and random variations in input data that are relevant for assessing different OD estimation methods, the computational platform, and the sampling LHC technique used. In the second part of this chapter, the methodology is demonstrated on a real urban corridor network for a well-known OD estimation method (entropy maximization estimation method) to illustrate which results can be obtained and how these can be used to benchmark different dynamic OD estimation methods. The chapter closes with a discussion of the results, conclusions and recommendations on further research and possible application directions.

3.2 The concept of benchmarking framework based on LHC

As reflected by equation (3.1), a particular OD estimation method h with a set of parameters and assumptions H , aims to infer the maximum probable OD matrix given the set of available inputs (data Y , prior and previous OD estimates X_{prior} and X_{t-1}, X_{t-2}, \dots and all the assumptions H used in the method, that is

$$X_t = h(X_{t-1}, X_{t-2}, \dots, X_{prior}, Y, H) \quad (3.1)$$

Typically, h is a highly nonlinear and dynamic function of its inputs. A typical question to be investigated would be, "How does X_t change when the inputs vary according to a certain assumed joint probability distribution?" Related questions are, "What is the expected value of X_t ?" and "What is the 99th percentile of X_t ?" This is even more important in case one wants to compare different OD estimation methods. Since OD estimation methods differ in many aspects (objective function, optimization method, single or bi-level, etc.), and since they may be assessed under many different circumstances (network type, with or without route choice, different traffic conditions, varying data availability, etc.), the number of simulations required for an exhaustive and statistically-sound comparison between different OD estimation methods will quickly become very large.

For example, if one wanted to assess a single OD estimation method in terms of its sensitivity to the availability of data Y (e.g. link traffic counts) in a network with say L links, each either equipped with a detector or not, the total number of data scenarios mounts up to $2L$. Furthermore, if the impact of data quality is to be considered, and one would assume that detectors may produce data, in example, 4 qualities (e.g. zero, 2%, 5% and 10% random errors), the total number of scenarios increases to $(1+4)L$. For a toy network with 15 links, this already corresponds to a ten-digit number. If data from combinations of different sources, each with different characteristics and prior error distributions is additionally considered, different parameter settings within this considered OD estimation method, different network types (e.g. grid networks versus freeway networks), and many of the other attributes which may affect the OD estimation problem, one would clearly require an enormous number of samples to come up with a statistically-sound estimate of the conditional distribution in equation (3.1).

The main purpose of the benchmarking framework is to execute a large number of simulations for varying OD estimation method inputs, reflecting the variability and uncertainty of both traffic demand and traffic data characteristics. Since the framework has been applicable to a whole range of OD estimation approaches, the workings of which are not explicitly considered, a simulation-based approach is deemed most adequate. In doing so, the OD estimator is considered as a black box, providing a certain outcome (OD estimate) given certain input.

The obvious choice for a simulation-based approach is to use some sort of constrained or stratified form of Monte Carlo sampling (Reuven (1981)), which does not scale with input dimensionality. There are various solutions to limit the large number of samples needed without fundamentally changing the overall idea. In this benchmarking framework, the solution strategy is to use so-called stratified sampling techniques such as the Latin hypercube (LHC) method.

3.2.1 The Latin hypercube sampling

The LHC method (McKay et al. (1979)) provides a computationally much more efficient alternative to random (Monte Carlo) sampling for estimating the conditional distribution in (3.1). The LHC is typically used to estimate multidimensional integrals for which a closed form solution is either very difficult to obtain or simply not available, such as many multivariate probability density functions arising from the simulation of physical processes. The LHC method has been used in traffic science before, for example for the estimation of a Mixed Logit model (Hess et al. (2006)) and in the calibration of microscopic traffic simulation software (Park & Schneeberger (2003)).

To illustrate the process of Latin hypercube sampling, a two-dimensional example is given in Figure 3.1.

The figure shows how random points are sampled from both (assumed known) cumulative distribution functions of inputs X_1 and X_2 from $N=5$ strata of equal probability, each encompassing 20% of the probability mass (Figure 3.1 top left and bottom right). The resulting samples are combined in a matrix (a 2-dimensional hypercube), which now contains 5 2-dimensional input samples (Figure 3.1 top right). The LHC method can be applied straightaway in the benchmarking framework given in the next section of this chapter. The requirement for applying this method is that knowledge of a priori distributions for the inputs is considered needed. For example, in case one wants to test OD methods under different degrees of measurement errors, one has to assume a particular distribution for these errors, for example a normal distribution or any other distribution deemed appropriate. This feature allows for comprehensive comparison between different dynamic OD estimation methods with different distribution assumptions on input data that exhibit certain properties and assumptions of these methods. The same holds for errors in the a priori OD matrices or variations in, for example, link capacities used in the assignment procedure for a particular OD estimation method. To keep the focus on the main components of benchmarking framework, the detailed mathematical derivation of LHC method is presented in Appendix A.

In Section 3.3.2, the application of the LHC method will be shown to generate input scenarios in terms of input data quality. In doing so, we will illustrate the specific issues encountered when applying the LHC method to model OD demand patterns and traffic data characteristics.

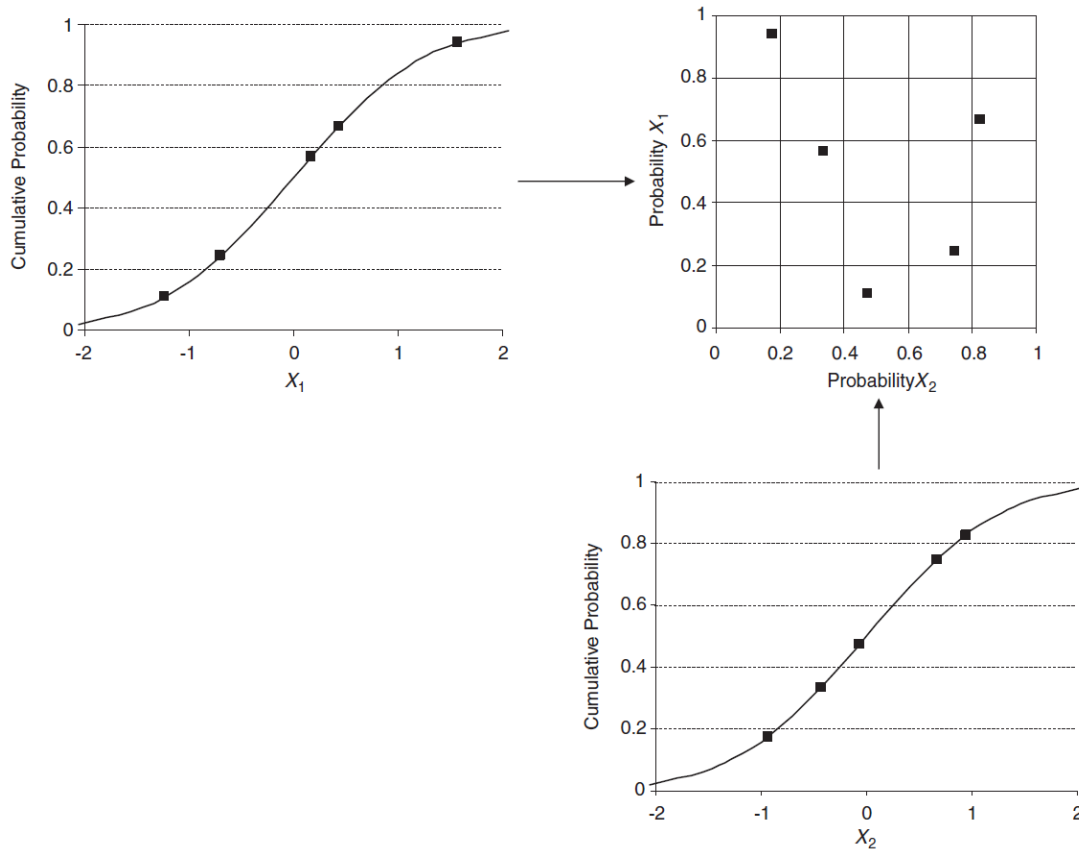


Figure 3.1: Example of LHC method for two variables with a normal distribution (Budiman & Alex (2006)).

3.3 Components in the benchmarking framework

In this section the overall framework will be discussed and some details on its constituting components will be provided. Figure 3.2 schematically outlines the proposed framework.

The benchmarking framework consists of three main components. Component 1 represents an **OD demand and traffic data generator** and essentially determines which input scenarios, varying in terms of network topology, traffic conditions, and data availability, need to be considered during the assessment. The OD demand and traffic data generator provides input to component 2, **computational platform**. The computational platform needs to ensure equal testing conditions for various OD demand estimation methods that would support fair comparison and an understanding of their relative merits. The main elements of the platform are: traffic simulator and dynamic OD demand estimation methods - selection and implementation of a single or multiple OD demand estimation methods to be assessed and compared. Component 3 represents an **output processor**. The final component in the benchmarking framework processes the outcomes of the OD demand estimations and calculates a series of performance indicators. It describes the output required for assessment, cross-comparison and analysis.

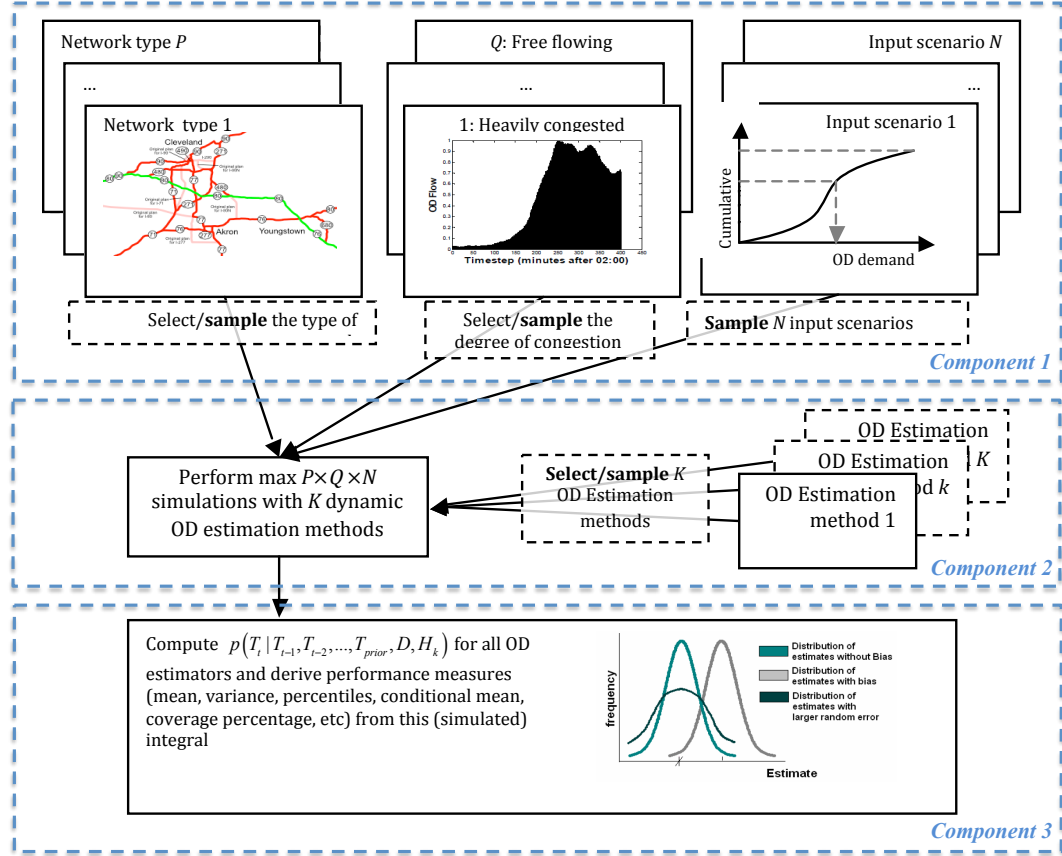


Figure 3.2: Framework for comparison of dynamic OD estimation methods.

In the following subsections more detail on the components and method used to sample the input scenarios will be provided.

3.3.1 Computational platform design and implementation

Computational platform needs to ensure equal testing conditions for various OD demand estimation methods that would support fair comparison and an understanding of their relative merits. The platform consists of two main elements:

- **Traffic simulator:** The traffic simulator must be able to replicate traffic flow dynamics and phenomena on network for given traffic demand. A short list of these simulators includes DYNASMART (Zhou & Mahmassani (2006), Tavara (2001)), Dynamaq (Cipriani et al. (2010)), MitSimLab (Balakrishna et al. (2007)), DynaMIT (Gupta (2003)), and Aimsun (Barcelo et al. (2010)). The selection of traffic simulator for benchmarking study depends on user experience and computational efficiency. For example, the mesoscopic models are substantially faster than the microscopic one. Thus, it allows for more elaborate testing and a richer experimental design.

- **OD demand estimation methods:** This element refers to selection and implementation of a single or multiple OD demand estimation algorithms to be compared.

A dynamic communication between the OD demand estimation methods and traffic simulator is necessary in order to execute a traffic simulation run within the OD demand estimation algorithm. Selected dynamic OD demand estimation methods receive following inputs from component 1, i.e., *demand and supply generator*:

- The demand pattern to be simulated in the form of OD flows per time interval;
- The time series of traffic data, e.g., link traffic counts, speeds, densities and occupancies at detectors;

Then, traffic simulator receives as input a dynamic OD demand generated by the dynamic OD estimation method and runs a new traffic simulation. When the simulation is finished, the platform runs component 3, i.e., *output data processor*, to collect and organize all the outputs.

Figure 3.3 presents a flowchart that shows the main elements of this computational platform. Within the dynamic OD demand estimation main function, whenever a simulation run is needed the traffic simulator is initiated. After the simulation runs have been completed, it imports the observed traffic data and the simulation outputs and runs output data processor.

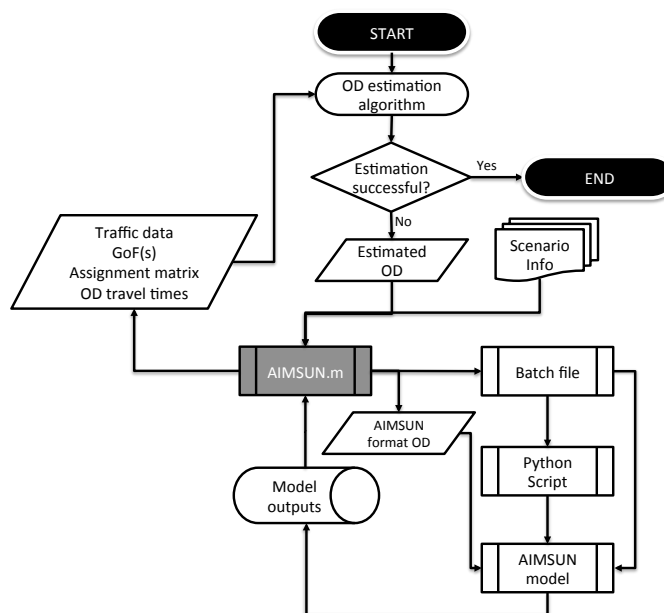


Figure 3.3: Flowchart with the main elements of the computational platform (Antoniou et al. (2014))

3.3.2 OD demand and traffic data generator

The OD demand and traffic data generator provides the input to the computational platform. A key requirement for the task of evaluating an OD demand estimation algorithms, and for comparison of multiple ones, is to test the performance under a range of different conditions and scenarios and to ensure that these conditions are consistent across algorithms. For that purpose, the OD demand and traffic data generator was developed taking into account multiple dimensions, including:

- Prior OD demand characteristics, including varying bias and random errors;
- Traffic data characteristics, including the number and location of sensors and the type of surveillance information, as well as its quality.

On the basis of various sources of systematic and random variations in input data, the stochastic generation of the OD demand and traffic data values proceeds in two steps:

1. First, random realizations of the different sources of variability are generated. For this generation, the LHC method is used. In the LHC simulation, data on the probabilities or frequencies of occurrence of different possible conditions are used. Important interdependences between the different sources of variability are taken into account by using conditional probability specifications.
2. Subsequently, the stochastically generated circumstances are translated into effects on the OD demand and traffic data, by using correction factor, α , as presented in following scenarios. By applying the correction factor on the ground truth values of OD demand and traffic data characteristics, the stochastic realizations of these OD demand and traffic data characteristics are generated.

Prior OD demand simulation scenarios

To estimate the dynamic OD matrix for a specific day and time period t , the prior matrix turns out to be an important source of information. Generally, the dynamic OD prior matrix $\tilde{\mathbf{x}}$ (with elements $\tilde{x}_{i,j,k}$) provides the base OD matrix which is matched and scaled on the basis of additional information (e.g. from loops or other data sources). The value of $\tilde{\mathbf{x}}$ may be determined from different sources of information (e.g. land-use models, travel surveys, or historical OD information) using different methods. In the ideal case, $\tilde{\mathbf{x}}$ reflects an OD matrix which is very close to the actual OD matrix, in the sense that it has a similar structure, for example, in terms of the distribution of trips over destinations, and the trip length distribution. In practice, however, $\tilde{\mathbf{x}}$ may differ substantially from real OD matrices, due to, for example, changed trip patterns or changes in within-day trip patterns.

By implementing the LHC method described in section 3.2.1, the fact that prior OD matrix may contain errors may be simulated. In following prior OD demand scenarios, the different sources of both structural and stochastic daily fluctuations present in within-day travel demands have been modeled.

1. **Scaled (total) demand variation scenario:** This scenario illustrates how the variation of demand flows across the sequence of departure intervals can influence the estimation results without changes in OD matrix pattern (the distribution over destinations in $\tilde{\mathbf{x}}$ is kept similar to the one in the ground truth OD matrix). To this end all demand values within the same departure interval are scaled with a factor α_k drawn from a random uniform distribution over the range $[0.5, 1.5]$, that is

$$\tilde{x}_{i,j,k} = x_{i,j,k} \times \alpha_k \quad \alpha_k \sim U(0.5, 1.5) \quad (3.2)$$

2. **Random demand variation scenario:** This scenario is based on the assumption that the prior OD matrix is the best estimate of the mean of the dynamic OD matrices. Any survey or off-line OD estimation procedure will utilize data from several days, essentially smoothing out any day to day variation present in the OD flows. In this case $\tilde{\mathbf{x}}$ is varied by adding uniformly random components to the ground truth OD matrix, representing the difference between the smoothed historical estimate and the particular daily realization:

$$\tilde{x}_{i,j,k} = x_{i,j,k} \times [0.8 + 0.4 \times \alpha_{i,j,k}] \quad \alpha_{i,j,k} \sim U(0, 1) \quad (3.3)$$

3. **Systematic demand variation scenario:** This scenario addresses situations where the prior OD demand might contain other, structural errors besides the random daily fluctuations. The structural errors might occur due to a structural deviation in the demand pattern resulting from unforeseen events, or from out-of-date surveys resulting in a structurally incorrect estimation of OD flows. The demand per each origin over destinations is generated from positively and negatively skewed mean values of distribution from a random demand scenario:

$$\tilde{x}_{i,j,k} = \gamma_{i,k} \times x_{i,j,k} \times [0.8 + 0.4 \times \alpha_{i,j,k}] \quad \alpha_{i,j,k} \sim U(0, 1) \quad \gamma_{i,k} = \pm 20\% \quad (3.4)$$

Traffic data simulation scenarios

The previous scenarios described realistic variations in the prior matrix. Next to the prior OD, the observed traffic data, \mathbf{y}_k , presents an important source of information. In general, OD estimation methods that use link traffic counts assume the availability of accurate link traffic counts from all links in a study network (Wu & Chang (1996), Hu et al. (2001)). In reality, detector availability is often limited. In recent years,

optimal detector locations aiming at maximizing observability of OD flows and the inherent uncertainty has been considered in several research studies (Gan et al. (2005), Antoniou et al. (2004), Oh & Ritchie (2005)). Depending on the optimization objective when determining detector layout, one may roughly classify the presented approaches into the following two categories:

1. Coverage of detectors with maximum link flows.
2. Coverage of detectors to separate as many OD pairs as possible (here an OD pair is regarded as separated when each of its feasible paths passes at least one of the counting links and hence all trips between any OD pair are observed for at least one link of their path).

For a comprehensive benchmark study it is important to take into account the real state and conditions (in terms of output, errors, location, etc.) of the detectors on the network. Additionally, consideration of the two aforementioned approaches to determining detector layouts can provide information on what kind of improvement can be gained if the availability of traffic counts is increased according to different link detector coverage rules. The discussion will therefore maintain the main idea of both categories without going into detail during the implementation of algorithms proposed in literature. Consequently, the random perturbations of traffic count measurements are assumed using the following two detector layouts:

1. Detector layout based on maximum coverage of all links in network: In this "best" case, the aim is to obtain benchmark results for the comparison of OD estimates with other detector layouts.
2. Detector layout based on the maximum link flows: Link traffic detectors are chosen according to maximum link flow measurements so that links with highest volume have priority. First, the links with highest volume are identified (in the study network these are the links on the main arterial road). Then, the detector layout is defined randomly removing the detectors from links to obtain tree scenarios with 80%, 60%, 40% and 20% coverage of the total number of links on network. This approach uphold the concept that link counts detectors must be located on the network so that a portion of trips between chosen OD pairs are observed passing at least one link on their paths.

A similar approach to capture variations in prior OD matrices was followed in developing scenarios regarding the quality of observed link traffic counts. For each of the prior OD demand levels described above, using the ground truth demand within traffic simulator, the ground truth link traffic counts, y_k , were calculated. These link traffic counts need to be perturbed with noise to mimic observation errors in the real world. In this experimental design, we consider only link traffic counts since selected OD demand estimation method uses this information in estimation process. However,

the proposed scenario can be applied to any other traffic condition data, e.g. densities, travel times, etc. The following scenarios are defined, applied to each observation point (i.e., detector location l and time interval k):

1. **Error-free observations scenario (EF):** This scenario assumes ideal operation of detectors on network which ensures observation of each vehicle that passed detector, that is

$$\mathbf{y}_k^{EF} = \mathbf{y}_k \quad (3.5)$$

2. **Low-error observations scenario (LE):** This scenario is based on the assumption that detectors fail to observe all vehicles. Therefore, the observed link traffic counts are generated for 90% of the ground truth link traffic counts and varied by adding uniformly random components in range of $\pm 10\%$, that is

$$y_{l,k}^{LE} = y_{l,k} \times [0.8 + 0.2 \times \alpha_{l,k}] \quad \alpha_{l,k} \sim U(0, 1) \quad (3.6)$$

3. **High-error observations scenario (HE):** This scenario addresses situations when detectors are more uncertain; e.g. in peak-hours, when congestion occurs on network. The observed link traffic counts are generated for 90% of the ground truth counts but varied by adding uniformly random components in range of $\pm 20\%$, that is

$$y_{l,k}^{HE} = y_{l,k} \times [0.7 + 0.4 \times \alpha_{l,k}] \quad \alpha_{l,k} \sim U(0, 1) \quad (3.7)$$

3.3.3 Output data processor

The final component in the benchmarking framework processes the outcomes of the OD demand estimations and calculates a series of performance indicators. The choice of performance indicators plays an important role in benchmark analysis of OD demand estimation methods under input uncertainty. In this benchmarking framework, we distinguish two levels of performance indicators:

1. Statistical performance indicators

A number of statistical measures can be used to evaluate the overall performance of OD estimation methods. They measure the divergence between estimated outputs, (e.g., OD demand, $\hat{\mathbf{x}}$, traffic counts, $\hat{\mathbf{y}}$, travel times, $\hat{\mathbf{T}}$) and their ground truth values (e.g., OD demand, \mathbf{x} , traffic counts, \mathbf{y} , travel times, \mathbf{T}). A short list of statistical measures includes: route mean square error (RMSE), mean absolute error (MAE), mean absolute normalized error (MANE), mean error (ME) and normalized mean error (NME).

These statistical performance indicators have their limitations, because each of them represents only part of the information contained in OD demand distribution. This problem can be dealt with only by considering multiple statistics of

these OD demand distributions. As a basis, of course, one needs to estimate (sufficiently accurate) the entire OD demand distribution, which is calculated for all (or user selection of) OD pairs. From this distribution, one can calculate the 90th percentile OD demand (TT^{90}), the average and median OD demand (TT_{mean} , TT^{50}), or any other statistic.

2. Computational efficiency

The one of research objectives in this theses is improvement in the computational efficiency of the dynamic OD estimation and prediction problem for real time applications. To quantify the computational efficiency, one can calculate the CPU computation time of selected OD demand estimation methods.

3.4 Case study

3.4.1 Network topology and method selection

A benchmarking exercise of several dynamic OD estimation methods requires familiarity with these methods to ensure fair and informative comparison. Thus, the purpose of this case study is to demonstrate the main features of the proposed benchmarking framework for well-known dynamic OD estimation method (maximum entropy estimation method (Van Zuylen & Willumsen (1980), Wu (1997), Tsekeris (2003))). The Section 2.3.3 provides a detailed explanation of maximum entropy OD demand estimation method.

The performance of the maximum entropy method is evaluated on a real corridor network, given in Figure 3.4. This network was chosen because of the availability and quality of an empirical detector data on the network, and because a calibrated OD demand was available in the microscopic version of the PTV VISSIM multi-modal traffic simulation model for reproducing the traffic propagation over the network. In this network, each node represents the origin and destination with a single route between them. The network consists of 214 OD pairs and 118 corresponding links. The evening peak hour reflecting the congested state at the network which was divided in 6 departure time intervals of 10 minutes is chosen for experiment. Consequently, the trips between some of the OD pairs are not completed within one interval. In this way, a vehicle entering the network during a particular departure time interval may need more than one time interval to reach a traffic detector where the departure time interval and detection time are different. In this study network, the maximum travel time between OD pairs observed on network takes two time intervals, $q = 2$. The true assignment matrix is derived from the assignment of true OD matrix in PTV VISSIM.

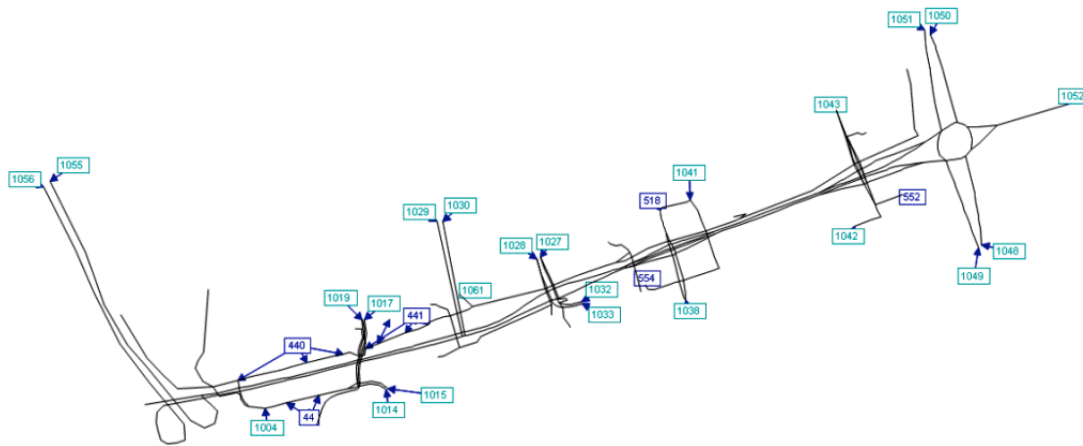


Figure 3.4: Urban corridor network, Kruithuisweg street, the Netherlands.

3.4.2 Considered scenarios and performance indicators

In terms of input data scenarios, the following selections are made:

1. Low-error observation scenario (LE) presented in Section 3.3.2 is selected in terms of traffic counts data quality and loop detectors availability.
2. All scenarios in terms of prior OD matrices presented in Section 3.3.2 are selected.

For the evaluation of the results of entropy maximization method from defined scenarios, the mean OD demand from resulting distributions is calculated as a performance indicator. This performance indicator choice stems from the fact that estimated dynamic OD demand may be compared with ground truth OD demand. In addition, performance indicator for link traffic counts is omitted because, it is possible that, even though the estimated link counts closely match the measured link counts, the estimated OD demand differ considerable from the ground truth OD demand. Also, two statistical metrics from the resulting distributions were calculated as performance indicators, the mean error (ME) and mean absolute error (MAE). In addition, CPU computation time is analyzed. The results are presented regarding total OD demand estimates, and followed by a more detailed analysis for separate OD pairs.

3.4.3 Results and discussion

Total demand analysis

The variations in OD flow over time as well as the uncertainty in the input data and their availability affects the reliability of OD matrix estimates. We show that this results in a considerable underestimation or overestimation of the total demand and/or between

specific OD pairs. Two criteria were used to check whether estimated OD matrices from the entropy method are on average equal to the true value, or show at least only a small variation around this value. The first criterion relates to the assessment of the total demand estimates under different scenarios of counting errors and errors in the prior OD matrix. Secondly, the assessment of certain OD pairs to the same error scenarios is examined.

The results are presented in a series of figures. Figure 3.5 presents the total estimated demand per departure time interval. The first column denotes the results obtained under influence of random prior OD matrices, while the second column denotes the results obtained from scaled prior OD matrices over departure time intervals. The two rows denote different scenarios with errors in link measurement for 20% and 40%.

The entropy method in presence of errors in link measurements on all links under-

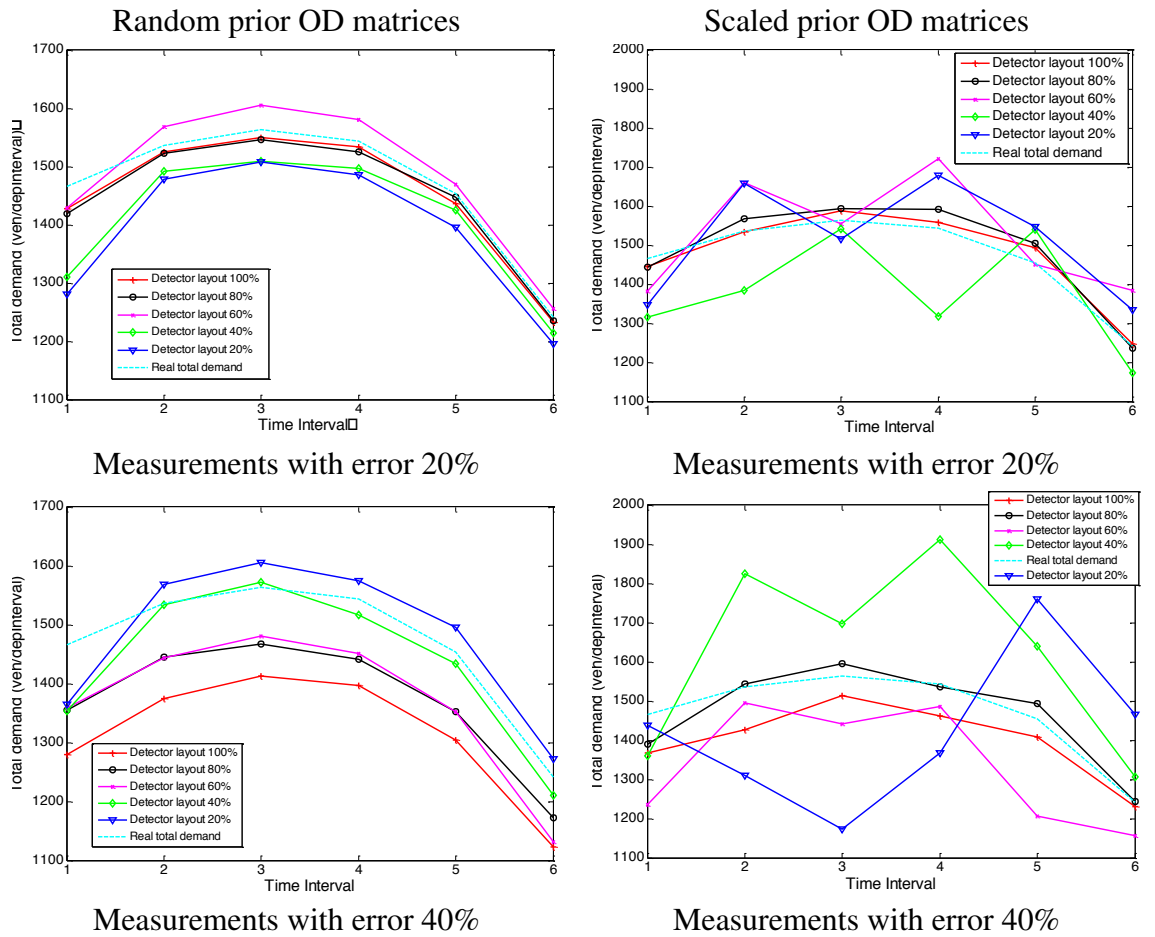


Figure 3.5: Average total demand estimates for different scenarios

estimates the total demand in case of random prior OD matrices. With a decreasing number of links that are measured, estimates of total demand are underestimated or overestimated, depending on the magnitude of errors imposed on the link measurements. It is also possible to observe that the prior OD information affects the reliability of the estimated OD matrices. The peaks in the estimated total demand curves shown

in the second column may be explained due to the fact that the demand in prior OD matrices is scaled over departure time intervals.

In scenarios with random demand variation in the prior OD matrices, for all assumed deviations it is observed that total demand estimates are very close to the real total demand, although for some OD pairs or even for specific origins, demands are overestimated or underestimated. The second criterion focuses on demand estimates for specific origins and analyzes the sensitivity of method under different scenarios.

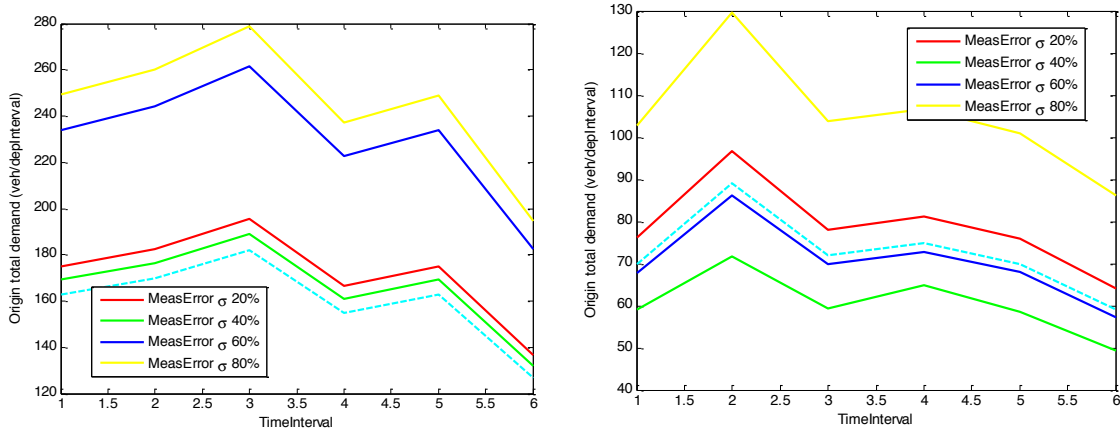


Figure 3.6: Average total demand for two origins with all links covered with detectors

The estimated total demands for two origins, one at the left inflow side (1056) and one at the middle of the network (1028), or the scenario with random perturbations of 20% in prior OD matrix, are given in the Figures 3.6 and 3.7. The reliability of estimated demand for the origin with higher demand volume (left figures) is affected by the availability of link counts data, which leads to underestimation and overestimation of the demand. This variation in estimates is also present in the origin whose demand is two times lower, in spite of the fact that the total demand in this scenario is very close to the real value.

The estimated total demand and total demand per origin for the scenarios with random prior OD matrices are given in the Figure 3.8 for total demand estimates and Figure 3.9 for origin 1030 (in the middle part of the network). The reliability of estimated demand (left figures) is not affected by the availability of link counts data and also with the size of random perturbations in prior OD. In contrary, the reliability of total demand estimates per origin is affected by the availability of link counts data which leads to underestimation of the origin demand. The entropy maximization method is robust against random perturbations in prior OD matrices if information from measurements from all links covered with the detector are used. The influence of higher random perturbations in prior matrices in prescience of more realistic 20% covered links with detectors leads to the underestimation and overestimation in estimates of total demand and total demand per origin. Less information from measurements leads to higher influence of prior matrices.

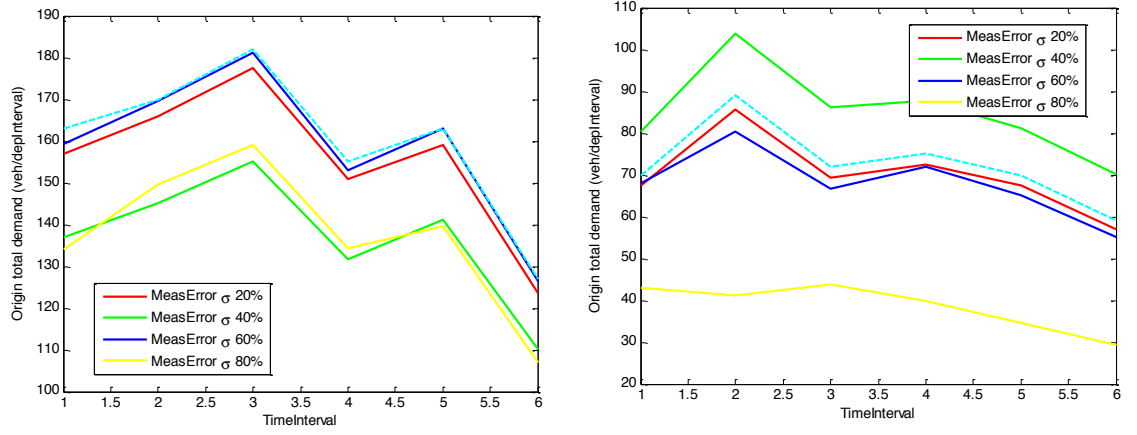


Figure 3.7: Average total demand for two origins with 60% links covered with detectors

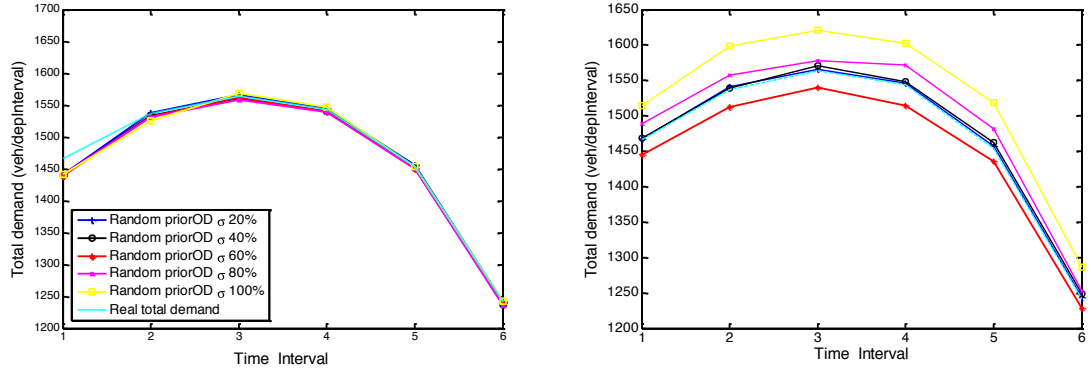


Figure 3.8: Total demand with all and 20% links covered with detectors

Since these variations are present among all origins and destinations, an additional analysis of OD demand estimates has been performed to obtain a deeper understanding of the efficiency and performance of the entropy method. The next subsection focuses on the influence of path length and OD demand volume per each OD pair.

OD pair analysis

Figure 3.10 shows the entire estimated OD demand distribution for two selected OD pairs. The first observation is the shape of the distribution, showing that for nearly all estimated OD matrices under each of the considered scenarios, there are always a few OD pairs that can be estimated very accurately, while at the same time, there are always some OD pairs with bias in estimates (Figure 3.10). In all tested scenarios it has been observed that the minimum absolute error for some OD pairs is close to zero while the maximum error is usually over 35%. It is thus interesting to test whether some OD pairs are more difficult to estimate than others, and to identify the properties of the paths and magnitude of the demand associated with such "problematic" OD pairs. Also, it seems worthwhile to compare the performance of different OD estimation

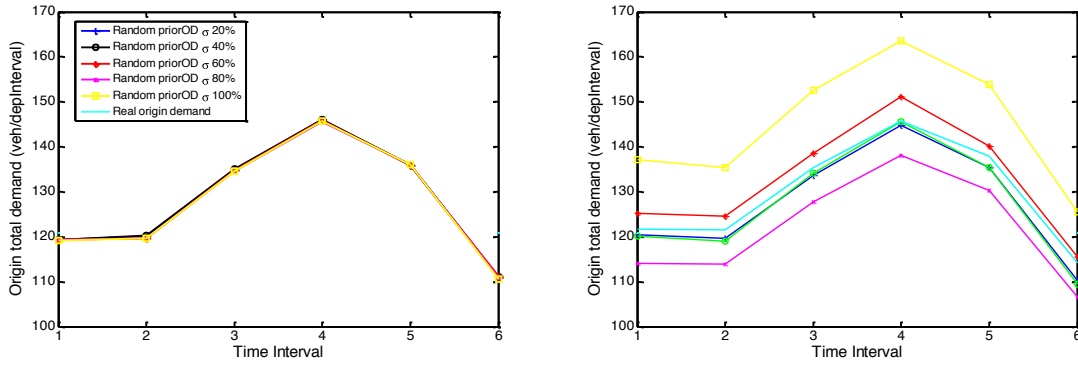


Figure 3.9: Total origin demand with all and 20% links covered with detectors

methods and their sensitivity on such OD pairs.

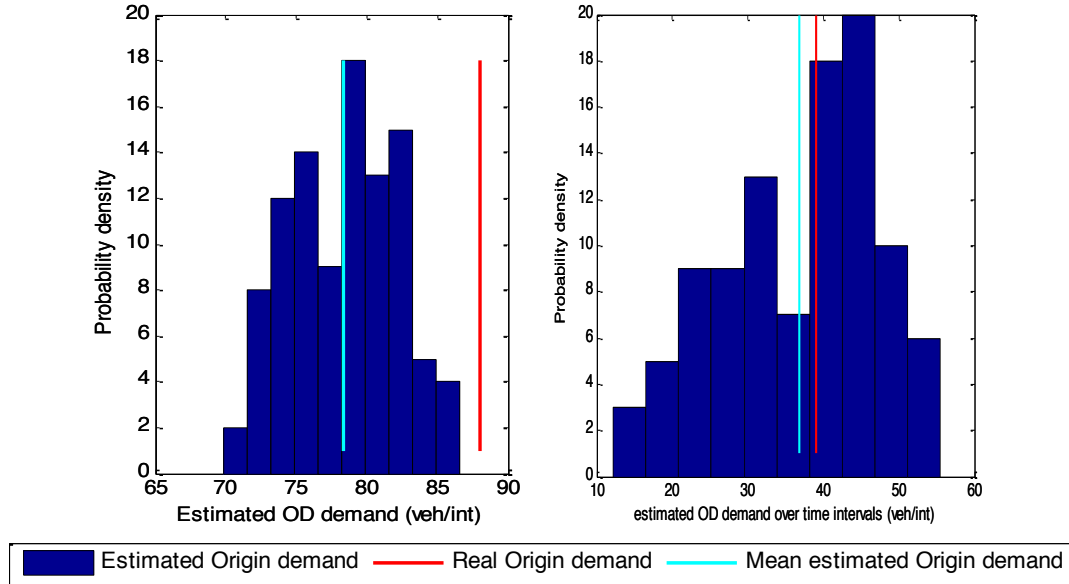


Figure 3.10: Simulated OD flow distribution for several OD pairs

The relationship between the average errors and the following two properties related to the topology of the network are explored:

1. The length of the path between origin and destination for an OD pair.
2. The magnitude of the OD demand for each OD pair.

The LHC method provides enough data to obtain meaningful error averages in all scenarios. In this case, the network depicted in Figure 3.4 is examined. The path lengths of all OD pairs are grouped from 1 to 7, where each group represents the average path length class between two succeeding OD pairs. On this network, this is close to 500m. All of the 214 OD pairs have been grouped into these 7 groups - one

for each possible path length - and computed the average error for all OD pairs in each group. Additionally, two different types of errors are considered: mean absolute error (MAE) and mean error (ME) between estimated and ground truth values. The last error provides more information on whether the flows underestimated or overestimated due to the path length. These data are presented in Figure 3.11 for scenario with different random perturbation of link measurements and 80% links covered with detectors.

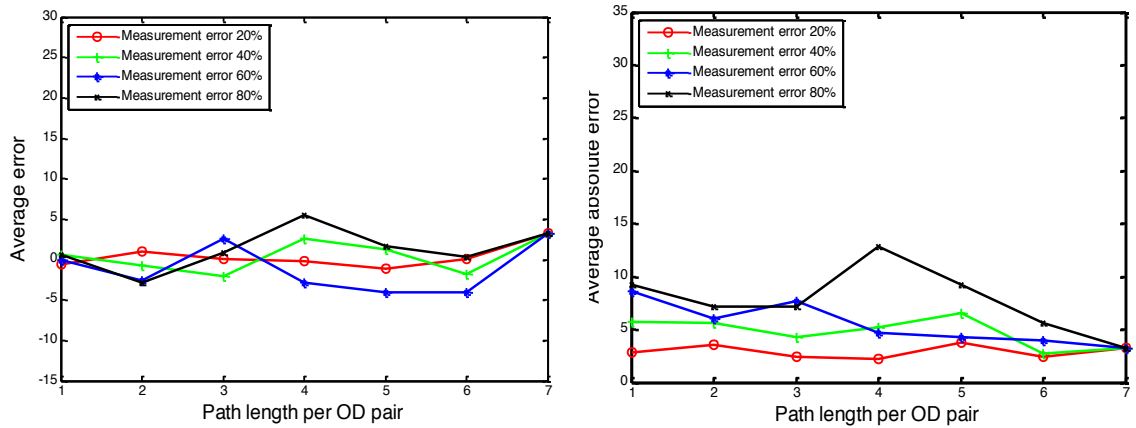


Figure 3.11: a) Mean error and b) Absolute mean error per path length

It can be seen that entropy method is more sensitive to path length, with increasing measurements errors leading to underestimation and overestimation of OD flows. However, the difference in errors among the path lengths is small, except for paths in group 7 which are around 4 km long. The dip in the curve for paths of length 2 and 4 may be explained by the fact that there were about twice as many OD pairs in these groups than in the other groups.

The performance of entropy method is also analyzed with respect to the demand volume between OD pairs. To examine the relation between ME and OD demand volume, a similar procedure is applied as was described above. In this case, the OD pairs are divided into the 7 groups based on the volume between OD pairs, with a difference of 5 vehicles per departure time interval between groups. The reason for such a small difference between groups is the very large number of OD pairs with low demand per departure interval. Subsequently, the average error for each group is computed. The results are shown in Figure 3.12 for the same scenario.

It is evident that the entropy method is sensitive to higher volumes of the OD demand and to increasing link measurements errors. Figure 3.12b) shows that the mean error decreases as the volume of OD demand increases. The smaller error values within group 7 reflect the case that the only 3 OD pairs have a higher demand of 45 vehicle/departure interval. Comparing the performance of the entropy method to path length and OD demand volumes, leads to conclusion that the entropy method is more sensitive to the level of demand than to path length regard to the increasing measure-

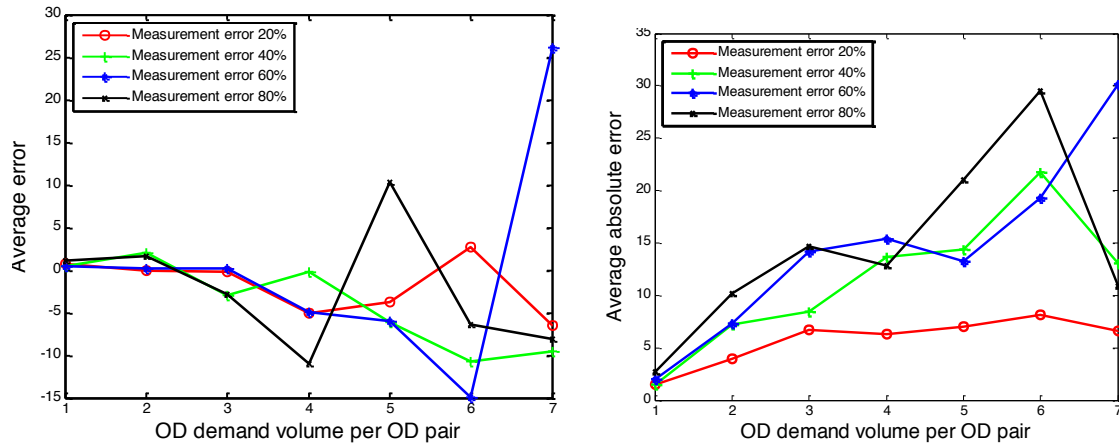


Figure 3.12: a) Mean error and b) Absolute mean error per OD demand volume

ments errors. In both figures, an overall error of around 5% is seen with low measurements errors, while an increase in measurement error leads to an underestimation of OD flows and increase of errors of around 15%.

Computational efficiency

Analysis of CPU computation time for maximum entropy method solved by iterative Revised Multiplicative Algebraic (RMART) algorithm shows that method needs around 32 minutes to simultaneously estimate OD demand with six departure time intervals for network with 214 OD pairs and loop detectors on all links. Although the RMART algorithm is designed to accelerate the convergence speed, it still requires extensive computation time. With a decreasing number of detectors on links, one can expect decrease in computational time but more uncertainty in OD demand estimates. However, the performance of entropy method is still computationally complex for real time applications.

3.5 Conclusions

This chapter demonstrates a new framework for assessing and comparing the performance of dynamic OD estimation methods to different types of uncertainty in the input data, and the parameters of the OD estimation methods. The analysis shows the dependence on the quality of the input, in terms of sensor configurations and accuracy, characteristics of data available, quality of the prior matrix, etc., that reveal confidence levels of the estimation outcomes under varying circumstances. Such insights are of high practical relevance, since the operational application of OD estimation methods, whether on-line or off-line, carries specific requirements regarding the accuracy and robustness of the OD estimators in relation to the characteristics of the input data.

Central to this benchmarking framework is a well-known simulation-based stratified sampling technique that allows assessing the performance without the need to perform an unfeasible number of simulations. The Latin Hypercube (LHC) method provides an efficient sampling method to effectively cover the whole range of the uncertainties defined in the scenarios. To the best of our knowledge the LHC method has not been applied to the OD estimation problem before and turned out to lead to a feasible approach to compute sensitivities using only a limited number of simulations.

To demonstrate the benchmarking framework we applied it successfully to the entropy maximization OD estimation method on a real urban test network. The application provided new insights into the performance of this estimation method with respect to the scale in the prior OD matrix. The entropy method furthermore was shown to be sensitive to path lengths and demand per OD pair. As a result, it can be concluded that it is harder to accurately estimate OD pairs that have longer paths or higher demand, since these show larger error in the link measurements. It was also found that flow estimates in case of higher demand per OD pairs were generally underestimated. Errors in link measurements have been shown to lead to a larger range of errors, and this results in underestimation and overestimation of total demand and per total origin demand. Finally, the results also showed that the use of less accurate link measurements can dominate the accuracy of entropy method especially when the link flow measurement availability is limited.

The contribution of the proposed benchmarking methodology is to enhance the understanding of the impact of the above-mentioned aspects on dynamic OD estimation methods in terms of robustness and accuracy. Using this benchmarking framework, both researchers and practitioners will gain insight into which OD estimation methods are suitable for both off-line applications, such as ex-ante simulation studies, and on-line applications, such as short-term traffic prediction, decision support systems, etc.

Future research should be aimed at applying the framework to a variety of OD estimation techniques in order to gain more insight into their characteristics, and to benchmark their performances in variety of situations. As motivated earlier in this Chapter, the statistical performance indicators found in practice and international literature all have their limitations, because each of them represents only part of information contained in dynamic OD demand estimates. The basic foundation of these performance indicators is that they are expressed as deviations in terms of OD demand or traffic data in respect to available ground truth data. Although the underling rational makes sense intuitively, the actual statistical measures in literature do not evaluate patterns and spatial correlations between OD pairs. Therefor, further research should be aimed at finding the new performance indicators as will be discussed in Chapter 6 of this thesis.

Chapter 4

Dimensionality reduction methods in OD demand estimation

Since OD matrices are high-dimensional, multivariate data structures, the estimation and prediction of such OD matrices is both methodologically and computationally cumbersome. This chapter explores the idea of dimensionality reduction and approximation of OD demand based on principal component analysis (PCA). In particular, the application perspectives of PCA for this purpose are considered.

First, by using PCA, the dimensionality of the time series of OD demand can indeed be significantly reduced. How PCA can be applied to linearly transform the high-dimensional OD matrices into the lower-dimensional space without significant loss of accuracy is shown. Moreover, the way in which the results from the PCA method can be used to reveal structure in the underlying temporal variability patterns in dynamic OD matrices is also described. The results indicate that it is possible to distinguish between three main patterns in dynamic OD matrices that follow structural, structural deviation and stochastic trends. Insight is provided into how these trends contribute to each OD pair and how this information can be used further in estimating and predicting dynamic OD matrices.

This chapter is an edited version of the articles:

Djukic, T., J.W.C. van Lint, S.P. Hoogendoorn, *Application of Principal Component Analysis to Predict Dynamic Origin-Destination Matrices*. Transportation Research Record: Journal of the Transportation Research Board 2283(1): 81-89, 2012.

4.1 Introduction

Much of the work in OD matrix estimation and prediction has focused on improving estimation and prediction of OD matrices with more sophisticated and less time consuming algorithms (Bierlaire & Crittin (2004), Kattan & Abdulhai (2006), Zhou & Mahmassani (2007)), and by including additional available data, ranging from traffic counts to automatic identification data (Antoniou et al. (2004), Asakura et al. (2000), Dixon & Rilett (2005)), and data from Bluetooth devices (Barcelo (2010)), to name a few. In this chapter focus will be on an important and often-overlooked aspect within OD estimation and prediction, namely the analysis of the information and structure contained in the (estimated, predicted or realized) OD demand. The use of quantitative methods to exploit this structure for various purposes in the estimation of dynamic OD demand shall be discussed, such as identifying underlying trends and correlations, and reducing the complexity of the estimation and prediction problem.

Since dynamic OD matrices are high dimensional multivariate data structures, the estimation and prediction of dynamic OD matrices is both methodologically and computationally cumbersome for real-time applications. There are three factors that increase the computational effort: a) the size of the state vector, b) the complexity of model components (e.g. assignment matrix and covariance matrices), and c) the number of measurements to be processed.

Firstly, solution algorithms typically scale $O(n^3)$ with the size of the state vector. For example, the Kalman filter algorithm is commonly used method to estimate and predict the OD matrices (Antoniou et al. (2006), Ashok et al. (1993), Barcelo (2010)). Since the computational complexity of the Kalman filter is typically in the order of $O(n^3)$ (Zhou & Mahmassani (2006)) where in the simplest case n is the total number of the OD pairs in the network, this can represent a potential computational bottleneck. In addition, each traveler takes a certain time to complete his/her trip in large-scale networks, and the resulting travel time can be very long depending on trip length between OD pairs and prevailing traffic conditions (i.e. congestion level in the network). The effect of time lag indicates that the traffic flow at the current observation time interval can include OD flows departing from a previous time interval, leading to an enormous computational strain. For example, if it is assumed that the number of lagged time intervals is h , then the state vector includes the state variables for departure time intervals $k, k-1, \dots, k-h$. Therefore, the size of the state vector ($n \times 1$) at time interval k increases by at least $((h+1) * n \times 1)$ and manipulation of vectors and corresponding covariances becomes cumbersome. Lately, an approximation of OD flows to handle the lagged OD flows has been proposed by Ashok & Ben-Akiva (2000). The approximation is based on the conjecture that much of the information about an OD flow related to redundancy over time is likely to be provided the first time it is counted. If this were true, OD flows corresponding to prior departure intervals could be held constant at their prior estimated values and only the flows for the current departure interval need to be estimated. Alternatively, a polynomial trend model proposed by

Zhou & Mahmassani (2007) can offer a compact representation of a time series of OD flows. However, the polynomial trend model is still very computationally intensive for large-scale networks due to high dimensionality of the state vector.

Secondly, the efficiency of solution algorithms depend on the complexity of its components. For example, the assignment matrices are high-dimensional and sparse, and most solution approaches require their multiplications and inversions, which is very time consuming.

Third, the computational cost of solving OD estimation problems in many cases scale with the number of available measurements. For example, the most critical factor affecting state observability is the ratio between link traffic counts, r , and number of OD pairs n . Therefore, for a given number of OD pairs, measuring more traffic counts or adding densities or speeds increases the chances of state observability being satisfied, which again increases the required computational effort.

Clearly, reducing the dimensionality of the state vector, is a way to improve computational efficiency. For example, let us assume that OD flows have been estimated for several previous days or months. These flows subsume various kinds of information, about trip making patterns and their spatial and temporal variations. Therefore, the key idea in this thesis' approach is to reduce the dimensionality of the OD matrix, in such way that the structural patterns are preserved. With this approach the computational cost can be speeded up dramatically, without significant loss of accuracy.

In other fields where large amounts of (dynamic) data are analyzed, Principal Component Analysis (Pearson (1901), Hotelling (1933)) or PCA, also known as the Karhunen-Loeve procedure, is a commonly used method. PCA extracts from data structures those components that explain most of the variance. In this chapter, PCA is applied to analyze times series of OD demand matrices on a real highway network. This shows that on a longer time scale (days to weeks), their structure can be well-captured using only a few dimensions, and that the principal components found can be categorized into three meaningful trend classes. These include a structural trend that captures the regular pattern; a large structural deviation trend that captures short-term fluctuations, and a stochastic trend that captures the random fluctuations. This taxonomy derived from the data itself resembles the characteristic trends in total OD demand observed in the work of Zhou & Mahmassani (2007), and provides a useful basis for researchers in the development of prediction models.

The chapter is organized as follows. The first part of the chapter demonstrates how high-dimensional OD demand can be transformed in low-dimensional space using PCA technique. In the second part of the chapter the structure of the experimental data set is defined, as is the way in which times series of OD matrices per departure time interval on a real highway network section are derived. Next, the chapter demonstrates how PCA can be used to reveal the underlying temporal distribution patterns in these data. The chapter closes with a discussion on further application perspectives of PCA in estimation and prediction of dynamic OD demand method presented in Chapter 5.

4.2 OD demand representation in low-dimensional space

This section gives a detailed explanation of the Principal Component Analysis (PCA) method, also known as the Karhunen-Loeve procedure, through transformation of dynamic OD demand into a space of lower-dimensionality. PCA method was first introduced by Pearson (1901) and Hotelling (1933) to describe the variation in a multivariate data set in terms of a set of uncorrelated variables. Since then, PCA has found application in many fields, such as image analysis, pattern recognition, image compression, and time-series prediction (Kong et al. (2005), Pentland et al. (1994), Jaruszewicz & Mandziuk (2002)). For example, Lakhina et al. (2004) has demonstrated application of PCA method to explore temporal variability of traffic flows between routers in internet traffic. It has also been used in traffic and transportation science before, for example, for the dimensionality reduction in traffic counts prediction (Sun et al. (2009)) and for dimensionality reduction in the calibration of travel demand from traffic counts (Flötteröd & Bierlaire (2009)).

Dynamic OD demand has a concise representation when expressed in terms of an orthonormal basis of eigenvectors $e_i, i = 1, 2, \dots, n$ that can be derived using PCA. The goal of the research presented in this section is to map vectors of the OD demand $X \in \mathbb{R}^n$ onto the new vector in an m -dimensional space, where $m < n$.

To transform OD demand in lower-dimensional space, the same rational is followed as the one presented in Jolliffe (2002) to describe the PCA procedure. Suppose that we have used a microsimulation-based demand model to generate a large sample of OD demand observations k (e.g. observations can represent daily OD demand, or OD demand per departure time interval) in a network, each being a realization of the n -dimensional OD demand vector $x_k = (x_1, x_2, \dots, x_n)$. Thus, there is $(k \times n)$ OD demand matrix X , where each row $i, i = 1, 2, \dots, k$ contains the vector of OD flows per time (e.g. for whole day, per departure time interval) and column $j, j = 1, 2, \dots, n$ denotes the realizations of the n -th OD pair over time. The OD demand data matrix X must be pre-processed before performing PCA. The columns of X are centered in such a way that the mean of each column is equal to zero by subtracting the mean of each OD pair over time. This will ensure that the cloud of data is centered on the origin of the eigenvectors. Once the covariance matrix of the OD demand matrix X is generated, PCA is applied to extract the eigenvectors $e_i, i = 1, 2, \dots, n$ and eigenvalues $\lambda_i, i = 1, 2, \dots, n$.

Now, transforming the OD demand data into the new coordinate system can proceed. Since the covariance matrix of X is real and symmetric, its eigenvectors e_1, e_2, \dots, e_n can be chosen as an orthonormal basis. Therefore, OD demand matrix X , or actually $(X - \bar{X})$, can be represented, without loss of generality, as a linear combination of a set of n orthonormal eigenvectors e_i

$$X - \bar{X} = c_1 e_1 + c_2 e_2 + \dots + c_n e_n = \sum_{i=1}^n c_i e_i \quad (4.1)$$

where the eigenvectors e_i satisfy the orthonormality relation

$$e_i^T e_j = \delta_{ij} \quad (4.2)$$

in which δ_{ij} is the Kronecker delta symbol. Explicit expressions for the coefficients c_i in (4.1) can be found by using (4.2) to give

$$c_i = (x_i - \bar{x}_i) e_i \quad (4.3)$$

which can be regarded as a simple rotation of the coordinate system from original x 's to a new set of coordinates given by c 's. An intuitive explanation of (4.3) is that the eigenvectors are used as weights on each of the original variables to compute the new set of variables: *the principal demand components* c_1, c_2, \dots, c_n . The representation of $(\hat{X} - \bar{X})$ on the orthonormal basis e_1, e_2, \dots, e_n is thus given by principal demand components c_1, c_2, \dots, c_n . The principal demand components c_i capture the contribution of each eigenvector e_i to the particular observations of OD demand. In turn, the eigenvectors e_i capture the common behavior of travelers over all OD pairs.

By sorting the eigenvectors in decreasing order by the size of the eigenvalue, the first m eigenvectors ($m \leq n$) can be retained, which captures the maximum data variance. However, since the covariance matrix of observed OD demand in general can be very large, it is inconvenient to evaluate and store all eigenvectors and eigenvalues explicitly. To avoid this, efficient algorithms can be used, which determine the m largest eigenvectors of the covariance matrix, such as the orthogonal iteration and power method (Golub & van Van Loan (1996)). The choice of first m eigenvectors depends on selection criteria. It can be based on simple heuristic observation of cumulative variance of eigenvalues, or it can be based on a more complex statistical and hypothesis test. The issue is discussed further in Section 4.3.2.

Once the m largest eigenvectors e_1, e_2, \dots, e_m are found, a new low-dimensional representation of the OD demand by using equation (4.1) can be expressed as follows:

$$\hat{X} - \bar{X} = \sum_{i=1}^m c_i e_i^T \quad (4.4)$$

where \hat{X} is the approximated OD demand constructed using the first m eigenvectors. Thus, the explicit expression for the approximated OD demand \hat{X} in low-dimensional space can be found by using equation (4.4) to give

$$\hat{X} = \sum_{i=1}^m c_i e_i^T + \bar{X} \quad (4.5)$$

To examine the intrinsic dimensionality of OD demand and to understand the underlying patterns in the time evolution of OD demand, the following subsection presents and explains the results obtained by applying PCA. First it is shown that only small set of principal components is necessary for a good approximation of high-dimensional

original data. Then, in order to discern the low dimensionality of OD demand, the temporal patterns in the dynamic OD demand are examined.

4.3 Exploring the temporal variability of dynamic OD matrices

4.3.1 Setting up OD matrices database

Following the process of performing PCA given in the previous section, the database of OD demand needs to be generated. This database is usually generated from available historical OD demand, off-line estimated dynamic OD demand or from synthetic OD demand data generated from a detailed travel demand microsimulation method (e.g. activity-based or trip-based methods). In addition, emerging data collection systems are receiving attention for their potential to provide direct observation of OD flow data. Although availability of these data is limited today, one may expect in near future large samples of OD flow data to be collected. To demonstrate application perspectives of PCA, the database of dynamic OD demand for network presented in Figure 4.1 is generated as follows.

A simplified freeway network consisting of 26 OD pairs $n = 26$, 27 nodes and 52 corresponding links (Figure 4.1) is considered. This network was chosen because of the availability and quality of empirical detector data, and turn fraction data for a long time period. Furthermore, a calibrated OD matrix is available. In this network, each on-ramp and off-ramp represents an origin and destination respectively, with a single route between them. Traffic counts at origin links and destination links are measured and aggregated in 15-minute intervals. In addition, available static OD matrices are transformed into time sliced OD matrices using the temporal evolution trends from collected traffic counts on the off-ramps and the on-ramps, yielding the dynamic OD matrix. The entire day (from 00:00 to 24:00) for five weekdays (from Monday to Friday) is considered.

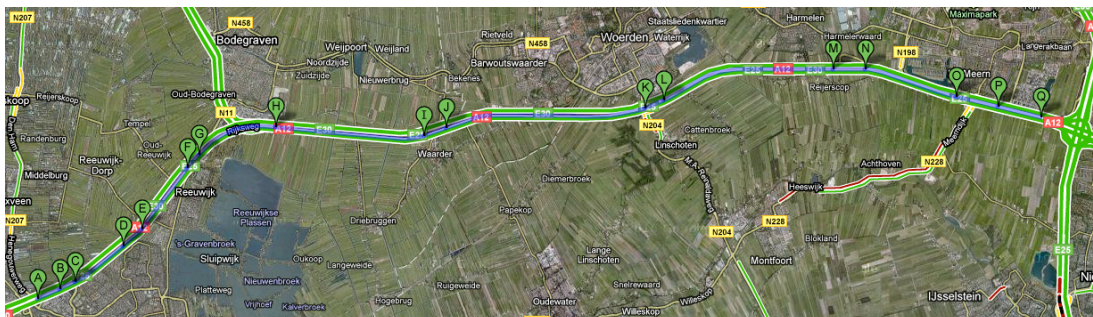


Figure 4.1: Freeway network, A12, Gouda Utrecht, the Netherlands.

4.3.2 Reducing the dimensionality of dynamic OD matrices

The original data set can be represented with relatively high accuracy by projection onto the first m - eigenvectors, ($m < n$) that have significantly larger eigenvalues, and contain the most useful information relating to the specific problem.

Criteria for selecting a limited number of principal components

The eigenvalues indicate how much variance is explained by each eigenvector according to the following equation:

$$\|Xe\| = e^T Se = \lambda e^T e = \lambda \quad (4.6)$$

where we have used the fact that since the eigenvectors are orthonormal, we have $E^T = E^{-1}$.

A number of approaches are available for determining how many eigenvectors should be retained in order to account for most of the variation in data X (Jolliffe (2002)). In general, the choice of how many eigenvectors should be retained is often made by means of visual examination of a number of different criteria. The simplest relates to plotting the eigenvalues sorted by size, i.e. by making a so-called scree plot and looking for an "elbow" in this plot. Figure 4.2 presents the scree plot of eigenvalues obtained from data set X for visual examination of eigenvalues magnitudes. A sharp elbow appears around the fifth eigenvalue, showing that the effective dimensionality of the time series of OD demand data is far smaller than the apparent dimensionality of the total number of OD pairs in network, n .

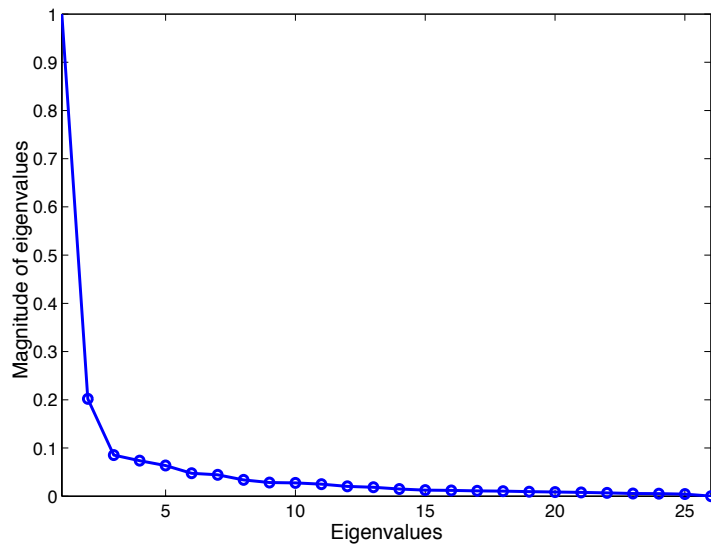


Figure 4.2: Scree plot of eigenvalues from centered OD matrix X .

A second criterion for choosing m eigenvectors is a predetermined total (cumulative) percentage of total variation explained, such as 95%. The results of applying this criterion are presented in Figure 4.3. In this case, both plots indicate that the choice of eigenvectors between 5 and 10 are enough to represent 95% of the total variance in data, where the first two eigenvectors represent over 70% of the total variance in data.

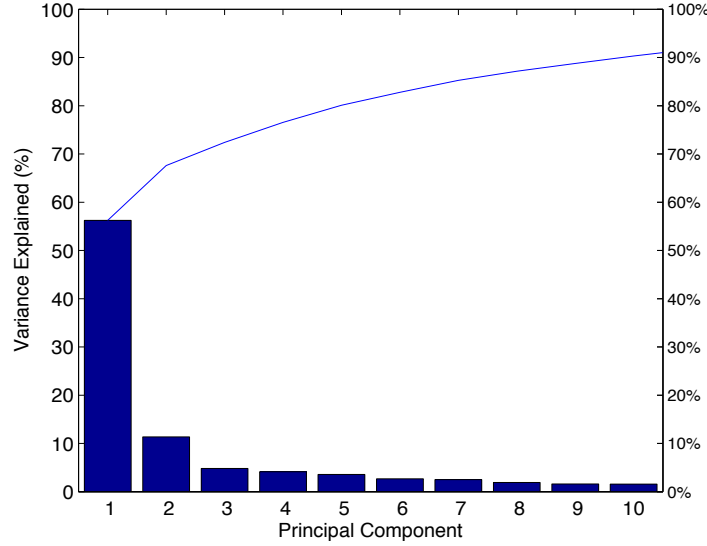


Figure 4.3: Cumulative percentage of total variation explained by eigenvalues.

Projecting the data onto a limited set of eigenvectors

Among all linear techniques, PCA provides the optimal approximation of original data X in terms of the quadratic reconstruction error $\|X - \hat{X}\|$; see (Sun et al. (2009)). The original OD matrix X can be approximated by projecting this data on new m -dimensional space applying equation (4.4).

To illustrate the observed low dimensionality of timeseries of OD matrices per departure time interval, graphical representation of a demand for two OD pairs with respect to the first five eigenvectors and original data is given in Figure 4.4. The figure shows that even if over 80% of dimensions are omitted from the original data X , the temporal variability of these OD flows can be captured well.

In order to explain the potential reasons for the low dimensionality of timeseries of OD matrices, the next subsection explores temporal variability patterns captured by principal demand components.

4.3.3 Finding temporal patterns in time series of OD matrices

This subsection first explores which principal demand components are most significant for each OD pair. Then, it shows the application of PCA to explore the temporal pat-

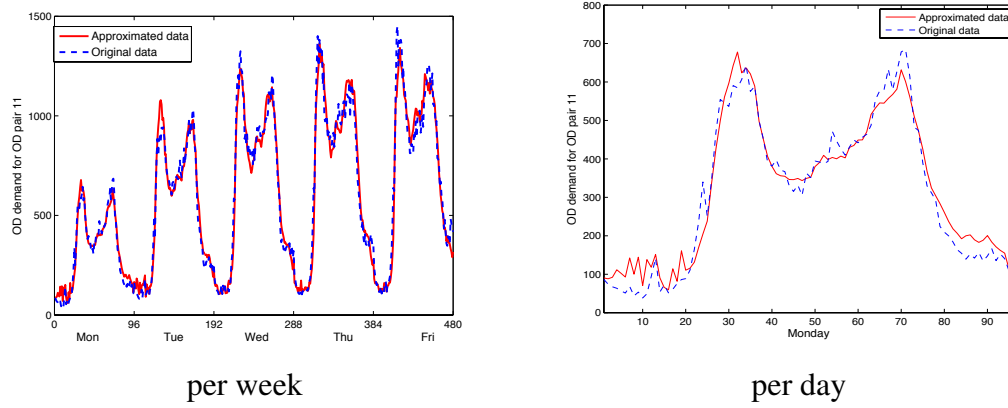


Figure 4.4: Original and approximated OD demand of one OD pair with five eigenvectors.

terns captured by principal demand components. It also demonstrates the application of PCA method to identify and separate different temporal variability trends within each OD pair.

Contribution of principal demand components to OD pairs

To understand how principal demand components contribute to each variable, and to each OD pair, the coefficients in eigenvector matrix E defined in the previous section are analyzed. The eigenvectors can be understood as the weights on each of the original variables used to compute the new set of principal demand components. Note that the row i of the eigenvector matrix E specifies the weights (referred to as *loading coefficients*) on each of the original variables used to compute the set of principal demand components (see equation (4.3)).

To explore the low dimensionality of time series of OD matrices, we will examine how each OD pair is composed of significant principal demand components and how the magnitude of the OD demand in each of OD pairs is related to these components. First, the number of loading coefficients in the eigenvector matrix that are significantly different from zero must be determined. This can be done by setting a threshold value $1/\sqrt{n}$, which implies a perfectly equal mixture of all principal demand components, taking into account that columns of E have a unit norm. Then, the number of significant principal demand components is obtained by counting how many loading coefficients in the rows of the eigenvector matrix E exceed this threshold in absolute value.

The cumulative distribution function of the number of loading coefficients per row of the eigenvector matrix E that exceed this threshold value is shown in Figure 4.5. The figure shows that most of the OD pairs are composed of five significant loading coefficients, and no OD pair has more than ten significant loading coefficients. From this result it can be concluded that each OD pair has only a small set of temporal variability features captured by principal demand components.

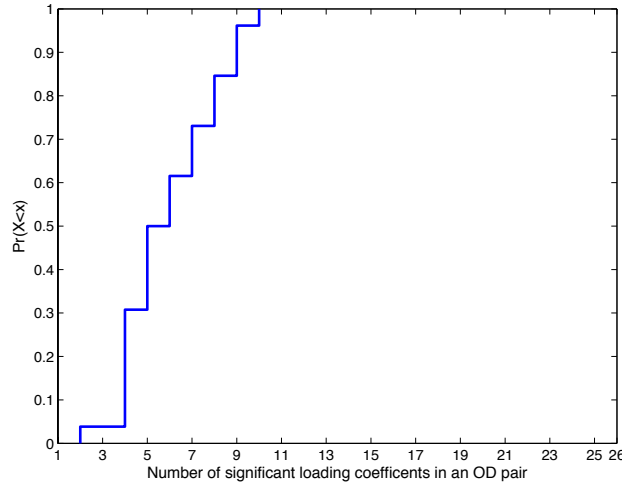


Figure 4.5: CDF of number of significant principal demand components that constitute each OD pair.

This should be followed by an examination of which loading coefficients are most significant per OD pair, and determining if they differ in terms of magnitude of OD flow. To analyze this, the point in each OD pair where loading coefficients occur above the threshold entries in eigenvector matrix E is inspected. Figure 4.6 shows the loading coefficients for each OD pair that exceed the threshold value as dots so that the OD pairs are presented in increasing order of magnitude of OD flow. Note that the columns of the eigenvector matrix E are organized in decreasing eigenvalue order, from the first eigenvector to the last. Thus, as could be expected, the OD pairs with highest OD flow consist mainly of the most significant principal demand components, while the smaller OD demands consist of less significant principal demand components. Therefore, the next subsection explores the temporal patterns captured by each principal demand component as playing a key role in understanding the properties of temporal variability of dynamic OD matrices.

Temporal variability patterns in time series of OD matrices

The previous subsection's analysis of the contribution of principal demand components to OD flows has emphasized the importance of understanding the temporal patterns captured by principal demand components. This temporal patterns can be derived by normalization of principal demand components to unit length by dividing by $\sqrt{\lambda_j}$, given as (Jolliffe (2002))

$$u_j = \frac{c_j}{\sqrt{\lambda_j}} \quad j = 1, \dots, d \quad (4.7)$$

The vectors u_j are vectors of size k and orthogonal by construction and are referred to as the temporal data vectors of X . The value of contribution is between 0 and 1 and, for a given principal demand component; the sum of the contributions of all observations over time is equal to 1. Thus temporal vector u_j captures the temporal

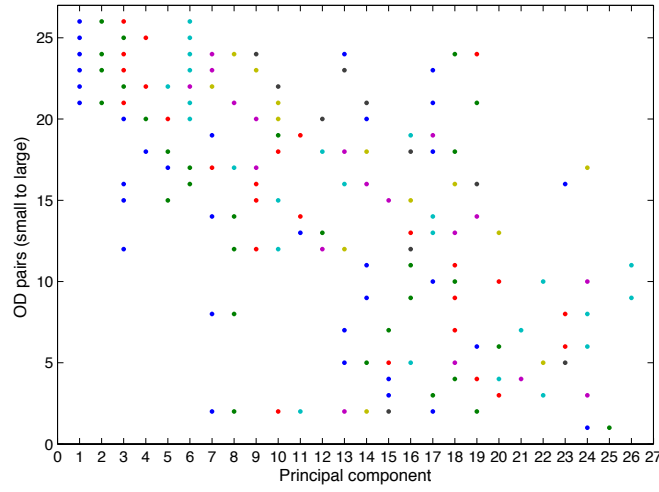


Figure 4.6: Correlation of OD flows volume and their principal demand components.

variability observed in all OD pairs along principal demand component c_j . Since the principal demand components are in order of contribution to the overall variance in OD demand, temporal vector u_1 captures the strongest temporal trend common to all OD flows, u_2 captures the next strongest, and so on. The set of temporal vectors u_j is stored in columns of a temporal pattern matrix U , which has size $(k \times n)$.

From an initial assessment of these temporal vectors, it turned out that there are always three distinctly different temporal patterns. Figure 4.7 depicts plots of temporal vectors on three principal demand components that capture the different patterns. In all tested temporal vectors, it was observed that each vector captures one of the three different temporal patterns that can be classified as follows:

1. **Structural patterns** The first plot in Figure 4.7 shows an example of a temporal vector that exhibits a strong structural pattern with peaks in morning and evening rush hours in weekdays. The observed periodicity clearly reflects average regular daily activities.
2. **Large structural deviation patterns** The second plot in Figure 4.7 shows an example of a temporal vector that exhibit a large structural deviation from the average pattern. The deviation component captures events that cause unforeseen deviations from the average pattern. Examples of such events could be large sporting events, days with extreme weather conditions, etc.
3. **Stochastic patterns** - The third plot in Figure 4.7 shows an example of a temporal vector describing the random deviations from the average pattern. These vectors capture the remaining random variation in OD demand and the majority of vectors u_j appear to be of this type.

This taxonomy, derived from PCA, can be viewed as parallel to characteristic trends in dynamic OD demand that have been found in other studies such as the work of

Zhou & Mahmassani (2007). Also, the statistical inference of the temporal and spatial correlation of OD demand has been studied by Martin L (2008). While previous studies have generally focused on describing these features for the purpose of modeling OD demand prediction methods, the above result show that the common patterns of OD demand variability can be determined entirely from the data without any modeling assumptions.

Decomposition of OD flows

Since the results show that each OD pair is composed of several significant principal demand components and based on the classification of temporal patterns captured by these components, it is interesting to investigate the decomposition of OD flows on these patterns. A clear benefit of classifying the trends in variability of OD demand would provide decomposition of each OD flow into its principal features.

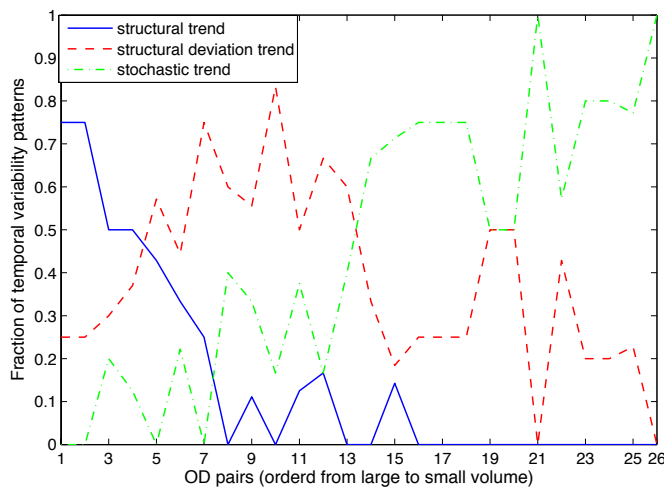
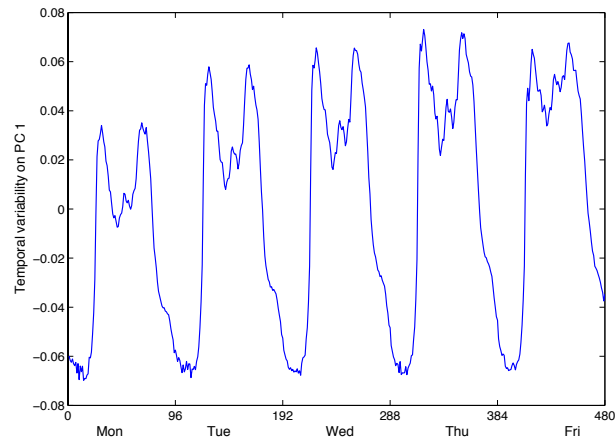
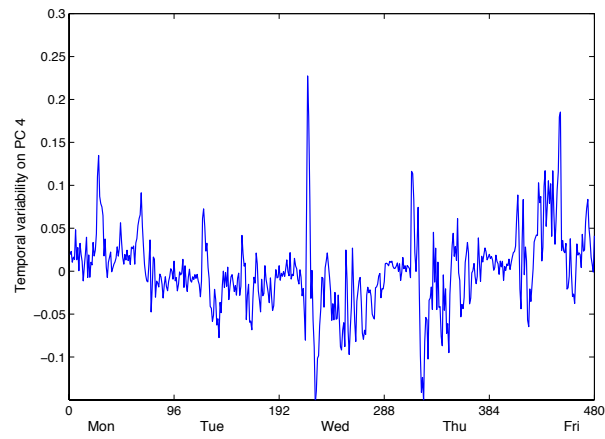


Figure 4.8: Decomposition of OD pairs into the temporal trends.

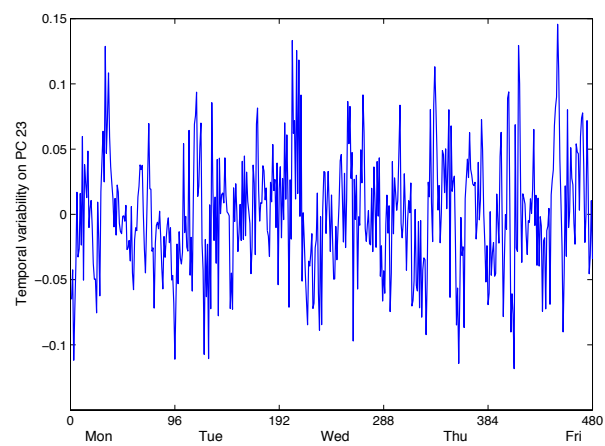
The relative contribution of each temporal pattern type to each OD pair has been determined. The results are shown in Figure 4.8, where the OD pairs on x-axis are sorted according to increasing order of demand volume per OD pair. For each OD pair we plotted the fraction of its variability contributed by three patterns: structural, large structural deviation and stochastic trends. It is possible to observe that high volume OD pairs are dominated by structural trends. As the demand volume tends to decrease, the relative contribution of the structural trend decreases, while the stochastic trend and large structural deviation pattern increase.



a) structural pattern



b) structural deviation pattern



c) stochastic pattern

Figure 4.7: Three different temporal trends captured by temporal vectors on principal demand components.

4.4 Conclusions

This chapter discusses the potential of using PCA to pre-process the time series of OD matrices per departure time for estimation of dynamic OD matrices. The contributions of this study include the following aspects. The application of PCA to the set of dynamic OD matrices over a long time period on freeway network section provides evidence of the applicability of PCA to reduce the high dimensionality of OD data. By using PCA, high-dimensional dynamic OD matrices can be accurately approximated over time using 20% of the original data.

Exploring the reasons for such low-dimensionality, by examining the temporal variability patterns in every principal demand component over OD pairs, the existence of three dominant patterns was presented. It was shown that the set of dynamic OD matrices consist of three patterns: structural, large structural deviation and stochastic trend. Furthermore, it was shown that each OD pair can be decomposed into these patterns. High volume OD pairs are dominated by structural trends. As the demand volume tends to decrease, the relative contribution of the structural trend decreases, while the stochastic trend and large structural deviation pattern increase. This result will be used further in Section 6, to determine structural correlation between OD pairs in matrix.

The next chapter shows how principal demand components are used as state variables instead of OD flows for real-time dynamic OD demand estimation and prediction. State space OD estimation model is formulated, where the eigenvectors e_i define the fixed structure of the OD matrices, and principal demand components are updated on-line from link traffic counts. Furthermore, some practical points, such as temporal correlation between traffic data observations introduced by dimensionality reduction of the state vector, will be discussed.

Chapter 5

Methodology for real time OD demand estimation and prediction

This chapter presents an methodology for solving the high dimensionality problem in real time OD demand estimation and prediction and shows the efficiency of the resulting approximation for large-scale networks. A way of applying principal component analysis (PCA) to linearly transform the high dimensional OD matrices into the lower dimensional space, without significant loss of accuracy has been achieved. Next, a new transformed set of variables (demand principal components) is defined, that is used to represent the OD demand in lower-dimensional space. These new variables are defined as state variables in a novel reduced-state space model for real-time estimation of OD demand. The quality improvement of OD estimates has been demonstrated, using this new formulation and a so-called, "colored" Kalman filter approach for OD estimation, in which correlated observation noise is taken into account. In this chapter, it is established that by significantly reducing the dimensionality of the OD data, in such a way that the structural patterns are preserved, the computational costs can be dramatically reduced. Moreover, the model performance and computational efficiency are thoroughly analyzed using real data from a large network, and method for obtaining a reduced set of state variables.

This chapter is an edited version of the articles:

Djukic, T., G. Flötteröd, H. van Lint, S. Hoogendoorn, *Efficient real time OD matrix estimation based on Principal Component Analysis*. Proceedings of the Intelligent Transportation Systems Conference, Anchorage, Alaska, 2012. and

Djukic, T., J. W. van Lint, S. Hoogendoorn, *Methodology for efficient real time OD demand estimation on large scale networks*. Transportation Research Board, Washington, DC, 2014.

5.1 Overview of methodology

This chapter focuses on the efficient estimation and prediction of OD matrices for large-scale networks, since they will be used for real-time applications such as dynamic traffic management. Here, efficient denotes the estimation and prediction of large-scale dynamic OD matrices with the least waste of computation time, to satisfy real-time application requirements. A convenient way of understanding the proposed methodology is to illustrate the proposed and OD demand estimation and prediction methodology in the generic way.

In general terms, all dynamic OD demand estimation and prediction methods aim to find the most probable OD matrix X_k , given previous within day estimates $X_{k-\tau}$, $\tau = 1, 2, \dots$, or historical OD matrices X_{prior} from day-to-day time period, the available (sensor) data Y , and all the other assumptions H (related to, for example, the assignment method and/or the assumed temporal evolution of the OD patterns). The common methodology with inputs and outputs into an OD matrix estimation and prediction is illustrated in Figure 5.1. The input data can be collected from different sources such as detector data (e.g. link traffic counts, speeds and densities) or from Bluetooth data, floating car data and licensee plate-recognition cameras (e.g. travel times and route choice). The most widely used sensor data are link traffic counts y_k , that are available for the entire analysis period (all the departure intervals) or at the end of each interval k . This generic methodology can be used for off-line and real-time applications. For off-line application, the entire set of link traffic counts for the analysis period can be used to simultaneously estimate OD matrices for all time intervals. For real-time application, at the end of each interval k , only the counts corresponding up to k^{th} time interval can be used to sequentially estimate the OD matrix for the current time interval. Finally, for real-time application, predictions of OD matrices are generated for intervals $k+1, k+2, \dots$ and the estimation and prediction process continues.

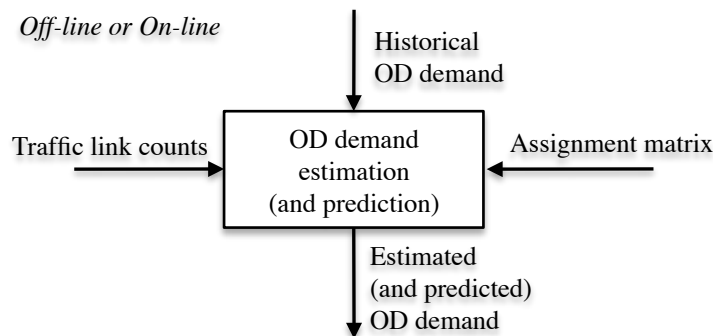


Figure 5.1: Overview of common OD estimation (and prediction) methodology

The OD estimation problem is computationally intensive because solution methods have to deal not only with high-dimensional structures of OD matrices, but also with

the computational complexity of these methods. One of the problems with high-dimensional datasets is that, in many cases, not all the measured variables are crucial to the understanding of the underlying phenomena of interest (e.g., temporal or spatial correlation between variables or associated patterns, etc.). In other words, one may postulate that high-dimensional data are multiple, indirect measurements of an underlying source. As a result, one possible solution approach to solve this issue of high-dimensionality is to map the high-dimensional OD matrices into a space of lower dimensionality, as presented in previous Chapter 4.

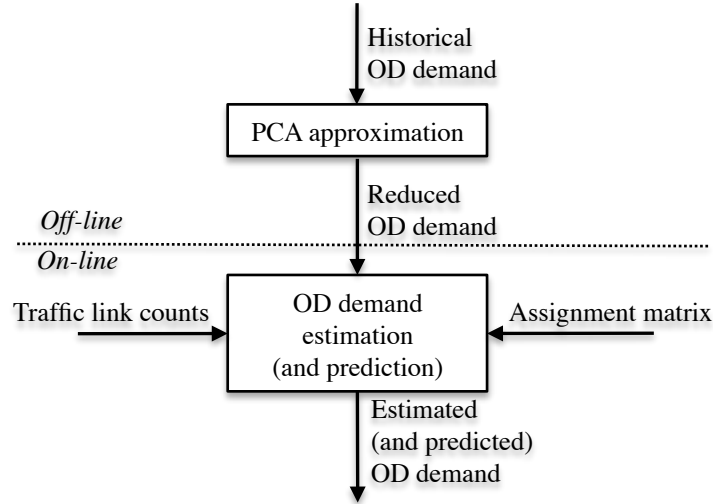


Figure 5.2: Overview of proposed OD estimation (and prediction) methodology

Figure 5.2 illustrates the proposed methodology for real-time OD demand estimation and prediction. This method, applies the Principal Component Analysis (PCA) Jolliffe (2002), to a data set of historical OD demands or synthetic OD demand data generated from a detailed travel demand microsimulation method (e.g. activity-based or trip-based methods). This historical OD dataset can be approximated as a linear combination of a set of only a few orthonormal vectors (eigenvectors) and principal demand components. These eigenvectors that capture the trip-making patterns and their spatial and temporal variations are extracted off-line, whereas the principal demand components capture the contribution of each eigenvector to the realization of a particular OD flow. These principal demand components are defined as state variables instead of the OD flows themselves. As a result, the dimensionality of the state is reduced substantially and the complexity of the estimation problem is likewise reduced. For real-time application, at the end of each interval k , the traffic counts corresponding up to the k^{th} time interval would be used to sequentially update principal demand components for the current time interval. Finally, the estimated principal demand components are used to obtain the estimates of OD matrix.

Reducing the dimensionality problem with PCA replaces the usual approach of using prior OD matrices by structural information on OD flows obtained either from historical OD demand data or from synthetic OD demand data generated from a detailed

demand microsimulation system (e.g. activity-based or trip-based methods). The importance and originality of this approach lies in the possibility of capturing the most important structural information without loss of accuracy and considerably decreasing the model dimensionality and computational complexity.

This chapter is organized as follows. In the next section, state space OD estimation model is formulated, where the eigenvectors e_i define the fixed structure of the OD matrices, and principal demand components are updated on-line from link traffic counts. Next, the properties of the colored noise Kalman filter to solve the proposed dynamic OD estimation and prediction method with time-correlated observations is presented. Finally, the proposed method is demonstrated on large-scale urban network (Vitoria, Spain).

5.2 A Reduced state space OD estimation and prediction model formulation

This section includes a demonstration of the use of the approximated OD demand presented in the Chapter 4 in a state space based formulation. The development of state space based models, first requires state variables to be defines. Below is a brief discussion of this definition in the context of dynamic OD demand estimation and prediction using the PCA approach presented previously.

Following the idea presented in Chapter 4, the chosen state is defined as a $(m \times 1)$ vector of *principal demand components*, c_k , where m represents the reduced number of variables in the state vector at time interval k . The principal demand components represent approximated OD demand, where each principal demand component c_i , for $i = 1, 2, \dots, m$ captures the contribution of each eigenvector e_i to the particular observations of OD demand. Therefore, the OD demand state in the network at time k is uniquely described by the vector of the principal demand components c_k in m -dimensional space, where $m < n$.

The state space model formulation consists of process and observation equations. Clearly, a process equation must be specified that captures the temporal evolution of the state, and an observation equation that uses whatever new information (i.e. observation) is available to estimate the state.

A process equation is defined as an autoregressive process on principal demand components, which provides a preliminary estimate of the OD flow. The process equation is defined as follows:

$$c_k = \sum_{q=k-\tau}^{k-1} \phi_k^q c_q + w_k \quad (5.1)$$

where ϕ_k^q , a $(m \times m * q)$ is the process matrix that represents the effects of previous states c_q on current state c_k , τ is a degree of the autoregressive process and w_k is a

vector of random variables capturing unobserved deviations in the process. The process noise vector w_k depicts a known Gaussian noise term defined with the following assumptions:

- mean $E[w_k] = 0$;
- variance $E[w_k^2] = \theta_k \delta_k$, where θ_k is a $m \times m$ variance covariance matrix, with eigenvalues on the diagonal stored in decreasing order, and the δ_k is the Kronecker symbol.

In addition to this, the state space model formulation uses an observation equation, defined as a linear relationship between the state variables (principal demand components) and the observations (e.g., traffic counts):

$$y_k = \sum_{h=k-\kappa}^k A_k^h x_h + v_k \quad (5.2)$$

where $y_k \in \mathbb{R}^r$ denotes a vector of link traffic counts for time interval k , and A_k^h is a $(r \times n * h)$ matrix, known as an assignment matrix, mapping OD flows that depart during intervals h to link traffic counts observed during interval k . Further, κ is the maximum number of time intervals needed to travel between any OD pair, and v_k is a vector of random variables capturing the observations error on detectors during interval k .

Following the lower-dimensional representation of OD demand by principal demand components and substituting (4.4) in (5.2), the observation equation (5.2) can be reformulated as:

$$\begin{aligned} y_k &= \sum_{h=k-\kappa}^k A_k^h (c_h E_h + \bar{x}_h) + v_k \\ &= \sum_{h=k-\kappa}^k H_k^h c_h + \bar{y}_h + v_k \end{aligned} \quad (5.3)$$

where $H_k^h = \sum_{h=k-\kappa}^k A_k^h E_h$ is a $(r \times m * h)$ matrix called observation matrix, mapping the principal demand components during intervals h to traffic counts observed during interval k . Note that the observation matrix H_k^h in equation (5.3) is not the same as the assignment matrix A_k^h given in (5.2). Finally, the matrix H_k^h is used for the linearization of the model; it equals the transformation of the assignment matrix A_k^h to the orthonormal basis matrix of eigenvectors E_h . The observation noise v_k depicts a known Gaussian noise term defined with following assumptions:

- mean $E[v_k] = 0$;
- variance $E[v_k^2] = R_k \delta_{km}$, where R_k is a $(r \times r)$ variance covariance matrix, and the δ_{km} is the Kronecker symbol.

In conclusion, note that this model uses the following input variables: process transition matrix ϕ_k^h , process error covariance matrix θ_k , and observation error covariance matrix R_k . These input data are usually derived from, for example, existing historical data on OD demand and observations. In transport modeling for the real-time applications, it is considered that data would be available over multiple days, making it possible to calibrate model inputs.

5.2.1 State Augmentation

Travel time on large-scale networks can be very long depending on trip length between OD pairs and prevailing traffic conditions (i.e. the congestion level in the network). The implication of trips with travel times that are significantly larger than the discrete time steps between consecutive OD departure intervals is that observations (traffic counts) made at the current time step k may include OD flows departing at a range of earlier time intervals $k-1, k-2, \dots, k-s$. This time lag effect can be dealt with through state augmentation - each state variable is estimated $s = \max(\kappa + 1, \tau)$ times. Therefore, in this section the process equation (5.1) and observation equation (5.3) are reformulated to capture time lag effect.

First, the state vector c_k is augmented to include additional state variables from previous time intervals as:

$$C_k = [c_k' \ c_{k-1}' \ \dots \ c_{k-s}']'$$

Thus, the process equation (5.1) can be written in augmented form as:

$$\begin{bmatrix} c_{k+1} \\ c_k \\ \vdots \\ c_{k-s+1} \end{bmatrix} = \begin{bmatrix} \phi_{k+1}^k & \dots & \phi_{k+1}^{k-s-1} & \phi_{k+1}^{k-s} \\ I & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & I & 0 \end{bmatrix} \times \begin{bmatrix} c_k \\ c_{k-1} \\ \vdots \\ c_{k-s} \end{bmatrix} + \begin{bmatrix} w_{k+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Then, the process equation (5.1) in matrix form is given as:

$$C_{k+1} = \Phi_k C_k + W_{k+1} \quad (5.4)$$

where

C_k is $(n(s+1) \times 1)$

Φ_k is $(n(s+1) \times n(s+1))$

W_{k+1} is $(n(s+1) \times 1)$

Now, the observation matrix is refined as

$$H_k = [h_k^k \ h_k^{k-1} \ \dots \ h_k^{k-s}]$$

where

H_k is $(r \times n(s+1))$

Then, the observation equation (5.3), expressed in terms of the augmented state vector, can be written as

$$y_k = H_k C_k + \bar{y}_k + v_k \quad (5.5)$$

Up to this point, a reduced state space OD estimation model has been formulated and the following section proceeds with solution approach.

5.3 Solution approach

For the sake of clarity, we will start with a reference to the idea of variables reduction in state vector and the state space model of Section 4.3. This is followed by the examination of a solution approach in which correlated observation noise is taken into account through the reduction of variables in the state vector.

5.3.1 Temporal correlation between observations introduced by dimensionality reduction

Equations (5.1) and (5.3) constitute a linear state space model. The solution approach of such a system of equations may seem fairly standard at first glance. However, since there are practical points which are not entirely obvious, they are illustrated here, before a solution algorithm is presented. The dimensionality reduction approach to approximate the dynamics of dynamic OD demand, X , given in equation (5.1) introduces an additional noise term that represents the variance of those OD flows that are filtered out by the PCA procedure. In order to explain the potential reasons for the temporal correlation between observations introduced by the dimensionality reduction of the state vector, the observation noise correlation has been analytically derived. Here, the effect of lagged time intervals κ has been omitted from observation equation (5.3), for the sake of simplicity.

The given observation equation (5.3) for the reduced number of state variables m over time interval k can be expressed as

$$\begin{aligned} y_k &= A_k \sum_{i=1}^m c_{i,k} e_{i,k} + A_k \sum_{i=m+1}^n c_{i,k} e_{i,k} + v_k \\ y_k &= A_k \sum_{i=1}^m c_{i,k} e_{i,k} + \xi_k \\ \xi_k &= A_k \sum_{i=m+1}^n c_{i,k} e_{i,k} + v_k \end{aligned} \quad (5.6)$$

where, ξ_k represent the observation noise that consists of additional noise introduced by dropped state variables from $m + 1$ till n at time interval k .

Further, observation equation (5.3) for the reduced number of state variables m for the next time interval $k + 1$ can be expressed as

$$\begin{aligned}
 y_{k+1} &= A_{k+1} \sum_{i=1}^m c_{i,k+1} e_{i,k+1} + A_{k+1} \sum_{i=m+1}^n c_{i,k+1} e_{i,k+1} + v_{k+1} \\
 y_{k+1} &= A_{k+1} \sum_{i=1}^m c_{i,k+1} e_{i,k+1} + \xi_{k+1} \\
 \xi_{k+1} &= A_{k+1} \sum_{i=m+1}^n (c_{i,k} + w_k) e_{i,k+1} + v_{k+1} \\
 \xi_{k+1} &= A_{k+1} \sum_{i=m+1}^n c_{i,k} e_{i,k+1} + A_{k+1} \sum_{i=m+1}^n w_k e_{i,k+1} + v_{k+1} \quad (5.7)
 \end{aligned}$$

where, ξ_{k+1} represent the observation noise at time interval $k + 1$ that consists of additional noise introduced by omitted state variables from $m + 1$ till n in previous time interval k . It follows that, ξ_k and ξ_{k+1} represent the temporal correlated observation noise. It is well known that this condition destroys the assumption of independency between process and observation noise that underlies the standard Kalman filter. The objective of this section is to find an effective method to deal with this kind of correlation.

5.3.2 Colored noise Kalman filter solution algorithm

If the observation noise is correlated, the new observation equation can be modeled as

$$y_k = H_k c_k + \xi_k \quad (5.8)$$

and the time-correlated observation noise is modeled as a first-order Gauss Markov process as follows:

$$\xi_{k+1} = \Psi \xi_k + v_k \quad (5.9)$$

where, with reference to equations (5.8) and (5.9), the correlation matrix Ψ is equivalent to the process transition matrix Φ_k for time correlated errors, and v_k is an observation noise vector assumed to be uncorrelated with the process noise vector w_k .

When the observation errors are temporally correlated, as has been shown in Subsection 5.3.1, the *time differencing approach*, first introduced in 1968 by Bryson & Henrikson (1968), is commonly applied as a way to model correlated observation noise in state-space model representations. The core idea behind this time differencing approach is the elimination of the time-correlated observation noise terms ξ_k , using a pseudo-observation equation z_k whose error is white. From equations (5.8) and (5.9),

we have

$$\begin{aligned}
 z_k &= y_{k+1} - \Psi y_k \\
 &= (H_{k+1}\Phi_k - \Psi_k H_{k+1})c_k + H_{k+1}w_k + v_k \\
 &= H_k^* c_k + v_k^*
 \end{aligned} \tag{5.10}$$

where $H_k^* = H_{k+1}\phi_k - \psi_k H_{k+1}$ and the new observation noise is given as

$$v_k^* = H_{k+1}w_k + v_k \tag{5.11}$$

Moreover, the mean and the variance of the new observation noise v_k^* are defined as follows:

- mean $E[v_k^*] = E[H_k w_k + v_k] = 0$;
- variance $E[v_k^* v_k^{*T}] = \text{var}(v_k^*) \sigma_{kj} = R_k^*$

where covariance matrix R_k^* is given as

$$\begin{aligned}
 R_k^* &= E[(H_{k+1}w_k + v_k)(H_{j+1}w_j + v_j)^T] \\
 &= H_{k+1}\theta_k H_{k+1}^T + R_k
 \end{aligned} \tag{5.12}$$

Further, the *decorrelation technique* presented in Bryson (2002) is applied to process equation (5.6) to eliminate the correlation that now exists between the new observation noise v_k^* (5.11) and the process noise w_k . A new process equation can be written as

$$\begin{aligned}
 c_k &= \phi_{k-1}c_{k-1} + w_{k-1} + J_{k-1}(z_{k-1} - H_{k-1}^*c_{k-1} - v_{k-1}^*) \\
 &= \phi_{k-1}^*c_{k-1} + J_{k-1}z_{k-1} + w_{k-1}^*
 \end{aligned} \tag{5.13}$$

where the new state process matrix is expressed as $\phi_k^* = \phi_{k-1} - J_{k-1}H_{k-1}^*$, and $J_{k-1}z_{k-1}$ is the control item of the new state-space equation system. The new process noise error is defined as

$$w_{k-1}^* = w_{k-1} - J_{k-1}v_{k-1}^* \tag{5.14}$$

To make the new process noise (5.14) and observation noise (5.11) uncorrelated, the covariance matrix is given as

$$E[w_k^* v_j^{*T}] = E[w_k v_j^T] - J_k E[v_k v_j^T] = (S_k - J_k R_k) \sigma_{kj} = 0 \tag{5.15}$$

Therefore, based on equation (5.15), the following condition should be satisfied:

$$J_k = S_k R_k^{*-1} \tag{5.16}$$

Moreover, the mean and the variance of the new state noise w_k^* are given as follows:

- mean $E[w_k^*] = E[w_k - J_k v_k^*] = 0$

- covariance $E[w_k^* w_j^{*T}] = \text{var}(w_k^*) \sigma_{kj} = \theta_k^*$

where the covariance matrix θ_k^* is given as

$$\begin{aligned}\theta_k^* &= E[(w_k w_k^T + J_k E[v_k v_k^T] J_k^T - E[w_k v_k^T] J_k^T - J_k E[v_k w_k^T])^T] \\ &= \theta_k + J_k R_k J_k^T - S_k J_k^T - J_k S_k^T \\ &= \theta_k + S_k R_k^{*-1} S_k^T\end{aligned}\quad (5.17)$$

For a more detailed derivation of a colored noise Kalman Filter, and derivation of covariance matrices θ^* , R^* and M^* refer to Bryson (2002).

At this time and for the given problem, a state space model has been depicted by equations (5.10) and (5.13), which satisfies the assumptions of the standard Kalman Filter. Clearly, the new process noise w_k^* and observation noise v_k^* are independent, zero-mean, Gaussian noise processes of covariance matrices θ_k^* and R_k^* respectively. Algorithm 1 summarizes the colored Kalman Filter equations as a solution of such a equation system.

Algorithm 1 The colored Kalman Filter

Initialization:

$$\hat{\mathbf{c}}_{0|0} = E[\mathbf{c}_{0|0}] \text{ and } P_{0|0} = E[\mathbf{c}_{0|0} - E[\mathbf{c}_{0|0}]]^T$$

When no additional information is available, $P_{0|0}$ is usually initialized as a matrix with large diagonal entries, reflecting the fact that the correctness of the initial estimate of $\hat{\mathbf{c}}_{0|0}$ is highly uncertain.

For $k = 1, 2, \dots$ **do:**

Compute the Kalman Gain:

$$K_k = P_{k|k} H_k^{*T} (H_k^* P_{k|k} H_k^{*T} + R_k^*)^{-1} \quad (5.18)$$

Correct mean and covariance:

$$c_{k-1|k} = c_{k-1|k-1} + K_k (z_k^* - H_k^* c_{k-1|k-1}) \quad (5.19)$$

$$P_{k-1|k} = (I - K_k H_k^*) P_{k-1|k-1} (I - K_k H_k^*)^T + K_k R_k^* K_k^T \quad (5.20)$$

Update mean and variance of state variables:

$$c_{k|k} = \phi_k^* c_{k-1|k} + J_k z_k \quad (5.21)$$

$$P_{k|k} = \phi_k^* P_{k-1|k} \phi_k^{*T} + \theta_k^* \quad (5.22)$$

End

Note that the time differencing solution algorithm uses one time interval latency in the

observation updating because the observation in time interval k must be used to update the state vector in the previous time interval, $k - 1$. Therefore, following the Kalman Filter terminology, $c_{k-1|k}$ denotes correction of the state variable for time interval $k - 1$, using the information from link traffic counts for interval k , and $P_{k-1|k}$ depicts the updated state error covariance matrix. The Kalman filter gain in equation (5.18) evaluates the importance of the new information obtained from link traffic counts at time interval k and can be interpreted as the weight given to the most recent information. Equations (5.19) and (5.20) reflect the corrected knowledge of the system state at time interval $k - 1$ with the obtained link traffic counts for interval k . In the update step, knowledge of the evolution of state and observations is used to update prior corrections. Therefore, the equations (5.21) and (5.22) reflect the best estimate of the system state $c_{k|k}$ and $P_{k|k}$ error covariance matrix at time interval k , including the information on link traffic counts for time interval k .

Finally, the result of the colored Kalman filter, the estimated a posterior state vector $c_{k|k}$, is used to estimate OD demand by applying equation (4.5). All that is required to extend the model to k -step prediction is to multiply the filtered vector by the appropriate ϕ_k matrix k times.

5.4 Numerical experiments

This section begins with a description of the input data used by proposed dynamic OD estimation and prediction method, e.g. historical OD demand generation and state variables reduction procedure. Two assessment scenarios are considered, in terms of number of variables in the state vector (i.e. with and without the reduction of state variables). These scenarios will be discussed in more detail below. Numerical experiments are performed on a large-scale network, (Vitoria, Basque Country, Spain) with real data to evaluate the performance of the proposed model and solution algorithm.

5.4.1 Network topology

Prior to method evaluation, a Vitoria network has been defined, consisting of 57 centroids, 3249 OD pairs with a 600km road network, 2800 intersections and 389 detectors, presented as black dots in Figure 5.3. This network is available in the mesoscopic version of the Aimsun TSS (2013) traffic simulation model for the reproduction of traffic propagation over the network. The true OD demand is available for this network, which allows analysts to assess the performance of the proposed method. The true assignment matrix and traffic counts on detectors are derived from the assignment of true OD matrix in Aimsun for the evening period from 19:00 to 20:00 reflecting a congested state of the network. The simulation period is divided into 15 minute time intervals with an additional warm-up time interval, $K = 5$. The link flows resulting from the assignment of the true OD demand are used to obtain the traffic count data per

observation time interval. The trips between some of the OD pairs are not completed within one time interval due to congestion on the network or the distance between OD pairs. In this way, a vehicle entering the network during a particular departure time interval might need more than one time interval to reach a traffic detector, where the departure time interval and detection time are different. In the chosen study network, the maximum travel time between OD pairs observed on the network takes four time intervals, which leads to very sparse assignment matrices, and the number of lagged time intervals $\kappa = 4$.



Figure 5.3: The Vitoria network, Basque Country, Spain

5.4.2 Simulating historical daily OD demand

A major problem with all method assessments is obtaining meaningful evaluations of the algorithm's results and performance, because the true sources of data are not available for comparison when working with real data. One solution is to use simulated OD demand data, where underlying sources and phenomena are known. To generate a simulated OD demand per departure time interval dataset for our case study, it is necessary to define an arbitrary model for OD demand generation, which represents a common spatial and temporal behavior of travelers.

Here, the Logit model is performed in sequence, in order to introduce the correlation to OD flows. First, the set of traveler's decisions before making a trip are defined, including decisions to make a trip or not, destination choice and departure time choice. Then, for each of these decisions, the set of alternatives available to travelers has been defined. The activity and traveling intentions of traveler t_r are presented in the Figure 5.4. The main principle of this model is that a large number of simulations are performed for varying model inputs, reflecting the variability in the travelers' behavior, and consequently in OD demand, based on Monte Carlo simulations.

The total number of trips per origin from available true OD matrix is assumed as an initial number of travelers per origin in simulations. Subsequently, 10.000 observations have been generated, each representing a realization of the 3249-dimensional

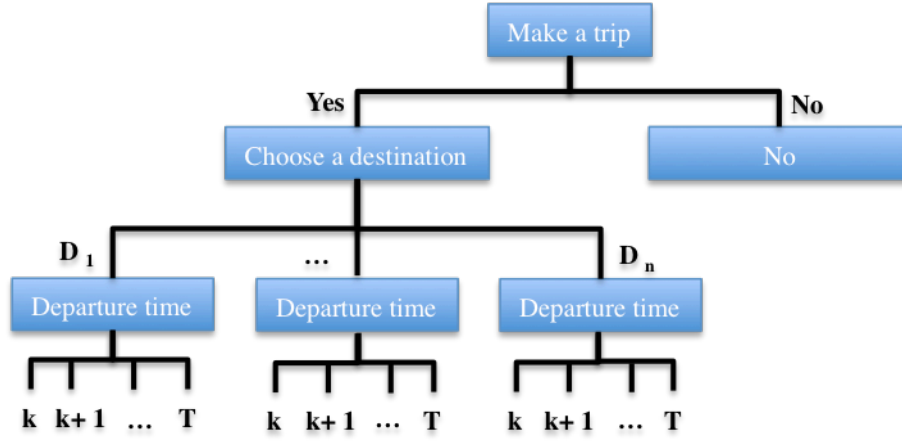


Figure 5.4: The set of decisions and alternatives for traveler

OD demand vector, where the total number of travelers per origin is equal to the true OD matrix, while their distribution over destinations is varied. Each generated OD demand vector per departure time interval is stored in the OD demand matrix, where each row represents one observation of OD demand, as defined in the section discussing the definition of state variables.

5.4.3 State vector reduction

To examine the effect of reducing the number of principal demand components in state vector, the PCA has been applied to the OD demand data matrix X_k over $K = 5$ departure time intervals. Once the PCA is performed, the set of eigenvectors $e_{i,k}$ for $i = 1, 2, \dots, 3249$ and eigenvalues $\lambda_{i,k}$ for $i = 1, 2, \dots, 3249$ per time interval k are obtained.

It has been shown in previous sections that eigenvalues can be used to explore the data reduction potential, for instance by considering the total (cumulative) percentage of total variation explained (e.g. 95%), Figure 5.5. It can be observed that 90% of the variance of the data is captured by the first 50 eigenvectors out of 3249. This result indicates that the state vector can be reduced by more than 90% and still capture the temporal and spatial variance in data.

The PCA has been performed on OD demand data set per time interval, so that it is possible in every time interval to identify a potential number of variables in state vector that describe the 95% of variance in the data set. Table 5.1 shows the number of state variables m that describe the 95% of variance in data set per departure time interval.

Table 5.1: The number of state variables that capture 95% of variance per time interval k

Departure time interval	1	2	3	4	5
The number of state variables	40	61	37	39	41

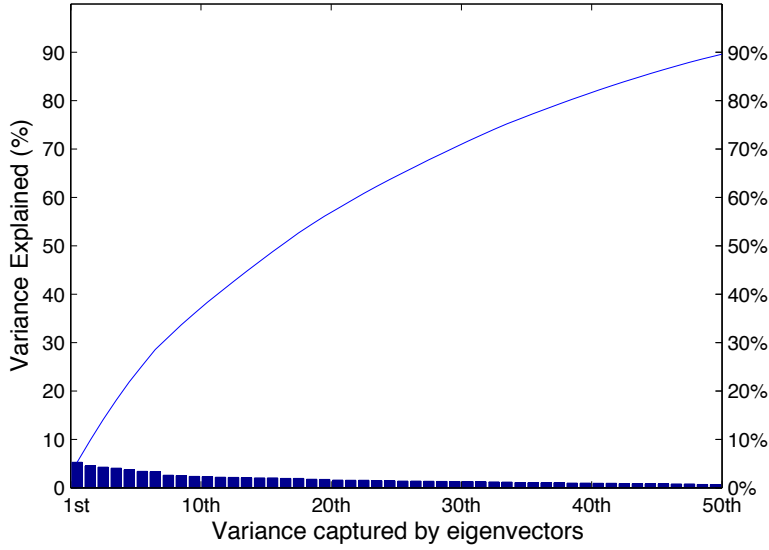


Figure 5.5: Cumulative percentage of total variation explained by eigenvalues for time interval k_2

It is evident from Table 5.1 that different values of variables in state vector over time intervals have been obtained. Therefore, the number of state variables is defined as $m = \max(m_k)$, for $k = 1, 2, \dots, K$, since omitting the principal demand components with highest captured variance in OD demand would lead to non-effective dimensionality reduction of the state vector. In the next subsection, the performance of the colored Kalman Filter for reduced variables in a state vector (for $m = 61$) will be compared to its performance when applied to a state vector with no reduced variables.

5.4.4 Method performance

Experiments have been conducted for a Vitoria network, Spain (given in Figure 5.3), for following two scenarios:

- Case 1: in this experiment run, the state variables (principal demand components) have been omitted from the state vector. Since the principal demand components in the state vector are arranged in decreasing order of eigenvalues, the principal demand components that capture the lowest variance have been removed, while the first $m = 61$ state variables remain;
- Case 2: in this experiment run, all state variables (principal demand components) in the state vector remain, so that $m = n = 3249$.

Table 5.2 presents: (1) the root mean square error (RMSE) per departure time interval and (2) mean absolute error (MAE) per departure time interval; that is

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\bar{x}_i - x_i)^2} \quad (5.23)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\bar{x}_i - x_i| \quad (5.24)$$

for \bar{x}_k estimated OD demand per time interval, depending on the number of the variables in state vector, m and n respectively.

Table 5.2: Error in the solution given by number of state variables in the state vector

	Time interval	Case 2: No reduction of stat.var.	Case 1: Reduction of stat.var
RMSE	k_1	6.803	14.674
RMSE	k_2	0.537	9.774
RMSE	k_3	0.493	10.026
RMSE	k_4	0.493	9.296
RMSE	k_5	0.527	9.553
MAE	k_1	3.039	8.186
MAE	k_2	0.288	4.330
MAE	k_3	0.243	4.501
MAE	k_4	0.233	4.268
MAE	k_5	0.278	4.409

It is clear from Table 5.2 that reducing the variables in state vector yields overestimation of OD demand. However, it is possible to observe that by reducing the dimensionality of the state vector by more than 90%, the colored Kalman Filter produces a reasonable reduction in accuracy. In real-time applications, it is always a question of trade-off between the computational efficiency and results' accuracy. Therefore, it is of interest to examine the optimal number of variables in a state vector, where the lower bound is defined as a minimum number of variables that capture 95% of the variance in the data set, while the upper bound is given by the computation time preferences. In addition, larger errors relate to the observability problem introduced by state variables reduction. Therefore, the state identifiability must be taken into account in the computation of the optimal number of state variables to achieve the Kalman Filter convergence.

Note that the initial idea of this research is to solve the computational complexity of the OD estimation problem for real-time applications. Therefore, in Table 5.3 the run time of colored Kalman Filter is shown for each scenario (e.g. no reduction of state variables in state vector and reduced number of state variables) on the Vitoria network.

Table 5.3: CPU time computations in seconds

Time interval	1	2	3	4	5	Total
Case 1	0.21	0.32	0.18	0.12	0.11	0.94
Case 2	5.89	12.34	9.74	8.37	5.92	42.26

Table 5.3 conveys that a significant reduction in CPU computation time can be achieved through the reduction of state variables. These times have been obtained by running MATLAB on a Dell computer with Intel Xeon, Quad Core processor, and 8GB (1600mHz) memory.

5.5 Conclusions

From the results presented in this chapter, it can be concluded that PCA can be used to linearly transform high-dimensional OD matrices into the lower-dimensional space without significant loss of estimation accuracy. A novel dynamic OD estimation and prediction method has been proposed, that uses the eigenvectors and principal demand components as state variables instead of OD flows. These variables can be used to construct a state space model that can be solved with recursive solution approaches such as the Kalman Filter. The proposed state space model, however, appears to be sensitive to the dimensionality reduction of the state vector due to the induced temporal measurement correlation. The chapter includes the derivation of an analytical solution for the so-called colored noise Kalman Filter algorithm that accounts for temporally-correlated measurement noise to avoid this limitation. It has been shown that a reduction of state variables in the proposed OD estimation and prediction model for large-scale networks will lead to computational efficiency with an acceptable degradation in results' accuracy.

An improvement of the algorithm presented in this chapter can be seen in two features: (1) the definition of the optimal number of principal demand components in the state vector so that the computational efficiency, results' accuracy and state observability are satisfied, and (2) adaptation of the model when additional data (i.e. speeds, density, travel times from different technological sources) can be considered to improve the quality of the estimated OD demand.

Chapter 6

Measures of performance in OD demand estimation

In this chapter we explore the application perspectives of so-called Structural SIMilarity (SSIM) index, a performance indicator that evaluates the structural similarity. By structural similarity, we mean that the spatial and temporal behavior of travelers reflected in OD trip patterns have a strong spatial and temporal correlation expressed by the OD pairs. To illustrate, if it is assumed that any prior OD matrix or available true OD matrix contains the best pattern information that we can think of, then the SSIM index can be viewed as an indication of the quality of the estimated OD matrix compared to the prior OD matrix or true OD matrix, respectively.

In the first part of the chapter we explain the concept of structural similarity in OD patterns. The theoretical background of SSIM is outlined and its main properties are explained. Furthermore, the properties of SSIM index are examined, compared to certain statistical measures. In doing so, it is shown that the statistical measures, such as MSE, are more sensitive in identifying and evaluating the structural pattern in OD matrices. Therefore, the SSIM index can be used as additional performance measure for benchmarking the dynamic OD estimation methods. According to the final objective of this section of the research - showing potential application of SSIM index as a new performance function, a new framework for the estimation of OD demand is presented.

This chapter is an edited version of the articles:

Djukic, T., S.P. Hoogendoorn, J.W.C. van Lint, *Reliability assessment of dynamic OD estimation methods based on structural similarity index*. Proceedings of the of the Transportation Research Board: 13p, 2013;

Djukic, T., J.W.C. van Lint, S.P. Hoogendoorn, *Incorporating the structural information into the estimation of OD matrices*. Proceedings of the TRISTAN conference: 5p, 2013.

6.1 The concept of structural similarity

OD matrices may be determined from different sources of information (e.g. land-use models, travel surveys, macroscopic demand simulation models, OD demand estimation methods) using various methods, and represent a common spatial and temporal behavior of travelers. These OD matrices subsume a wealth of information about relationships that affect trip making, and about their variation over space and time. The OD demand pattern is a function of the spatial and temporal distribution of travelers' activities as well as characteristics of the transportation system (Ashok et al. (1993)). Therefore, the structural relationships in OD demand can generally be explained by two types of patterns in OD demand, i.e. the temporal and spatial pattern. The temporal OD demand pattern represents the correlation of the OD demands for the same OD pair over different time periods (e.g., 8:00 - 8:15 - 8:30 - 8:45 - 9:00 a.m.). The spatial OD demand pattern refers to the correlation of the OD demands during the same time period between different OD pairs (e.g., at 8:00, at 8:15, at 8:30, at 8:45 or at 9:00 a.m.). Rules that yield the spatial correlation between OD pairs can be explained on a simple network example, outlined in Figure 6.1.

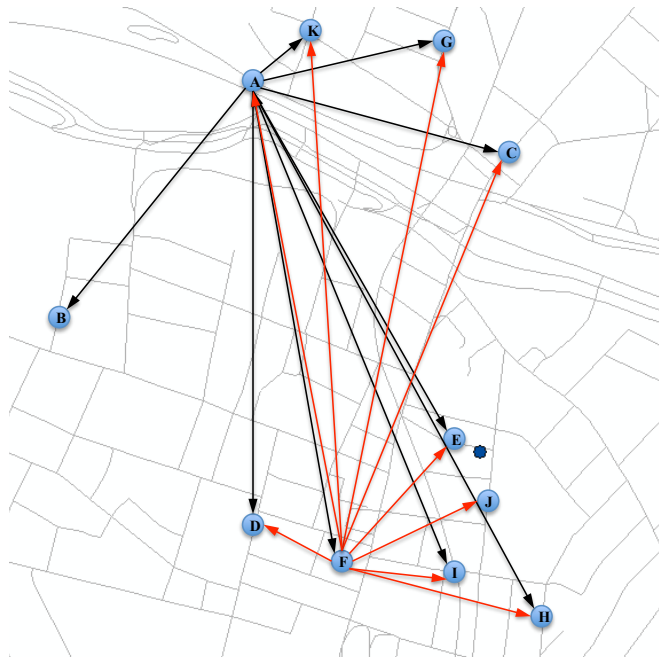
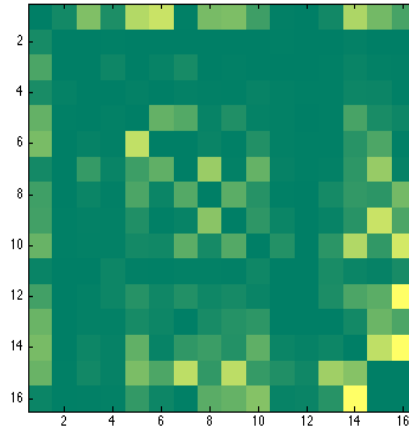


Figure 6.1: Illustration of spatial interaction between OD pairs

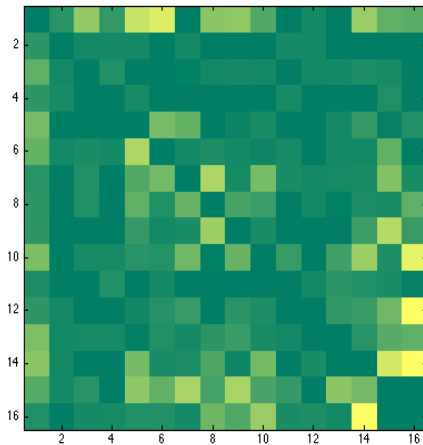
In the case of Figure 6.1, a residential area is depicted by nodes A, B, C, K, G, H and a workplace area is depicted by nodes D, F, I, E, J. The gravity model illustrates the macroscopic relationships between locations (homes and workplaces) to predict spatial interaction patterns. It has long been postulated that the interaction between two locations (e.g., A and F) declines with increasing distance, time, and cost between them, but at the same time, the interaction is positively associated with the amount of activity at each location. Furthermore, it has been recognized that the nature of an

individual's activities from which the travel demand derives is interdependent among events and spatio-temporal features of the transportation network, as an essential research foundation in activity-based travel demand modeling. In view of the above OD demand research foundations, it is clear that the correlation between OD pairs should not be overlooked. Thus, structural similarity refers to the strong spatial and temporal correlation of the spatial and temporal behavior of travelers reflected in OD matrices.

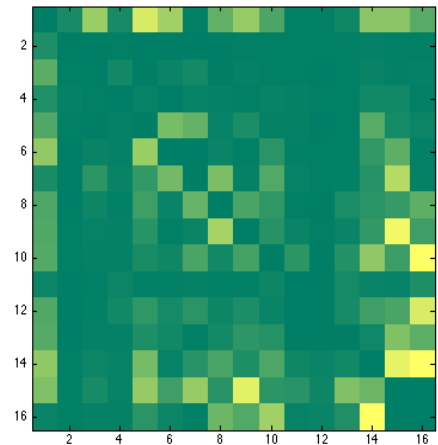
Statistical measures that ignore the spatial correlation between OD pairs in OD matrices may fail to provide effective and accurate quality measures. To show this, the mean square error (MSE) is used as a performance indicator in following example. In Figures 6.2(b) and 6.2(c), the two estimated OD matrices are compared at one time interval (at 8:00 a.m.) using two different OD estimation methods with the available ground truth OD matrix (Figure 6.2(a)).



((a)) Real OD matrix



((b)) MSE=69, SSIM=0.8724



((c)) MSE=69, SSIM=0.9702

Figure 6.2: Comparison of patterns in real and estimated OD matrices: a) real OD matrix; b) and c): estimated OD matrices that have the same MSE with respect to the real OD matrix, but different structural patterns.

For better visual examination of the pattern in the OD matrices, where the origin zones are given in rows and destinations in columns, they are represented as images where the number of trips per OD pair is used as an index to the colormap that determines the color for each OD pair. For example, if u represents the number of trips for OD pair (1,2), $x_{1,2} = u$, (Figure 6.2(a)), then the color of OD pair $x_{1,2}$ is the color represented by row u of the color map. In Figure 6.2, the OD pair with demand of 20 trips has a dark green color, and OD pair with demand of 400 trips has a light yellow color. The MSE between the ground truth OD matrix and both of the estimated OD matrices are exactly the same. However, the visual examination of the two estimated OD matrices clearly indicate that they capture different structural patterns. For example, a close look at the top left corner in Figures 6.2(b) and 6.2(a) shows that not only the number of trips, but also the spatial correlation between OD pairs in the estimated matrix is destroyed, compared to the ground truth OD matrix. Contrary to this, a pattern in the top left corner of Figure 6.2(c) resembles the same spatial correlation between OD pairs as in the ground truth OD matrix. This indicates that such statistical measures are only designed to find the distances between a pair of attributes in a data set or overall distance amongst all data.

One possible solution approach to tackle this issue is to propose a new metric that will incorporate the structural correlation between OD pairs. In this chapter, the application perspectives of the so-called Structural Similarity (SSIM) Index (Wang et al. (2004)) are explored, which is used to quantify the similarity between two images, based on the degradation of the structural information in an image compared to a reference image. Originating from image processing and analysis, the SSIM approach was motivated by the observation that images are highly structured, meaning that samples of natural images have strong dependencies. These dependencies carry important information about the structures of the objects in the visual scene. The key idea is that if the OD demand is represented in the form of a matrix, the OD pairs can be seen as pixels in an image (see Figure 6.2). The spatial correlation between OD pairs is preserved and carries information on dependencies between OD pairs. Such OD matrices are thus highly structured and stem from the combination of various kinds of information, such as OD matrix demand volume and the spatio-temporal correlation between OD pairs. When applying SSIM index to evaluate OD matrices, one needs to consider ordering of OD pairs and their physical meaning in OD matrix, and hence if it is valuable measure to compare OD matrices. Detailed analysis of SSIM index properties is required to demonstrate potential application in evaluation of OD demand.

In the next sections a definition and detailed explanation of the SSIM index and its application to OD matrices are presented. The measure ensures that the amount of structural information in a reference OD matrix is preserved in an estimated OD matrix. The suitability of the MSE and SSIM index under different scenarios are demonstrated and compared, to illustrate the key benefits of the SSIM. In addition, sensitivity of SSIM index on ordering of OD pairs in OD matrix is demonstrated.

6.2 Theoretical background of the SSIM

The SSIM index is typically used as a method for measuring the similarity between two images, based on the degradation of the structural information in one image compared to its reference. In general, this method has been employed for image and video processing in three ways. First, it can be used to monitor image quality for quality control systems. Second, it can be applied to benchmark image processing systems and algorithms. Third, it can also be embedded into image processing systems to optimize algorithms and parameter settings (Brunet et al. (2012)). Note that Wang et al. (2004) introduced the SSIM index in the context of measuring similarity to explore and compare the structural information between images, a problem that is in many respects similar to exploring the structures in OD matrices. An exciting consideration is the possibility of numerous extended applications beyond image processing, since the SSIM index does not rely on specific image or visual models. The generic definition of SSIM suggests that it should find broad applicability.

To explain the SSIM metric, a similar rationale is followed as in (Wang et al. (2004)), to represent its application within the OD demand context. Assume that the OD demand for a particular time interval k is defined by the form of the matrix where the rows of matrix represents the origins $i, i = 1, 2, \dots, I$, and columns represent the destinations j , with $j = 1, 2, \dots, J$, of trips. The SSIM index is computed within a local $N \times N$ square box, which moves cell-by-cell from the top-left to the bottom-right corner of the OD matrix, as is shown in Figure 6.3. This results in computing a SSIM index for each square box.

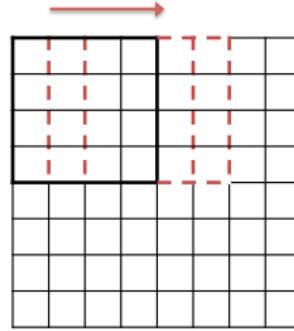


Figure 6.3: Computation of local SSIM index per sliding $N \times N$ square box

To evaluate the structural similarity between two OD vectors, let $\mathbf{x} = \{\mathbf{x}_n | n = 1, 2, \dots, 2N\}$ and $\hat{\mathbf{x}} = \{\hat{\mathbf{x}}_n | n = 1, 2, \dots, 2N\}$ be two vectors that have been extracted from the same spatial location of the square box from reference OD matrix \mathbf{X} and estimated OD matrix $\hat{\mathbf{X}}$ (as shown in Figure 6.4).

The most general form of the metric that is used to measure the structural similarity between two vectors \mathbf{x} and $\hat{\mathbf{x}}$ consists of *three main components* and is given as

$$SSIM(d, \hat{d}) = [l(d, \hat{d})^\alpha][c(d, \hat{d})^\beta][s(d, \hat{d})^\gamma] \quad (6.1)$$

	D ₁	D ₂	.	.	.	D _n
O ₁	0	20	129	30	180	202
O ₂	24	0	0	2	0	0
.	76	1	0	15	0	8
.	26	6	1	0	5	0
.	101	2	0	4	0	101
O _n	122	0	7	3	194	0

((a)) Reference OD matrix

	D ₁	D ₂	.	.	.	D _n
O ₁	0	25	99	35	185	187
O ₂	31	0	0	7	0	0
.	71	1	0	10	0	3
.	26	6	1	0	5	0
.	91	2	0	4	0	111
O _n	112	0	13	2	184	0

((b)) Estimated OD matrix

Figure 6.4: Computation of local SSIM index per sliding $N \times N$ square box in OD matrices

In this equation l is used as a distance metric to compare the mean values of the two matrices, c compares the standard deviation of the matrices, and finally, s compares the matrix structure. Now let us look at each of the components in detail. As said, the term $l(x, \hat{x})$ compares the mean values of the vectors \mathbf{x} and $\hat{\mathbf{x}}$, $\mu_x = \frac{1}{N} \sum_{n=1}^N x_n$, and is defined by the following expression

$$l(x, \hat{x}) = l(\mu_x, \mu_{\hat{x}}) = \frac{2\mu_x\mu_{\hat{x}} + C_1}{\mu_x^2 + \mu_{\hat{x}}^2 + C_1} \quad (6.2)$$

The term $c(x, \hat{x})$ compares the standard deviation (the square root of variances) of the vectors, $\sigma_x = \sqrt{\frac{1}{N} \sum_{n=1}^N (x_n - \mu_x)^2}$, and takes the similar form given by

$$c(x, \hat{x}) = c(\sigma_x, \sigma_{\hat{x}}) = \frac{2\sigma_x\sigma_{\hat{x}} + C_2}{\sigma_x^2 + \sigma_{\hat{x}}^2 + C_2} \quad (6.3)$$

Finally, the structure term $s(x, \hat{x})$ is defined as the correlation (inner product) between the normalized OD demand vectors \mathbf{x} and $\hat{\mathbf{x}}$, $\frac{x - \mu_x}{\sigma_x}$ and $\frac{\hat{x} - \mu_{\hat{x}}}{\sigma_{\hat{x}}}$, and is an effective metric for quantifying structural similarity. This is equivalent to the correlation coefficient which measures the degree of linear correlation between vectors \mathbf{x} and $\hat{\mathbf{x}}$. Geometrically, $s(x, \hat{x})$ correspond to the cosine of the angle between two vectors $x - \mu_x$ and $\hat{x} - \mu_{\hat{x}}$, independent of the lengths of these vectors. Thus, the structure term $s(x, \hat{x})$ is defined as follows:

$$s(x, \hat{x}) = s\left(\frac{x - \mu_x}{\sigma_x}, \frac{\hat{x} - \mu_{\hat{x}}}{\sigma_{\hat{x}}}\right) = \frac{\sigma_{x\hat{x}} + C_3}{\sigma_x\sigma_{\hat{x}} + C_3} \quad (6.4)$$

where $\sigma_{x\hat{x}} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(\hat{x}_i - \mu_{\hat{x}})$. The structure term $s(x, \hat{x})$ reflects the similarity between two OD demand vectors. It equals one if and only if the structures of the two demand vectors being compared are exactly the same.

The constants C_1, C_2, C_3 in equations (6.2), (6.3) and (6.4) are used to stabilize the met-

ric for the case where the means and variances become close to zero. The parameters in equation (6.1), $\alpha > 0$, $\beta > 0$ and $\gamma > 0$, are used to adjust the relative importance of the three components. In order to simplify the expression, as is recommended in (Zhou & List (2004)), the following is set: $\alpha = \beta = \gamma = 1$, and $C_3 = C_2/2$. This results in a final form of the SSIM index between two OD matrices

$$SSIM(x, \hat{x}) = \frac{(2\mu_x\mu_{\hat{x}} + C_1)(2\sigma_{x\hat{x}} + C_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + C_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + C_2)} \quad (6.5)$$

Finally, at each step the local statistics μ_x , σ_x , $\sigma_{x\hat{x}}$ and SSIM index within the square box are calculated. The measure of overall quality of the entire estimated OD matrix is given as a mean of the local SSIM indexes as

$$MSSIM(X, \hat{X}) = \frac{1}{M} \sum_{m=1}^M SSIM(x_m, \hat{x}_m) \quad (6.6)$$

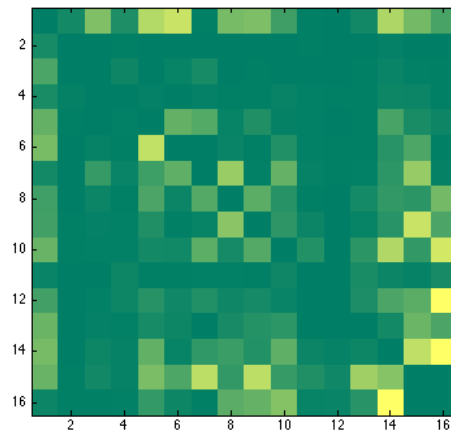
where \mathbf{X} and $\hat{\mathbf{X}}$ are the reference and the estimated OD matrices, respectively, \mathbf{x}_m and $\hat{\mathbf{x}}_m$ are the vectors of OD matrix contents at the m_{th} local square box; and M is the number of local square boxes of the entire OD matrix.

The main properties of the SSIM index can be listed as follows:

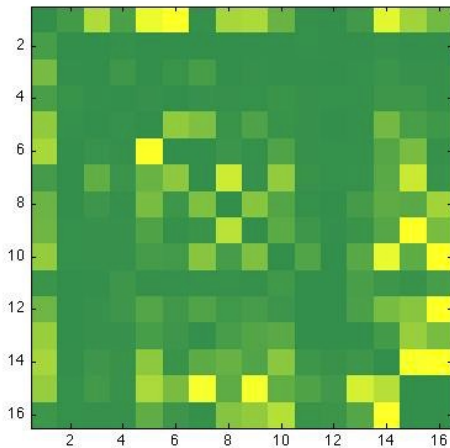
- **Symmetry:** $SSIM(x, \hat{x}) = SSIM(\hat{x}, x)$. The two OD matrices being compared give the same index value regardless of their ordering; i.e., comparing the reference OD matrix to estimated OD matrix or estimated OD matrix to reference OD matrix does not affect the resulting similarity measurement;
- **Boundedness:** $-1 \leq SSIM(x, \hat{x}) \leq 1$. This is a useful property for a metric since an upper bound can serve as an indication of how close the two OD matrices are to being perfectly identical. This is in contrast with most statistical types of measurements, which are typically unbounded;
- **Unique maximum:** The maximum value $SSIM(x, \hat{x}) = 1$ is achieved if and only if two OD matrices are exactly the same, $\mathbf{X} = \hat{\mathbf{X}}$. In other words, the SSIM measure can quantify any variations that may exist between the two OD matrices. For example, the value $SSIM(x, \hat{x}) = 0$ represents that the estimated OD matrix does not capture the spatial correlation between OD pairs as is given in the reference OD matrix.
- **Ordering:** the order of origins in rows and destinations in columns needs to reflect common properties between OD pairs to be evaluated. The order of OD pairs in reference and estimated OD matrix must be the same. Otherwise, the SSIM index will give a biased result.

6.3 The sensitivity of the MSE error and SSIM index

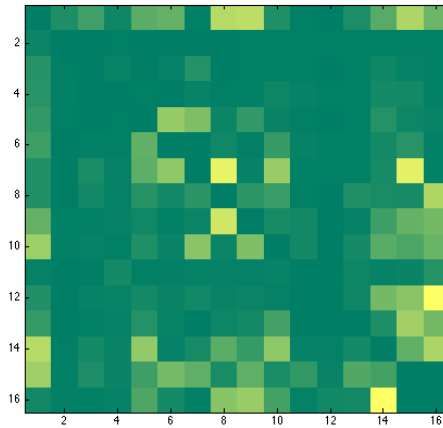
To illustrate the properties and the advantages of the SSIM index over more traditional statistical measures that are often used in sensitivity analysis of OD estimation methods or in the optimization process, the following several scenarios that reflect the importance of assumptions that an engineer is making when she/he decides to use the mean square error (MSE) are examined. For better visual examination of the structure in the OD matrices, they are represented as images where the values of the OD flows are used as indices into the color map that determine the color for each OD pair (see section 6.1).



((a)) Reference OD matrix



((b)) MSE=200, SSIM=0.9405



((c)) MSE=200, SSIM=0.7932

Figure 6.5: Comparison of patterns in reference and generated OD matrices: a) ground truth OD matrix; b and c): generated OD matrices that have the same MSE with respect to the reference OD matrix, but different structural patterns

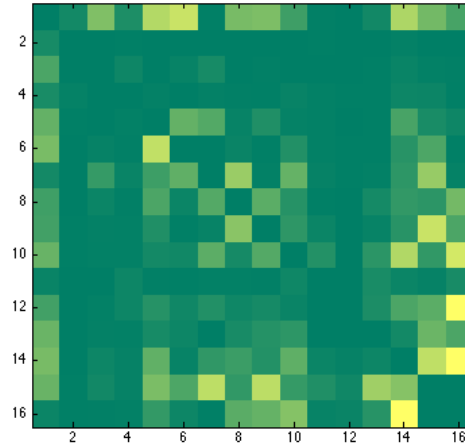
This example will show that use of the MSE error is not sufficient for researchers and practitioners to pinpoint the strengths and weakness of different OD estimation

methods. For example, the insensitivity of the statistical measure MSE to evaluate the effect of the sign, i.e., the increase or decrease in OD demand over OD pairs is demonstrated. In Figure 6.5, the first perturbed OD matrix (Figure 6.5(b)) was obtained by adding a constant value to all cells in the ground truth OD matrix (Figure 6.5(a)). The second perturbed OD matrix (Figure 6.5(c)) was generated by the same method, except that the signs of the constant were randomly chosen to be positive or negative. By means of visual inspection it is clear that the two generated OD matrices are clearly different. The first perturbed matrix (Figure 6.5(b)) resembles the ground truth matrix much more closely than the second one (Figure 6.5(c)). Yet, the MSE ignores the effect of signs, and reports the same value for both perturbed OD matrices, while the SSIM index captures the structural difference in the matrices. This result indicates that only the MSE error had been chosen to assess the performance of the two OD estimation methods, it would be possible to conclude that both methods perform similarly. In contrast, if the SSIM index was chosen as a performance criteria, one might conclude that the method with a higher SSIM index value performs better than the other.

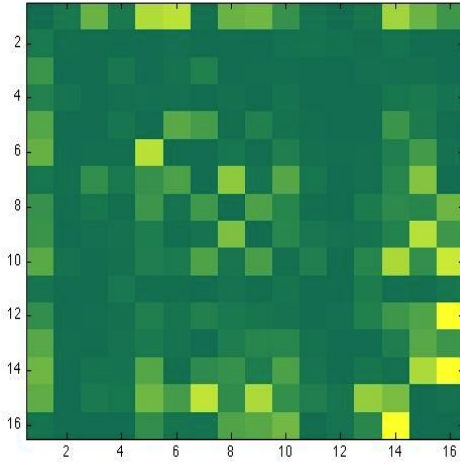
Let us consider another example, where the OD matrices have different MSE values but very similar patterns. In the Figure 6.6, both perturbed OD matrices (Figure 6.6(b)) and (Figure 6.6(c)) were obtained by adding an independent Gaussian noise to the ground truth OD matrix (Figure 6.6(a)). Apparently, the OD matrices that undergo this small geometrical modification have very large MSE values relative to the ground truth OD matrix, yet show a negligible loss of structural information. In case that OD matrices (Figure 6.6(b)) and (Figure 6.6(c)) are outputs of two different OD demand estimation methods, results indicate that the method with an estimated OD matrix (Figure 6.6(b)) performs better than method with an estimated OD matrix (Figure 6.6(c)), which is consistent with the MSE value. Therefore, the SSIM index can be used as additional criteria in performance assessments of OD estimation methods or as a performance function. The examples presented in Figure 6.5 and Figure 6.6, indicate that simplified OD estimation models with the associated assumptions about violated or ignored structural correlation may fail to provide efficient and accurate estimates of OD demand.

6.4 The sensitivity of the SSIM index on OD pairs ordering

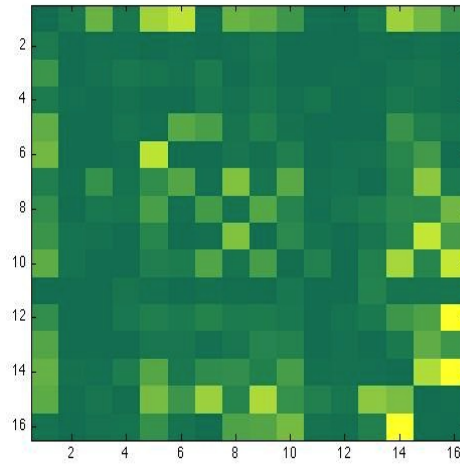
Potential limitation of SSIM index in equation (6.5) to evaluate similarity between two images is that the ordering of the image samples carries important perceptual structural information. Such implicit underling assumption is a critical issue behind the philosophy of the SSIM index approach, which attempts to distinguish structural and non-structural deviations between closely related pixels. For example, any one pixel in an image is likely to be closely related to the six pixels that surround it, but that pixel is unlikely to be related to one which is a long distance away. Clearly, this cri-



((a)) Reference OD matrix



((b)) MSE=157, SSIM=0.904



((c)) MSE=326, SSIM=0.904

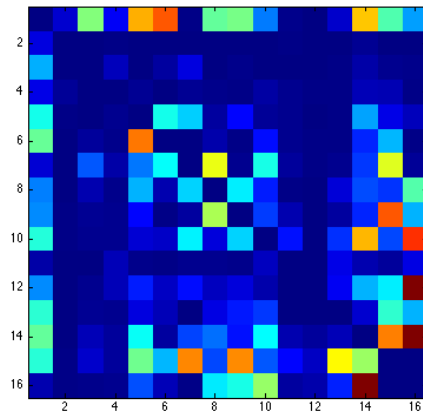
Figure 6.6: Comparison of patterns in reference and estimated OD matrices: a) "ground truth" OD matrix; b and c): estimated OD matrices that have the same SSIM with respect to the reference OD matrix, but different MSE values.

teria substantially differs from criteria used to represent OD demand per time interval in a matrix form whose rows and columns represent origin and destination nodes and elements on diagonal are equal to zero (i.e., $x_{i,j} = 0$, for $i = j$). In this case, the one OD pair is not correlated only with OD pairs that surround it, but with other OD pairs in a whole OD matrix.

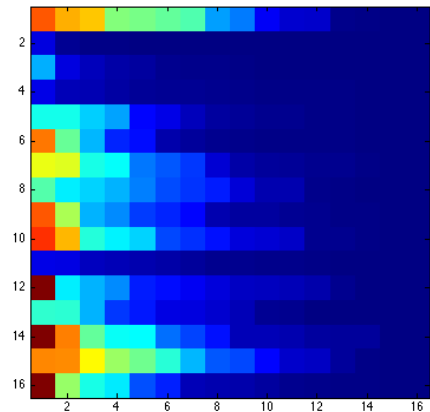
To illustrate the sensitivity of the SSIM index on OD pairs ordering in OD matrix, the following several scenarios that reflect sensitivity of the SSIM index on ordering of OD pairs and size of the sliding window $N \times N$ are examined. In this case, the OD matrices are represented as images for better visual examination as described in Section 6.1.

In Figure 6.7, the bottom-left prior OD matrix was generated by adding structural and random perturbations to the ground truth OD matrix (Figure 6.7(a)) as described in

Section 3.3.2. In the top-right image (Figure 6.7(b)), the spatial ordering of the OD pairs in ground truth OD matrix was changed by sorting OD pairs flow volume in descending order. This rational is also complementary with findings in Chapter 4, where OD pairs with high volume are dominated by structural temporal trends. As the OD demand volume per OD pair tends to decrease, the contribution of stochastic temporal trend increase. In this way, OD pairs that are closely related share similar properties (e.g. attraction, production, travel costs, distance) reflected in OD flow volume. The bottom-right image (Figure 6.7(d)) was obtained by applying the same reordering procedure to the prior OD matrix (bottom-left).



((a)) Reference OD matrix



((b)) Reordered reference OD matrix

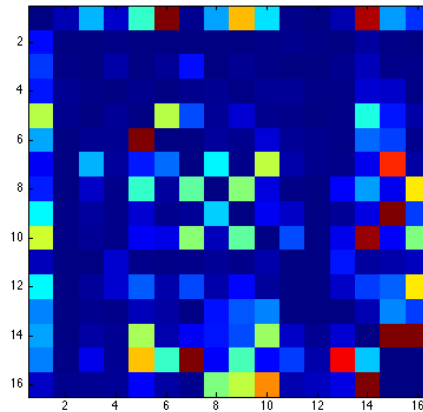
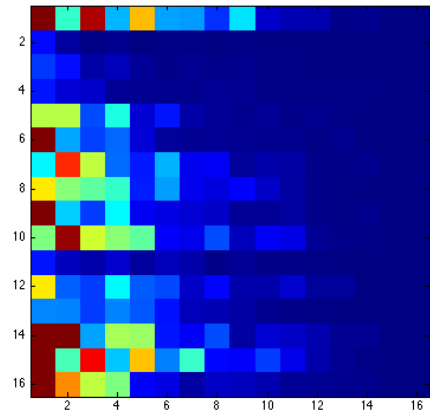
((c)) $MSE=638$, $SSIM_{6 \times 6}=0.8981$ ((d)) $MSE=638$, $SSIM_{6 \times 6}=0.8615$

Figure 6.7: Sensitivity of SSIM index on OD pairs ordering: a) "ground truth" OD matrix; b) reordered "ground truth" OD matrix; c) prior OD matrix and d) reordered prior OD matrix

As presented in previous section, the MSE error fails to take into account the spatial dependences between OD pairs. Clearly, in this example, the MSE error between two left OD matrices and two right OD matrices is same. However, the SSIM index

value obtained for window size 6×6 between two left OD matrices and two right OD matrices is slightly different (see Figure 6.7). This result indicates sensitivity of SSIM index on OD pairs ordering in OD matrix and potential shortcoming for its application as a quality measure. However, potential shortcoming of SSIM index may be overcome by the choice of the $N \times N$ sliding window size or by ordering of OD pairs such that closely related OD pairs in OD matrix share common properties as presented above.

The choice of the $N \times N$ sliding window size provides a balance between SSIM's ability to adapt to local image statistics and its ability to accurately compute the statistics within an image patch (Brooks et al. (2008)). Since the aim is to evaluate the correlation between all OD pairs, one can choose sliding window size such that $N = I$ to evaluate structural similarity between two OD matrices. Results presented in Table 6.1 indicate that window size defined by the size of the OD matrix allows accurate statistical estimation at the cost of being less sensitive to fine correlation distortions in OD matrices captured by $SSIM_{6 \times 6}$. Thus, with SSIM index one can, for example, determine the size of the sliding window as $N = I$ or consider the ordering of OD pairs in matrix such that closely related OD pairs share common properties. Since structure term $s(x, \hat{x})$ defines similarity between OD matrices and holds same properties as correlation measure, the performance of SSIM index is compared with normalized correlation between reference OD matrix and prior OD matrix. Clearly, results given in Table 6.1 indicate potential application of this measure in evaluation of correlation between OD pairs.

Table 6.1: Sensitivity of SSIM index on OD pairs ordering in OD matrix

	MSE	$SSIM_{6 \times 6}$	$SSIM_{16 \times 16}$	NormCorrelation
Reference OD vs prior OD	638	0.8981	0.9074	1.1128
Reordered reference OD vs prior OD	638	0.8615	0.9074	1.1128

The next section presents the application perspectives of the SSIM index as a performance indicator in addition to statistical measures for performance assessment of dynamic OD demand estimation methods. The second part of the section provides insight into how the conventional dynamic OD demand estimation methods can be further extended and enhanced by applying normalized correlation measure formulated in SSIM index equation (6.5).

6.5 Towards a new research directions

In this section, SSIM index directions for application and research are outlined, with a particular focus on SSIM as a performance indicator, while also incorporate the index into proposals for new measures of error.

6.5.1 SSIM index as a performance indicator

An important and often-overlooked aspect in the benchmarking of different OD estimation methods is the selection of an appropriate performance indicator or set of criteria that can be used to evaluate the quality of estimated OD matrices. The choice of performance criteria plays an important role in benchmarking of dynamic OD demand estimation methods under input data uncertainty. Few studies have focused on evaluation of the reliability and accuracy of the estimated OD matrices in the absence of a ground truth OD matrix (Bierlaire (2002), Yang et al. (1992)), and with an available ground truth OD matrix (Marzano et al. (2009), Djukic et al. (2011)).

A number of statistical measures have been proposed and used in literature to evaluate the quality of an OD estimator, such as root mean square error, normalized root mean square error and mean percentage error. These statistical measures are widely used because they are simple to calculate and have clear meanings. However, the basic foundation of these performance indicators is that they are based on pointwise deviations in terms of OD demand or traffic counts in respect to available ground truth data. Although the underlying rationale makes sense intuitively, the actual statistical measures in literature do not capture the important aspect of the structural similarity of the estimated and reference OD matrix. This sensitivity of statistical measures may fail to provide effective and accurate quality measures.

Therefore, the key idea underpinning the present approach is to define a quality metric, in such way that the spatial correlation in the estimated OD matrix is quantified adequately. The previous section showed that the SSIM index can be used to quantify how well the estimated OD matrices capture the pattern structure of the true OD matrix. To illustrate, if we assume that any prior OD matrix or available true OD matrix contains the best pattern information that we can think of, then the SSIM index can be viewed as an indication of the quality of the estimated OD matrix compared to the prior OD matrix or true OD matrix, respectively. Using the SSIM index as a new performance indicator for performance assessment of dynamic OD estimation methods is proposed. When applying the SSIM index one needs, for example, to determine the size of the sliding window as $N = I$ or to consider the ordering of OD pairs in matrix such that closely related OD pairs share common properties. The new metric would provide researchers and practitioners with better insight into how to assess the quality of the estimated OD matrix, and how to draw strict conclusions about the quality and efficiency of OD estimation methods. More specifically, it is argued that traffic engineers must rethink whether statistical measures such as MSE or RMSE are the most useful criteria of choice in their comparative studies and applications.

6.5.2 SSIM index as the error measure

In conventional dynamic OD demand estimation methods, the common approach is to select the optimal OD matrix from multiple solution candidates in terms of certain

selection of error measures; thus, final estimation and predictions are then based on the selected error measure. As indicated in the literature review in Chapter 2, these measures of error depend on a particular estimation framework, such as maximum likelihood, generalized least square, Bayesian inference, etc. For example, the generalized least square function (Bell (1991), Cascetta et al. (1993), Antoniou et al. (2006)) is equal to the Mahalanobis distance measure. This is in turn often transformed into a Euclidean distance measure under common simplifying assumptions, such as the assumption that the variance covariance matrix is equal or diagonal to the identity matrix. The models that aim to minimize the Euclidean distance between the prior OD and estimated OD matrix can provide unrealistic estimation results because too little information is taken from the prior OD matrix. Simplified models with the associated assumptions about violated or ignored spatial correlation between different OD pairs may fail to provide effective and accurate estimates of OD demand. Ashok et al. (1993) first addressed the structural information problem in their research, and proposed the use of deviations from OD flows instead of the actual OD flows as the state variables.

Figures 6.6 and Figure 6.5 demonstrate that an Euclidean distance measure as a performance function is only designed to find the distances between a pair of attributes in a data set, or overall distance among all data. These measures are unable to find and distinguish different correlation structures in data, and cannot provide an in-depth explanation of the patterns in OD demand.

In this subsection we are particularly interested in formulating a concept for a new performance function to estimate dynamic OD matrices, which would ensure that the amount of structural information in the prior OD matrix is preserved in the estimated OD matrix. One possible solution approach is to use a performance function that incorporates the spatial correlation between OD pairs within one time interval, and that allows the modeler to control the trade-off between simplicity of the model and the level of realism. The potential application in forming a background to the present derivation is the generic formulation of an dynamic OD demand estimation problem that incorporates the spatial correlation in a performance function as a penalty factor. For example, if an dynamic OD demand estimation method given by equation (2.5) it is assumed to exist, the workings of which are not explicitly considered, the new performance function can be interpreted as consisting of two elements:

$$\hat{x}_k = \underset{x \geq 0}{\operatorname{argmin}} g(\sigma_{\tilde{x}_k}, \sigma_{x_k}) [\alpha f(x_k, \tilde{x}_k) + (1 - \alpha) f(\hat{A}_h^k x_k, y_k)] \quad (6.7)$$

where the key feature is the function $g(\sigma_{\tilde{x}_k}, \sigma_{x_k})$ added in front of equation (2.5). This function should serve the propose of identifying and distinguishing the spatial correlations between vectors of prior and estimated OD demand, so as to scale the objective function f in a way that spatial correlation distortions are penalized more. Motivated by the normalized correlation formulation in SSIM index equation (6.5), the function

$g(\sigma_{\tilde{x}_k}, \sigma_{x_k})$ can be formulated as

$$g(\sigma_{\tilde{x}_k}, \sigma_{x_k}) = \frac{(\sigma_{\tilde{x}_k}^2 + \sigma_{x_k}^2 + C)}{(2\sigma_{\tilde{x}_k x_k} + C)} \quad (6.8)$$

where $\sigma_{\tilde{x}_k x_k}$ represents correlation between the normalized OD demand vectors \tilde{x}_k and x_k , and the constant C is included to avoid instability when correlation index is close to 0. This function is lower-bounded by 1, when $\sigma_{\tilde{x}_k}$ and σ_{x_k} are fully correlated, or in other words, when correlation between OD pairs in estimated OD demand is same as in prior OD demand. The estimated OD matrix that has a significantly different structure than the prior OD matrix has a higher $g(\sigma_{\tilde{x}_k}, \sigma_{x_k})$ and therefore gives more penalty to performance function in equation (6.7).

The formulation of the new performance function for the estimation of dynamic OD demand given in equation (6.7) has several advantages over existing dynamic OD estimation and prediction methods: (1) most importantly, it introduces additional information in the estimation process on the basis of structural patterns in the OD matrix. This allows the selection of the most likely best estimated OD matrix, taking into account both the estimation error as well as the structural similarity between OD matrices; (2) information on the spatial correlation between OD pairs is included when estimating the OD demand to rule out unrealistic estimation results caused by too little information being incorporated from prior OD demand; (3) the approach can be used in combination with a weighted performance function, where the weighting value is determined by the analyst's relative confidence in quality of prior OD demand or other traffic data.

Note that the purpose of this section is to demonstrate the potential application of SSIM index properties as an error measure and describe the main features of the novel performance function in estimation of dynamic OD matrices. The presented framework remains academic in nature and must be interpreted as presentation of a concept. The possibility of accomplishing dynamic OD demand estimation defined in equation (6.7) exposes important further research questions.

6.6 Discussion

In this chapter, an emerging alternative performance measure, SSIM index, has been proposed and reviewed, and its potential application to a wide variety of problems in OD demand estimation and prediction has been discussed.

The chapter includes a discussion of the potential of using the SSIM index as a quality measure that quantifies the similarity between two OD matrices such as between an OD matrix estimate and a reference OD (e.g., the prior OD matrix, or ground truth OD matrix). The most important feature of the new metric is that it includes additional information regarding structural patterns of OD matrices, both in a spatial and temporal

sense. This quality metric has been shown to have several advantages over existing statistical measures that measure pointwise deviations between two OD matrices. As an example, SSIM appears to be more sensitive in capturing the structural correlation between OD pairs; it ignores the effect of the signs of the error in estimated OD matrix. Further, the SSIM index has been proposed as a performance indicator, in addition to existing statistical measures, in benchmarking studies.

Since the final objective of this section of the research was to demonstrate the potential application of SSIM index as a new performance function, a new framework for the estimation of dynamic OD demand was presented. This allows the optimal OD matrix to be determined, taking into account both the estimation error as well as the structural similarity between the OD matrix to be estimated and prior OD matrix. The presented framework is still theoretical in nature and must be interpreted as such. In particular, the mathematical characteristics of the measure calls for further investigation (e.g., non-convexity) and its implications for estimation need to be formulated and possibly modified. New results in more realistic settings will be obtained in future research to ascertain that the method performs well in practice.

Chapter 7

Conclusions and future work directions

The motivation for the research conducted in this thesis is twofold. First, in order to obtain more efficient dynamic OD demand estimation and prediction in real time, a lower-dimensionality of the OD demand should be taken into account without significant loss of accuracy. Secondly, policy makers and ITS operators are interested in the effects of dynamic traffic management measures, e.g., to achieve more efficient large-scale transport networks at a within-day level, through better individualized travel advice. They like to know what will happen to the performance of the transport network if efficient OD demand tools are used in traffic management, and what the confidence bounds are around estimated and predicted OD demand. For this purpose, this thesis presents a real-time dynamic OD demand estimation and prediction method for dynamic traffic management applications.

This chapter begins with an assessment of the contributions of this research to the dynamic OD demand modeling approaches outlined in Section 7.1.1. Conclusions resulting from this research concerning its practical implications are presented in Section 7.1.2. Section 7.2 outlines some interesting topics and suggestions for further research.

7.1 Conclusions

This summary of the main results of the research established in the previous chapters distinguishes findings and contributions achieved in the field of OD demand estimation and prediction as well as implications for the field of traffic management.

7.1.1 Contributions achieved for dynamic OD demand modeling

In this section, the focus is on the main contributions of the research to dynamic OD demand estimation and prediction. Specifically,

- The new categorization approach based on modeling-steps with which, dynamic OD demand estimation and prediction methods are described, has been proposed. A rich variety of these methods developed so far and in use today has been reviewed. The proposed categorization approach provides better insight how various challenges within each modeling step have been tackled and how different methods relate to each other.
- The proposed dimensionality reduction method can effectively reveal structure in the underlying temporal variability patterns in dynamic OD matrices, i.e. structural, structural deviation and stochastic trends. In addition, this method offers valuable opportunities to identify how these trends contribute to each OD pair, and to significantly reduce the dimensionality of dynamic OD matrices. This method provides an effective mechanism to linearly transform the high-dimensional OD demand data into a lower-dimensional space, without significant loss of accuracy.
- The developed real-time OD estimation and prediction method provides key capability to utilize representation of OD demand in lower dimensional space, and produces a state space representation for modeling reduced OD demand flows without a complex solution approach. The proposed method represents a significant advantage over existing OD estimation and prediction methods where the size of the state vector is defined by number of OD pairs, and thus dependent on the network scale. This state vector reduction formulation significantly decreases computational complexity in both time and space dimensions and provides an efficient solution for large-scale networks.
- The benchmark task formulation provides a generic framework to account for the inherent diversity in dynamic OD demand estimation methods, where the OD estimator is considered a black box, providing a certain outcome given certain input. The simulation-based benchmark framework with the efficient sampling method can account for a wide range of different circumstances related to input

data availability and quality, and network layout. It provides valuable opportunities to pinpoint the strengths and weaknesses of existing OD demand estimation and prediction methods, and to trace a path for further improvements.

- This research provides a performance indicator to evaluate structural similarities between OD matrices. This measure quantifies dependencies between OD pairs that carry important information about the structures in OD matrices, unlike more commonly used measures of performance such as mean square error. The application of performance measure is twofold. It can be applied as an additional performance measure for performance assessment of dynamic OD estimation and prediction methods, or as a new measure of error in OD demand problem formulation.

7.1.2 Implications for dynamic traffic management

The focus of the discussion in this section are the main contributions to dynamic traffic management. Specifically,

- The modeling approach enables the consideration of different network scales, although it was developed for specific applications on large-scale networks. Given the large size of most real transport networks, computation time plays an important role in any practical implementation. In general, the computational costs of implementing dynamic OD demand estimation and prediction method proposed in this thesis seem to be a function primarily of the following parameters: a) number of state variables and traffic data observations, b) spatial distribution of the network and c) congestion level and degree of autoregressive process. The number of measurements dictate the size of the assignment matrix to be inverted in the estimation process. The spatial distribution of network and congestion level requires augmented state vector, due to the effect of travel times between OD pairs on the network. The higher degree of autoregressive process has the same effect, and implies an augmented state of higher dimension. Thus, the main advantage of the proposed method for improving computational efficiency is achieved by reduction of the state vector, i.e. state variables are defined as projections of OD demand in low-dimensional space.
- The elements of assignment matrix are obtained through iterative application of the OD estimation and prediction method and simulation-based DTA model. We note that in the event of an iterative solution technique, though convergence cannot be guaranteed, empirical study indicates that the estimation procedure is fairly robust with respect to the quality of assignment matrices obtained from simulation-based DTA model. Also, whenever more reach traffic data can be observed (e.g., travel times) or computational resources permit, one of the stochastic assignment matrix-based methods proposed by Ashok & Ben-Akiva (2002) should be used to avoid iterative process.

- The choice of length of estimation interval is defined by the domain of application (e.g., dynamic traffic management) that makes use of proposed dynamic OD estimation and prediction method.
- The insights into the performance of dynamic OD demand estimation methods resulting from qualitative assessment are of great benefit for practitioners. The benchmark tool provides and supports comparison in a variety of settings and conditions in order to help determine the particular situations and conditions under which one dynamic OD estimation method might behave more favorably than another. This will be very helpful in calibrating and validating simulation models using OD demand estimations.

The principal objective of a method solving a large-scale real-time problem for dynamic traffic management is to be able to compute a reliable solution in a small given time interval, i.e. within computational burdens. In this thesis several new methods have been proposed representing progress towards this overreaching goal. The proposed methods offer important theoretical and practical advances, as listed above, and it is envisaged that both practitioners and researchers will benefit from the results presented in this thesis.

Besides these benefits, there remain, however, unresolved issues to be addressed in the future. These problems provide topics for further research, discussed in the next section.

7.2 Future research directions

There are multiple interesting issues that require further research. Follow-up research regarding desired extensions of the proposed methods to increase their general applicability and lead to their producing more realistic outcomes, come together in one set of research directions. A second direction of research may lead to the integration of the proposed methods with new fundamental insights from advanced traffic data and theoretical advances on travel behavior. Here are some recommendations for further research:

Prediction of OD demand: As has been shown in this thesis, PCA provides tools for identification and selection of the variables (OD pairs) that explain most of the deterministic pattern in historical OD demand. The prediction of OD flows can be further extended through the application of PCA to pre-process available (e.g., estimated, historical, observed) OD demand data in following steps:

1. Identification step: first, the temporal variability patterns in historical OD demand matrices are examined and they are classified into three trend classes: a structural trend that captures the regular pattern, a large structural deviation trend

that captures the short occasional fluctuations, and a stochastic trend that capture the random fluctuations, as described in the previous section. In this manner, the principal demand components that capture the structural trend, or even OD pairs that are dominated by structural trend may be identified. For example, when the principal components with the growth trend are selected, one can identify the OD pairs that may generate a bigger demand on a network.

2. Reduction step: in this step, selection criteria of principal demand components or OD pairs are proposed, based on the trend that dominated them. For example, only the OD pairs or principal components having a structural trend and structural deviations trend will be selected.
3. Prediction step: finally, prediction methods are applied on the decreased subset of principal demand components as independent variables or on the chosen subset of OD pairs to predict OD demand for future time intervals.

In this way, a reduction of computational load in the OD demand prediction is accomplished. Further development of OD prediction models on these independent variables is recommended.

The mapping between OD flows and traffic data: The properties of the DTA simulation model used to compute the elements of assignment matrix could contribute to bias OD demand estimates and predictions, increasing computational costs. Additional forms should be investigated for the mapping of relationship between OD flows and observed traffic data that describe the temporal evolution of assignment fractions. It might also be possible to use information from traffic condition data observations in estimating, for example, assignment fractions and route-choice fractions. Integration of the theoretical and modeling advances in DTA should be incorporated into further OD demand estimation advancements.

Evaluation of the existing dynamic OD demand estimation methods: Another important set of further research directions to be addressed relate to comprehensive benchmark study of different dynamic OD demand estimation and prediction methods. Given the generic benchmark methodology presented in this thesis, one way of evaluating the accuracy and performance of the OD demand estimation methods is a joint test of their performance, i.e. comparing the estimates and predictions of link flows, travel times, etc., from the OD methods with those observed by surveillance system or generated by simulation method. A promising approach to evaluating different dynamic OD demand estimation and prediction methods jointly with traffic simulator is the use of a simulation benchmark platform. Antoniou et al. (2014) developed a benchmark platform that use the mesoscopic version of the Aimsun simulation model TSS (2013) as the common traffic model. The benchmarking platform is designed to ensure equal testing conditions for various dynamic OD demand estimation methods and in doing so supports fair comparison and an understanding of their relative merits. Such a benchmark framework can be directly applied to indicate the sensitivity of the performance of dynamic OD demand estimation and prediction methods.

Appendices

Appendix A

The Latin hypercube sampling method

To explain the method, the same rationale is followed as in (McKay et al. (1979)). Assuming there is a multivariate nonlinear model (e.g. a computer model), that calculates output $Y = h(X)$ as a function of M stochastic input variables $X = (X_1, X_2, \dots, X_M)$ with known distributions. In this case, h denotes the OD estimation approach described by equation (3.1).

The LHC method uses the following steps to approximate the posterior distribution $p(Y | X)$ of the output Y :

1. First the assigned distribution of each input is partitioned into N equal probability intervals, where N represents the desired number of generated input datasets (simulation runs). For example, if one chooses $N=10$, each input distribution will be subdivided into 10 bins according to $P(X_i \leq 0.1)$, $P(X_i > 0.1 \cap X_i \leq 0.2), \dots, P(X_i > 0.9)$. Typically N is in the order of 10-50 (resulting in a large reduction in the number of simulation runs compared to random Monte Carlo sampling, in which the number of runs is 1 or 2 orders larger).
2. Next, a uniformly random sequence of N values $U_{i1}, \dots, U_{in}, \dots, U_{iN}$ is generated for each considered input variable i . Then, the probability values P_{in} for each input $i = 1, 2, \dots, M$; and interval $n = 1, 2, \dots, N$ are computed by

$$P_{in} = (U_{in} + n - 1)/N \quad (\text{A.1})$$

This procedure ensures that exactly one probability P_{in} will fall within each of the N intervals.

3. The next step in obtaining the Latin hypercube samples is to generate the sequence of sampled values for each input X_i using the probabilities P_{in} and an inverse distribution function F_n^{-1} , which may be different for each input:

$$X_{in} = F_i^{-1}(P_{in}) \quad \forall i = 1, 2, \dots, M \quad \forall n = 1, 2, \dots, N \quad (\text{A.2})$$

4. Finally, the selected values of each input X_{in} are randomly permuted to form the N required M -dimensional input sets so that each "probability stratum" n is used only once in every input sample. This implies that a set of N Latin hypercube sample points in M -dimensional Euclidean space contains one point in each of the intervals for each of the M variables;
5. By simulating and evaluating $Y_n = h(X_n)$, $n = 1, 2, \dots, N$ for all N generated input sets, the posterior distribution in equation (3.1) can be approximated.

Notation

The main abbreviations and symbols that are used in this thesis are presented as follows:

List of abbreviations

<i>ITS</i>	Intelligent transportation systems
<i>DTMS</i>	Dynamic traffic management systems
<i>OD</i>	Origin-destination
<i>DTA</i>	Dynamic traffic assignment
<i>DNL</i>	Dynamic network loading model
<i>RC</i>	Route choice model
<i>DUE</i>	Dynamic user equilibrium
<i>SUE</i>	Stochastic user equilibrium
<i>ARIMA</i>	Auto Regressive Integrated Moving Averages
<i>PCA</i>	Principal component analysis
<i>MSE</i>	Mean square error
<i>SSIM</i>	Structural similarity index
<i>LHC</i>	Latin Hypercube method
<i>AVI</i>	Automatic vehicle identification data
<i>AVL</i>	Automatic vehicle location data
<i>FCD</i>	Floating car data
<i>FIFO</i>	First-in-first-out
<i>GLS</i>	Generalized Least Square
<i>KF</i>	Kalman filter

List of symbols

$G = (U, L)$	Directed graph with set of nodes N and links L
L	Directed links indexed by $l \in L$
\hat{L}	Directed links equipped with detectors $\hat{L} \subseteq L$
U	Set of nodes indexed by $u \in U$
O	Origin nodes indexed by $o \in O$
D	Destination nodes indexed by $d \in D$
n	Set of all OD pairs $\Omega \subseteq U \times U$
\hat{L}	Set of r links with detectors $\hat{L} \subseteq L$
K	Departure time intervals indexed by $k \in K$
H	Set of time-dependent link travel times
$\bar{G}(U \times H, \bar{L})$	time-space extension of G based on H

$\mathbf{x}_{i,k}$	Demand volume for OD pair $i \in \mathbb{R}^n$ for departure time interval k
$\tilde{\mathbf{x}}_{i,k}$	Prior demand volume for OD pair $i \in \mathbb{R}^n$ for departure time interval k
$\hat{\mathbf{x}}_{i,k}$	Estimated demand volume for OD pair i for departure time interval k
$\mathbf{y}_{\hat{L},k}$	Observed traffic counts on links $lin\mathbb{R}^r$ at the observation time interval k
$\hat{\mathbf{y}}_{\hat{L},h}$	Simulated traffic counts on link $lin\mathbb{R}^r$ at the observation time interval h
$\mathbf{A}_{i,h}^{l,k}$	Traffic assignment matrix, that is proportion of vehicles on link l at observation time interval k , coming from OD pair i departed during the time intervals h
$e_{i,k}$	Eigenvector indexed by $i \in \mathbb{R}^n$ for departure time interval k
$\lambda_{i,k}$	Eigenvalue indexed by $i \in \mathbb{R}^n$ for departure time interval k
$c_{i,k}$	Principal demand component, represent the OD demand $x_{i,k}$ projected to the orthonormal basis matrix of eigenvectors, e_i
m	Reduced number of variables in state vector, $m < n$
$\hat{\mathbf{x}}_{i,k}$	Approximated demand volume for OD pair i during the departure time interval k
$H_{i,h}^{l,k}$	Traffic observation matrix, that is transform of the traffic assignment matrix $A_{i,h}^{l,k}$ to the orthonormal basis matrix of eigenvectors, e_i
t	Prediction departure time intervals
p'	Maximum number of time intervals needed to travel between any OD pair
q'	Degree of autoregressive process

Bibliography

- Anderson, B., J. Moore (1979) *Optimal Filtering*, Englewood Cliffs, New Jersey: Prentice-Hall. 35
- Antoniou, C., M. Ben-Akiva, H. Koutsopoulos (2004) Incorporating automated vehicle identification data into origin-destination estimation, *Transportation Research Record: Journal of the Transportation Research Board*, 1882(1), pp. 37–44. 16, 53, 66
- Antoniou, C., M. Ben-Akiva, H. N. Koutsopoulos (2006) Dynamic traffic demand prediction using conventional and emerging data sources, *Intelligent Transport Systems, IEEE Proceedings*, 153(1), pp. 97–104. 18, 21, 22, 36, 66, 108
- Antoniou, C., B. Ciuffo, L. Montero, J. Casas, J. Barcelo, E. Cipriani, T. Djukic, V. Marzano, M. Nigro, M. Bullejos, J. Perarnau, M. Breen, V. Punzo, T. Toledo (2014) A framework for the benchmarking of od estimation and prediction algorithms, *Proceedings of the 93rd Annual Meeting of the Transportation Research Board, Washington, D.C.* xi, 50, 115
- Asakura, Y., E. Hato, M. Kashiwadani (2000) Origin-destination matrices estimation model using automatic vehicle identification data and its application to the han-shin expressway network, *Transportation*, 27(4), pp. 419–438. 21, 66
- Ashok, K. (1996) *Dynamic origin-destination matrix estimation and prediction*, PhD thesis, Massachusetts Institute of Technology. 16
- Ashok, K., M. E. Ben-Akiva (2000) Alternative approaches for real-time estimation and prediction of time-dependent origin-destination flows, *Transportation science*, 34(1), pp. 21–36. 21, 36, 66
- Ashok, K., M. E. Ben-Akiva (2002) Estimation and prediction of time-dependent origin-destination flows with a stochastic mapping to path flows and link flows, *Transportation science*, 36(2), pp. 184–198. 18, 25, 26, 36, 39, 113
- Ashok, K., M. E. Ben-Akiva, M. I. of Technology (1993) Dynamic origin-destination matrix estimation and prediction for real-time traffic management systems, *Transportation and traffic theory*. 34, 66, 96, 108

- Balakrishna, R. (2006) *Off-Line Calibration of Dynamic Traffic Assignment Models*, PhD thesis, Massachusetts Institute of Technology. 18, 23, 33
- Balakrishna, R., M. Ben-Akiva, H. Koutsopoulos (2007) Off-line calibration of dynamic traffic assignment: Simultaneous demand and supply estimation, *Transportation Research Record*, (2003), pp. 50–58. 26, 32, 49
- Balakrishna, R., H. Koutsopoulos (2008) Incorporating within-day transitions in the simultaneous off-line estimation of dynamic origin-destination flows without assignment matrices, *Transportation Research Record*, (2085), pp. 31–38. 33
- Barcelo, e. a., J. (2010) A kalman-filter approach for dynamic od estimation in corridors based on bluetooth and wi-fi data collection, *12th World Conference on Transportation Research WCTR*. 18, 23, 36, 66
- Barcelo, J., L. Montero, L. Marquos, C. Carmona (2010) Travel time forecasting and dynamic origin-destination estimation for freeways based on bluetooth traffic monitoring, *Transportation Research Record: Journal of the Transportation Research Board*, 2175(1), pp. 19–27. 26, 36, 49
- Bell, M. G. (1991) The real time estimation of origin-destination flows in the presence of platoon dispersion, *Transportation Research Part B: Methodological*, 25(23), pp. 115 – 125. 18, 30, 36, 108
- Bert, E. (2010) *Dynamic urban origin-destination matrix estimation methodology*, PhD thesis, Swiss Federal Institute of Technology in Lausanne. 16
- Bierlaire, M. (2002) The total demand scale: a new measure of quality for static and dynamic origin-destination trip tables, *Transportation Research Part B: Methodological*, 36(9), pp. 837–850. 44, 107
- Bierlaire, M., F. Crittin (2004) An efficient algorithm for real-time estimation and prediction of dynamic od tables, *Operations Research*, 52(1), pp. 116–127. 18, 30, 66
- Brooks, A. C., X. Zhao, T. N. Pappas (2008) Structural similarity quality metrics in a coding context: exploring the space of realistic distortions, *Image Processing, IEEE Transactions on*, 17(8), pp. 1261–1273. 106
- Brunet, D., E. R. Vrscay, Z. Wang (2012) On the mathematical properties of the structural similarity index, *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*, 21(4), pp. 1488–1499. 99
- Bryson, A. E. (2002) Applied linear optimal control, *Cambridge University Press, Cambridge, UK*, p. pp. 310312. 87, 88
- Bryson, A. E., L. J. Henrikson (1968) Estimation using sampled data containing sequentially correlated noise, *J. Spacecr. Rockets*, 56(pp.662 - 665). 86

- Budiman, M., B. M. Alex (2006) A conditioned latin hypercube method for sampling in the presence of ancillary information, *Computer geosciences*, 32(9), pp. 1378–1388. xi, 48
- Cascetta, E., D. Inaudi, G. Marquis (1993) Dynamic estimators of origin-destination matrices using traffic counts, *Transportation Science*, 27(4), pp. 363–373. 11, 18, 24, 30, 108
- Cascetta, E., M. N. Postorino (2001) Fixed point approaches to the estimation of o/d matrices using traffic counts on congested networks, *Transportation Science*, 35(2), pp. 134–147. 28
- Chang, G.-L., X. Tao (1996) Estimation of dynamic od distributions for urban networks, in: *Transportation and traffic theory: Proceeding of the 13th international symposium on transportation and traffic theory*. 18, 35
- Chang, G.-L., J. Wu (1994) Recursive estimation of time-varying origin-destination flows from traffic counts in freeway corridors, *Transportation Research Part B: Methodological*, 28(2), pp. 141–160. 18, 35, 38
- Cipriani, E., M. Florian, M. Mahut, M. Nigro (2010) Investigating the efficiency of a gradient approximation approach for the solution of dynamic demand estimation problems, in: Tampere, V. F. I. L., C.M.J., ed., *New Developments in Transport Planning : Advances in Dynamic Transport Assignment*, Edward Elgar, Cheltenham UK and Northampton, MA, USA. 26, 49
- Cipriani, E., M. Florian, M. Mahut, M. Nigro (2011) A gradient approximation approach for adjusting temporal origindestination matrices, *Transportation Research Part C: Emerging Technologies*, 19(2), pp. 270 – 282. 18, 23, 33
- Cremer, M., H. Keller (1987) A new class of dynamic methods for the identification of origin-destination flows, *Transportation Research Part B: Methodological*, 21(2), pp. 117–132. 11, 30, 38
- Dixon, M. P., L. R. Rilett (2002) Real-time od estimation using automatic vehicle identification and traffic count data, *Computer-Aided Civil and Infrastructure Engineering*, 17(1), pp. 7–21. 18, 21, 30, 36
- Dixon, M. P., L. R. Rilett (2005) Population origin-destination estimation using automatic vehicle identification and volume data, *Journal of Transportation Engineering*, 131(2), pp. 75–82. 66
- Djukic, T., J. van Lint, S. Hoogendoorn (2011) Efficient methodology for benchmarking dynamic origin-destination demand estimation methods, *Transportation Research Record: Journal of the Transportation Research Board*, 2263(1), pp. 35–44. 107

- Flötteröd, G., M. Bierlaire (2009) Improved estimation of travel demand from traffic counts by a new linearization of the network loading map, *Proceedings of the European Transport Conference (ETC)*, Netherlands. 27, 68
- Flötteröd, G., M. Bierlaire, K. Nagel (2011) Bayesian demand calibration for dynamic traffic simulations, *Transportation Science*, 45(4), pp. 541–561. 6
- Frederix, R., C. M. Tampere, F. Viti, L. H. Immers (2010) The effect of dynamic network loading models on dta-based od estimation, *New developments in transport planning: advances in dynamic traffic assignment*, Edward Elgar, Cheltenham UK and Northampton, MA, USA. 26
- Frederix, R., F. Viti, R. Corthout, C. M. Tampère (2011) New gradient approximation method for dynamic origin-destination matrix estimation on congested networks, *Transportation Research Record: Journal of the Transportation Research Board*, 2263(1), pp. 19–25. 18, 23, 27, 33
- Gan, L., H. Yang, S. C. Wong (2005) Traffic counting location and error bound in origin-destination matrix estimation problems, *Journal of Transportation Engineering*, 131(7), pp. 524–534. 53
- Genin, Y. (1970) Gaussian estimates and kalman filtering, in: C.T.Leondes, ed., *Theory & Applications of Kalman Filtering*, 139, NATO AGARDograph. 33
- Golub, G., C. van Van Loan (1996) *Matrix Computations (Johns Hopkins Studies in Mathematical Sciences)(3rd Edition)*, The Johns Hopkins University Press. 69
- Grewal, M., A. Andrews (2008) *Kalman Filtering: Theory and Practice Using MATLAB*, Wiley. 35
- Gupta, A. (2003) *Observability of Origin-Destination Matrices for Dynamic TRaffic Assignment*, Master's thesis, MIT. 26, 49
- Hess, S., K. E. Train, J. W. Polak (2006) On the use of a modified latin hypercube sampling (mlhs) method in the estimation of a mixed logit model for vehicle choice, *Transportation Research Part B: Methodological*, 40(2), pp. 147 – 163. 47
- Hotelling, H. (1933) Analysis of a complex of statistical variables into principal components, *Journal of Educational Psychology*, 24(6), pp. 417–441. 67, 68
- Hu, S.-R., C.-B. Chen (2004) Dynamic estimation of freeway origin-destination demand and travel time using extended kalman filtering algorithm, in: *Networking, Sensing and Control, 2004 IEEE International Conference on*, vol. 2, pp. 1329–1334. 36
- Hu, S.-R., S. M. Madanat, J. V. Krogmeier, S. Peeta (2001) Estimation of dynamic assignment matrices and od demands using adaptive kalman filtering, *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 6(3), pp. 281 – 300. 52

- Jaruszewicz, M., J. Mandziuk (2002) Application of pca method to weather prediction task, in: *Neural Information Processing, 2002. ICONIP '02. Proceedings of the 9th International Conference on*, vol. 5, pp. 2359–2363. 68
- Jolliffe, I. T. (2002) *Principal Component Analysis*, Springer. 68, 71, 74, 81
- Kalman, R. E. (1960) A new approach to linear filtering and prediction problems, *Transactions of the ASME - Journal of Basic Engineering*, (82 (Series D)), pp. 35–45. 34
- Kattan, L., B. Abdulhai (2006) Noniterative approach to dynamic traffic origin-destination estimation with parallel evolutionary algorithms, *Transportation Research Record: Journal of the Transportation Research Board*, 1964(1), pp. 201–210. 33, 66
- Kong, H., L. Wang, E. K. Teoh, X. Li, J.-G. Wang, R. Venkateswarlu (2005) Generalized 2d principal component analysis for face image representation and recognition, *Neural Networks*, 18(5-6), pp. 585–594. 68
- Kwon, J., P. Varaiya (2005) Real-time estimation of origin-destination matrices with partial trajectories from electronic toll collection tag data, *Transportation Research Record: Journal of the Transportation Research Board*, 1923. 18
- Lakhina, A., K. Papagiannaki, M. Crovella, C. Diot, E. D. Kolaczyk, N. Taft (2004) Structural analysis of network traffic flows, *SIGMETRICS Perform. Eval. Rev.*, 32(1), pp. 61–72. 68
- Lindveld, K., I. Netherlands Research School for Transport, Logistics (2003) *Dynamic O-D matrix estimation: a behavioural approach*, TRAIL thesis series, Netherlands TRAIL Research School. 6, 18, 30
- Martin L, H. (2008) Statistical inference for time varying origindestination matrices, *Transportation Research Part B: Methodological*, 42(6), pp. 542–552. 76
- Marzano, V., A. Papola, F. Simonelli (2009) Limits and perspectives of effective o-d matrix correction using traffic counts, *Transportation Research Part C: Emerging Technologies*, 17(2), pp. 120–132, doi: DOI: 10.1016/j.trc.2008.09.001. 44, 107
- McKay, M. D., R. J. Beckman, W. J. Conover (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, 21(2), pp. 239–245. 47, 119
- Mishalani, R., B. Coifman, D. Gopalakrishna (2002) *Evaluating Real-Time Origin-Destination Flow Estimation Using Remote Sensing-Based Surveillance Data*, chap. 80, pp. 640–647. 18, 22
- N. Caceres, F. B., J.P. Wideberg (2007) Deriving origindestination data from a mobile phone network, *IET Intelligent Transport Systems*, 1, pp. 15–26. 21

- Nguyen, S. (1977) Estimating an od matrix from network data: A network equilibrium approach, *University of Montreal Publication*, (60). 23
- Nie, Y. (2006) *A Variational Inequality approach for inferring Dynamic Origin-Destination travel demands*, PhD thesis, University of California, Davis. 16
- Nie, Y., H. M. Zhang, W. W. Recker (2005) Inferring origindestination trip matrices with a decoupled gls path flow estimator, *Transportation Research Part B: Methodological*, 39(6), pp. 497–518. 32
- Oh, C., S. Ritchie (2005) Development of methodology to design advanced traffic surveillance systems for traffic information based on origin-destination, *Transportation Research Record: Journal of the Transportation Research Board*, 1935(1), pp. 37–46. 53
- Okutani, I., Y. J. Stephanedes (1984) Dynamic prediction of traffic volume through kalman filtering theory, *Transportation Research Part B: Methodological*, 18(1), pp. 1 – 11. 34, 39
- Paige, C. C., M. A. Saunders (1982) Lsqr: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Software*, pp. 43–71. 30
- Park, B., J. Schneeberger (2003) Microscopic simulation model calibration and validation: Case study of vissim simulation model for a coordinated actuated signal system, *Transportation Research Record: Journal of the Transportation Research Board*, 1856(1), pp. 185–192. 33, 47
- Parry, K., M. L. Hazelton (2012) Estimation of origindestination matrices from link counts and sporadic routing data, *Transportation Research Part B: Methodological*, 46(1), pp. 175 – 188. 22
- Pearson, K. (1901) On lines and planes of closest fit to systems of points in space, *Philosophical Magazine*, 2(6), pp. 559–572. 67, 68
- Peeta, S., A. Ziliaskopoulos (2001) Foundations of dynamic traffic assignment: The past, the present and the future, *Networks and Spatial Economics*, 1(3), pp. 233–265. 25, 26
- Pentland, A., B. Moghaddam, T. Starner (1994) View-based and modular eigenspaces for face recognition, in: *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, pp. 84–91. 68
- Peterson, A. (2007) The origin-destination matrix estimation problem : Analysis and computations. 16
- Reuven, Y. R. (1981) *Simulation and the Monte Carlo Method*, John Wiley & Sons, Inc., 304 pp., 539488. 47

- Sherali, H. D., T. Park (2001) Estimation of dynamic origin-destination trip tables for a general network, *Transportation Research Part B: Methodological*, 35(3), pp. 217 – 235. 18, 31
- Sohn, K., D. Kim (2008) Dynamic origin destination flow estimation using cellular communication system, *Vehicular Technology, IEEE Transactions on*, 57(5), pp. 2703–2713. 36
- Spall, J. C., S. Member (1992) Multivariate stochastic approximation using a simultaneous perturbation gradient approximation, *IEEE Transactions on Automatic Control*, 37, pp. 332–341. 33
- Stathopoulos, A., T. Tsekeris (2004) Hybrid meta-heuristic algorithm for the simultaneous optimization of the od trip matrix estimation, *Computer-Aided Civil and Infrastructure Engineering*, 19(6), pp. 421–435. 33
- Sun, J., Y. Feng (2011) A novel od estimation method based on automatic vehicle identification data, in: Chen, R., ed., *Intelligent Computing and Information Science*, vol. 135 of *Communications in Computer and Information Science*, Springer Berlin Heidelberg, pp. 461–470. 22
- Sun, Z., P. Lingshan, Z. Zhang, L. Zhang, D. Qingyan (2009) Traffic flow forecast based on combining principal component analysis with support vector machine, *Journal of Highway and Transportation Research and Development*. 68, 72
- Tavana, M. H., H. (2001) Estimation of dynamic origin-destination flows from sensor data using bi-level optimization method, *Presented at the 80th annual meeting of the Transportation Research Board, Washington DC, USA*. 18, 26, 30, 31, 49
- Tsekeris, S. A., Theodore (2003) Real-time dynamic origin-destination matrix adjustment with simulated and actual link flows in urban networks, *Transportation Research Record: Journal of the Transportation Research Board*, 1857(1), pp. 117–127. 18, 37, 55
- TSS (2006) *Aimsun planner manual*, version 1.1 ed., Transport Simulation Systems, Barcelona, Spain. 20
- TSS (2013) *Aimsun 7 dynamic simulator User's Guide*, version 1.1 ed., Transport Simulation Systems, Barcelona, Spain. 89, 115
- Van Aerde, M., B. Hellings, L. Yu, H. Rakha (1993) Vehicle probes as real-time atms sources of dynamic od and travel time data, *Proceedings of the ATMS Conference*. 21
- Van Der Zijpp, N. (1997) Dynamic origin-destination matrix estimation from traffic counts and automated vehicle identification data, *Transportation Research Record: Journal of the Transportation Research Board*, 1607(1), pp. 87–94. 25

- Van Der Zijpp, N., E. De Romph (1997) A dynamic traffic forecasting application on the amsterdam beltway, *International Journal of Forecasting*, 13(1), pp. 87–103. 18, 22
- Van der Zijpp, N., C. Lindveld (2001) Estimation of origin-destination demand for dynamic assignment with simultaneous route and departure time choice, *Transportation Research Record: Journal of the Transportation Research Board*, 1771, pp. 75–82. 30
- Van Zuylen, H. J. (1978) The information minimising method: validity and applicability to transport planning., *New Developments in Modelling Travel Demand and Urban Systems (Edited by G. R. M. Jansen et al.)*, Saxon, Farnborough. 23
- Van Zuylen, H. J., L. G. Willumsen (1980) The most likely trip matrix estimated from traffic counts, *Transportation Research Part B: Methodological*, 14(3), pp. 281–293. 55
- Wang, Z., A. C. Bovik, H. R. Sheikh, E. P. Simoncelli (2004) Image quality assessment: From error visibility to structural similarity, *IEEE Transactions on image processing*, 13(4), pp. 600–612. 98, 99
- Willumsen, L. (1978) Estimation of an o-d matrix from traffic counts: a review., *Institute for Transport Studies Working Paper*, (99). 23
- Wu, J. (1997) A real-time origin-destination matrix updating algorithm for on-line applications, *Transportation Research Part B: Methodological*, 31(5), pp. 381–396. 18, 37, 55
- Wu, J., G.-L. Chang (1996) Estimation of time-varying origin-destination distributions with dynamic screenline flows, *Transportation Research Part B: Methodological*, 30(4), pp. 277–290. 52
- Yamamoto, T., T. Miwa, T. Takeshita, T. Morikawa (2009) Updating dynamic origin-destination matrices using observed link travel speed by probe vehicles, in: Lam, W. H. K., S. C. Wong, H. K. Lo, eds., *Transportation and Traffic Theory 2009: Golden Jubilee*, Springer US, pp. 723–738. 37
- Yang, H., Y. Iida, T. Sasaki (1991) An analysis of the reliability of an origin-destination trip matrix estimated from traffic counts, *Transportation Research Part B: Methodological*, 25(5), pp. 351–363. 44
- Yang, H., T. Sasaki, Y. Iida, Y. Asakura (1992) Estimation of origin-destination matrices from link traffic counts on congested networks, *Transportation Research Part B: Methodological*, 26(6), pp. 417–434. 107
- Zhou, X. (2004) *Dynamic Origin-Destination demand estimation and prediction for off-line and on-line dynamic traffic assignment operations*, PhD thesis, University of Maryland, College part. 16

- Zhou, X., G. F. List (2004) An information-theoretic sensor location model for traffic origin-destination demand estimation applications, *Transportation science*, 44(2), pp. 254–273. 101
- Zhou, X., H. S. Mahmassani (2006) Dynamic origin-destination demand estimation using automatic vehicle identification data, *Intelligent Transportation Systems, IEEE Transactions on*, 7(1), pp. 105–114. 18, 26, 31, 49, 66
- Zhou, X., H. S. Mahmassani (2007) A structural state space model for real-time traffic origin-destination demand estimation and prediction in a day-to-day learning framework, *Transportation Research Part B: Methodological*, 41(8), pp. 823–840. 18, 22, 35, 39, 40, 66, 67, 76
- Zhou, X., X. Qin, H. S. Mahmassani (2003) Dynamic origin-destination demand estimation with multiday link traffic counts for planning applications, *Transportation Research Record: Journal of the Transportation Research Board*, 1831(1), pp. 30–38. 30

About the author

Tamara Djukic was born in Cacak, Serbia, in 1980, where she completed primary school and Gymnasium. She moved to Belgrade, Serbia, in 1999 to study traffic and transportation science with the focus on urban and road transport and traffic at Belgrade University, Faculty of Transportation and Traffic Engineering. She holds a diploma and M.Sc. in urban and road traffic engineering from Belgrade University (2005). In 2005 Tamara began research in the scientific field of dynamic traffic management and control as a research associate at the Faculty of Transport and Traffic engineering, University of Belgrade.



In September 2008 Tamara started her PhD research on dynamic OD matrix estimation and prediction at the department Transport & Planning, Faculty of Civil Engineering and Geosciences of Delft University of Technology (TU Delft), the Netherlands. This project was founded by ITS Edulab, the Dutch traffic and transportation laboratory for students, a cooperation between Rijkswaterstaat center for Transport and Navigation and TU Delft. During PhD program, Tamara has participated in MULTITUDE project (COST Action TU0903) supported by the European Union COST office. Furthermore, she was a visiting scholar at Royal Institute of Technology (KTH) in Sweden. Her research interests also include the modeling and simulation of transportation systems, intelligent transport systems, and calibration and optimization applications. In her eight years of experience she has been involved in a large number of research projects in Europe. She is an active reviewer for various journals and international conferences.

Publications

Journal articles

Djukic, T., J.W.C. van Lint, S.P. Hoogendoorn (2012) Application of principal component analysis to predict dynamic origin-destination matrices, *Transportation Research Record: Journal of the Transportation research Board*, 2283(1), pp. 81-89.

Djukic, T., J.W.C. van Lint, S.P. Hoogendoorn (2011) Efficient methodology for benchmarking dynamic origin-destination demand estimation methods, *Transportation Research Record: Journal of the Transportation research Board*, 2263, pp. 5-44.

Peer reviewed conference papers

Antonioni, A., Ciuffo, B., Montero, L., Casas, J., Barcelo, J., Cipriani, E., Djukic, T., Marzano, V., Nigro, M., Bullejos, M., Perarnau, J., Breen, M., Punzo, V. & Toledo, T. (2014) A framework for the benchmarking of OD estimation and prediction algorithms. In s.n. (Ed.), *Proceedings of the 93rd Annual Meeting of the Transportation Research Board*, Washington DC, pp. 1-14.

Djukic, T., G., Flötteröd, J.W.C. van Lint, S.P. Hoogendoorn (2012) Efficient real time OD matrix estimation based on principal component analysis. In s.n. (Ed.), *Proceedings of the 2012 15th IEEE conference on intelligent transportation systems (ITSC 2012)*, Anchorage, Alaska, pp. 115-121.

Djukic, T., J.W.C. van Lint, S.P. Hoogendoorn (2013) Incorporating the structural information into the estimation of OD matrices. In s.n. (Ed.), *TRISTAN VIII, Triennial symposium on transportation analysis*, Atacama, Chile, pp. 1-5.

Djukic, T., J.W.C. van Lint, S.P. Hoogendoorn (2011) Accuracy and robustness of the dynamic OD estimation methods under input uncertainties. In s.n. (Ed.), *2nd International Conference on Models and technologies for ITS*, Leuven, pp. 1-5.

Djukic, T., J.W.C. van Lint, S.P. Hoogendoorn (2014) Methodology for efficient real time OD demand estimation on large scale networks. In s.n. (Ed.), *Transportation Research Board 93rd Annual Meeting Compendium of Papers (DVD)*, Washington DC, pp. 1-16.

Djukic, T., S.P. Hoogendoorn, J.W.C. van Lint (2013) Reliability assessment of dynamic OD estimation methods based on structural similarity index. In s.n. (Ed.), *Transportation Research Board 92nd Annual Meeting Compendium of Papers (DVD)*, Washington DC, pp. 1-13.

Djukic, T., J.W.C. van Lint, S.P. Hoogendoorn (2012) Exploring application perspectives of principal component analysis in predicting dynamic OD matrices. In s.n. (Ed.), *Transportation Research Board 91st Annual Meeting Compendium of Papers (DVD)*, Washington DC, pp. 1-14.

Djukic, T., J.W.C. van Lint, S.P. Hoogendoorn (2011) An efficient methodology for benchmarking dynamic OD demand estimation methods. In s.n. (Ed.), *Transportation*

Research Board 90th Annual Meeting Compendium of Papers (DVD), Washington DC, pp. 1-15.

Djukic, T., J.W.C. van Lint, S.P. Hoogendoorn, H.J. van Zuylen (2010) Benchmark framework for sensitivity analysis of dynamic OD demand estimation methods. In s.n. (Ed.), *Proceedings 11th TRAIL Congress*, Delft: TRAIL Research School, available online, pp. 1-6.

Djukic, T., G., Flötteröd, J.W.C. van Lint, S.P. Hoogendoorn (2012) Efficient real time OD matrix estimation based on principal component analysis. In s.n. (Ed.), *Proceedings 12th TRAIL Congress*, Delft: TRAIL Research School, available online, pp. 1-6.

The following conference papers are under review:

Djukic, T., J. Barceló, L. Montero, M. Bullejos, E. Cipriani, J.W.C. van Lint, S.P. Hoogendoorn. Advanced traffic data for dynamic OD demand estimation: The state of the art and benchmark study, in *94th Annual Meeting of the Transportation Research Board*, Washington DC, 2015. (accepted for presentation)

Book chapters

Lint, J.W.C. van, T. Djukic (2012) Applications of Kalman filtering in traffic management and control. In P. Mirchandani, & J. Cole Smith (Eds.), *New directions in informatics, optimization, logistics, and production*, Hannover: Informs, pp. 59-91.

Manuscript in preparation

Djukic, T., J.W.C. van Lint, S.P. Hoogendoorn (in progress) Dynamic OD matrix estimation and prediction: state-of-the-art.

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 100 titles see the TRAIL website: www.rsTRAIL.nl. The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on transport, infrastructure and logistics.

Djukic, T., *Dynamic OD Demand Estimation and Prediction for Dynamic Traffic Management*, T2014/9, November 2014, TRAIL Thesis Series, the Netherlands

Chen, C., *Task Complexity and Time Pressure: Impacts on activity-travel choices*, T2014/8, November 2014, TRAIL Thesis Series, the Netherlands

Wang, Y., *Optimal Trajectory Planning and Train Scheduling for Railway Systems*, T2014/7, November 2014, TRAIL Thesis Series, the Netherlands

Wang, M., *Generic Model Predictive Control Framework for Advanced Driver Assistance Systems*, T2014/6, October 2014, TRAIL Thesis Series, the Netherlands

Kecman, P., *Models for Predictive Railway Traffic Management*, T2014/5, October 2014, TRAIL Thesis Series, the Netherlands

Davarynejad, M., *Deploying Evolutionary Metaheuristics for Global Optimization*, T2014/4, June 2014, TRAIL Thesis Series, the Netherlands

Li, J., *Characteristics of Chinese Driver Behavior*, T2014/3, June 2014, TRAIL Thesis Series, the Netherlands

Mouter, N., *Cost-Benefit Analysis in Practice: A study of the way Cost-Benefit Analysis is perceived by key actors in the Dutch appraisal practice for spatial-infrastructure projects*, T2014/2, June 2014, TRAIL Thesis Series, the Netherlands

Ohazulike, A., *Road Pricing mechanism: A game theoretic and multi-level approach*, T2014/1, January 2014, TRAIL Thesis Series, the Netherlands

Cranenburgh, S. van, *Vacation Travel Behaviour in a Very Different Future*, T2013/12, November 2013, TRAIL Thesis Series, the Netherlands

Samsura, D.A.A., *Games and the City: Applying game-theoretical approaches to land and property development analysis*, T2013/11, November 2013, TRAIL Thesis Series, the Netherlands

- Huijts, N., *Sustainable Energy Technology Acceptance: A psychological perspective*, T2013/10, September 2013, TRAIL Thesis Series, the Netherlands
- Zhang, Mo, *A Freight Transport Model for Integrated Network, Service, and Policy Design*, T2013/9, August 2013, TRAIL Thesis Series, the Netherlands
- Wijnen, R., *Decision Support for Collaborative Airport Planning*, T2013/8, April 2013, TRAIL Thesis Series, the Netherlands
- Wageningen-Kessels, F.L.M. van, *Multi-Class Continuum Traffic Flow Models: Analysis and simulation methods*, T2013/7, March 2013, TRAIL Thesis Series, the Netherlands
- Taneja, P., *The Flexible Port*, T2013/6, March 2013, TRAIL Thesis Series, the Netherlands
- Yuan, Y., *Lagrangian Multi-Class Traffic State Estimation*, T2013/5, March 2013, TRAIL Thesis Series, the Netherlands
- Schreiter, Th., *Vehicle-Class Specific Control of Freeway Traffic*, T2013/4, March 2013, TRAIL Thesis Series, the Netherlands
- Zaerpour, N., *Efficient Management of Compact Storage Systems*, T2013/3, February 2013, TRAIL Thesis Series, the Netherlands
- Huibregtse, O.L., *Robust Model-Based Optimization of Evacuation Guidance*, T2013/2, February 2013, TRAIL Thesis Series, the Netherlands
- Fortuijn, L.G.H., *Turborotonde en turboplein: ontwerp, capaciteit en veiligheid*, T2013/1, January 2013, TRAIL Thesis Series, the Netherlands
- Gharehgozli, A.H., *Developing New Methods for Efficient Container Stacking Operations*, T2012/7, November 2012, TRAIL Thesis Series, the Netherlands
- Duin, R. van, *Logistics Concept Development in Multi-Actor Environments: Aligning stakeholders for successful development of public/private logistics systems by increased awareness of multi-actor objectives and perceptions*, T2012/6, October 2012, TRAIL Thesis Series, the Netherlands
- Dicke-Ogenia, M., *Psychological Aspects of Travel Information Presentation: A psychological and ergonomic view on travellers' response to travel information*, T2012/5, October 2012, TRAIL Thesis Series, the Netherlands
- Wismans, L.J.J., *Towards Sustainable Dynamic Traffic Management*, T2012/4, September 2012, TRAIL Thesis Series, the Netherlands
- Hoogendoorn, R.G., *Swiftly before the World Collapses: Empirics and Modeling of Longitudinal Driving Behavior under Adverse Conditions*, T2012/3, July 2012, TRAIL Thesis Series, the Netherlands
- Carmona Benitez, R., *The Design of a Large Scale Airline Network*, T2012/2, June 2012, TRAIL Thesis Series, the Netherlands



Summary

Dynamic origin-destination (OD) demand is important input to many simulation models applied within dynamic traffic management systems (DTMS) for predicting traffic states on the network. The inability to provide high-quality dynamic OD demand estimates makes prediction with simulation models simply impossible, irrespective of how well these models have been calibrated. This thesis presents methods regarding the provision of efficient and reliable dynamic OD demand information for DTMS applications.

About the Author

Tamara Djukic conducted her PhD research at Delft University of Technology. She holds a MSc degree in Civil Engineering with specialization in road traffic and transport. Her research interests include the traffic state estimation and prediction and data processing.

TRAIL Research School ISBN 978-90-5584-179-0