MP: $\hat{T} = \min_{\Lambda} z(T(\Lambda))$

EC: $\langle G(F(T)), [F(T) - F^*(T)] \rangle + \langle J(T), [T - T^*] \rangle \geq 0 \quad \forall F, T \in \Omega$
Stellingen
Dynamic O-D matrix estimation; a behavioural approach.
Ch.D.R. Lindveld, 2 september 2003

1. Hoewel verkeerssystemen in de praktijk nooit in hun evenwichtstoestand zullen worden aangetroffen, leveren dynamische en statische evenwichtsmodellen toch essentiële inzichten voor de bestudering van dergelijke systemen. (hoofdstuk 2)

2. Voor het schatten van dynamische herkomst-bestemmings matrices voor planningsdoeleinden is een toedeling noodzakelijk die uitgaat van evenwicht in de simultane keuze van route en vertrektdistheid. (paragraaf 3.4)

3. Voor het schatten van dynamische herkomst-bestemmings matrices is de kalibratie van de hiervoor benodigde dynamische verkeerstoedelingen, en speciaal de reistijdmodellering, even belangrijk als de kalibratie van de routespreiding. (paragraaf 5.2)

4. Het onderbepaalde schattingsprobleem van een dynamische herkomst-bestemmings matrix kan, door parametrisatie van onbekende herkomst-bestemmings matrix door middel van een apriori matrix en een parametervector, worden teruggebracht tot het goedbepaalde probleem van schatting van de parametervector. (paragraaf 4.4.3)

5. Het verdient aanbeveling om in vertrektdistipkeuzemodellen niet alleen rekening te houden met het gewenste aankomsttijdstip maar ook met het gewenste vertrektdistijdstip

6. Conceptuele en wiskundige modellen vertegenwoordigen de enige bron van kennis over latente grootheden in het verkeerssysteem en verdienen alleen al op die grond waardering.

7. Voor een juiste toepassing van uitkomsten van kwantitatieve modellen ter onderbouwing van besluiten is het van belang om naast de modeluitkomst een indicatie van de onzekerheid van die uitkomst te geven, bijvoorbeeld door een betrouwbaarheidsinterval.

8. Bestaande discrete-keuzemodellen binnen de vervoersmodellering kunnen worden verbeterd door onbetrovwaarder en onzekerheid van aspecten van keuzealternatieven op te nemen als verklarende variabele in de nutsfunctie.

9. De kwaliteit en kosten-effectiviteit van wetgeving kunnen significant verbeterd worden door het beoogde doel, gedragsverandering na invoering ervan, en de kosten van invoering en handhaving vooraf te bestuderen met kwantitatieve methoden, zoals thans gebruikelijk is bij verkeerskundige ingrepen.

10. Het verbod op het doorbreken van digitale beveiligingen, zoals geformuleerd in de "Digital Millennium Copyright Act", de hierdoor geïnspireerde Europese richtlijn inzake copyright, en de neerslag daarvan in het voorstel tot wijziging van de "Auteurswet 1912" slaan te ver door in de bescherming van auteursrechthebbenden vanwege de vele mogelijkheden tot oneigenlijk gebruik van dit verbod en de feitelijke ontkrachting van het recht van "fair use".

Deze stellingen worden verdedigbaar geacht, en zijn als zodanig goedgekeurd door de promotor Prof. dr. ir. P.H.L. Bovy.

---


Propositions
Dynamic O-D matrix estimation; a behavioural approach.
Ch.D.R. Lindveld, 2 September 2003

1. Although in practice transportation systems will never be observed in their equilibrium state, dynamic and static equilibrium models nonetheless provide essential insights into such systems. (chapter 2)

2. In order to estimate dynamic origin-destination matrices for transportation planning purposes, an equilibrium assignment with respect to simultaneous choice of route and departure time is required. (section 3.4)

3. For the estimation of dynamic origin-destination matrices, the calibration of the required dynamic assignment and, specifically, its travel time functions, is as important as the calibration of the route choice model. (section 5.2)

4. The underdetermined dynamic origin-destination matrix estimation problem can, through parametrisation of the unknown origin-destination matrix by an apriori matrix and a parameter vector, be reduced to the well-posed problem of estimating the parameter vector. (section 4.4.3)

5. It is advisable for departure time choice models to account for both the preferred time of arrival and the preferred time of departure3.

6. As conceptual and mathematical models constitute the only source of knowledge on latent variables in the transportation system, they deserve recognition on that ground alone.

7. Correct application of the results of quantitative models to support policy decisions is served by providing a measure of uncertainty of the model outcomes in addition to the model outcomes themselves, for example, through confidence intervals.

8. Existing discrete choice models for transportation modelling may be improved by incorporating a measure of the uncertainty and unreliability of aspects of the choice alternatives as explanatory variables into the utility function.

9. Both the quality and the cost-effectiveness of legislation may be significantly improved by studying the intended effect, behavioural change after introduction of the legislation, as well as the cost of enforcement, beforehand through quantitative methods, as is usual for transportation facilities.

10. The ban on circumvention of digital protection measures, as formulated in the "Digital Millennium Copyright Act", the European copyright directive inspired by that Act, and the reflection thereof in the proposal to change the "Auteurswet 1912" goes too far in its protection the interests of copyright owners, in view of the many opportunities for misuse of this ban and its de-facto negation of the right of "fair use".

These propositions are considered tenable and have been approved by the thesis supervisor professor dr. ir. P.H.L. Bovy.

---


Dynamic O-D matrix estimation
A behavioural approach
Dynamic O-D matrix estimation
A behavioural approach

Proefschrift

ter verkrijging van de graad van doctor aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op dinsdag 2 september 2003 om 13.00 uur
door Charles Dirk Rudolf LINDVELD
doctorandus in de wiskunde
egenomen te Eindhoven
Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. P.H.L. Bovy

Samenstelling promotiecommissie:
Rector Magnificus, voorzitter
Prof. dr. ir. P.H.L. Bovy, Technische Universiteit Delft, promotor
Prof. dr. H.J. van Zuylen, Technische Universiteit Delft
Prof. dr. I. Chabini, Massachusetts Institute of Technology, USA
Prof. dr. ir. R. Dekker, Erasmus Universiteit Rotterdam
Prof. A. Daly, University of Leeds, UK
Prof. dr. M.G.H. Bell, Imperial College London, UK
Prof. dr. ir. M.F.A.M. Van Maarseveen, Universiteit Twente

Trail Thesis Series, The Netherlands TRAIL Research School
This thesis is the result of a Ph.D. study carried out from 1998 to 2002 at Delft University of Technology, Faculty of Civil Engineering and Geosciences, Transportation Planning and Traffic Engineering Section.

ISBN 90-5166-980-1

Published and distributed by:
Uitgeverij Eburon
Postbus 2867
2601 CW Delft
info@eburon.nl
www.eburon.nl

© 2003 by Charles Lindveld. Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij electronisch, mechanisch, door fotocopieën, opnamen of op enig andere manier, zonder voorafgaande schriftelijke toestemming van de rechthebbende(n).

© 2003 by Charles Lindveld. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission in writing from the copyright holder.

Printed in The Netherlands
Preface

I would like to express my thanks and gratitude to my thesis supervisor, Piet Bovy, for giving me the opportunity to undertake this work, and for his patience and support during its completion. I am indebted to several members of the Transport and Traffic Engineering section of the Delft University of Technology for discussions and advice.

I gratefully acknowledge the financial support of the Regiolab project, sponsored by the Adviesdienst Verkeer en Vervoer, during the final year of my thesis research.

In addition I would like to express my gratitude to prof. Ismail Chabini for giving me access to the software for dynamic traffic assignment developed by him at MIT. Furthermore I gratefully acknowledge the permission of the Adviesdienst Verkeer en Vervoer to use the results of the survey and the model estimation carried out to update the time-of-day choice module of the Dutch National Model system (the "LMS").
# CONTENTS

**Preface** v

**Notation** xvii

1 **Introduction** 1
   1.1 Context and general considerations .......................... 1
   1.2 Issues involved ............................................. 2
      1.2.1 Conceptual hierarchy ................................... 2
      1.2.2 Discussion of issues involved .......................... 3
   1.3 Delineation of the issues addressed in this thesis .......... 7
      1.3.1 Issues addressed in this thesis ......................... 7
      1.3.2 Issues not addressed in this thesis .................... 8
   1.4 Contributions of this thesis and practical significance ... 9
      1.4.1 Main contributions ........................................ 9
      1.4.2 Secondary contributions ................................... 9
      1.4.3 Practical significance .................................... 10
   1.5 Thesis outline ............................................... 10

2 **Context and relevance of dynamic O-D matrix estimation** 13
   2.1 Introduction .................................................. 13
   2.2 The context of dynamic O-D matrix estimation ............... 13
      2.2.1 Quantitative transportation models .................... 14
      2.2.2 Applications and examples of transportation models ... 14
   2.3 Dynamic versus static transportation models ................ 15
      2.3.1 Dynamics in transportation modelling .................... 15
      2.3.2 Dynamic O-D matrix estimation ........................... 21
   2.4 Equilibrium modelling and existence of the O-D matrix .... 22
      2.4.1 Equilibrium and game theory ............................. 22
      2.4.2 Equilibrium and projected dynamical systems .......... 24
      2.4.3 Consequences for the equilibrium concept ............. 24
      2.4.4 Equilibria and reality ................................... 24
      2.4.5 The relevance of equilibrium models .................... 27
   2.5 Summary ...................................................... 28

3 **A behavioural framework for dynamic O-D matrix estimation** 31
   3.1 Introduction .................................................. 31
   3.2 State of the art ............................................... 32
      3.2.1 Literature ................................................ 32
3.2.2 Existing dynamic O-D matrix estimation methods .......................... 32
3.2.3 Time discretisation ........................................................................ 34
3.3 Transportation modelling framework used ...................................... 35
3.4 Estimation of dynamic O-D matrices from observations ................... 38
3.5 Dynamic O-D matrix estimation using a behavioural model structure .... 41
3.6 Rigid and elastic demand ................................................................... 43
3.7 Grouped observations ........................................................................ 44
3.7.1 The static case .............................................................................. 44
3.7.2 Extension of grouped observations to the dynamic case ................. 44
3.7.3 Definitions and notation ................................................................. 46
3.7.4 Fixed datagroups .......................................................................... 48
3.7.5 Assignment-dependent datagroups ............................................... 48
3.8 Unification of mode-choice and departure time choice .................. 49
3.8.1 STENs ....................................................................................... 50
3.8.2 Modelling departure time choice as route choice in an augmented
STEN ................................................................................................. 52
3.9 Summary .......................................................................................... 53

4 Dynamic O-D matrix estimation as a statistical problem ................. 55
4.1 Introduction ....................................................................................... 55
4.2 Literature review ............................................................................... 55
4.3 Statistical structure of the problem .................................................. 56
4.3.1 Intrinsic stochasticity .................................................................... 56
4.3.2 Predicting the expectation $S_{ik}$ of stochastic group observation $S_{ik}$ 57
4.3.3 Stochasticity accounted for in the estimation ................................ 58
4.3.4 Distributional assumptions ........................................................... 59
4.3.5 Probability distribution of $S_{ik}$ and derivation of the objective
function ............................................................................................. 61
4.3.6 Incorporating additional information ......................................... 62
4.4 The objective function used for the upper-level O-D matrix estimation
problem ............................................................................................... 64
4.4.1 First simplification ........................................................................ 64
4.4.2 Parametrisation of the O-D matrix: matrix models ..................... 65
4.4.3 Final objective function after parametrisation and further simplification ................................. 68
4.5 Summary .......................................................................................... 69

5 Dynamic O-D matrix estimation as an optimisation problem ........... 71
5.1 Introduction ....................................................................................... 71
5.2 Mathematical formulation of the dynamic matrix estimation problem
with deterministic route choice ............................................................. 72
5.2.1 The dynamic assignment map ...................................................... 72
5.2.2 The objective function ................................................................. 73
5.2.3 Formulation of the dynamic O-D matrix estimation problem .......... 74
5.2.4 The MPEC formulation and game theory .................................... 75
5.3 The dynamic matrix estimation problem with stochastic route choice ... 76
6 Dynamic user equilibrium assignment

6.1 Introduction .......................................................... 79
6.2 Literature survey on User Equilibrium assignment models .......... 81
6.3 Distinguishing between route choice equilibrium and departure time equilibrium .................................................. 83
6.4 Dynamic Equilibrium Conditions .................................... 84
  6.4.1 Equilibrium conditions for users with random taste variation and perception errors: DUE-STOCH equilibrium ....................... 85
6.5 VIP formulation of DUE-STOCH assignment ........................ 86
  6.5.1 Characterisation by equilibrium conditions .................... 86
  6.5.2 Path costs for the VIP formulation ............................. 88
6.6 Dynamic Network Loading ............................................. 90
  6.6.1 The discrete-time Dynamic Network Loading model .......... 91
  6.6.2 Algorithm for discrete-time dynamic network loading .......... 94
  6.6.3 Dynamic assignment fractions .................................. 95
6.7 Derivatives of traffic flow with respect to dynamic O-D flows ....... 96
  6.7.1 Calculating the derivatives .................................... 96
  6.7.2 Direct calculation of derivatives ............................. 97
  6.7.3 Approximations through truncated derivatives ................. 99
6.8 Combined path choice and departure time choice .................... 103
  6.8.1 Solution algorithm for the combined path choice and departure time choice assignment .............................................. 104
6.9 Miscellaneous ......................................................... 105
  6.9.1 DUE traffic assignment and STENs ............................ 105
  6.9.2 Path enumeration for stochastic assignment models ........... 106
6.10 Summary ............................................................. 106

7 Behavioural choice models for route and departure time ............. 109

7.1 Introduction .......................................................... 109
7.2 Determining appropriate choice models ................................ 109
  7.2.1 Role of the models ............................................. 110
  7.2.2 Model requirements ............................................. 110
7.3 Logit models for route choice ....................................... 111
  7.3.1 The C-logit model ............................................. 111
  7.3.2 The path-size model .......................................... 112
  7.3.3 The Logit-kernel model ...................................... 112
  7.3.4 Discussion ..................................................... 114
7.4 Logit models for departure time choice ................................ 114
  7.4.1 Utility function .............................................. 114
  7.4.2 Model form .................................................... 115
  7.4.3 Application issues ............................................ 115
7.5 Departure time adjustment modelling .................................. 116
  7.5.1 Application issues: determining \( t_{PD} \) and \( t_{PA} \) ............ 118
# 8 Solution methods

8.1 Introduction ......................................................... 121
8.2 Solution algorithms ................................................. 121
  8.2.1 Descent algorithms using approximate derivatives ........... 124
  8.2.2 Maher's heuristic ............................................... 125
8.3 Choice of a solution algorithm ................................... 126
8.4 Summary ............................................................. 127

# 9 Implementation and verification of the model

9.1 Introduction ......................................................... 129
  9.1.1 Verification: aspects considered, criteria, norm ............ 129
9.2 The software ....................................................... 130
9.3 Verification of the software ....................................... 131
9.4 The test-case ....................................................... 131
  9.4.1 The network .................................................... 132
  9.4.2 The apriori demand matrix ................................... 133
9.5 The observations ................................................... 133
9.6 Estimation results with only route choice ......................... 134
  9.6.1 Apriori agrees with group observations ....................... 134
  9.6.2 Apriori 20% higher than group observations ................... 134
  9.6.3 Apriori 20% lower than group observations ..................... 135
9.7 Estimation results with simultaneous choice of route and departure time 135
  9.7.1 Apriori agrees with group observations ....................... 136
  9.7.2 Apriori 20% higher than group observations ................... 136
  9.7.3 Apriori 20% lower than group observations ..................... 137
9.8 Conclusions .......................................................... 137

# 10 Case study: Kruithuisweg

10.1 Introduction ....................................................... 139
10.2 Description of the study area .................................... 139
  10.2.1 Dynamic observations ....................................... 141
  10.2.2 Dynamic demand ............................................. 142
10.3 Model description ................................................ 143
  10.3.1 Modelling of intersections in analytic dynamic traffic assignment 143
  10.3.2 Link-performance functions ................................ 145
  10.3.3 Route choice ............................................... 145
10.4 Experimental setup ............................................... 145
10.5 Estimation results ................................................. 146
  10.5.1 Comments on computational aspects .......................... 146
  10.5.2 Apriori agrees with group observations ....................... 146
  10.5.3 Apriori 20% higher than group observations ................... 148
  10.5.4 Apriori 20% lower than group observations ..................... 149
  10.5.5 Randomised apriori matrix ................................... 151
  10.5.6 Conclusions .................................................. 152
11 Extensions and applications 153
11.1 Introduction ........................................... 153
11.2 Refinement of the solution algorithms ............... 154
11.3 Improvement of the choice models used .......... 154
11.4 Extension of the model to more general cases .... 154
  11.4.1 Extension to multiple userclasses ............. 155
  11.4.2 Extension to a multimodal setting .......... 155
  11.4.3 Extension to elastic demand ................. 156
  11.4.4 Estimation of dispersion parameters ....... 158
  11.4.5 Estimation of optimal tolls and optimal traffic control 158
11.5 Summary ............................................ 159

12 Conclusions 161
12.1 Introduction ........................................... 161
12.2 The problem and its characteristics ............... 162
12.3 Summary of the work done .......................... 163
  12.3.1 Behavioural framework of the estimator .... 163
  12.3.2 Parameterisations to solve the underspecification 163
  12.3.3 Statistical estimation ......................... 163
  12.3.4 Solution approach and use of derivatives .. 164
  12.3.5 Development and implementation of a DTA procedure 164
12.4 Contributions of this thesis ........................ 164
12.5 Avenues for further work ........................... 165
12.6 Evaluation ........................................... 165
  12.6.1 Verification ................................... 165
  12.6.2 Practical significance ......................... 165
  12.6.3 Advantages of the approach used ........... 166
  12.6.4 Reflection on the assumptions made ........ 166

Bibliography 169

A Calculation of dynamic assignment fractions 181

B Calculation of dynamic assignment fractions for attraction 189

C Derivatives of the matrix w.r.t. to parameters 191
  C.1 Derivative of the matrix with respect to parameter $\Lambda$ 191
  C.2 Calculation of derivatives $\frac{\partial s_{ik}}{\partial \lambda_{mn}}$ 191

Index 195

Summary 197

Samenvatting 199

About the author 203
List of Figures

1.1 A conceptual hierarchy for ordering the O-D matrix estimation issues 3
1.2 Structure of this thesis ............................................. 11

2.1 Generic transportation modelling framework for analysis of policy and control measures ....................................................... 14
2.2 Day-to-day and within-day variation ........................................... 16
2.3 Number of trips (all modes) by purpose and departure time (weighted) (Source: OVG 1998) ....................................................... 18
2.4 Number of trips (car driver) by purpose and departure time (weighted) (Source: OVG 1998) ....................................................... 18
2.5 Generic form of a dynamic transportation model .................................... 19
2.6 A mechanism for day-to-day dynamics in the traffic pattern .................. 20

3.1 Relationship between O-D matrix and assignment in the static and dynamic cases (rigid demand) ....................................................... 35
3.2 Relationships for a Stochastic User Equilibrium (static, rigid demand) .... 36
3.3 Relationships for Dynamic Stochastic User Equilibrium Assignment (dynamic, rigid demand) ....................................................... 37
3.4 Relationships between objects in dynamic O-D matrix estimation (without departure time adjustment) ....................................................... 39
3.5 Relationships between objects in dynamic O-D matrix estimation (with departure time adjustment) ....................................................... 41
3.6 Integration of the dynamic O-D matrix estimation into the behavioural model structure ................................................................. 42
3.7 Relationship between internal timestep $k$ and external timestep $\kappa$ .... 46
3.8 Relationship between $S_M$ and the O-D matrix flows ................................... 47
3.9 Space Time Expanded Network (STEN) corresponding to a physical network. (x) denotes time period $x$ ....................................................... 51
3.10 STEN with a virtual destination ....................................................... 51
3.11 Augmented STEN with schedule delay costs that depend on the preferred departure time ................................................................. 52

4.1 Hypothesised relationship between stochastic O-D matrix and stochastic group observations ................................................................. 57
4.2 Modelled stochastic relationships ....................................................... 58
4.3 Stationarity in the static case ....................................................... 59
4.4 Relationship between stochastic and modelled group observations for estimation ................................................................. 59
4.5 Extending the assumption of stationarity of $T_{k}^{r,s \_d,}$ ........................................... 60
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.6</td>
<td>Number of O-D matrix cells and parameters needed for parametric model</td>
<td>67</td>
</tr>
<tr>
<td>5.1</td>
<td>Interaction between objective function and constraints in the dynamic bilevel formulation</td>
<td>75</td>
</tr>
<tr>
<td>6.1</td>
<td>Hierarchical relationship between DTAC-STOCH and DUE-STOCH</td>
<td>84</td>
</tr>
<tr>
<td>6.2</td>
<td>Cost components associated with STEN links</td>
<td>88</td>
</tr>
<tr>
<td>6.3</td>
<td>Influence diagram for the discrete-time DTA algorithm</td>
<td>91</td>
</tr>
<tr>
<td>6.4</td>
<td>Small and large time intervals</td>
<td>92</td>
</tr>
<tr>
<td>6.5</td>
<td>Influence diagram for the C-load discrete-time DNL algorithm</td>
<td>94</td>
</tr>
<tr>
<td>6.6</td>
<td>Influence diagram for DTA-DNL algorithm</td>
<td>97</td>
</tr>
<tr>
<td>6.7</td>
<td>Influence diagram when the DNL and the DTA are not considered</td>
<td>99</td>
</tr>
<tr>
<td>6.8</td>
<td>Influence diagram when only the effect of the DNL is taken into consideration</td>
<td>100</td>
</tr>
<tr>
<td>6.9</td>
<td>Influence diagram when the effects of both DNL and DTA are accounted for</td>
<td>102</td>
</tr>
<tr>
<td>6.10</td>
<td>MSA solution algorithm for the combined DTA / DTC equilibrium problem</td>
<td>104</td>
</tr>
<tr>
<td>7.1</td>
<td>Departure time, travel time, and arrival time</td>
<td>114</td>
</tr>
<tr>
<td>7.2</td>
<td>Preferred departure time, preferred arrival time and travel time</td>
<td>117</td>
</tr>
<tr>
<td>7.3</td>
<td>Minimisation problem for a utility maximising traveller</td>
<td>118</td>
</tr>
<tr>
<td>8.1</td>
<td>Iterative method for O-D matrix estimation</td>
<td>122</td>
</tr>
<tr>
<td>8.2</td>
<td>The generic repeated line-search algorithm</td>
<td>123</td>
</tr>
<tr>
<td>8.3</td>
<td>Flow diagram of Maher’s heuristic for O-D matrix estimation adapted to the dynamic situation</td>
<td>126</td>
</tr>
<tr>
<td>9.1</td>
<td>Toplevel structure of the software implementation</td>
<td>130</td>
</tr>
<tr>
<td>9.2</td>
<td>Simple test-network</td>
<td>132</td>
</tr>
<tr>
<td>9.3</td>
<td>Estimation results with perfect apriori</td>
<td>134</td>
</tr>
<tr>
<td>9.4</td>
<td>Estimation results when apriori is too high</td>
<td>135</td>
</tr>
<tr>
<td>9.5</td>
<td>Estimation results when apriori is too low</td>
<td>135</td>
</tr>
<tr>
<td>9.6</td>
<td>Estimation results with perfect apriori</td>
<td>136</td>
</tr>
<tr>
<td>9.7</td>
<td>Estimation results when apriori is too high</td>
<td>136</td>
</tr>
<tr>
<td>9.8</td>
<td>Estimation results when apriori is too low</td>
<td>137</td>
</tr>
<tr>
<td>10.1</td>
<td>The Kruithuisweg study area</td>
<td>140</td>
</tr>
<tr>
<td>10.2</td>
<td>Intersections on the Kruithuisweg</td>
<td>140</td>
</tr>
<tr>
<td>10.3</td>
<td>Kruithuisweg: transportation network</td>
<td>141</td>
</tr>
<tr>
<td>10.4</td>
<td>Hypothetical (static) 2-hour O-D matrix for the Kruithuisweg</td>
<td>142</td>
</tr>
<tr>
<td>10.5</td>
<td>Visualisation of the hypothetical static O-D matrix</td>
<td>142</td>
</tr>
<tr>
<td>10.6</td>
<td>Hypothetical departure-time profile per O-D matrix cell</td>
<td>143</td>
</tr>
<tr>
<td>10.7</td>
<td>Estimation results with perfect apriori (O-D matrix cells)</td>
<td>147</td>
</tr>
<tr>
<td>10.8</td>
<td>Estimation results with perfect apriori (Group observations)</td>
<td>147</td>
</tr>
<tr>
<td>10.9</td>
<td>Objective function z by iteration number</td>
<td>147</td>
</tr>
<tr>
<td>10.10</td>
<td>Estimation results with inflated apriori (O-D matrix cell values)</td>
<td>148</td>
</tr>
<tr>
<td>10.11</td>
<td>Estimation results with inflated apriori (group observations)</td>
<td>148</td>
</tr>
</tbody>
</table>
10.12 Objective function $z$ by iteration number .................................................. 149
10.13 Estimation results with deflated apriori matrix (O-D matrix cells) ........ 149
10.14 Estimation results with deflated apriori matrix (group observations) ... 150
10.15 Decrease of the objective function .............................................................. 150
10.16 Estimation results with randomly perturbed apriori matrix (O-D matrix
   cells) .............................................................................................................. 151
10.17 Estimation results with a randomly perturbed apriori matrix (group ob-
   servations) .................................................................................................... 151
10.18 Objective function $z$ by iteration number ................................................... 151

11.1 Adaptation of the model structure to incorporate multimodal transporta-
    tion on supernetworks ..................................................................................... 156
11.2 Extension of the behavioural framework with control measures ............. 158

A.1 Six possibilities for incidence of traffic departing during time interval $h$
    on link $a$ during time interval $k$ .................................................................. 182
A.2 Calculation of assignment fractions .............................................................. 184
A.3 Calculation of dynamic assignment fractions .............................................. 187
A.4 Calculation of assignment fractions: close-up ............................................ 188
## List of Tables

<table>
<thead>
<tr>
<th>Table Number</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Classification of setups for O-D matrix estimation</td>
<td>21</td>
</tr>
<tr>
<td>3.1</td>
<td>Approaches to dynamic O-D matrix estimation</td>
<td>33</td>
</tr>
<tr>
<td>3.2</td>
<td>List of approaches to dynamic O-D matrix estimation</td>
<td>34</td>
</tr>
<tr>
<td>3.3</td>
<td>Distinctions between rigid and completely elastic demand</td>
<td>43</td>
</tr>
<tr>
<td>6.1</td>
<td>List of acronyms for User Equilibrium assignments</td>
<td>80</td>
</tr>
<tr>
<td>6.2</td>
<td>Literature on the dynamic VIP formulations of UE</td>
<td>81</td>
</tr>
<tr>
<td>6.3</td>
<td>Literature on sensitivity analysis of VIP formulations of static UE</td>
<td>83</td>
</tr>
<tr>
<td>8.1</td>
<td>Comparison of MPEC solution methods</td>
<td>123</td>
</tr>
<tr>
<td>9.1</td>
<td>Setup for model verification</td>
<td>129</td>
</tr>
<tr>
<td>9.2</td>
<td>Link flows</td>
<td>132</td>
</tr>
<tr>
<td>9.3</td>
<td>O-D flows</td>
<td>132</td>
</tr>
<tr>
<td>9.4</td>
<td>The O-D matrix and its margin totals</td>
<td>133</td>
</tr>
<tr>
<td>9.5</td>
<td>Table Caption</td>
<td>133</td>
</tr>
<tr>
<td>10.1</td>
<td>Zone centroid names</td>
<td>141</td>
</tr>
</tbody>
</table>
Notation

This section lists the symbols used and page number where they are defined.

General notation

~ : observed value (p.41)
* : equilibrium value (p. 73)
^ : estimated value (p.19)
\( x \) : stochastic variable \( x \) (p.56)
\( a \) : directed arc \( a \) in a graph (also called link) (p.14)
\( A \) : set of links (arcs) in a graph (p.19)
\( n \) : node or vertex of a graph
\( N \) : set of nodes (p.19)
\( G = G(N,A) \) : directed graph consisting of node set \( N \) and arc set \( A \)
\( h \) : departure time interval index (small intervals) \( h \in H \)
\( H \) : the set of all time intervals \( h \)
\( k \) : general time interval index (small intervals) \( k \in K \)
\( K \) : the set of all time intervals \( k \)
\( \kappa \) : big time interval related to group observations, \( \kappa \in \mathcal{K} \)
\( P_{rs} \) : set of all paths \( p \) between O-D pair \( (r,s) \) (p.14)
\( l \) : data group
\( p \) : path through a graph; \( p \in P_{rs} \) (p.14)
\( r \) : trip origin (p.14)
\( (r,s) \) : O-D pair (p.14)
\( s \) : trip destination (p. 14)
\( Z \) : set of zones. \( r,s \in Z \)
\( \tilde{S}_{lk} \) : observed group count for datagroup \( l \) during external time period \( K \)
\( S_{lk} \) : modelled group count for datagroup \( l \) during external time period \( K \)
\( T_{h}^{rs} \) : O-D-t matrix; indices \( r,s,h \) added for emphasis
\( t_{h}^{rs} \) : O-D-t matrix cell: origin \( r \), destination \( s \), time period \( h \)
\( \Lambda_{lh} \) : parameter vector for the matrix model, with indices \( l \) (data group) and time interval \( h \)
\( \lambda_{lh} \) : parameter value for the matrix model, for data group \( l \) and time period \( h \)
\( \Theta_{h} \) : parameter vector for a departure time rescheduling model for time interval \( h \)
Path variables

\[ C_{ak} \] : matrix with (generalised) link cost for link \( a \) and time period \( k \)
\[ c_{ak} \] : (generalised) link cost on link \( a \) during time period \( k \)
\[ \Delta_{aphk}^{rs} \] : the dynamic link-path incidence matrix
\[ \delta_{aphk}^{rs} \] : cell value of the dynamic link-path incidence matrix: 1 if traffic from O-D pair \( rs \), setting out during period \( h \) on path \( p \) reaches link \( a \) during time period \( k \) (p.72)
\[ \Phi_{aphk}^{rs} \] : matrix of assignment fractions
\[ \varphi_{aphk}^{rs} \] : assignment fraction of the cell of the dynamic O-D matrix for O-D pair \( rs \), setting out during period \( h \) that leaves link \( a \) during period \( k \).(p.181)
\[ \mathcal{F} \] : Set of feasible path flows \( f \)
\[ F_{ph}^{rs} \] : path flows \( f_{ph}^{rs} \)
\[ f_{ph}^{rs} \] : departure flow rate on path \( p \) of traffic from origin \( r \) to destination \( s \) during time interval \( h \)
\[ f_{k}^{rs} = \sum_{p \in P_{rs}} f_{ph}^{rs} \] : the departure flow rate for O-D pair \( (r,s) \) during time interval \( k \)
\[ \bar{f}_{l} \] : sampling fraction for observation group \( l \).
\[ c_{ph}^{rs} \] : the experienced generalised travel cost over path \( p \) from origin \( r \) to destination \( s \) by the flow departing during time interval \( h \)
\[ \pi_{ph}^{rs} = \min \{ \pi_{ph}^{rs}(h) \} \] : minimum experienced generalised travel cost over paths \( p \) from origin \( r \) to destination \( s \) for flows departing during time interval \( h \)
\[ P_{ph}^{rs} \] : Matrix with the route-choice proportions for O-D pair \( (r,s) \) starting during time period \( h \)
\[ p_{ph}^{rs} \] : element of matrix \( P_{ph}^{rs} \)
\[ \tau_{pa}^{rs}(h) \] : travel time on path \( p \) from \( r \) to \( s \) starting during time period \( h \) until link \( a \) is reached

Link variables

\[ u_{ak} \] : link entrance flow rate during period \( k \)
\[ U_{ak} \] : cumulative entrance flow at the beginning of time period \( k \)
\[ v_{ak} \] : link exit flow rate during period \( k \)
\[ V_{ak} \] : cumulative exit flow at the beginning of time period \( k \)
\[ X_{ak} \] : number of vehicles of link \( a \) during time period \( k \)
\[ T_{ak} \] : travel time over link \( a \) for flows entering link \( a \) during time period \( k \)
\[ s_{ak} \] : exit time from link \( a \) for flows entering link \( a \) during time period \( k \)
\[ c_{ak} \] : link cost on link \( a \) during time period \( k \)
Time variables

\([0, T]\) : O-D traffic demand period
\([0, \Delta]\) : DTA analysis period (the period from the time when flows enter the network to the time all flows exit the network)
\(\Delta\) : minimum free flow link travel time over all links
\(\delta = \frac{\Delta}{M}\), \(M\) is a positive integer
\(k\) : index for general time interval \([k\delta, k\delta + \delta)\)
\(h\) : index for departure time interval \([h\delta, h\delta + \delta)\)

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNL</td>
<td>Dynamic Network Loading (p.37)</td>
</tr>
<tr>
<td>DTA</td>
<td>Dynamic Traffic Assignment (p.12)</td>
</tr>
<tr>
<td>GBLP</td>
<td>Generalised Bi-level Program (p.74)</td>
</tr>
<tr>
<td>GLS</td>
<td>Generalised Least Squares (p.56)</td>
</tr>
<tr>
<td>MNL</td>
<td>MultiNomial Logit (p.8)</td>
</tr>
<tr>
<td>MPEC</td>
<td>Mathematical Problem with Equilibrium Constraints (p. 50)</td>
</tr>
<tr>
<td>NCP</td>
<td>Nonlinear Complementarity Problem (p. 81)</td>
</tr>
<tr>
<td>O-D</td>
<td>Origin-Destination matrix (p.1)</td>
</tr>
<tr>
<td>DTAC-STOCH</td>
<td>Stochastic Departure Time Adaptation Choice (p.103)</td>
</tr>
<tr>
<td>DUE-STOCH</td>
<td>Dynamic Stochastic User Equilibrium (p. 86)</td>
</tr>
<tr>
<td>STEN</td>
<td>Space Time Extended Network (p. 31)</td>
</tr>
<tr>
<td>TLD</td>
<td>Trip Length Distribution (p. 49)</td>
</tr>
<tr>
<td>VIP</td>
<td>Variational Inequality Problem (p. 72)</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Context and general considerations

This thesis deals with the estimation of dynamic Origin-Destination (O-D) matrices. Such matrices quantify the time-varying travel demand between geographic zones in a study area. These inter-zonal flows are not directly observable but can be estimated from various observable quantities. In general this estimation problem is severely underdetermined, as the number of unknown flows far exceeds the number of independent observations. This is unfortunate as O-D matrices are a key input for transport models. The classical approach to this problem (which we will follow) is to use a model of the transport system to reduce the solution space to the point where the estimation problem is no longer underdetermined.

The state of the art in transportation modelling is to view transport systems as markets characterised by (travel) demand and (infrastructure) supply. The relevance of any infrastructural element can then be represented by its impact on supply, whilst behavioural patterns (or changes in them) can be represented in terms of how they affect demand. This framework allows the study of changes in either factor in a quantitative way through its effect on the outcome of the demand-supply interaction models (equilibrium or otherwise).

The supply-side is usually represented through a transportation network, most characteristics of which (such as topology, dimensions, and to some extent also capacity and reaction to traffic loads) can be obtained from maps and from engineering documentation.

The interaction between demand and supply is modelled through so-called traffic assignment models, which calculate the resulting quantities (i.e. traffic flows) throughout the network and their cost (mostly in terms of travel time). The way in which a zone-to-zone travel demand distributes itself across available resources (such as alternative routes and time-periods) depends on user-behaviour, so that the determination of the supply-demand equilibrium should take account of the characteristics of this user behaviour. On the other hand the effect of the distributed demand on the level of service (i.e. the price) of specific resources is of a more mechanical nature.
In static transportation models, the supply-demand interaction within the transportation system considered is approximated by a steady state system during the period to which the O-D matrix pertains. This approximation may be reasonable when demand is much lower than capacity, but it breaks down when a non-negligible fraction of capacity is used. Even when demand temporarily exceeds capacity the steady-state approximation fails because of the non-linear response of the travel-time on infrastructure elements to demand. Congestion and traffic jams clearly violate the steady-state assumption because they build up, level off, and disappear within a few hours. Infrastructural capacity is (generally) sufficient on a 24-hour basis but is scarce during certain peak periods. This fact has fuelled interest in dynamic transportation models, i.e. models in which the time-varying nature of transportation phenomena is explicitly represented and the steady-state hypothesis is relaxed.

Extending the market-approach to the dynamic case, we are required to specify the time-dynamic representations of infrastructure and travel demand and their interaction. With respect to travel demand, this is equivalent to the determination of a time-dependent O-D matrix, typically in the absence of direct O-D observations, which is the issue that is addressed in this thesis.

We note that since different user classes sharing the same road can interact in different ways (e.g. lorries face different speed limits than passenger cars on most motorways, their route choice and departure time choice behaviour may be completely different), it is both possible and useful to separate the userclasses in the demand-supply calculations and to estimate different dynamic O-D matrices for these userclasses. In this thesis we will focus on a behavioural approach towards the estimation of dynamic O-D matrices, accounting for certain behavioural aspects of transportation demand on the position of the equilibrium. Although we will focus on the demand estimation for a single vehicle type (passenger cars), the impact of multiple vehicle types in the calculation of supply-demand interaction can be easily accommodated, and the estimation may be extended to cover multiple vehicle types, as will be discussed in Chapter 12. Although vehicles of the same type may cover different user-classes, which differ in their route choice behaviour, we will assume a single user-class throughout.

1.2 Issues involved

1.2.1 Conceptual hierarchy

As O-D matrix estimation consists of finding a demand-supply equilibrium that is consistent with observations, it has close ties with both demand modelling and demand-supply interaction modelling and is affected by the issues pertaining to both. We briefly list the issues involved and order them according to the conceptual level at which they occur. Because of the complexity of the modelling and algorithmic issues involved take a substantial amount of effort to solve, and any treatment of O-D matrix estimation takes on a technical character. In fact O-D matrix estimation depends on the ability to repeatedly solve a transport equilibration model for different inputs (demand) until it matches observations and in this way is a form of model calibration.
We will use the six-step hierarchy shown in Figure 1.1 to distinguish conceptual levels by which the issues involved can be ordered.

![Hierarchy Diagram]

Figure 1.1: A conceptual hierarchy for ordering the O-D matrix estimation issues

The first level shown is that of physical reality (the domain of discourse), of which we construct an (informal) conceptual model by identifying important objects (e.g. "traveller", "zone", "route", "travel time") and relationships between them at the conceptual level. At the third level theories are formulated that formalise relationships between the objects distinguished at the conceptual level. Such theories can be operationalised in various mathematical models, where the choice of which mathematical models to use occurs at the fourth level. Once the mathematical models are stated, the next level contains the algorithms used to calculate the model response to certain inputs. At the sixth level the algorithms are operationalised in computer software and applied.

The relationship between subsequent levels is drawn as a cone to emphasise that at each upwards step through the hierarchy, multiple possibilities exist for mapping the lower-level concepts onto higher level concepts.

1.2.2 Discussion of issues involved

We will now present the major issues involved in the dynamic O-D matrix estimation problem according to their level in this six-step hierarchy.

Reality level

Observations such as traffic counts reside at this level. A special property of the observations is that they contain noise (whether due to measurement error or fluctuations in the process), the modelling and treatment of which is an issue for higher levels.

An issue that naturally arises in conjunction with reality is whether we should regard reality as being in equilibrium. This has far-reaching consequences for the theory and models that we may apply. For example O-D matrix estimation has been mostly
approached without assuming equilibrium (see e.g. the discussion of the ME2 method in Ortuzar and Willumsen (2001)), but more recent methods incorporate it, as noted in Cascetta (2001).

Conceptual level

The first issue at the conceptual level is which concepts to include, and on which concepts to focus. A summary of travel demand by geographical zone may result in an O-D matrix of trips, tours (return-trips), or even to chained trips. In addition, O-D flows might be differentiated by mode, travel purpose, gender, age, etc.; infrastructure may thought of as a continuum instead of a network. Further issues are whether behavioural aspects should be explicitly dealt with or whether the O-D matrix estimation problem should be cast in purely descriptive, statistical terms (as in real-time O-D matrix estimation with Kalman filters). Strictly speaking, we have a number of users of the transportation system who repeatedly choose their transport mode, route, and departure time, and who cause interference for each other. The natural conceptual framework used for this situation is that of game-theory. Then there is the issue of whether to assume a demand-supply equilibrium (regardless of whether it exists in reality) and how to formulate it when it is used at all.

Also at the conceptual level is the issue of under-determinedness: as a rule, many more zonal traffic flows are possible than independent observations are available. Finally there is the issue of statistical discrimination between candidate O-D matrices on basis of the available evidence (observations): given the large number of free parameters and the often scarce, expensive, correlated, and noisy observational data, it is often hard (if not impossible) to discriminate between two candidate O-D matrices on basis of observational evidence alone, so that using apriori information (e.g. from demand models) may be required.

Theory level

At the theory level, there is the issue of what relationships to include in a formal framework that relates the unknown quantity (the O-D matrix) to the ones we can observe (such as: traffic counts, trip length distributions, margin totals, observed partial matrices), how to characterise the traffic equilibria, and what theoretical basis (if any) to use for approaching the behavioural aspects. Possibilities include: Random Utility Maximisation (RUM), and a psychological approach involving prospect theory, attitudes, knowledge and opportunities. Furthermore it is debatable which aspects of the O-D matrix should be assumed to be endogenous and which should be considered exogenous. For example, distribution effects, modal shift, total number of trips generated may all be modelled, and could therefore be included in a framework for O-D matrix estimation.

The theoretical development of classical static O-D matrix estimation has culminated in the view that a static O-D matrix can be estimated through the optimisation of a criterion function based on the difference between the estimated matrix and an apriori matrix, and the difference between observed and estimated values of e.g. traffic counts.
(see e.g. Cascetta (2001)). Known criterion functions include likelihood functions, generalised least squares objectives, and entropy criteria. The link between the O-D matrix and observable quantities is then represented by a demand-supply equilibrium calculated through a static user-equilibrium assignment, where the equilibrium is over route choice. The resulting mathematical formulation is a bilevel optimisation problem, and more specifically an MPEC. The question then is how and to what extent the classical static approach can be generalised to the dynamic case.

In the context of dynamic O-D matrix estimation many of the classical ideas reappear in a different form. The observations now carry a time-dimension, and the static user-equilibrium assignment is replaced by dynamic user-equilibrium assignment. However, there are important differences in the sense that the supply-demand equilibrium is now time-dependent, and that users have an additional degree of freedom: they can choose their departure time in addition to their route.

With departure time choice we mean the entire complex of considerations (such as penalties for late arrival, travel time uncertainty) that results in a particular departure time being chosen. As with route choice, departure time choice can be linked to the (generalised) cost associated with possible departure-times, of which travel time is a component. From this point of view it is clear that departure time will be influenced by the amount of traffic at a particular time, analogous to route choice. For consistency it is thus necessary to find a simultaneous demand-supply equilibrium with respect to both departure time choice and route choice. As a consequence, a mechanism for departure time adjustment has to be made endogenous to the dynamic traffic equilibrium assignment. The question then is whether one should regard departure time as rigid, fully adaptable (as e.g. route choice) or adaptable only within certain margins.

If one accepts departure time adjustment as endogenous, a further issue is that a chain of influences can be distinguished that lead up to the actual departure time. The actual departure time depends on the desired departure time, which is strongly influenced by the desired arrival time, which in turn is strongly dependent on the choice of activities that an individual has scheduled for the day.

Modelling level

At the modelling level we face the issue of which model formulation to use to capture the relationships specified at the theory level. If one decides on modelling behavioural aspects through RUM, a number of possible models exist such as the well-known logit and probit based discrete choice models. Although several RUM models are available for route-choice, and one or two for departure time choice, there is the question of how to let them act together in a simultaneous route and departure time choice model. Noting that user equilibrium conditions for flows in networks is covered by existing behavioural theory, the question arises if it is possible to incorporate departure time choice into such a framework and if so in what way.

The demand-supply interaction can be modelled analytically or using simulation methods, and can be carried out with a wide variety of link cost functions. The O-D matrix
itself can be represented with an independent variable for each cell, it can be parameterised through a distribution model, or it can be parameterised with a free parameter per independent observation.

The dynamic O-D matrix estimation problem is more complicated than in the static case, because in order to relate observations to the correct O-D cells the time when traffic is observed is now as important as the location where it is observed. Therefore the issue of spatio-temporal consistency enters the model, which in practice boils down to the requirement of correctly modelling travel-times.

Then there is the problem of how to handle the noise that is present in the observations at the modelling level, or more precisely the signal-to-noise ratio in the observations. In the dynamic case the number of independent observations increases at the same rate as the number of time periods, but the total amount of observed traffic remains the same, so that the number of vehicles observed per period decreases. Therefore the amount of observations per dynamic O-D matrix cell decreases, and the influence of noise in the data increases. This issue can be dealt with in various ways, e.g. by increasing the number of observations, using a priori information, and by imposing an a priori structure on the O-D matrix.

Algorithmic level

Once the models are specified, an issue at the algorithmic level is how to extract explicit (numerical) results from the implicit formulation of the models specified at the modelling level. Typical issues are convergence, computation time, accuracy, memory and storage requirements, and data structures to be used. For O-D matrix estimation, the size of the problem involved in real-life applications is so large that some thought about the efficiency of the algorithms is necessary to ensure scalability.

The solution of the dynamic O-D matrix estimation problem involves minimisation of an objective function conditional on the outcome of the supply-demand interaction model. As any change in the O-D matrix made to improve the objective function also shifts the predicted solution of the supply-demand model, it may well result in a net deterioration of the objective function. This is an important issue, which boils down to the problem of finding a descent direction for the minimisation problem after the reaction of the equilibrium condition has been taken into account. A way to address this issue is to calculate derivatives of the objective function with respect to changes in the O-D matrix, which necessarily involves the calculation of derivatives of the solution of the supply-demand model with respect to changes in the O-D matrix.

A further issue is the nature of the optimisation problem with respect to issues such as local optima and non-convexity of the problem. Both issues are the focus of ongoing research in mathematical optimisation, and fundamentally complicate the search for efficient and convergent solution algorithms. In acknowledgment of the importance of high quality solution algorithms, we will strive to make our solution algorithms as rigorous as possible. Considering however that the aim of this thesis is to contribute to the field of transportation modelling (an engineering discipline) rather than to mathematics, we will focus on setting up the model structure while deferring some of the mathematical issues for further work.
Application level

At the application level the issues relate to the properties of the software: can the software scale to large-size real-world problems, is it inter-operative with other software packages, can the required inputs be collected, can the output be interpreted? A medium-size transportation model may easily contain about 300 geographical zones, which (theoretically) leads to 90,000 O-D flows. When a time-step of six seconds is used and a 120-min time period is modelled, this results in 1200 time periods and hence $1.08E8$ dynamic O-D matrix cells, which has implications for practicality. The first requirement for software able to handle such situations is a scalable algorithm, and the second requirement is a careful implementation.

A frequently overlooked issue is how to correctly apply the software. Due to the nature of the available data (and especially with respect to the random and systematic errors it contains, the conflicting observations, its scarcity) the actual estimation of O-D matrices needs both engineering judgement as a statistical approach in selecting and weighing the data sources to be used and dealing with noise and the inevitable inconsistencies between available data-sources.

### 1.3 Delineation of the issues addressed in this thesis

#### 1.3.1 Issues addressed in this thesis

This thesis addresses a subset of the issues addressed in Section 1.2.2. The first issue (to be addressed in chapter 2) concerns the appropriateness of assuming an equilibrium at the conceptual level. We argue that the equilibrium assumption can be justified to a large extent from a game-theoretic point of view, and that due to its internal consistency the equilibrium state forms an appropriate yardstick with which to compare policy options even when equilibrium is not strictly attained.

At the conceptual level, a classical set of concepts (O-D matrix, network, equilibrium, travel time, travel cost, route, departure time) and relationships (supply-demand equilibrium) is adopted for use in chapter 3, to which the distinction between departure time choice and departure time adjustment due to traffic conditions is added.

At the theory level, behavioural aspects such as route choice and departure time choice are explicitly considered and the assumption of a user-equilibrium regarding these is incorporated into the framework. The departure time adjustment process is therefore made endogenous in our approach.

The issue of under-determinedness is tackled by using apriori information and by explicitly parameterising the O-D matrix by a one-to-one correspondence with observations. The framework proposed in chapter 3 specifies that destination choice and mode choice are exogenous, and that the level of demand is assumed to be independent of the level of service (i.e. inelastic demand). The theoretical approach used to approach human behaviour is that of Random Utility Maximisation, as specified in chapter 7. The statistical approach taken towards O-D matrix estimation is reported in chapter 4, and
follows that of Hazelton (2000), who shows that the theoretically preferred (but computationally intractable) true Maximum-Likelihood estimator can be approximated by classical CGLS estimators, which results in a tractable quadratic objective function that the O-D matrix estimate must minimise. The issue of how to minimise this function with a traffic equilibrium as side constraint is discussed in chapter 5, where the problem is recognised as an MPEC. To assist in solving this MPEC, issue of calculating derivatives of the lower level problem (the DTA) was addressed.

At the modelling level the path-size and C-logit models are selected to model route choice behaviour, while the Multinomial Logit (MNL) model is used to model departure time choice, as discussed in chapter 7. The demand-supply interaction is addressed through a user-equilibrium formulation and dynamic network loading.

At the algorithmic level a pragmatic approach towards the solution of the MPEC problem was taken, as described in chapter 8. The datastructures were set up as pointer-structures with a view towards scalability.

At the application level, the issue of the required inputs is largely resolved because the network requirements are fairly standard, the parameters of the choice models will be obtained from existing studies, and the required apriori matrices will be obtained from existing static transportation models. In view of the large amount of information that must be passed between the equilibrium calculation and the O-D matrix estimation proper, it was decided to combine the O-D matrix estimator and the supply-demand interaction model in a single program.

1.3.2 Issues not addressed in this thesis

As dynamic O-D matrix estimation is at the cross-roads of supply-demand interaction, any issue pertaining to either transportation demand modelling or supply modelling might legitimately be studied, insofar as it affects the O-D matrix estimate. Of the issues not addressed, we mention the estimation of a tour-based matrix, or a multimodal matrix at the conceptual level. At the theoretical level we note that elastic demand is not considered and that the route choice and departure time choice theories exist that call for an explicit account of individual knowledge, appraisal of uncertainty, and learning processes. Also day-to-day dynamics is not considered. At the modelling level more sophisticated choice models exist than are used in this thesis. Further unaddressed issues are how to model queues in the network, how to deal with correlation between observations such as traffic counts, screenline counts, triplength distributions, how to calculate error bounds on the matrix estimates, treatment of elastic demand, sophisticated choice models of route and departure time. In summary we can say that there is scope for follow-up research.
1.4 Contributions of this thesis and practical significance

1.4.1 Main contributions

This thesis makes the following contributions to the subject of dynamic O-D matrix estimation:

1. A behavioural framework is presented in which dynamic O-D matrix estimation can be formulated.

2. The classical approach of equilibrium-constrained O-D matrix estimation is extended to the dynamic case, and a valid and tractable estimation method for dynamic O-D matrix estimation is developed and implemented in software.

3. The dynamic O-D matrix estimation problem is formulated as a Mathematical Program under Equilibrium Constraints, which can be seen as an extension of the MPEC formulation for static O-D matrix estimation. In addition this formulation is very general, and offers rich opportunities for other applications.

4. An operational dynamic O-D matrix estimator is presented in which a limited form of departure time choice is made endogenous through a departure time adjustment model.

5. A distinction is made between departure time choice (as resulting from lifestyle and planned activities) and departure time adjustment (in the sense of coping with traffic conditions without change in activities). Within the bandwidth of the travel time variations considered, the preferred departure times can be expected to remain constant. Given the preferred departure times, there remains a limited trade-off between departure time, travel-time, and schedule delay (i.e. deviations from desired departure time and arrival time), analogous to the relationship between direct and indirect utility.

6. A conceptually unified treatment of departure time choice and route choice is presented on basis of Space-Time extended networks.

7. Analytical derivatives of the dynamic traffic assignment model are calculated, which opens an important avenue for solving the dynamic O-D matrix estimation problem. This was achieved by using an analytic assignment method to solve the dynamic O-D matrix estimation problem instead of following the current trend which emphasises simulation models.

1.4.2 Secondary contributions

Apart from the main contributions to the central problem of O-D matrix estimation, we list a number of secondary contributions that have been established in the course of the dissertation research.
1. An existing parametrisation of the O-D matrix in the static case based on "group observations" is extended to the dynamic case. In this formulation, a multiplicative model based on Maximum Entropy considerations is used to map each independent observation onto a single parameter, which leads to an identifiable estimation problem.

2. An existing statistical approach towards static O-D matrix estimation is adapted to the dynamic case. To this end the static approach is applied to each individual time-slice of the matrix which for this purpose are viewed as statistically independent. Next a single parameterised version of the dynamic O-D matrix is fitted on basis of the objective function derived from the individual time-slices. By using a parametrisation that combines an apriori matrix that has consistent temporal structure with observation-specific parameters, it is ensured that the smallest necessary change is made to the apriori matrix and that the end result is a single coherent dynamic O-D matrix.

3. An existing analytical Dynamic Traffic Assignment approach was restructured and adapted in various ways. Among others, the adapted DTA explicitly calculates and stores the dynamic assignment map with which aggregate observations can be disentangled and related to individual O-D flows. In addition the C-logit and path-size route choice models were added as part of this DTA.

4. A way is proposed to estimate the latent demand matrix (i.e. the dynamic O-D matrix if no congestion were present) by reversing the departure time choice process.

1.4.3 Practical significance

The direct practical significance of this thesis is still limited at this stage of development, as the proposed estimation method has not yet been demonstrated on a full-size (e.g. 400-zone) network. It does however extend the classical static O-D matrix estimation method for general networks and general observations to dynamic O-D matrices, thereby providing an avenue of development towards full-size implementation.

1.5 Thesis outline

An overview of the structure of this thesis is given in Figure 1.2. In the introduction (chapter 1) the problem definition is stated (the estimation of dynamic O-D matrices given that this estimation problem is underdetermined) and motivated through its relevance for transportation modelling. In addition the issues that need to be resolved for this estimation are noted, and the contributions of this thesis are listed.

The question of existence and relevance of a dynamic O-D matrix is discussed in chapter 2. There the context and application of dynamic O-D matrices presented in chapter 1 is elaborated upon, and the question of existence and relevance of a dynamic O-D matrix is discussed in the context of repeated games with incomplete information and user equilibria in transportation systems.
Upon concluding that although both the concept of "the" dynamic O-D matrix and the user-equilibrium are modelling artifacts, they are sufficiently useful to adopt, a behavioural equilibrium framework for O-D matrix estimation is proposed in chapter 3. This framework incorporates route choice and departure time choice equilibria, the formulation of the O-D matrix estimation problem as a bi-level optimisation problem, the role of dynamic traffic assignment as a link between the unobserved matrix and observations, and the incorporation of behavioural aspects through choice models. All four of these elements have their own mathematical formulation, which will be presented in chapters 4 to 7. In addition, existing ideas on the interpretation of any traffic observation in terms of linear combinations of O-D flows as "group observations" and the idea of alleviating the indeterminacy of the statistical estimation problem through the use of apriori information are incorporated and applied to the estimation problem at hand.

Figure 1.2: Structure of this thesis
The prevailing view of O-D matrix estimation is a statistical one in view of the inevitable errors in real-world observations and the intrinsic fluctuations in day-to-day travel demand. Therefore one should cast the O-D matrix problem as a statistical estimation problem. This is done in chapter 4 where modification of an existing formulation for static O-D matrix estimation to handle dynamic O-D matrix estimation leads to a tractable objective function that an O-D matrix estimate should minimise.

The minimisation of this objective function under the assumption of a user-equilibrium with respect to route choice and departure time choice is recognised as a bi-level optimisation problem, and (more specifically) as a Mathematical Program with Equilibrium Constraints (MPEC) and discussed from an optimisation point of view in chapter 5. The upper level of this problem consists of finding an O-D matrix that best fits the objective function. The objective function however is constrained by a user equilibrium in the bottom level of the bi-level formulation.

Satisfying the equilibrium constraints consists of finding a solution to the equilibrium conditions on basis of a given O-D matrix. Solving for the equilibrium is known as the Dynamic Traffic Assignment (DTA) problem, which is covered in chapter 6. The DTA maps the O-D matrix to dynamic traffic flows. By inverting this mapping, observations such as counts and observed arrival distributions can be linked to the unobserved O-D flows, which is needed in order to solve the upper level problem. In order to be able to invert this mapping, assignment fractions must be calculated, stored, and manipulated, for which an existing DTA implementation was extended. Both the dynamic route choice equilibrium determined by the DTA and the departure time choice equilibrium are reflections of user behaviour. The models used to express this user-behaviour are presented in chapter 7.

The MPEC formulation in which the O-D matrix estimation is cast needs special solution algorithms and can be solved using mathematically exact solution methods or heuristics. In chapter 8 solution methods for the mathematical formulation (MPEC) of the O-D matrix estimation problem are compared and discussed, and a heuristic solution method is adopted.

Basic properties of the implementation of the dynamic O-D matrix-estimator are verified on small example networks in chapter 9, thereby providing a check on the validity of the model itself. In addition the software is applied to the Kruiithuisweg urban motorway in chapter 10.

The formulation of the O-D matrix estimation problem used (i.e. as an MPEC) is extremely general, and both the modelling framework and the software developed can be extended with little effort to solve a much wider range of problems than the one for which they were originally designed. Chapter 12 contains a number of promising extensions to the standard O-D matrix estimation problem, along with other possible applications of the behavioural framework that was presented in chapter 3. Our findings and conclusions are presented in chapter 12.
Chapter 2

Context and relevance of dynamic O-D matrix estimation

2.1 Introduction

In the previous chapter we introduced the question of dynamic O-D matrix estimation. In this chapter we will briefly review the context of dynamic O-D matrix estimation (transportation models), and set out to find the O-D matrix estimate that is best suited to prediction. We then turn to the question of whether our problem is well-posed from a theoretical point of view, i.e. whether we can actually expect a dynamic O-D matrix to exist given that the transport system is in a constant state of flux. We will link the question of existence to equilibrium transport modelling and game theory, and point to recent findings in the literature on game theory that indicate that to a certain extent this question can be answered affirmatively.

2.2 The context of dynamic O-D matrix estimation

Origin/Destination (O-D) matrix estimation (whether dynamic or static) is an essential part of transportation modelling as applied to transportation planning. Such models of transportation systems are used to e.g. support infrastructure investment decisions, support what-if analyses on policies, to set fare structures, etc. A description can be found in Cascetta (2001) and Daly (2000). The key element is that in such cases the O-D matrix is used for prediction of future (possibly very near future) travel demand rather than the estimation or reconstruction (e.g. in real-time) of current demand. This is an important distinction in view of the work by Hazelton (2001), who proved that the Bayesian estimators for the mean O-D matrix could be arbitrarily different from those needed for the best possible reconstruction of the current O-D matrix (e.g. the one that gives rise to a certain traffic flow pattern). We will focus on reconstructing the O-D matrix, thereby providing a pivot point on which growth models may be applied to predict future demand.
2.2.1 Quantitative transportation models

A description of the methods used to model transportation systems is given in chapter 1 of Cascetta (2001) and in Ortuzar and Willumsen (2001), and has the structure shown in Figure 2.1.

Shown in the top half is the real world with its travel demand, the transportation network, the traffic situation that results from the interaction of the two, and a set of policy / control measures. In order to assess policy/control measures in a quantitative way one often uses a transportation model (shown in the lower half of the figure) in which the major components: travel demand and the transportation network. Travel demand is primarily driven by autonomous development (e.g. socio-economic and demographic developments, all of which are exogenous to transportation models). Travel demand is represented as a static Origin-Destination matrix $T_h^{rs}$ that gives the travel flows between geographical zones $r$ and $s$. The transportation network is represented as a graph $G(N, A)$ with nodes $N$ and arcs $A$. Each arc has several attributes, such as length, width, free-flow speed, capacity, road type, etc. The traffic pattern that results from the interaction between $T_h^{rs}$ and $G(N, A)$ can be calculated quantitatively, and is represented by a triple consisting of traffic flow $v_a$ on link $a$, travel time $t_a$, and travel cost $C_p^r$ on path $p$ linking the O-D pair $(r, s)$. The set of all paths without cycles between O-D pair $(r, s)$ is denoted as $P_{rs}$.

![Figure 2.1: Generic transportation modelling framework for analysis of policy and control measures](image)

Figure 2.1: Generic transportation modelling framework for analysis of policy and control measures

2.2.2 Applications and examples of transportation models

On basis of such a quantitative representation of the traffic state one can calculate quantitative figures of merit based on predefined evaluation criteria. In this way one
can assess the impact of mobility growth, and compare e.g. policy alternatives, traffic control schemes, and analyse the effect of network changes. The value of the outcome of such computational exercises depends on the faithfulness of the model with respect to the measures being analysed. Many factors affect the faithfulness of the model, and an important one is to have a sufficiently accurate representation of the travel demand. In transportation forecasting many aspects of the travel demand are modelled so that they can respond to changing conditions. One of the most important factors, affecting all aspects of travel demand, is travel time; a large proportion of which is caused by traffic congestion. As the amount of time lost in congestion is a nonlinear function of the demand level (and the infrastructure), the effects are scale-dependent. Almost paradoxically therefore, in view of the many elements that are modelled, the accuracy of the initial level and spatial distribution of travel demand is important.

O-D matrix estimation is the task of obtaining a sufficiently accurate estimate of this demand, both in terms of structure and amount of traffic. Much research effort has been expended on the determination of O-D matrices for the static case, and as a result commercial software packages exist to estimate static O-D matrices (e.g. the well-known packages EMME/2 (see INRO (2002)) and TRIPS (see MVA (2002))).

2.3 Dynamic versus static transportation models

For approximately 40 years, transportation models were static in the sense that they did not reflect the time-dynamics of the transportation system. The past 10 years have seen efforts to move towards time-dynamic transportation models. With respect to the demand side of the models this allows an improved differentiation between the travel patterns during the day, and (given appropriate behavioural models such as discussed in chapter 3) to assess their impact on route choice and departure time. With respect to the transportation network, dynamic modelling allows one to correctly account for the fact that the response of transportation networks to traffic load is a dynamic one (sometimes causing the network to act as a network of waiting queues), and permits the evaluation of time-dependent policy (or traffic control) measures. With respect to the traffic situation, dynamic modelling permits one to improve the modelling of congestion effects such as travel delays, emission levels, noise levels. As a result dynamic assignment models are now becoming operational. One of the building blocks which is still lacking from the toolkit of dynamic transportation models is a flexible and efficient way of estimating O-D matrices for use in dynamic transportation models.

2.3.1 Dynamics in transportation modelling

Dynamics in transportation modelling addresses the time-varying aspects of the transport system, and can either refer to day-to-day dynamics, in which account is taken of the evolution of a traffic system over several days (or months), or it can refer to within-day dynamics where it refers to the changing pattern of use during the day (see also Cascetta (2001), Section 1.2.3).
The within-day dynamics tends to repeat itself day after day. From traffic counts we know that considerable day-to-day variation in traffic flows exists: seasonal variation, variation between weekdays, beginning and end of the holiday season, school holidays. The day-to-day dynamics and the within-day dynamics are illustrated in Figure 2.2.

**Figure 2.2: Day-to-day and within-day variation**

**Definition 1** The O-D demand $T_{h,d,m}^{rs}$ is defined as the number of trips from $r$ to $s$ that departs during time interval $h$ on day $d$ during month $m$.

**Definition 2** The average O-D demand $T_{h}^{rs}$ is defined as the number of trips from $r$ to $s$ that departs during time interval $h$.

This leads us to the question: should we seek an O-D matrix for each period $T_{h,d,m}^{rs}$ (with: $d$ day of week, $m$ month) or should we seek a representative average $T_{h}^{rs}$? Note that we use the time interval $h$ to indicate the time of path entry; general time interval denoted with $k$.

Whilst in practice departure-time choice and scheduling are the result of experience collected over several days, we will consider such choice as exogenous and will not model them. We will take the view that the largest amount of variation in traffic flow patterns and the most interesting dynamic effects such as congestion occur within a single day, and that day-to-day fluctuations can be accounted for in other ways, or can be neglected. We will also assume that we can assume that the daily traffic pattern is sufficiently stable to allow us to model departure time adjustments based on within-day dynamics instead of on day-to-day dynamics.
With these assumptions, we can focus on modelling a single working day. At the worst we will have to model several types of working day, but this is not different in principle from modelling an average working day. From a statistical point of view, we will assume stationarity of the time-series $T^*_h = E(T^*_{hd} | h)$ given the time period $h$, and take the view that:

$$
T^*_{hd} = T^*_h + (T^*_{hd} - T^*_h)
$$

$$
= T^*_h + \varepsilon
$$

with $E\varepsilon = 0$, and focus on within-day dynamic matrix estimation for working days as it applies to matrices for transportation planning models.

This is a fairly strong assumption, but a useful one which in our opinion is justified if one is interested in the long-term properties of the transportation system. We note that, if necessary, the error committed in this way can be reduced including seasonal effects, and repeating the estimation for each on several datasets, each representing a particular type of day.

This choice has the following implication: only a *representative average* (if one exists) of the true dynamic O-D matrix is sought. This is tantamount to an assumption of a sufficient degree of stationarity of the traffic process for the type of day that is estimated. Therefore, as in static O-D matrix estimation, the day-to-day variation will be reflected in the estimation process as an *error term*. When datasets of observations from different days are used this error term will likely result in *inconsistencies between these datasets*, and adversely affect the quality of the estimation.

**Within-day dynamic transportation models**

Although the existence of peak hours is nothing new, we will show some empirical evidence on the form of the peaks, to illustrate that even with the current flexibility in working hours, demand management, teleworking, and multi-model transport, demand is still peaked and that various travel purposes peak at different times and that the largest contribution to .

**Empirical evidence** The within-day profile sketched in Figure 2.2 corresponds to the profile found in empirical data, as shown in Figure 2.3 for all transport modes based on the Dutch Annual household survey OVG 1998. (see e.g. Excl (1997) for a description). Clearly travel to work forms the largest fraction of the distance-weighted trips (veh*Km) between 06:00 hr and 09:00 hr., and that travel to home forms the largest fraction between 15:30 hr and 18:00 hr with a secondary peak around 22:00 hr. The same holds for travel mode car driver, as shown in Figure 2.4).
Figure 2.3: Number of trips (all modes) by purpose and departure time (weighted) (Source: OVG 1998)

Figure 2.4: Number of trips (car driver) by purpose and departure time (weighted) (Source: OVG 1998)

Modelling consequences  Clearly the within-day variation of travel demand should be accounted for in transportation models. We will assume discretised time throughout.
With the incorporation of dynamic travel demand $T_{rs}^h$ (the travel demand from zone $r$ to zone $s$, departing during time period $h$), the model in the lower half of Figure 2.1 becomes

\[ T_{rs}^h \rightarrow \text{Simulated policy / control measures} \rightarrow G(N,A,k) \]

\[ \text{Modelled traffic situation} \ (v_a(k), \tau_a(k), C_p^r(h))_{a \in A, h \in H, k \in K, rs \in Z \times Z, p \in P_a} \]

\[ \text{Evaluation criteria} \]

Figure 2.5: Generic form of a dynamic transportation model

The transportation model shown in 2.5 is identical to that in the lower half of Figure 2.1 except for the addition of the variables $h$ and $k$, which stand for sufficiently small discrete time periods. Travel demand is represented as a dynamic Origin-Destination matrix $T_{rs}^h$. The transportation network is represented as a graph $G(N,A,k)$ with nodes $N$ arcs $A$, related to time period $k$. Within-trip rerouting is not modelled. The traffic pattern that results from the interaction between $T_{rs}^h$ and $G(N,A,k)$ can be calculated, and is represented by a triplet consisting of: 1) traffic flows $v_a(k)$ on links $a \in A$ during time period $k \in H$, 2) travel times $\tau_a(k)$, and 3) travel costs $C_p^r(h)$ on paths $p \in P_{rs}$ linking the O-D pairs $rs \in Z \times Z$.

Day-to-day dynamics in transportation models

As noted, there is a considerable fluctuation in day-to-day traffic flows; we will assume that we have day-dependent travel demand, and denote the total travel demand matrix on day $d_i$ as $T_{rs,d_i}$. We will not present mechanisms for the generation of $T_{rs,d_i}$ but assume that it is given (or sufficiently predictable). Given a total daily travel demand $T_{rs,d_i}$, the traffic state is determined by the choice of departure time, and the choice of route.

We will briefly describe the relationship between the static matrix $T_{rs,d_i}$ and the dynamic matrix $T_{rs,d_i}$. In the literature (see Cascetta and Cantarella (1991), Cascetta (1989), and Cantarella and Cascetta (1995)) we find the following mechanism for day-to-day dynamics. Both route choice and departure time choice are based on the perception $C_p^{rs,d_i}(h)$ of the time-dependent generalised transport cost based on the experienced cost $C_p^{rs,d_i}(h)$ on day $i$ by each individual user of the transport system. The departure time choice leads to a dynamic O-D demand $T_{rs,d_i}$, which leads to time-dependent path flows $F_p^{rs,d_i}(h)$ for each departure time interval $h$, each path $p$ between $r$ and $s$ entering the network. The traffic process that results from the flows $F_p^{rs,d_i}(h)$ induces objective generalised costs $C_p^{rs,d_i}(h)$ which are experienced by the transport network users, and
serve to update their perception $\tilde{C}_{p}^{rs,d_i+1}(h)$ of these costs for the following day. This is illustrated in Figure 2.2.

![Diagram](image)

**Figure 2.6: A mechanism for day-to-day dynamics in the traffic pattern**

Clearly we have a relationship between the perceived transport costs $\tilde{C}_{p}^{rs,d_i}(h)$ and the route flows $F_{p}^{rs,d_i}(h)$ on day $d_i$ and the perceived cost $\tilde{C}_{p}^{rs,d_i+1}(h)$ of the following day. It turns out (as shown in Cantarella and Cascetta (1995)) that the stability of the system depends on the updating mechanism: if travellers don't update their cost perception too abruptly (e.g. in a way that allows memory of past situations to dampen the response to yesterday's experiences) then we may expect an equilibrium, or a band of oscillations around this equilibrium. If travellers respond violently to their most recent experience we may expect an oscillatory or chaotic system (as shown in Cascetta and Cantarella (1991)).

**Consistency and equilibrium**

In transportation models one typically demands *consistency* between variables. Applied to the mechanism for day-to-day variation in Figure 2.6 this means that we seek within-day consistency between path costs and path flows and between link travel times and link flows. Regarding the between-day dynamics, we can either take account of the day-to-day evolution of the experienced costs (which we will not do), or we can assume consistency between the experienced costs on day $i$ and those on day $i+1$. Mathematically this means that in both cases we demand the existence of a fixed point in the map relating the *experienced* within-day dynamic cost and in the resulting traffic flows. In both cases the assumption of a fixed point of these maps is equivalent to postulating an equilibrium in the game-theoretic sense (see Section 2.4.1).
2.3.2 Dynamic O-D matrix estimation

The aim of dynamic O-D matrix estimation is to find a dynamic O-D matrix $T^{rs}_h$ that is a sufficiently accurate. In practice this translates into two separate requirements on $T^{rs}_h$:

- it should reproduce observed flows after assignment,
- its structure should not deviate too much from reality in the sense that its margin totals should match the (observed) margin totals, and that the interaction pattern between zones should match reality.

The first requirement ensures that the overall level of traffic is correct, and is important when congestion is significant. It is well known that the first requirement generally does not "fix" the matrix (typically many matrices can be found that satisfy this requirement), and therefore the second requirement is needed. The second requirement is even more important for the analysis of measures that have an impact on only part of the network. This is because only the people that are affected by the measure in question will react, each according to his or her specific situation, which strongly depends on their origin $r$ and destination $s$. As the only information on the spatial distribution of traffic is carried by the O-D matrix, the degree to which the affected population can be correctly identified depends on the correctness of the O-D matrix.

Classification of relevant O-D matrix estimation methods

Matrix estimation methods can be classified in various ways; one of them is by comparing the assignment technique that they are designed for, and the type of matrix that they estimate, as shown in table 2.1.

<table>
<thead>
<tr>
<th>Estimated O-D matrix</th>
<th>Assignment method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static</td>
<td>Static</td>
</tr>
<tr>
<td>Dynamic</td>
<td>Dynamic</td>
</tr>
<tr>
<td></td>
<td>Fully static</td>
</tr>
<tr>
<td></td>
<td>Extended static</td>
</tr>
<tr>
<td></td>
<td>Repeated static</td>
</tr>
<tr>
<td></td>
<td>Dynamic</td>
</tr>
</tbody>
</table>

Table 2.1: Classification of setups for O-D matrix estimation

The combinations are as follows:

**Fully Static** This is the classical case; both the estimated matrix and the assignment are static. See Cascetta (2001) for a description.

**Extended static** In this case the estimated static O-D matrix is divided into equal slices, one for each time interval, resulting in a steady flow of traffic with constant intensity. The network response to this demand is dynamic.
Repeated static In this case the estimated O-D matrix consists of several time-slices, all of which are assigned statically (possibly with queues being carried over from one period to another). A software package that supports this type of assignment is SATURN.

Dynamic This is the case that we will consider: both the estimated matrix and the assignment method are dynamic.

Sequential and simultaneous estimation

Given that we will concentrate on dynamic O-D matrix estimation with dynamic assignment, a further subdivision into sequential and simultaneous estimation methods can be made. In sequential estimation the slices of the O-D matrix are estimated one after another, using observations that are available up to (and including) the time period of the matrix slice, and often taking the estimation result of the previous time slice as a starting point for the estimation. The estimates of individual time slices are not revised after they are estimated. This method is the only option in on-line settings, but is also applicable in off-line settings.

In simultaneous estimation, the entire dynamic O-D matrix is estimated at once, which is only possible off-line due to its data requirements. The sequential estimation results in smaller estimation problems, but the simultaneous estimation is more appealing from a theoretical point of view. We will focus on simultaneous O-D matrix estimation.

2.4 Equilibrium modelling and existence of the O-D matrix

As one of the fundamental assumptions of O-D matrix estimation (whether dynamic or static) is that of the existence of an O-D matrix (at least to the extent that it is a useful approximation of the transportation system), this existence claim should be considered. As will be shown in Section 2.4.1, game-theory provides useful results in this respect and it turns out that for a subset of games, so-called potential games, the literature contains proofs that the repeated game will converge to its equilibrium. In addition it will be argued that the convergence results found in the literature on so-called potential games provide some justification for assuming the existence of an equilibrium state (at least as a modelling device).

2.4.1 Equilibrium and game theory

Game-theory has been used to provide a framework to describe the connection between the macroscopic state of a transportation system and the collective choice of transportation system users ever since the Wardrop principle was formulated (Wardrop (1952)) given a certain demand pattern. The essence of a network equilibrium was given in Wardrop (1952), who proposed an equilibrium in route choice such that no individual driver could unilaterally improve his travel time. This formulation is cast in the form
of a n-person non-cooperative game in which drivers are the participants, route choice is the strategy, and route travel time is the pay-off of a strategy. By taking a route drivers cause the travel time on that route to increase. The traffic equilibrium (if one exists) is then an instance of a Nash equilibrium.

It can also be used to provide a framework for the evolution of that demand pattern as this situation can be interpreted as a repeated game (see Basar and Olsder (1999) for a definition, and Sandholm (2001) for an application, and Altman and Wynter (2003) for a recent overview).

From this point of view we can regard the day-to-day evolution of demand as the outcome of a repeated game. The within-day consequence for the state of the transportation network is traditionally modelled as a one-shot game, but can also be seen as part of the repeated game. Viewed in this framework, the question of existence of network equilibrium conditions and their appropriateness to represent reality becomes one of modelling accuracy, and not one of principle.

Convergence results

The evolution of a traffic system with learning users converges to equilibrium if it is a so-called "potential game" (see Sandholm (2001)). Classical game theory presupposes that each participant in the game is aware of the payoff of each strategy, meaning that each driver is assumed to know the traffic situation on the entire network. In view of the discussion above this is not realistic. A partial solution is presented by recent findings in game theory which show the following:

- a system is stable under bounded rationality if and only if it is structurally stable (see Anderlini and Canning (2001)). Structural stability means that a small change in the system parameters leads to a small change of the equilibrium set. Given the observed robustness of traffic levels, it seems reasonable to assume that the traffic system is structurally stable, and thus stable under bounded rationality.

- the continuity of the dependence of the solution follows from the theory of network equilibria as projected dynamical systems, as proposed in Nagurney and Zhang (1998), who also prove that the equilibrium solution is a "monotone attractor" provided that the cost function is continuous and locally monotone. The stronger the assumptions on the cost function (in terms of the monotonicity of the cost function), the stronger the results on the "attractorship" of the equilibrium solution (under sufficiently strong conditions of monotonicity the equilibrium solution even becomes a fixed-time attractor).

On this basis we conclude that although in reality we can't expect true convergence to equilibrium, there is some justification in expecting stochastic convergence to an equilibrium solution. This stochastic convergence is not adversely affected by bounded rationality (i.e. limited knowledge), provided that the location of the equilibrium depends in a continuous way on the parameters of the system. As we can model the entire
choice hierarchy as route choice through a suitably defined network, the existence of an O-D matrix (in a stochastic sense) seems assured.

Unfortunately multiple equilibria may exist in the presence of heterogeneous populations as shown in Konishi (2003) and in the presence of adaptive traffic control as shown in Van Zuylen and Taale (2000). It seems that adaptive traffic controllers (i.e. controllers that can extend the green time when the flow increases) destroy the monotonicity of the cost function. It remains to be seen to what extent traffic control and populations that are heterogeneous in a "smooth" way can cause instability in the mode-choice, departure-time choice, and activity scheduling processes.

2.4.2 Equilibrium and projected dynamical systems

Network equilibria (both dynamic and static) studied in transportation can be formulated as variational inequalities. In Nagurney and Zhang (1998) an overview and synthesis is given, showing that with every equilibrium formulated as a VI problem, a dynamic system can be associated that has the equilibrium as its stationary point. The mathematical framework is called projected dynamical systems.

Within this framework, Nagurney and Zhang (1998) show the stability and asymptotic stability of traffic equilibria with elastic demand under a general route choice adjustment procedure, provided that the link cost functions and the demand functions are strictly monotone.

2.4.3 Consequences for the equilibrium concept

As noted in sections 2.4.1 and 2.4.2, game theory and the theory of projected dynamical systems provide some justification for the idea that under fairly strong conditions a true steady-state solution to the day-to-day adjustment process exists, and that under mild conditions an expected average state of the transportation system emerges from the day-to-day adjustment process in transportation systems.

If the perception of transportation systems users is based on a damped learning process, i.e. that it does not respond too abruptly to the most recent experience, then Cascetta and Cantarella (1991) show that the state of the transportation system and also users' perception of its level of service stabilises, or takes on a periodic solution (they also show that if user's perception does respond strongly to the most recent experience then the state of the transport system may develop in a chaotic fashion).

The existence of such an expected average state means that e.g. choice of departure time and route (which strictly speaking results from the day-to-day adjustment process) can be modelled as if they were dependent only on the within-day conditions of the transportation system.

2.4.4 Equilibria and reality

The relationship between equilibrium solutions of transportation models and the reality of the transport system is a complicated one. On the one hand equilibrium modelling
is an elegant, powerful, and expressive simplification of reality, but on the other hand there are good reasons to view the claim that the equilibrium situation represents reality with suspicion. In fact, the whole concept of O-D matrix estimation points to a weakness in the current generation of demand models: if they were perfect then O-D matrix estimation would be redundant whenever a demand model is available.

**Critique of equilibrium models**

Strictly speaking the idea that equilibrium states are the system states that one would expect to see in reality during a randomly selected day is false. This is argued by a number of authors, e.g. Goodwin (1998), Ortuzar and Willumsen (2001) (p.385), Bell and Iida (1997) (p.2). Even if one assumes the existence of a system that tends to equilibrium, in reality such systems are always subject to unpredictable external influences, so that even as the system adapts it is no longer in equilibrium.

Furthermore different levels of the choice hierarchy in transportation modelling (as proposed in e.g. Ben-Akiva and Lerman (1989)) have different time-scales at which decisions can be adapted to changing circumstances: long time-scales (decades) for urban development decisions, medium time-scales (years) for mobility decisions such as the choice of workplace, residential location, car ownership levels, and short time-scales (months or weeks) for travel decisions such as travel mode, route, and departure-time. If changes occur, the slower-reacting levels will lag behind -by definition-, which may result in a state of disequilibrium over an appreciable period of time.

Counter-examples to the equilibrium hypothesis are:

- the disparity in the actual travel demand between Zoetermeer and The Hague and what would be expected on basis of their respective socio-economic characteristics

- disequilibrium due to the difference in time-scales between the adjustment of travel demand and that of office locations (see Van Wee and Van Der Hoorn (2002))

- supply in the case of long-term decisions such as land-use, as exemplified by the current housing market in The Netherlands where demand exceeds supply in various segments.

In addition the assumption of complete rationality together with full information of all participants in the transportation system is obviously false in the case of road users (as indicated by the dissemination of traffic information for motorways) and users of public transport systems (e.g. when train service is interrupted due to an accident, a breakdown of rolling stock or the tracks).

The article by Gunn and Van der Hoorn (1998) shows significant discrepancies between the evolution predicted the LMS and observed effects after 15 years, which couldn't all be unequivocally explained by a change in the inputs to the LMS.

Furthermore one may question the realism of equilibrium modelling in the presence of multiple equilibria, a stylised example of which was presented in Van Zuylen and Taale (2000).
Last but not least, if one considers the day-to-day evolution of a transportation system and accepts the existence of certain natural mechanisms such as updating of users' choice based on travel experience (see Cascetta (2001), chapter 5.8), then one finds that the attractors of the resulting dynamic system may be fixed points, k-periodic, quasi-periodic, or even a-periodic, depending on the speed of information dissemination and the strength of the response of decision makers in reaction to new information Cascetta and Cantarella (1991). In view of these considerations, fundamental doubt arises as to the theoretical credibility of equilibria in transport systems as well as to the practical value of equilibria.

**Defense of equilibrium models**

A number of arguments can be put forward in the defense of the realism of equilibrium modelling as applied to transport.

The empirically observed variance in daily traffic volumes on ordinary roads and motorways is not large enough to reject the hypothesis of equilibrium states with an overlay of noise, provided one accounts for predictable and understandable variation (e.g. related to day of the week season through month of the year). This is substantiated by findings in e.g. Aldrin (1995) who decomposed traffic flow time-series into their long-term, seasonal, and daily components and investigated the influence of special days and unaccounted errors for a toll cordon around Oslo, in Stamatiadis and Allen (1997), and Davis (1997) who investigated the accuracy of the Annual Average Daily Traffic (AADT) of several studies in the US, and in Ritchie (1986) who estimated a model for the AADT in the USA. Their findings support the idea that the unaccounted error is much smaller (typically 11%-15% of the AADT) than the AADT itself. Although these findings relate to traffic counts, we assume that the level of service and the traffic state as a whole vary proportionally to the observed traffic counts, and therefore have the same level of variance. This supports the assumption of the existence of a steady-state (or equilibrium) traffic state, at least to a fair approximation.

The current state of the art in static O-D matrix estimation treats O-D matrix estimation as a statistical estimation problem (see e.g. Cascetta (2001), chapter 8). The stochastic nature of the problem can be attributed to a measurement-related component induced by sampling procedures and measurement error, and an intrinsic component which can be interpreted as the O-D matrix itself having a stochastic character. As we cannot differentiate between the two on basis of available data, this assumption leads to the same statistical estimation problem and is to be preferred in view of the day-to-day variability.

In addition, there is the evidence from discrete-choice modelling studies which suggest that in applications such as mode choice, destination choice, route choice, most of the variance in individuals' behaviour can be understood and explained in terms of utility maximisation. If we accept this, then we have a clear adaptation mechanism which, at individual level, prompts people to select alternatives with maximum utility, or in the case of transport, with maximum indirect utility. Of course in reality more factors influence the choice process than simple utility maximisation. Examples in the area
of route choice can be found in Bovy and Stern (1990), and an argument for bounded rationality can be found in Asakura and Hato (2000).

Last but not least, the equilibrium states of a repeated Nash game constitute a characteristic of that game, comparable to the relationship between a stochastic variable and its expectation and its mean. The fact that one cannot count on observing the mean in any particular case does not detract from its usefulness as a characterisation of the stochastic variable, although one would want to know the variance as well.

We will therefore take the position that equilibrium states do have significance in relation to the transportation system, but that one should be cautious in interpreting the outcomes of such equilibria.

2.4.5 The relevance of equilibrium models

Apart from the realism of equilibrium models, there is an important reason to use them e.g. in forecasting: consistency. An equilibrium model guarantees consistency of all endogenous variables within the model (in so far as these are linked through feedback loops within the model), and it guarantees consistency of the endogenous variables with the exogenous variables.

This is crucial in view of the multitude of interactions and feedback loops that have been identified in transportation systems, because one needs a quantitative way of dealing with counteracting influences in order to assess the outcome of their joint action. Thereby it allows the equilibrium model framework to be used as a yardstick with which various policies and/or network variants can be compared, in the certainty that "all other things are equal" (at least the feedback mechanisms are). The predictions may well be wrong in an absolute sense, but they will generally be wrong by the same amount and in the same direction for all policy variants considered.

Internal consistency, feedback, and fixed points

Generally, demand models will be required to be internally consistent, in the sense that the level of service induced by assigning modelled demand to the network will not lead to changes in that demand. In other words, a fixed point or an equilibrium is assumed in the choice behaviour. This assumption implies a dynamic equilibrium at each level of the choice hierarchy. Note that the identity of the decision makers making the choices is not specified, and in principle the model is thus indifferent to people changing their choice on a daily basis. However, this indifference greatly depends on the amount of detail in the model (origin, destination, departure time, purpose, age group, gender, income class etc.). Mathematically, this is expressed by demanding that the predicted demand should be a fixed point of the model system. This type of constraint is used in Cascetta and Cantarella (1991), and more recently a variational inequality approach was used in the ESTRAUS model for Santiago (see Florian et al. (2000), ) to calculate the fixed point of collections of interacting models.

Regardless of the question whether such model systems are complete with respect to feedback, some equilibria will, according to the choice hierarchy, be reached quicker
than others. (familiarity with alternatives, within-household constraints, contractual obligations, location of home, job, etc.), so that it is often not clear if the system is in equilibrium at all!

Usually an equilibrium assignment is assumed as a model for the interaction between traffic demand and level of service, and equilibrium between transport mode and level of service.

**Relevance for matrix estimation: snapshots**

The relevance of the foregoing discourse on equilibria and dynamical systems for matrix estimation is that:

1. in practice one will not estimate a matrix on a lengthy time-series of day-to-day traffic observations, but on a *snapshot* of dynamic traffic observations within a single day, augmented by e.g. yearly workday averages for other quantities such as trip margins and triplength distributions. This is needed to ensure different from the practice. Applying such snapshots to a general day implies the assumption that the snapshots are representative of the state of the system. This in turn requires the assumption that the system is (more or less) in equilibrium.

2. the assumptions made (sometimes implicitly) in matrix estimation are made explicit by presenting a transport model with those assumptions built in.

3. usually the majority of the matrix cells are unobserved so that a model is needed to estimate the unobserved cells.

4. that when the matrix is to be used for prediction, a model is needed to estimate the changes.

Even if pivot point models are used, part of matrix estimation can be done by estimating parameters of the transportation models used. This is already done when distribution functions are estimated, but this approach can be extended.

**Modelling focus**

With respect to matrix estimation, the focus of the modelling should probably be on the lower three components of the choice hierarchy: trip generation, mode choice, and route choice, and feedback. For dynamic matrix estimation this includes departure time estimation.

**2.5 Summary**

In this chapter we noted how O-D matrix estimation fits into a generic transportation model. We have also noted the difference between within-day dynamics and day-to-day dynamics, and motivated our choice to focus on within-day dynamics. We have
noted the relationship between the assumption of internal consistency and equilibrium assumptions, and stressed the role of equilibrium modelling. The constant state of flux of the transportation system in which traffic conditions change from day to day leads to the question of the validity and relevance of equilibrium modelling for this type of system. As the concept of a daily O-D matrix is closely linked to the existence of an equilibrium state, this is an issue that should be faced.

Theoretical results from the literature on game theory, day-to-day dynamics, and projected dynamic systems provide some justification for the assumption that the concept of an equilibrium state has some reality value. In addition, there is empirical evidence that the amount of within-day variation in the Annual Average Daily Traffic is much smaller than its mean.

We will use these results as a justification for the assumption that an equilibrium state and therefore its accompanying average dynamic O-D matrix exist, and that departure-time adjustments can be modelled as if they were dependent on within-day traffic conditions instead of emerging from a day-to-day process. We will also take these results as a "license" in the next chapter to formulate a behavioural framework based on supply-demand interaction and market equilibrium into which dynamic O-D matrix estimation can be embedded and formulated.
Chapter 3

A behavioural framework for dynamic O-D matrix estimation

3.1 Introduction

In the previous chapter we have selected an approach towards dynamic O-D matrix estimation based on supply-demand interaction and market equilibrium, and sketched the transportation model to which dynamic O-D matrix estimation contributes. We distinguished between within-day-dynamics and between-day dynamics, and selected the within-day dynamics as the focus of our work. We noted the role of equilibrium modelling and discussed its consequences and its justification, which implies that an equilibrium O-D matrix can be assumed to exist.

This chapter presents an equilibrium-based behavioural analytical framework which we will adopt for dynamic O-D matrix estimation. This framework is a synthesis of elements found in the literature which identifies links between observable quantities and commonly available observations of the transportation system and adds a parametrisation for the O-D matrix. It uses the idea that all observations refer to linear combinations of O-D flows and extends this to the dynamic case. Furthermore it unifies the choices that users can make in the course of their trip (departure time, travel mode, and route) as route choice in an extended network. This network is called a supornetwork when e.g. multiple travel modes are represented in the network, and a Space Time Extended Network (STEN) when the network explicitly incorporates time. The matrix estimation will be formulated as a minimisation problem under equilibrium constraints in Section 3.5, and part of the contribution of this chapter is to show that departure time choice can be grouped with route choice as an equilibrium problem on a STEN in Section 3.8.
3.2 State of the art

3.2.1 Literature

Dynamic O-D matrix estimation builds on static O-D matrix estimation. The issue of static O-D matrix estimation is a classical one, and considerable work has been carried out on static O-D matrix estimation. For an overview of static O-D matrix estimation we refer to Abrahamsson (1998), Cascetta and Nguyen (1988), Cascetta (1984), and Cascetta (2001). As a recent development we mention Yang et al. (2001), who simultaneously estimate an O-D matrix and the route-dispersion coefficient of a SUESTOCH assignment using a non-linear programming approach. Particular attention is paid to the statistical aspects of static O-D matrix estimation in Hazelton (2000). A formulation using bi-level programming is presented in Florian and Chen (1991) and Yang et al. (1994). A general approach using fixed-point formulations is presented in Cascetta and Postorino (2001). Interestingly the dynamic O-D matrix estimation problem also occurs in the design of datacommunication networks (e.g. the Internet), as shown in Cao et al. (2000), and Medina et al. (2002), although their assumption of fixed routing is less appropriate to transportation networks.

A framework for general dynamic O-D matrix estimation is given in Cascetta et al. (1993). An estimation approach based solely on traffic counts for general networks is given in Sherali et al. (2001). For the special case in which the entire network on which the O-D matrix is to be estimated several sources are available such as Kremer and Keller (1987), Nihan and Davis (1987), Bell (1991), and Van Der Zijpp (1996).

3.2.2 Existing dynamic O-D matrix estimation methods

Existing dynamic O-D matrix estimation methods can be broadly categorised as on-line methods (in which new information on the evolution of the O-D matrix under consideration arrives piecemeal), and off-line approaches (in which all data pertaining to the dynamic O-D matrix is available at once). A framework for dynamic O-D matrix estimation was presented in Cascetta et al. (1993), in which a distinction is made between simultaneous dynamic O-D matrix estimation (in which the entire dynamic O-D matrix is estimated in one stroke) and sequential O-D matrix estimation (in which the dynamic O-D matrix is estimated separately by time-slice). It is clear that for on-line applications a sequential estimation method is the only way forward. We aim for an off-line O-D matrix estimator, and can therefore use both sequential and simultaneous methods.

We will add distinctions between behavioural and purely statistical approaches, between corridor networks and general networks, and between methods that have been formulated only for corridors versus models that have been formulated for general networks, between methods that use only counts and methods that use more general observations, methods that use apriori matrices (implicitly or explicitly), the type of objective function used, and whether deterministic or stochastic networks are used. This gives us ten axes along which we can group approaches to dynamic O-D matrix estimation as summarized in 3.1:
<table>
<thead>
<tr>
<th>Nr.</th>
<th>Attribute</th>
<th>Attribute levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Time-frame</td>
<td>on-line</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⇐⇒ off-line</td>
</tr>
<tr>
<td>2</td>
<td>Problem formulation</td>
<td>Simultaneous, sequential</td>
</tr>
<tr>
<td>3</td>
<td>Model approach</td>
<td>statistical, behavioural</td>
</tr>
<tr>
<td>4</td>
<td>Type of network</td>
<td>closed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⇐⇒ general</td>
</tr>
<tr>
<td>5</td>
<td>Route-alternatives</td>
<td>present</td>
</tr>
<tr>
<td></td>
<td></td>
<td>⇐⇒ absent</td>
</tr>
<tr>
<td>6</td>
<td>Use of data</td>
<td>counts only, any grouped observation</td>
</tr>
<tr>
<td>7</td>
<td>Use of apriori matrix</td>
<td>optional, needed</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>&quot;explicit&quot;, &quot;implicit&quot;</td>
</tr>
<tr>
<td>9</td>
<td>Objective function</td>
<td>ML, ME, quadratic</td>
</tr>
<tr>
<td>10</td>
<td>Network characteristics</td>
<td>deterministic, stochastic</td>
</tr>
</tbody>
</table>

Table 3.1: Approaches to dynamic O-D matrix estimation

The first axis of distinction is on-line versus off-line estimators. On-line estimators must work in real-time and therefore have limited data availability and must work faster than real-time. For off-line applications one can distinguish between (see 2) simultaneous estimation, in which the entire dynamic O-D matrix is estimated at once, and sequential dynamic O-D matrix estimation, in which the O-D matrix is estimated timeslice by timeslice. The model approach can either be a purely statistical one (see 3) in which scant attention is paid to behavioural aspects of the underlying transportation phenomena) or can incorporate behavioural elements. The type of network (see 4) can either be a closed one (in which it is essential that all entries and exits to the network are continuously monitored) or a general one, in which this assumption is relaxed. This is important if the approach is to handle real-world situations, in which a "closed" system is the exception rather than the rule. The presence or absence of route-alternatives (see 5) in a network determines whether the problem of relating aggregate traffic observations (such as counts) to individual O-D flows has a spatial component. If it has, disentanglement is more complicated, and partially depends on the quality of the route choice models. Then (see 6) one can distinguish between estimation methods that exclusively rely on traffic counts and those that can incorporate other data sources. Another characteristic (see 7) is the presence or absence of the use of apriori matrices, whether this use is explicit (see 8) apriori matrix given) or implicit (e.g. selection of the most likely matrix through the Maximum Entropy principle). Also (see 9) the objective function can vary from the methodologically correct Likelihood function (ML) and Minimum Information estimators (see e.g. Van Zuylen (1983)) to the more pragmatic quadratic and Maximum Entropy methods (see also Van Zuylen (1983)). Finally (see 10) one can consider networks that are deterministic (all path costs are known deterministic functions of the traffic load) and networks that are stochastic (e.g. because link travel times are have a stochastic character over and above the impact of traffic, such as road works or incidents).

We note that stochastic networks can be reflected in the standard behavioural models by incorporating e.g. the variance of the stochastic travel costs as an additional variable and segmenting the population into risk averse, risk-prone, and risk-neutral segments. In this way the aspect of stochastic networks can be incorporated into the standard framework of transportation modelling.
Table 3.2 summarises the following applications of dynamic O-D matrix estimators in the literature; we have added our model for reference:

<table>
<thead>
<tr>
<th>Reference</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kremer and Keller (1987)</td>
<td>on-l.</td>
<td>sequential</td>
<td>statistical</td>
<td>closed</td>
<td>no</td>
<td>counts</td>
<td>optional</td>
</tr>
<tr>
<td>Camus et al. (1997)</td>
<td>on-l.</td>
<td>sequential</td>
<td>both</td>
<td>general</td>
<td>yes</td>
<td>counts</td>
<td>optional</td>
</tr>
<tr>
<td>Ashok (1996), Sherali et al. (2001)</td>
<td>both</td>
<td>sequential</td>
<td>statistical</td>
<td>general</td>
<td>yes</td>
<td>counts</td>
<td>optional</td>
</tr>
<tr>
<td>Van Der Zijpp (1996)</td>
<td>off-l.</td>
<td>simultaneous</td>
<td>statistical</td>
<td>closed</td>
<td>no</td>
<td>counts</td>
<td>optional</td>
</tr>
<tr>
<td>Tavana and Mahnasaani (2001)</td>
<td>off-l.</td>
<td>simultaneous</td>
<td>routechoice</td>
<td>general</td>
<td>yes</td>
<td>counts</td>
<td>needed</td>
</tr>
<tr>
<td>this thesis</td>
<td>off-l.</td>
<td>simultaneous</td>
<td>behavioural</td>
<td>general</td>
<td>yes</td>
<td>group</td>
<td>needed</td>
</tr>
</tbody>
</table>

Table 3.2: List of approaches to dynamic O-D matrix estimation

The column headings in Table 3.2 are: (1) on-line / off-line, (2) sequential versus simultaneous estimation, (3) statistical or behavioural approach used, (4) closed (i.e. all entries and exits fully observed) or general networks, (5) route choice incorporated, (6) type of data that can be used, and (7) whether an apriori matrix is needed or optional.

In this thesis we will present a behavioural off-line method for simultaneous estimation which is suitable for general networks and can handle any grouped observations including traffic counts.

3.2.3 Time discretisation

Throughout this thesis we use a discrete-time formulation for the DUE-STOCH and DTAC-STOCH for practical reasons.

In the discrete-time formulation we distinguish three different discretisations, leading to three sets of intervals:

- a coarse external discretisation with big intervals (typically related to the time-discretisation of the externally given apriori O-D-t matrix and e.g. the time-resolution of external data sources such as traffic counts, production-attraction etc.. Large intervals are denoted by a different font style: $\mathfrak{H} \in \mathfrak{F}$ for the large time interval in which traffic enters the network, and $\mathfrak{T} \in \mathfrak{R}$ for the large interval during which an observable quantity is considered (e.g. flows that correspond to counts, or to production-attraction observations).

- a medium-size internal discretisation, with big intervals defined by the minimum free-flow link travel time $\Delta$ ($\Delta = \min_{a \in A} t_{free}^{a}$).

- a fine internal discretisation with small intervals of width $\delta$, which are related to the internal structure of the model by the relationship $\delta = \frac{1}{M}$, with $M \geq 1$ given. Small intervals are denoted by $h$ (reserved for the small time interval during which traffic enters the network, and $k$ which is reserved for small time periods in which network quantities are considered.
Internally the O-D-t matrix is stored as demand per small time interval; this is calculated (and where necessary interpolated) from the time discretisation of the external O-D-t matrix.

### 3.3 Transportation modelling framework used

In this chapter we will sketch the transportation modelling framework that we will use, followed by a general approach towards matrix estimation. We assume that the O-D matrices are fixed (i.e. we have rigid demand), are unobserved, and that their estimation from observed quantities is strongly overparametrised (i.e. the number of unknowns exceeds the number of independent observations by a large margin).

Figure 3.1 shows the causal structure of static and dynamic (discrete-time) transportation models when rigid demand is assumed.

![Figure 3.1: Relationship between O-D matrix and assignment in the static and dynamic cases (rigid demand)](image)

In the **static case** a vector $T^{r+}$ of zonal trip production is split across destinations $s$ by a distribution model given by its destination-choice probabilities $P_{sl|r}$ into an apriori O-D matrix $T^{rs,0}$ (only one travel mode is considered). The resulting travel demand represented by matrix $T^{rs,0}$ is assigned to the transport network, thereby causing path flows $F_p$, link traffic flows $\{u_a\}_{a \in A}$ on every link $a$ of the transport network, which induce link travel times $\{\tau_a\}_{a \in A}$. The path travel costs $C_p(h)$ are affected by the link travel times on the individual links of the path, and influence the path choice and hence the path flows $F_p$. Therefore a circular relationship ($F_p \rightarrow u_a \rightarrow \tau_a \rightarrow C_p \rightarrow F_p$) exists.

In the **dynamic case**, a dynamic travel demand matrix $T^{rs}_h$ can be thought of as the superposition of a time dimension given by a departure time choice model $P_{hrs}$ giving
the probability of departing during time interval $k$ given that origin $r$ and destination $s$, on the static apriori matrix $T^s r$. On assignment this matrix $T^s r$ leads to time-dependent path flows $F^p_h$, which in turn lead to time-dependent link flows $\{v^a_h(k)\}_{a \in A, k \in K}$. The link flows determine link travel times $\{\tau^a_h(k)\}_{a \in A, k \in K}$, which affect the time-dependent path travel costs $C^p_h$ on each path $p$ during departure time interval $h$, which in turn determine the path flows $F^p_h$. As with static models, one can solve for consistency between path flows and the transportation costs that they induce.

**User equilibrium assignment**

A user equilibrium is assumed between all modelled traffic flows and transport costs (such that all costs on used routes are equal, and that link flows and the resulting link costs are mutually consistent). The relationships between path flow and path cost are shown in Figure 3.2.

$$\sum_a c^i a \delta^a = c^p \quad f^p_v = p^v \tau^v$$

$$c^i a (u_a (u_a)) = c^i_a \quad u_a = \sum_{a \in A, h \in Z} \delta^a f^p_v$$

**Figure 3.2:** Relationships for a Stochastic User Equilibrium (static, rigid demand)

The variables shown in Figure 3.2 are the O-D matrix $T^s r$ giving the travel demand between $r$ and $s$, the individual route choice probabilities $p^s_r$, the individual path flows $f^s_p$ that result from the choice process, the individual link flows $u_a$ that result from application of the link-path incidence matrix $A_{ap}$ to the vector $\{f^s_p\}_{p \in P}$ of path flows, the link travel time $\tau_a$ that results from the link flow $u_a$ and the corresponding (generalised) link cost $C^i_a$, the path cost $C^p$ that results from multiplication of the link-path incidence matrix $A_{ap}$ by the vector $C^i_a$, and finally the route choice probabilities that result from the path costs $C^p$.

In mathematical terms the solution of this set of interrelationships corresponds to the solution of a fixed-point problem, which can be formulated (see e.g. Cascetta (2001), Patriksson (1994)) for cost (in fact any of the variables in the loop may be selected as invariants) as in eqn. 3.1:

$$C^p = C^i a (\tau_a (A_{ap} p^s_r (C^p) T^s r)) A_{ap}$$

(3.1)

The relationships between flows and costs for the dynamic case are shown in 3.3. Note again that we use the time interval $h$ to indicate the time of path entry; general time intervals are denoted by $k$. 
The variables shown in Figure 3.3 are: the O-D matrix $T_{rs}^h$ giving the travel demand between $r$ and $s$, the route choice probabilities $p_{rs}^h$, the path flows $f_{ph}^s$ that result from the route choice process. The relationship between $h$ and the time interval $k$ at which a link is affected by the flow entering during time interval $h$ is determined by the path $p$ and the dynamic link-path incidence matrix $\Delta_{phka}^r$. This matrix, determines the periods $k$ during which a path flow $f_{ph}^s$ that set out during period $h$ on path $p$ will touch link $a$. It is determined by static link-path incidence matrix $\Delta_{ap}^r$ (which represents which links are on which path) and the link travel time matrix $\tau_{ak}$. The link travel times $\tau_{ak}$ in turn depend on the link inflows $u_{ak}$ that result from application of the path flows $f_{ph}^s$ and the dynamic link-path incidence matrix $\Delta_{phka}^r$. The link inflow is related to the path flows as $u_{ak} = \sum_{r,h,p} \delta_{phka} f_{ph}^s$. 

Given a consistent triple of path flows, link flows, and dynamic link-path-incidence matrix \( \{ f_{ph}^s, u_{ak}, \Delta_{phka}^r \} \), the path travel times $\tau_{ph}$ can be calculated from the link travel times $\tau_{ak}$ as $\tau_{ph} = \sum_{r,h,p} \delta_{phka} \tau_{ak}$, which contribute to the (generalised) path cost vector $C_{ph}$. Finally the route choice probabilities $P_{ph}^s$ depend on the path cost vector $C_{ph}$.

**Fixed point formulation of dynamic equilibrium assignment**

Although similar in idea to the static assignment, the dynamic case presents special complications. The variables are: the dynamic O-D matrix giving the travel demand between $r$ and $s$ starting during time period $h$: $T_{rs}^h$, the dynamic route choice probabilities for time period $h$: $p_{rs}^h$, the dynamic path flows $f_{ph}^s$ that result from the route choice process, the dynamic link flows $u_{ak}$ that result from application of the dynamic link-path assignment fractions $\delta_{phka}^s$ that indicate if traffic leaving origin $r$ during time period $h$ en route to destination $s$ via path $p$ will exit link $a$ during time period $k$. The link travel time $\tau_{ak}$ that results from link flow $u_{ak}$ can be calculated. At this point however a difficulty arises: the link outflow $u_{ak}$ depends on the assignment fractions $\delta_{phka}^s$ that in turn need the travel times to be calculated, which need the link outflows. Solving this fixed-point problem is known as Dynamic Network Loading (DNL), which
is discussed in more detail in chapter 6. Thus in the dynamic case we have two simultaneous fixed-point problems instead of one. Assuming that these two coupled fixed-point problems can be solved, the (generalised) path cost matrix \( C_{ph} \) can be calculated from the link travel times using the dynamic link-path incidence matrix \( \Delta_{apk} \), and finally the dynamic route choice probabilities \( P_{pk}^{r} \) that result from the path costs \( c_{pk} \).

The fixed-point problem corresponding to dynamic traffic assignment is shown in eqn. 3.2:

\[
\tau_{ak} = \tau_{ak} (u_{ak}) \quad (3.2)
\]

\[
= \tau_{ak} \left( \sum_{rshp} \delta_{phka}^{r} f_{ph}^{r} \right) \quad (3.3)
\]

\[
= \tau_{ak} \left( \sum_{rshp} [\delta_{phka}^{r} (\Delta_{apk}, \tau_{ak}) P_{pk}^{r} (C_{ph}, \Delta_{apk}, \tau_{ak}) T_{h}^{r}] \right) \quad (3.4)
\]

3.4 Estimation of dynamic O-D matrices from observations

In the estimation of O-D matrices a similar parallel exists: viewed on a sufficiently abstract level both problems share a common structure in that the demand is unobserved and has to be estimated from observable quantities (e.g. traffic flows) on basis of a model that maps the unobserved demand to observable quantities. In both cases the O-D matrix is estimated as a matrix that, when mapped onto the observed values of the observable quantities, replicates those observations. This can be interpreted as inverting this mapping. From the point of view of constructing a solution method however, the dynamic case presents complications that require a different approach.

Discussion of the method

The method sketched above leans heavily on the existence of models to produce a synthetic O-D matrix, and on the possibility of parameterising the O-D matrix through a model (called the matrix model) with respect to some parameter vector \( \Lambda \).

Often (see e.g. Cascetta (2001)), the matrix model is absent and the matrix is estimated directly: i.e. the parameter vector \( \Lambda \) simply consists of all matrix elements (an exception is presented in Gunn et al. (1997)). While this approach may be feasible in the static case, in the dynamic case we have many more matrix elements that must be estimated. As observed in the literature (see Cascetta et al. (1993), Van Der Zijpp (1996), Camus et al. (1997)), the structure of the O-D matrix generally does not change much from one time period to the other, so that we can expect benefit from a more parsimonious parametrisation of the matrix as this reduces the number of free parameters.

Examples are the case where a matrix is given as a set of tripends and a distribution model (in which case the \( \Lambda \) are the parameters of the distribution model), or a multiplicative model involving an apriori matrix and balancing factors (in which case the \( \Lambda \) are the balancing factors).
The essence of the method is the assumption that the latent O-D matrix gives rise to various derivative quantities, such as production, attraction, traffic counts, trip length distributions etc. which are observable, but whose observations may contain errors and must therefore be expected to contradict each other to some degree. This transforms the matrix estimation into a statistical estimation problem of the parameters Λ that produce a matrix with the best possible fit to the observations.

The advantage of this method is its flexibility: it can deal with overparametrised problems by seeking a vector Λ that produces a best fit in some way, with underparametrised problems (through the use of the apriori matrix), with noisy data of all descriptions, and it can be used to incorporate apriori knowledge on the structure of the matrix. In fact the use of apriori matrices makes the procedure a Bayesian estimator, with the synthetic models furnishing the apriori matrix.

**Model structure for dynamic O-D matrix estimation without departure time feedback**

The relationships between the objects in dynamic matrix estimation without departure time adjustment are shown in Figure 3.4.

![Diagram](attachment:image.png)

Figure 3.4: Relationships between objects in dynamic O-D matrix estimation (without departure time adjustment)

Shown top left is the output of a static synthetic model for the apriori O-D matrix $T^{\tau A,0}$. A departure time choice model fits this matrix with a time dimension, yielding an apriori dynamic O-D matrix $T^{\tau A,0}_h$. 
Shown at the top center is a parameter vector $\Lambda$ that parameterises an O-D matrix model (by choosing $\Lambda$ equal to the set of O-D flows, the unparametrised situation is reproduced; otherwise one might use a gravity model combined with K-factors, a distribution model based on a discrete-choice model and have $\Lambda$ represent the choice model coefficients). Shown at the top right is a parameter vector $\Phi$ of parameters for a route choice model.

Shown in the solid box in the centre are the matrix model $T^{rs}_h \left( T^{rs,0}_h, \Lambda \right)$ and the assignment model. The matrix model takes the apriori matrix and a parameter vector $\Lambda$ and returns an adjusted O-D matrix $\hat{T}^{rs}_h$. The adjusted O-D matrix $\hat{T}^{rs}_h$, the route costs $C_p(k)$, and the parameter vector $\Phi$ feed into the path choice model that outputs path choice probabilities $P^rs_p(C_p, \Phi)$. The application of the path choice probabilities on the adjusted matrix gives the dynamic path flows $f^rs_p(k)$ for every O-D pair $rs$ and every path $p$ between them. These path flows give rise to link flows $v_a(k)$ on each link $a$ of the network, which determine the link travel times $\tau_a(k)$, which in turn (co-)determine the dynamic path costs $C^rs_p(k)$. As noted in Section 3.3, dynamic user equilibrium is sought between the variables $C^rs_p(h)$, $P^rs_p(C_p, \Phi)$, $f^rs_p(h)$, $v_a(k)$, and $\tau_a(k)$. The dynamic user equilibrium assignment is described in more detail in chapter 6. The issue of consistency of link flows is discussed in Section 3.3 and chapter 6.

Dynamic observable quantities $S_{lk}(T^{rs}_h)$, dynamic (observable) link flows $u_{ak}$, and (possibly) link travel times $\tau_a(k)$ are calculated by the traffic model. In practice the internal time-steps at which the demand-supply model operates are usually smaller than those for which observations are available. Therefore we will allow observed flows to be related to a larger time interval $\kappa$ than the interval $k$ used to describe the demand-supply interaction and aggregate the calculated group observations to the time interval $\kappa$ for which external observations are available: $S_{lk}(T^{rs}_h) = \sum_{k \in \kappa} S_{lk}(T^{rs}_k)$, $u_{ak} = \sum_{k \in \kappa} u_{ak}$, $\tau_{ak} = \frac{1}{\kappa} \sum_{k \in \kappa} \tau_{ak}$ in order to make them comparable with the observed group totals $\bar{S}_{lk}$, $\bar{u}_{ak}$, $\bar{\tau}_{ak}$.

Model structure for dynamic O-D matrix estimation with departure time feedback

The relationships between the objects in dynamic matrix estimation with departure time choice are shown in Figure 3.5. The only difference with Figure 3.4 is the addition of a departure time adjustment mechanism (shown inside the shaded box).
Figure 3.5: Relationships between objects in dynamic O-D matrix estimation (with departure time adjustment)

Shown inside the shaded box is the departure time adjustment model. This model operates on the O-D matrix and shifts each departure time from starting interval $h$ to starting interval $h'$. The model calculates the probability $P_{h' | h}(C_p, \Theta)$ of the O-D relationship $rs$ to depart during time $h'$ given that the preferred departure time interval was $h$ under the influence of path travel costs $C_p(h)$ and parameter vector $\Theta$. We propose to use a departure time rescheduling model (as opposed to the departure time choice model used in the construction of the dynamic apriori matrix).

A similar model was proposed in Marchal and de Palma (2001), which however does not distinguish between departure time choice and departure time adjustment.

### 3.5 Dynamic O-D matrix estimation using a behavioural model structure

In this section we will show how the matrix estimation can be added to the model structure shown in Figure 3.5.

The matrix estimation is carried out by adjusting a parameter vector $\Lambda_{ik}$ to obtain the best possible fit between the observed flows $\bar{S}_{ik}, \bar{v}_{ak}$ and the observable flows $S_{ik}(T_h^s(\Lambda_{ik}))$, $v_{ak}(T_h^s(\Lambda_{ik}))$. For this we will need an objective function $z : \mathbb{R}^{|S_{od}|} \times$
\[ R^{[\text{obs}]} \rightarrow R \text{ in terms of link flows } v_{\text{kr}} \text{ and other observable quantities } S_{\text{lk}} \text{ that measures} \]

the fit of the observable quantities derived from the matrix estimate to the observed quantities. In this way the matrix estimation problem can be stated as:

\[ \hat{T}_{\text{min}}^s = z \left( S_{\text{lk}} (T_h^s (\Lambda_{lk})), \tilde{S}_{\text{lk}}, v_{\text{ak}} (T_h^s (\Lambda_{lk})), \tilde{v}_{\text{ak}} \right) \]  

(3.5)

Note that: \( \hat{T}_{\text{lk}}^s = T_h^s (\Lambda_{lk}) \).

The matrix estimation is carried out by adjusting the parameter vector \( \Lambda_{lk} \) to obtain the best possible fit between the observed flows \( \tilde{S}_{\text{lk}} \) and observable flows \( S_{\text{lk}} (T_h^s) \).

We will use the structure shown in Figure 3.6.

Figure 3.6: Integration of the dynamic O-D matrix estimation into the behavioural model structure

Shown is the same model structure as in Figure 3.5 with the addition (shown shaded) of the objective function \( z \) that measures the difference between the observable and observed quantities. By adjusting the parameter vector \( \Lambda_{lk} \) the matrix estimate \( T_h^s (\Lambda_{lk}) \) changes. Using this revised demand an equilibrium is sought in terms of route choice and departure time choice through the model structure presented in the previous section. When equilibrium is reached, a new set of internally consistent derived quantities results that enter the objective function together with the observed quantities.

This estimation problem constitutes a Mathematical Program with Equilibrium Constraints (MPEC), which can also be interpreted as a bi-level programming problem. The MPEC can be solved exactly as shown in Patriksson and Rockafeller (2002).
MPEC and the formulation of the matrix estimation as an optimisation problem including the parametrisation of the matrix through a parameter vector will be discussed in more detail in Section 5.

### 3.6 Rigid and elastic demand

The dynamic O-D matrix whose estimation is the subject of this thesis can be seen as a case between rigid and elastic demand, with consequences for the formulation of the assignment algorithms and the combined route choice departure time choice problem. In order to clarify this issue we list various examples of constraints in table 3.3.

The dynamic O-D matrix is denoted as $T_{h}^{rs}$, and we use a shorthand notation to represent the constraints. When constraints are imposed on partial sums of $T_{h}^{rs}$, these are denoted as $T_{h}^{rs} = \sum_{h} T_{h}^{rs}$. When constraints on the partial sums of $T_{h}^{rs}$ are absent, this is denoted as $T_{h}^{rs}$ meaning that $\sum_{h} T_{h}^{rs}$ is modelled, and depends on the network conditions.

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Designation</th>
<th>Constrained</th>
<th>Modelled</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rigid dynamic</td>
<td>$T_{h}^{rs}$</td>
<td>route choice</td>
</tr>
<tr>
<td>2</td>
<td>Time elastic, rigid</td>
<td>$T_{r}^{rs}$, $T_{h}^{rs}$, $T_{r}^{rs}$</td>
<td>route choice, $T_{h}^{rs}$</td>
</tr>
<tr>
<td>3</td>
<td>Time elastic, demand</td>
<td>$T_{r}^{rs}$, $T_{r}^{rs}$</td>
<td>route, departure time, total traffic volume</td>
</tr>
<tr>
<td>4</td>
<td>Space-time elastic</td>
<td>$T_{+}^{rs}$</td>
<td>route, $T_{r}^{rs}$, $T_{h}^{rs}$</td>
</tr>
<tr>
<td>5</td>
<td>Completely elastic</td>
<td>$T_{h}^{rs}$</td>
<td>everything</td>
</tr>
</tbody>
</table>

Table 3.3: Distinctions between rigid and completely elastic demand

In case Nr. 1, all matrix elements $T_{h}^{rs}$ are fixed, and only route choice is modelled (by the DTA procedure); this is called rigid demand. This is the sort of constraint that one would impose for short-term predictions. The reason is that users are slow to change their destination choice (which would affect partial sums such as $T_{r}^{rs}$) because it is related to constraints that are expensive and therefore slow to change (jobs, house etc.). Also mode-choice is generally found to be less sensitive to changes than route choice (e.g. involving vehicle ownership, car competition within households, season tickets etc.). For very short-term forecasts one would therefore want to keep such things fixed, and only allow route choice.

In case Nr. 2, the total amount of traffic $T_{r}^{rs}$ per O-D pair during the entire period is fixed (for the same reasons as in the first case) but departure time distribution is modelled along with route choice in a combined departure time route choice assignment. This would correspond to the situation where people had enough time to reconsider their departure time and their route choice, but would still be subject to high inertia in changing e.g. their mode or destination.

Case Nr. 3 would correspond to the case where the amount and departure time of traffic $T_{r}^{rs}$ leaving zones is kept fixed (e.g. for a residential area where the number of workers is known), but the choice of destination may vary. This would be the case for forecasts in which people can be expected to switch jobs, but not to move house, or
the case where new shopping centres are constructed but people are not expected to change house.

In case 4, only the total amount of traffic in the matrix is kept fixed; both the structure of the O-D matrix and the departure times are modelled. This would correspond to the case where a working population is expected to change house, jobs, departure time, but not its mobility.

In case Nr. 5, every aspect of the demand is modelled; this would e.g. be the case when a complete transportation forecasting model is run for long-term forecasts of traffic conditions, so that people’s mobility (depending e.g. on demographic characteristics), their choice of origin (i.e. their house), destination (i.e. primarily their jobs, but also shopping, social or recreational destinations), their travel mode and everything else may be expected to be subject to change and should therefore be modelled.

We will use case 2, where the total amount of traffic in the matrix is independent of the level of service, and both route choice and departure time adjustment are modelled.

### 3.7 Grouped observations

#### 3.7.1 The static case

As noted in Gunn (1977), every static traffic count and every observation of a partial-matrix, zonal production or attraction, etc. can be interpreted as the observation of a group $S_l$ of weighted linear combination of O-D flows with weights $\phi_i^{rs}$.

$$S_l = \sum_{rs \in l} \phi_i^{rs} T_h$$

In this way any observation of the traffic system that is directly related to the O-D matrix as in eqn. 3.6 (such as traffic counts, screenline counts, production, attraction, trip-length distribution) can be treated in a uniform way. The central question in static O-D matrix estimation is how to relate observations to matrix cells, which amounts to determining the cell weights $\phi_i^{rs}$. For traffic counts, these are given by the assignment fractions.

#### 3.7.2 Extension of grouped observations to the dynamic case

The idea of grouped observations extends to the dynamic case, and can also be extended to include multiple userclasses.

**Definition 3** The observable flows $S_{lk}$ are calculated as in eqn. 3.7 as a weighted sum of O-D matrix cells.

$$S_{lk} = \sum_{rs \in l} \sum_{h \in I^*_h} \phi_i^{rs} T_h$$
In this definition we have the dynamic O-D matrix $T_h^{rs}$, with the index $h$ relating to the departure time period, and the dynamic group assignment fraction $\phi_{ikh}^{rs} \in [0, 1]$ denoting what fraction of O-D flows $(r, s)$ with departure period $h$ contributes to the dynamic observation group $l$ during observation time period $k$. The summation index for O-D pairs $(r, s)$ runs over all O-D pairs in $I_i$, the index set of observation group $l$. To emphasise the time-dependency, the index for time periods $h$ runs over all time periods in $I_{tgh}$, the index set of time periods for which the O-D cells contribute to the group observation $S_{ik}$ at time $t$.

The group assignment fraction $\phi_{ikh}^{rs}$ is related to the dynamic assignment fraction $\varphi_{apkh}^{rs}$ (see Section 5.2.1) that gives the fraction of O-D flow $(r, s)$ with departure period $h$ that enters link $a$ on path $p$ during observation time period $k$, as will be detailed below. Note that an extension to multiple userclasses would not affect the structure of the definition.

The extension to dynamic group observations has an unexpected effect: the time period to which the observation relates (i.e. interval $k$) is not the same as the time period (interval $h$) of the O-D-t matrix cells that contribute to the observation. This arises from the fact that the time interval to which the matrix cells relate is the time interval during which these matrix cells enter the network; the observation relates to O-D-t flows that are already on the network, and may have entered several time periods earlier. For any $\phi_{ikh}^{rs} > 0$, the difference $\tau = k - h$ is the travel time from the origin to the location where the matrix cell $T_h^{rs}$ contributes to observation group $l$.

**Definition 4** We will define a dynamic group observation $S_{ik}$ of group $l$ relating to observation time period $\kappa$:

$$S_{ik} = \sum_{h \in k} S_{ih}$$  \hspace{1cm} (3.8)

We will distinguish between internal and external time intervals. The distinction between the observation period $\kappa$ and the smaller time periods $k$ is necessary because the statistical estimation problem to which the O-D-t matrix estimation will be reduced associates one free parameter with each independent observation $S_{ik}$ whilst the model must work internally with a finer time discretisation in time intervals $k$ which is related to the smallest free-flow link travel time in the network. This is illustrated in Figure 3.7.
3.7.3 Definitions and notation

We will refer to observations that reflect on groups of O-D-t matrix cells as datagroups, distinguished by their datagroup number \( l \in L \), and refer to the set of datagroups as \( L \). The general form of the description of the datagroups as they enter the calculation of the objective function is given in (3.9):

\[
S = \{ S_l \}_{l \in L} = \left\{ l, \left\{ \left( k, \bar{S}_{lk}, h \right), \{(r, s, h), \phi_{lk}^{ra}\}_{(r,s) \in I_{lk}J_{lk}H} \right\}_{k \in K} \right\}_{l \in L} \tag{3.9}
\]

with:

- \( l = \) group number, member of group list \( L \)
- \( \kappa = \) observation time interval, \( \kappa \in \mathcal{K} \)
- \( \bar{S}_{lk} = \) observed group total for group \( l \), and time period \( k \)
- \( f_l = \) sampling fraction (1 for counts, <1 for survey-based observations)
- \( (r, s, h) = \) ordered triples of origin, destination and departure time interval in time-dependent index set \( I_{lk}J_{lk}H \)
- \( i = \) origin, member of a set \( I_{lk} \) of origins for this combination of \( k \) and \( l \). \( I_{lk} \subseteq I \), the set of all origins
- \( j = \) destination, member of \( s \) set \( J_{lk} \) of destinations for this combination of \( k \) and \( l \). \( J_{lk} \subseteq J \), the set of all destinations
- \( \phi_{lk}^{ra} = \) fraction of matrix cell \( T_{h}^{ra} \) that contributes to group observation \( l \) during time period \( k \)

All datagroups are assumed to be dynamic in that they provide a complete list of observed group totals and sampling fractions \( \left( \bar{S}_{lk}, f_l \right)_{l \in L}, \forall \kappa \in \mathcal{K} \). The datagroup totals relate to the value of the O-D flows as in (3.10).
\[ S_{lk} = \sum_{(r,s,y) \in I_{lk}} \phi_{lkh}^s T_{h}^r \]  \hspace{1cm} (3.10)

Note that in general we scale the estimated observed flow \[ \sum_{(r,s,y) \in I_{lk}} \phi_{lkh}^s T_{h}^r \] down to the level of the observation \( \bar{S}_{lk} \) by multiplying with the sampling fraction \( l < 1 \), instead of grossing up the observed flow to the level of the estimated flow. This is done for statistical reasons.

In the estimation, the objective is to match each observed datagroup total \( \bar{S}_{lk} \) with synthetic totals \( S_{hk} \).

The relationship between the observation group \( S_{lk} \) and the O-D matrix flows is illustrated in Figure 3.8 for the case of traffic counts, of total production and attraction, and screenline counts.

\[ S_{lk} = \sum_{(l,h,k,p) \in I_{lk}} T_{h}^r \phi_{lkh}^r \]

\( S_{2,k} \): attraction

\( S_{3,k} \): production

\( S_{4,k} \): screenline count

Figure 3.8: Relationship between \( S_{kl} \) and the O-D matrix flows

Note that for any \( \phi_{lkh}^r > 0 \), the difference \( \tau = k - h \) is the travel time from the origin to the location where the matrix cell \( T_{h}^r \) contributes to observation group \( l \). The datagroup-related assignment fractions are related to the assignment fractions \( \varphi_{ophk}^r \) through:

\[ \varphi_{lkh}^r = \sum_{p \in P_{rs}} \varphi_{ophk}^r \]  \hspace{1cm} (3.11)

We can then distinguish between two types of datagroups: fixed datagroups, and assignment-dependent datagroups depending on the amount of information relating the matrix cells to the group totals.
3.7.4 Fixed datagroups

For fixed datagroups, the index sets $I_{lk}, J_{lk}Y_{kl}$ are given and $\phi_{lkh}^s = 1$. Examples are trip production, screenline counts, cordon counts, and observed partial matrices on cordon.

Trip production

For data on dynamic trip production, we assume that for each $l$ and $k$ the data contains the trip production and the sampling fractions $f_l$:

$$\left( \tilde{S}_{lk}, f_l \right)_{(r,s,y) \in I_{lk}J_{lk}Y_{kl}}$$

In this case, $I_{lk} = \{r_l\}$, and $J_{lk} = \{s \in J\}$ indicate the set of all origins and destinations respectively. Production data are often obtained from household surveys, and therefore have an associated sampling fraction $f_l$ given by the survey documentation. The group totals relate to the O-D matrix as: $\tilde{S}_{lk} = f_lT_{lk}^* = f_l \sum_{s \in J} T_{lk}^s$. Note that the approach is to scale the synthetic matrix totals down by the sampling fraction instead of grossing up the observations. This reflects the true statistical weight of the observations.

Observed partial matrix for roadside survey on a cordon

For partial matrices observed on cordon, we assume that for each $l$ and $k$ the data contains the observed partial matrix, so that the sampling fraction $f_l$ is given, with $I_{lk} = \{r_n, n \in N_{kl}\}, J_{lk} = \{s_n, n \in N_{kl}\}$, and the group total $\tilde{S}_{lk} = N_{kl}$.

From a statistical point of view the set of observations of individual O-D flows is a sample of the set of all O-D movements during time period $k$ crossing the cordon. Ideally the probability of appearing in the sample is proportional to the amount of traffic in the O-D cell, and can therefore be used to estimate the structure and the level of the (partial) O-D matrix. In practice however we may have a finite capacity for intercepting vehicles (either by stopping them or by noting their license plate), which results in a sampling rate that varies with the total traffic intensity across the cordon. In either case the statistics of the sampling determine the reliability of this impression of the O-D matrix. If one lacks confidence in the reliability of this snapshot, the group total $\tilde{S}_{lk}$ can be used in conjunction with the observed index sets $I_{lk}$, and $J_{lk}$ (or even with expected index sets) as an ordinary group observation.

3.7.5 Assignment-dependent datagroups

For assignment-dependent datagroups both the index sets $IJ_{kl}$ and the assignment fractions $\phi_{lkh}^s$ are not given in advance, but must be determined from the DTA.

Trip attraction

In contrast to the static case, in the dynamic case trip attraction observations are not simply equivalent to production observations. The reason is that attraction observations at time $k$ relate to O-D-t matrix cells at time $k - \tau (h)$, with $\tau (h)$ the experienced
travel time for the contributing O-D pair departing in time interval \( h \). Note that in the special case that \( \tau (h) \) is observed directly (as assumed e.g. in Van Der Zijpp (1996)) and no route choice is possible, this group regains its independency from the assignment. For data on dynamic trip production, we assume that for each \( l \) and \( k \) the data contains the production and the sampling fraction: \( \left( \bar{S}_{lky}, f_{lky} \right)_{(r,s,h) \in I_kJ_kY_{kl}} \).

In this case, \( I_k = \{ r \in I \} \) indicates the set of all origins and \( J_k = \{ s_k \} \) the set of destinations. Attraction data are often obtained from household surveys, having an associated sampling fraction \( f_j \) which is given by the survey documentation. The group totals relate to the O-D matrix as: \( \bar{S}_{lky} = f_{lky} T_{lky} = f_{lky} \sum_{r \in l} T_{lry} \).

**Undifferentiated counts**

For traffic counts (undifferentiated by vehicle type) our observations are simply \( \left\{ \bar{S}_{k} \right\}_{k \in K} \). Any information regarding the cells that the group total relates to: \( \{(r, s, h), \phi_{rsk}^l \}_{(r,s,h) \in I_kJ_kH_{kl}} \) must be derived from the DTA.

**Observed trip length distribution**

Trip length distributions (TLD's) are often derived from household surveys. They describe the distribution of the number of trips across cost bands (e.g. travel time, travel distance, monetary cost) occurring in reality. Whilst in principle TLD's can be differentiated by geographic region, one often sees a single TLD being used for a complete study area. This is the case which we will address here; when TLD's are differentiated by geographic region, one simply gets more TLD-type datasets with corresponding index sets.

The data given in a TLD \( l \) subdivided by time interval \( k \), consists of a sampling fraction \( f_l \) for the survey, and a set of cost bands \( d_{kl} \) each with a number of trips in them: \( \left\{ d_{kl}, f_l, \bar{S}_{lk} \right\}_{k \in K} \).

The information regarding which matrix cells correspond to which cost band must be derived from the DTA because e.g. route choice will affect the average trip length and cost, and congestion will affect the average travel time. This leads to a collection of \( m \) index sets: \( \left\{ d_{m}, (r, s, k) \left( (r,s,k) \in I_{km}J_{km}Y_{km} \right) \right\}_{km} \), one index set \( I_{km}J_{km}Y_{km} \) per cost band \( m \).

**3.8 Unification of mode-choice and departure time choice**

The literature on static O-D matrix estimation contains instances where the O-D matrix estimation is formulated as a bi-level problem, as noted in chapter 5. The advantage is that it separates the ("first-level", or "outer") problem of O-D matrix estimation from that of determining the assignment map (the "second-level" or "inner") problem. In doing this it provides an avenue for generalisation of the assignment procedure. In Luo
et al. (1997), we find a particular type of bi-level problem called an Mathematical Program with Equilibrium Constraints (MPEC, to be explained in chapter 5). In an MPEC, the second-level problem is given by a variational inequality (VI) instead of an optimisation problem, which allows us to apply it to dynamic assignment (see also chapter 6). This will enable us to formulate the dynamic O-D matrix estimation as an MPEC.

However, in the case of Dynamic O-D matrix estimation we have departure time choice in addition to route choice. If we can interpret route-choice and departure-time choice as a simultaneous route choice on some extended network, then the whole framework of equilibrium assignments applies including its existence and uniqueness results. It turns out that the literature contains the concept of space-time extended networks with which one can unify route choice and departure time choice as route choice as route choice on an extended network.

In sections 3.8.1-3.8.2 we will introduce the concept of Space Time Extended Networks (STEN) and use these as a vehicle with which to express departure time choice as a sort of route choice on a STEN, as was done in Lindveld and Van Der Zijpp (2000) and Van Der Zijpp and Lindveld (2000). We note that although this attempt to express the dynamic assignment is possible it has the drawback that the STEN becomes one of the unknowns, which is no improvement on the formulation of dynamic assignment as a VIP (see Section 6). Given that one can view the STEN as the outcome of a dynamic assignment procedure, we will include the STEN-based approach as a way of visualising departure time choice by interpreting it as route choice on a STEN and to illustrate the difference between departure time choice and departure time adjustment, but will not use it to carry out any actual calculations.

3.8.1 STENs

A Space-Time Extended Network (STEN) (see Bell and Iida (1997)) explicitly represents time by having a complete layer of all nodes of the physical network per time period. The travel time on STEN links corresponds to the time difference between the time periods of the start node and end node of the link. An example of a STEN is given in Figure 3.9. Shown on top is the original (two-link) network, with origin r and destination s, and one intermediate node.

The STEN is built by replicating the nodes as many times as there are time periods, and adding STEN links (shown as sloping solid lines). The links that correspond to physical network are shown as dotted lines.
Figure 3.9: Space Time Expanded Network (STEN) corresponding to a physical network. (x) denotes time period x

Traffic entering a link’s a-node arrives at its b-node at time \( t + \tau \), with \( t \) the time of entry, and \( \tau \) the link travel time. STEN links connect nodes that belong to different time periods (shown as sloping links in Figure 3.9). We will assume that the time discretisation is chosen such that \( t \) and \( t + \tau \) are never in the same time period, i.e. that there are no horizontal links in Figure 3.9.

To link a dynamic O-D-t matrix \( T_{ht}^s \) with the STEN, we can add a virtual destination \( s \), and virtual links connecting it to the STEN, as shown in Figure 3.10.

Figure 3.10: STEN with a virtual destination

Note the absence of links between the physical origin \( r \) and the time-dependent origins \( r(1), \ldots, r(4) \). This highlights the fact that the matrix \( T_{ht}^s \) specifies the departure times.
3.8.2 Modelling departure time choice as route choice in an augmented STEN

In Figure 3.10 it is assumed that the dynamic O-D-t matrix specifies a fixed travel demand per period that is not influenced by the level of service in the network. We will distinguish between preferred departure time (defined as the departure time that would have been chosen under free-flow conditions) and actual departure time that results from departure time rescheduling.

Departure time rescheduling was cast into a random utility maximisation framework by Small (1987). In this framework travellers are assumed to have a preferred arrival time at their destination; they incur disutility when travelling, when arriving early, and when arriving late. Their choice set consists of a number of possible departure times with an associated disutility according to travel time and schedule delay (early or late arrival). Travellers are assumed to choose the alternative with the least expected disutility, thereby trading off travel time against schedule delay costs.

It turns out that departure time choice can be modelled as route choice through an appropriately defined STEN. Travel time that varies with departure time can be modelled using a STEN such as shown. The problem is how to associate schedule delay costs with links.

This can be achieved by extending the STEN with artificial nodes that represent the 'desired' departure time, and connecting these nodes with the STEN using links that represent the schedule delay costs, see Figure 3.11.

![Diagram](image)

Figure 3.11: Augmented STEN with schedule delay costs that depend on the preferred departure time
In Figure 3.11 an extra set of virtual origin nodes ($r_1(1)$, $r_1(2)$, $r_1(3)$, $r_1(4)$) has been added to the network. These nodes correspond to preferred departure time of travellers from origin $r$. Given their preferred departure time, travellers will reschedule their departure to an actual departure time under the influence of e.g. congestion. The links from $r$ to $r_1(k)$ denote the choice of a preferred departure time (based e.g. on an activity scheduling model); these links are drawn as dotted lines to emphasize that they are exogenous to the matrix estimation. The links from $r(t)$ and $r(t)$ denote a choice of an actual departure given the preferred departure time, and carry a disutility that corresponds to the departure time shift. The arrival times are the time slices at which the chosen routes enter one of the destination nodes $s_t(1)...s_t(4)$. A schedule delay proportional to the difference between the actual arrival time (i.e. departure time plus travel time) and the preferred arrival time is associated with the links from the nodes $s_t(1),...,s_t(k)$ to the virtual destination $s$. In this way both the rescheduling penalty and a time-dependent schedule delay cost can be associated with the links of the STEN modelled for all traffic that enters the network from origin $r$, according to its preferred and actual departure time. As a result, the dynamic user optimum departure time and route choice (DUO-D&R) problem is reformulated as a dynamic user optimum (DUO) assignment problem with respect to the augmented STEN. It is assumed that appropriate utility functions are known for departure time choice, and for departure time rescheduling.

3.9 Summary

In this chapter we have proposed a behavioural approach towards dynamic O-D matrix estimation. In this approach we have extended the existing bilevel approach for static O-D matrix estimation to the dynamic case, thereby retaining the split of the O-D matrix estimation problem into an upper-level problem (the O-D matrix estimation proper) and a lower level problem (dynamic equilibrium traffic assignment (DTA)). As a result the two problems can be studied separately.

The upper level problem attempts to minimise a criterion function of the O-D matrix to be estimated, conditional on constraints imposed by the lower level problem (the DTA). The lower level problem will be studied in chapter 6.

It turns out that in the step from static to dynamic O-D matrix estimation, departure time choice emerges as a natural complement of the lower-level problem next to route-choice. In the course of formulating the combined departure-time and route-choice problem as an assignment on a STEN, a natural distinction between departure time choice and departure time adjustment emerged. Although the STEN-based approach to the combined departure-time and route choice equilibrium assignment does not provide a good avenue for calculations, it proved to be of conceptual value.

In addition an existing formulation of generalised costs was generalised to the dynamic case, which suggested an obvious approach towards parameterising the O-D matrix, in the sense that one free parameter is associated with each available generalised cost. This leads to a statistical estimation problem that is no longer overparametrised.
In chapter 4 we will examine the statistical properties of the O-D matrix estimation problem, and the parametrisation of the O-D matrix to obtain an objective function in terms of the O-D matrix parametrisation.
Chapter 4

Dynamic O-D matrix estimation as a statistical problem

4.1 Introduction

In chapter 3 we presented a behavioural framework for dynamic O-D matrix estimation, in which the O-D matrix estimation problem was cast as a bi-level optimisation problem. In this chapter we will study this problem with the aim of deriving an objective function. We will argue that the dynamic O-D matrix estimation should be approached as a statistical problem rather than a fitting problem for two reasons: first of all the dynamic O-D matrix itself is a stochastic variable, and secondly all our observations contain noise, part of which consists of measurement errors and part of which is related to the day-to-day variability of the traffic system. This aspect has already been assimilated to some extent into the O-D matrix estimation framework by formulating it as a "best-fit", but the difference is that in a statistical estimation problem proper account is taken of the error structures and the probability distributions of the observed data whereas in a fitting problem no account is taken of such considerations.

In this chapter we will apply the method proposed in Hazelton (2000) to the framework presented in chapter 3 to derive the objective function needed in chapter 5. In doing this we extend the work of Hazelton (2000) to dynamic O-D matrix estimation and apply it to congested networks by modelling the route choice probabilities separately. It turns out that the chosen objective function forms an approximation of a rigorous (but thoroughly intractable) statistical approach to the O-D matrix estimation problem.

4.2 Literature review

Due to the strong similarity in structure between static and dynamic O-D matrix estimation, we will refer to the literature on static O-D matrix estimation for information on the statistical structure of O-D matrix estimation. The issue of static O-D matrix estimation as a statistical problem has been addressed by Cascetta and Nguyen (1988), Cascetta (1984), Van Zuylen and Willumsen (1980), Van Zuylen (1983), Bell and Hida
(1997) who use a generalised least squares approach, and Maher (1983) who assumes normally distributed O-D flows and uses a Bayesian approach towards the estimation.

A rigorous statistical treatment of static O-D matrix estimation is presented in Hazelton (2000), who notes that the use of GLS estimators has firm statistical foundations but leads to sub-optimal estimators under the natural assumption of Poisson distributed O-D demands. The work of Hazelton (2000) shows how the Poisson distribution leads to unwieldy distributions for link counts (in our context group counts), which however can be approximated by a Normal distribution with the same first and second moments as the counts distribution. This results in Multi Variate Normal (MVN) estimators that are difficult to use because of the relationship between the first and second moments. This difficulty can be resolved using the so-called Expectation Maximisation (EM) algorithm, and circumvented by coupling the first and second moments, while calculating the variance-covariance matrix from the observations. This last stratagem results in a GLS type estimator formulated in terms of known quantities.

The issue of the statistical aspects of dynamic O-D matrix estimation have been addressed in Cascetta et al. (1993), Van Der Zijpp (1996), and Ashok (1996).

4.3 Statistical structure of the problem

In this section we will first consider the intrinsic stochasticity, followed by a specification of what stochasticity will be represented in the model.

4.3.1 Intrinsic stochasticity

Our observational data will contain measurement errors and genuine variation of demand. We can recognise several sources of measurement error in our observational data: sampling errors, stochastic counting errors, systematic counting errors. We note that in static O-D matrix estimation a substantial amount of within-day variance in the O-D matrix (and in the traffic flows) can only be accounted for as "noise". This is especially true if one estimates a static 24-hour matrix, and somewhat less if one estimates peak-period matrices. We also note that some data sources will contain variation due to the time-frame in which they were collected (e.g. trip length distributions taken from household surveys may refer to several months or even several years).

We will use the modelling framework set out in chapter 3 to link the quantities that play a role in dynamic O-D matrix estimation and identify the sources of random variation. We will assume that our framework is correct in the sense that all the elements present really exist and that their interrelationships in reality are the same as in our framework. We will focus on the grouped observations, as traffic counts are a special case of grouped observation. This leads us to the relationships shown in Figure 4.1.

Notation

In this chapter, quantities of a stochastic nature are underlined; if "x" is stochastic it is denoted as \( \underline{x} \), and its expectation is denoted with the same symbol without underlining: \( E(\underline{x}) = x \).
\[
\begin{align*}
\begin{array}{c}
\mathcal{P}_{p_{hi\text{rs}}} \quad P_r^p(h) \
\downarrow \quad \downarrow \\
T^r_{hi} \quad T^r_h \quad f^r_p(h) \\
\delta^r_{ph} \quad \delta^r_{ph} \\
\sum_{k \in K} \tilde{S}_l^i(k) \quad \xi_l
\end{array}
\end{align*}
\]

Figure 4.1: Hypothesised relationship between stochastic O-D matrix and stochastic group observations

The static O-D matrix \( T^r_{hi} \) has a stochastic nature due to its day-to-day variations (corresponding to the day-to-day dynamics shown in Figure 2.6 ), has expectation \( E(T^r_{hi}) = T^r_{hi} \), and is subject to departure choice symbolised by the departure time choice probabilities \( \mathcal{P}_{p_{hi\text{rs}}} \), resulting in a dynamic O-D matrix \( T^r_h \). The stochastic elements are the variability in the total daily demand incorporated by \( T^r_{hi} \) and the variability in the departure time choice.

The travel demand represented by the matrix \( T^r_h \) takes to the network and selects a route \( p \), symbolised by the path choice fractions \( \mathcal{P}_p^r(h) \) with \( E(\mathcal{P}_p^r(h)) = P_p^r(h) \) to give stochastic route flows \( f^r_p(h) = P_p^r(h)T^r_h \). The route flows traverse their paths, and contribute to individual group observations \( \delta^r_{ph} \) depending on the travel time on their path; this is represented symbolically by the assignment fractions \( \delta^r_{ph} \). The travel time and hence the time interval \( k \) during which the travel demand that entered the network during time period \( h \) (represented by the path flows \( f^r_p(h) \)) contributes to group \( l \) is again stochastic because of the well-known variability of the daily traffic situation. The result is a set of stochastic group flows \( \delta^r_{ph} \), which are not in themselves observable. Finally the observation process aggregates the \( \delta^r_{ph} \) to observation time interval \( \kappa \) and induces a certain measurement error, say \( \xi_l \), resulting in stochastic group observations \( \tilde{S}_{lk} = \sum_{k \in \kappa} \delta^r_{ph} + \xi_l \).

### 4.3.2 Predicting the expectation \( S_{lk} \) of stochastic group observation \( \bar{S}_{lk} \)

In the O-D matrix estimation we will model the relationships of Section 4.3, as shown in Figure 4.2. This model links the quantities that play a role in dynamic O-D matrix estimation and will be used to apply the results presented in Hazleton (2000) to our context of grouped observations.

The dynamic traffic assignment model adopted works with deterministic quantities instead of stochastic ones. We assume that these quantities represent the expectation of
the stochastic quantities. Since all relationships (except those involving the calculation
of link travel time) are linear, we will obtain the expectation of the observed flows when
we apply the model to the expectation of the input quantities involved (assuming the
model is correct). Regarding the crucial (and non-linear) link-performance functions,
we assume that they are set up and calibrated so that they give unbiased estimates of
travel time when applied to the expectation of traffic densities and flows.

\[
\begin{align*}
T^{rs} & \rightarrow P_{h_{rs}} T^{rs} = T^{rs}_h \rightarrow F^{rs}_p(h) = P^{rs}_p(h) T^{rs}_h \\
\delta^{rs}_{lpk} & \rightarrow S_i(k) = \sum_{r_{sp}h} f^{rs}_p(h) \delta^{rs}_{lpk} \\
S_{ik} & = \sum_{kex} S_i(k) \\
\bar{S}_{ik}
\end{align*}
\]

Figure 4.2: Modelled stochastic relationships

The expectation (or the average) \( T^{rs} \) of the stochastic O-D matrix \( T^{rs} \) is multiplied by
the expected departure time fractions \( P_{h_{rs}} \) (assumed fixed and given) resulting in the
expectation \( T^{rs}_h \) of the stochastic dynamic O-D matrix \( T^{rs}_h \). This matrix is multiplied by
the path choice fractions \( P^{rs}_p(h) \) (assumed fixed) to give the expectation \( F^{rs}_p(h) \) of
the stochastic route flows \( F^{rs}_p(h) \). Multiplied by the assignment fractions related to the
group observations, these flows form the expectation \( S_{ik} \) of the stochastic group flows
\( S_{ik} \). The group flows are not in themselves observable: observable are the perturbed
group flows \( \bar{S}_{ik} \) that result through the addition of an observation error \( \varepsilon_l \).

4.3.3 Stochasticity accounted for in the estimation

In the matrix estimation the stochasticity of the system is partially modelled as only
the first two moments of the stochastic group counts are taken into account.

The expectation of the counts is calculated by the dynamic traffic assignment model,
as described in Section 4.3.2.

An estimate of the covariance matrix of the group counts is obtained from the distri-
butional assumptions on the (stochastic) demand matrix \( T^{rs}_x \) and the (expectation
of the) matrix of route choice probabilities \( P^{rs}_x \) and the dynamic link path incidence
matrix \( \delta^{rs}_{lpk} \), as shown in Figure 4.3. The calculation of the covariance matrix of \( S_{ik} \) is
specified in Section 4.3.5.
4.3.4 Distributional assumptions

We will assume that the matrix cells are independently distributed Poisson variables. One of the strong points of the GLS estimator proposed in Cascetta (1984) is that it does not depend on distributional assumptions. However such estimators are sub-optimal when we do know the distribution. We will make distributional assumptions, and find that we can arrive at a GLS-estimator with a known variance-covariance matrix.

In the literature on static O-D matrix estimation (e.g. Hazelton (2000), Gunn et al. (1997)) we generally find the assumption that the cells of the static O-D matrix $T^{rs}$ have independent Poisson distributions with (fixed) mean $T^{rs}$. This means that the daily amount of traffic in matrix cell $r,s$ is given by a Poisson process with intensity $T^{rs}$, of which each day is a realisation. The implication is that seasonal trends are assumed to be either absent or accounted for and modelled. As a static O-D matrix is always related to a specific time period, the additional implication is that during this time period we have a constant rate of travel demand $T^{rs}$. This is illustrated in Figure 4.4. Shown are the (stationary) mean of the poisson process for O-D flows $T^{rs}$ (dotted line), and the daily average mean $T^{rs,d}$ of the matrix cell (dashed line).

We will extend the analysis of Hazelton (2000) to the matrix cells of each individual
time period $h$, hypothesising that the period-specific traffic flow levels can be modelled as a draw from an independent Poisson statistic $T_{h}^{r,s}$ with fixed intensity $T_{h}^{r,s}$. This means that we assume that the same daily O-D pattern is repeated over and over as illustrated in Figure 4.5 (dotted line).

![Figure 4.5: Extending the assumption of stationarity of $T_{h}^{r,s,d_{i}}$](image)

Shown in Figure 4.5 is the stationary average O-D pattern $T_{h}^{r,s}$ (dotted line) and the day-to-day evolution of travel demand $T_{h}^{r,s,d_{i}}$ (dashed line).

Our assumption differs from the ones used in e.g. Van Der Zijpp (1996) and Ashok and Ben-Akiva (2000) (which place strong emphasis on the correlation between time periods) in that we assume independence between the O-D flows $T_{h}^{r,s}$ and $T_{h+1}^{r,s}$ in successive time periods $h$ and $h+1$ (given the means $T_{h}^{r,s}$ of $T_{h}^{r,s,d_{i}}$). However we also assume that the mean O-D flows $T_{h}^{r,s}$ follow a pattern, that resembles the pattern that is present in the apriori matrix $T_{h}^{r,s,0}$.

We assumed that the O-D flows $I_{h}^{r,s}$ have independent Poisson distributions with mean $t_{h}^{r,s}$. If we further assume that the conditional expectation of the route flows $E\left(I_{p}^{r,s}(h) | I_{h}^{r,s}\right)$ has a binomial distribution with parameters: $n = I_{p}^{r,s}(h) t_{h}^{r,s}$ and $p = p_{p}^{r,s}(h)$, then the route flows $I_{p}^{r,s}(h)$ are also independently Poisson distributed with mean

$$E_I^{r,s}(h) = I_{p}^{r,s}(h) = p_{p}^{r,s}(h) t_{h}^{r,s}$$

and covariance

$$\sigma^{F,1}_{ij} = \text{cov} \left( I_{p_{i}}^{r,s}(h), I_{p_{j}}^{r,s}(h) \right) = \begin{cases} p_{p}^{r,s}(h) t_{h}^{r,s} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

where $\sigma^{F}_{ij}$ denoted the covariance matrix of the flows.

If we assume that the route-flow is distributed following a multinomial distribution, then the route flows $E_{p}^{r,s}(h)$ are no longer independent but are distributed with mean $E_{p}^{r,s}(h) = I_{p}^{r,s}(h) = p_{p}^{r,s}(h) t_{h}^{r,s}$ and covariance

$$\sigma^{F,2}_{ij} = \text{cov} \left( E_{p_{i}}^{r,s}(h), E_{p_{j}}^{r,s}(h) \right) = \begin{cases} p_{p}^{r,s}(h) \left(1 - p_{p}^{r,s}(h)\right) t_{h}^{r,s} & \text{if } i = j \\ -p_{p}^{r,s}(h) p_{p}^{r,s}(h) t_{h}^{r,s} & \text{if } i \neq j \end{cases}$$

The covariance matrix $\sigma^{F}_{ij}$ will be used later on to calculate the covariance matrix of the group flows.
4.3.5 Probability distribution of $S_{ik}$ and derivation of the objective function

The group flows are all of the form: $s_{lk} = \sum_{rshp} \delta_{lphk}^{rs} f_{p}^{rs} (h)$ and are therefore not independent. In matrix form: $S = \Delta F$. The probability of observing a certain group value $s_{lk}$ for $S_{lk}$ is:

$$P(S_{lk} = s_{lk}) = \sum_{F:s_{lk} = \Delta F} P(S_{lk} = s_{lk} | F = F) \cdot P(F = F)$$  \hspace{1cm} (4.4)

As the group flow is determined by the path flows, we have:

$$P(S_{lk} = s_{lk} | F = F) = \begin{cases} 1 & \text{if } s_{lk} = \Delta f. \\ 0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (4.5)

Combining 4.4 and 4.5 we get:

$$P(S_{lk} = s_{lk}) = \sum_{F:s_{lk} = \Delta f} P(F = F)$$  \hspace{1cm} (4.6)

As noted in Hazelton (2000) the calculation of this probability is tractable only for the simplest networks because of the number of path flows that can give rise to a specified group flow, but the first and second moments can be readily calculated; both for the assumption of independent poisson distributions of the route flows and for the assumption of multinomially distributed flows.

Under the assumption of independent poisson distributed flows we get:

$$E(S_{lk}) = \Delta_{phka}^{rs} F_{ph}^{rs} = \Delta_{phka}^{rs} P_{ph}^{rs} T_{h}^{rs}$$  \hspace{1cm} (4.7)

and

$$\sigma_{ij}^{S,1} = \Delta_{phka}^{rs} \sigma_{ij}^{F,1} \cdot (\Delta_{phka}^{rs})^{T} = \Delta_{phka}^{rs} \text{diag} (P_{p}^{rs} (h) T_{h}^{rs}) (\Delta_{phka}^{rs})^{T}$$  \hspace{1cm} (4.8)

Under the assumption of multinomial flow distribution, we get the same expected flow:

$$E(S_{lk}) = \Delta_{phka}^{rs} F_{ph}^{rs} = \Delta_{phka}^{rs} P_{ph}^{rs} T_{h}^{rs}$$  \hspace{1cm} (4.9)

and a different covariance matrix, which is not diagonal:

$$\sigma_{ij}^{S,2} = \Delta_{phka}^{rs} \sigma_{ij}^{F,2} (\Delta_{phka}^{rs})^{T}$$  \hspace{1cm} (4.10)
Log-likelihood of group flows

The log-likelihood of an observed set of group flows is expressed in terms of the stochastic values $s_{ik}$ of the group flows, and the observed values of the group flows $s^{obs}_{ik}$ as:

$$L(T, P) = \ln \left( \prod_{ik} P \left( s_{ik} = s^{obs}_{ik} \right) \right)$$

(4.11)

$$= \sum_{ik} \ln \left( P \left( s_{ik} = s^{obs}_{ik} \right) \right)$$

(4.12)

In view of the difficulties in calculating $P \left( s_{ik} = s^{obs}_{ik} \right)$, it is useful to approximate the distribution of $s_{ik}$ by a MVN normal distribution:

$$S_{ik} \sim N \left( S_{ik}, \sigma^{2}_{ik} \right)$$

(4.13)

The corresponding log-likelihood is (and denoting the transpose of matrix $X$ as $X^T$):

$$L(T, P) = -\text{const} - \frac{\#lk}{2} \ln \left( |\sigma^S_{ij}| \right) - \frac{1}{2} \left( S^{obs}_{ik} - S_{ik} \right) \left( \sigma^S_{ik} \right)^{-1} \left( S^{obs}_{ik} - S_{ik} \right)^T$$

(4.14)

The use of this log-likelihood is again difficult (see Hazelton (2000)), so that further simplification is useful. The true variance-covariance matrix $\sigma^S_{ik}$ can be replaced by the sample covariance matrix $\sigma^{S, obs}_{ij}$:

$$\sigma^{S, obs}_{ij} = (\#lk - 1)^{-1} \left( S^{obs}_{ik} - \bar{S}_{ik} \right) \left( S^{obs}_{ik} - \bar{S}_{ik} \right)^T$$

(4.15)

This gives an approximate log-likelihood function (omitting constants):

$$\tilde{L}(T, P) = -\frac{1}{2} \sum_{ik} \left( S^{obs}_{ik} - \bar{S}_{ik} \right) \left( \sigma^{S, obs}_{ij} \right)^{-1} \left( S^{obs}_{ik} - \bar{S}_{ik} \right)^T$$

(4.16)

which is a sum of squares, so that maximising the objective function produces a type of GLS estimator as used in the static case by Bell (1991) and Cascetta (1984), but now for the dynamic case. We will use this approximate log-likelihood function to construct the objective function.

### 4.3.6 Incorporating additional information

As noted in Hazelton (2000) this framework can incorporate additional sources of information such as partial matrices and especially *apriori matrices*, so that it covers group flows defined in 3.7. This results in additional terms in the log-likelihood function in eqn. 4.16.
Observed partial matrices

In this section we will indicate how observed partial matrices can be incorporated into the estimation. Sometimes it is possible to observe parts of the O-D matrix, e.g. through roadside surveys. Such surveys give observations $t^{rs}_h$ of a sample of the O-D flows $T^{rs}_h$, e.g. in the form of a list $\mathcal{L} = \left\{ t^{(rs)}_h \right\}_{j \in \mathcal{J}, h \in \mathcal{H}}$. With the distributional assumptions in Section 4.3.4, the exact likelihood of the list $\mathcal{L}$ of dynamic flows is:

$$L_{\mathcal{L}}(T) = \sum_{j \in \mathcal{J}} \sum_{h \in \mathcal{H}} t^{(rs)}_h \ln \left( T^{(rs)}_h \right) - T^{(rs)}_h$$  \hspace{1cm} (4.17)

In some cases (such as with cordon surveys or screenline surveys) the survey set-up allows to identify a set of O-D relationships $\mathcal{M}$ that must pass the survey points if they contain a non-zero amount of travel demand, irrespective of route choice. For such O-D pairs the absence from the set of observed O-D relationships means a valid observation of 0 trips for all O-D cells $t^{rs}_h$ that correspond to O-D pairs $(rs) \in \mathcal{M} - \mathcal{L}$. Of course one should ensure that the number of flows $|\mathcal{L}|$ is greater than the number of O-D pairs for this to work. If this is true, one can extend the list $\mathcal{L}$ with zero flows for those cells: $\mathcal{L}' = \mathcal{L} \cup \left\{ 0^{rs}_h \right\} \{(rs) \in \mathcal{M} - \mathcal{L}\}$.

With this extra set of flows, the log-likelihood in eqn 4.11 becomes:

$$L^*(T,P) = L(T,P) + L_{OD}(T,P)$$  \hspace{1cm} (4.18)

Apriori matrices

In view of the under-determinedness of the estimation problem we need to make extensive use of apriori information. We note that Cascetta et al. (1993) incorporate this into their GLS framework by minimising two sums of squares: one for the apriori and one for the counts (= observable groups). In this formulation the apriori matrix enters as a dataset, and one must specify the relative weight of the sum of squares of the apriori and the sum of squares of the grouped counts.

One can also reduce the under-determinedness by seeking the "closest" matrix to the apriori that reproduces the observations by parameterising the unknown matrix in some way by the prior and free coefficients. The question is then what parameterisation to choose.

Apriori matrices can be formally incorporated through a Bayesian framework by maximising the joint likelihood of the observations given the apriori and the likelihood of observing the observations:

$$\ln \left[ f \left( T, P | \mathcal{D} \right) \right] = l^* \left( T, P \right) + \ln h \left( T, P \right)$$  \hspace{1cm} (4.19)

The log-likelihood estimates are those values of $T'$ and $P'$ that maximise eqn. 4.19. The incorporation of apriori matrices is noted in Hazleton (2000) to be able to have a
profound effect on the curvature of the log-likelihood function, especially if the likelihood is based only on counts. This reflects that the estimation problem with counts alone has difficulty in distinguishing between matrices, a point that disappears if an apriori is present. The issue of coming up with the right apriori remains however.

We will use the approach in eqn. 4.19 in order to derive the objective function in Section 4.4.

Incorporation of measurement errors

Although Hazelton (2000) also gives a way of including measurement errors into the objective function, we will reserve this for further work as it does not change the essence of the objective function.

4.4 The objective function used for the upper-level O-D matrix estimation problem

The Maximum-likelihood objective function given in eqn. 4.19 combines the log-likelihood of observing the flows as a function of the O-D matrix with the likelihood of having and O-D matrix given the apriori matrix (and could have been used directly), but still has two drawbacks: it is complicated, and it is overparametrised as the entire O-D matrix appears as a free parameter. First we will simplify the objective function in the way proposed in Hazelton (2000), and obtain a quadratic objective function. We note that we will have to constrain the values of the O-D flows to be non-negative. It turns out that we can remove the constraints and reduce the number of free parameters at the same time by parameterising the O-D matrix.

Based on the objective function in eqn. 4.19 and the approximation of its first term in eqn. 4.16 in the previous sections, we will use the following general objective function as a starting point:

\[ I^* (T, P) = w_1 \sum_{lk} \left( s_{lk} - \Delta PT \right) S^{-1} \left( s_{lk} - \Delta PT \right)^T + w_2 \sum_{j \in J} \sum_{h \in H} t_h^{(rs)} \ln \left( T_h^{(rs)} \right) - T_h^{(rs)}, \tag{4.20} \]

The weights \( w_1 \) and \( w_2 \) are introduced to compensate for the change of scale due to the approximation of the real log-likelihood.

4.4.1 First simplification

We simplify eqn. 4.20 by approximating the term \( \sum_{j \in J} \sum_{h \in H} t_h^{(rs)} \ln \left( T_h^{(rs)} \right) - T_h^{(rs)} \) by \( \sum_{j \in J} \sum_{h \in H} \left( t_h^{(rs)} - T_h^{(rs)} \right)^2 \), and changing the weights \( w_1 \) and \( w_2 \) on the terms as appropriate. This gives the following GLS problem:
\[
z(T) = w_1 \sum_{l,k} (s_{lk} - \Delta PT) S^{-1} (s_{lk} - \Delta PT)^T \\
+ w_2 \sum_{j \in J} \sum_{h \in H} (t^{(rs)}_{hi} - T^{(rs)}_{hi})^2 \\
= w_1 z_1(T) + w_2 z_2(T, r^{rs,0})
\]  \hspace{1cm} (4.22)

that is split into two parts: one \(z_1\) for the match with the observations, and one \(z_2\) for the match with the apriori matrix. This is a classical form of objective function for O-D matrix estimation, which has been extensively researched. We will use this objective function as an intermediate step, omitting the term \(z_2\) and substituting a parametrisation of \(T\) into \(z_1\).

### 4.4.2 Parametrisation of the O-D matrix: matrix models

For practical reasons, we will assume that we estimate intrazonal flows separately (if at all) as these tend to perturb the estimation due to their size and poor observability, so that we are left with \(n(n - 1)\) O-D matrix cells to estimate. In principle we could regard every (of-diagonal) cell of the dynamic O-D matrix \(T_{hi}^{rs}\) as a free parameter to be determined in the estimation, or equivalently that we estimate K-factors (see Ortuzar and Willumsen (2001)) that scale the apriori matrix up to the observations. The disadvantage of this approach is that even with the use of an apriori matrix, the resulting estimation problem is seriously over-parameterised.

A number of approaches have been considered to resolve this issue. One is to use a distribution model, e.g. one based on destination-choice modelling (see Cascetta (2001)), and estimate its parameters. This approach is particularly appropriate as we assume that we will work in the context of a planning environment, so that we should have access to static distribution models and their output. We believe that using distribution models to provide the apriori matrix improves the quality of the estimation, as distribution models used in transportation planning tend to incorporate most if not all of the available data on the transportation system in question, often in considerable detail. Therefore, such models (when properly calibrated) can be expected to give the best apriori estimate of the O-D matrix given available transportation planning data, such as household surveys, road surveys, cordon counts, triplength distributions, etc.

In addition using distribution models to generate the apriori matrix "seeds" the apriori matrix for those cells that contain little traffic.

We also note that the estimation of latent parameters in such models could be used to improve the distribution models per se, especially in cases where otherwise K-factors (see Ortuzar and Willumsen (2001)) would be applied.

The downside of this approach is that typically the distribution models involved are quite large and often require disaggregate input, so that we prefer a simpler solution. We will assume that we will have the output of a static distribution model, and use a simple non-behavioural model to parameterise the dynamic matrix.
We have adopted a simple model inspired by the one used in the static base-matrix estimation (in Gunn et al. (1997) and Lindveld (1997), which in turn draw on Ben-Akiva (1997)). The form of this model was derived from the maximum entropy formulation of adapting a matrix to counts (see e.g. Ortuzar and Willumsen (2001)):

\[ T^{rs} = T^{rs,0} \prod_{l=1}^{L} \mu_l^{f_l P^{rs}}_l = T^{rs,0} \prod_{l=1}^{L} e^{\lambda_l f_l P^{rs}}_l \]

\[ \mu_l = e^{\lambda_l}, \ l = 1 \ldots L \]  

(4.23)

with:

- \( T^{rs} \) the modelled apriori matrix
- \( T^{rs,0} \) an apriori matrix
- \( P^{rs}_l \) assignment fraction
- \( \mu_l \) parameters in the multiplicative model

**Extension to the dynamic case**

Extending eqn. 4.23 to the dynamic case we obtain a parameter vector \( \Lambda_{lk} \) with a time dimension \( \kappa \) and the following parametrisation:

\[ T^{rs}_h = T^{rs,0}_h \prod_{lk} e^{\lambda_{lk} f^l P^{rs}_{lk}}_l \]

\[ l = 1 \ldots L, \ \kappa \in k \]  

(4.24)

with:

- \( T^{rs}_h \) the modelled dynamic O-D matrix
- \( T^{rs,0}_h \) a dynamic apriori matrix
- \( f^l \) sampling fraction \( f^l \in [0,1] \), 1 for counts but <1 for surveys
- \( P^{rs}_{lk} \) assignment fraction
- \( \lambda_{lk} \) elements of the parameter vector \( \Lambda_{lk} \)

Note that the parameter vector \( \Lambda_{lk} \) is given with respect to the time intervals \( \kappa \) for which observations \( S_{lk} \) are given.

**Observations on the parametrisation**

As noted in Section 3.7.2, there is a one-to-one correspondence between parameters and group observations, which ensures identifiability. The proposed parametrisation also requires much less parameters than there are O-D matrix cells, as shown below.

If we only have observations for zonal productions, zonal attractions and traffic counts, the number of parameters is:

\[ N_{\text{parametrised}} = \left( \frac{2N_{zones} + N_{counts}}{H} \right) T \]  

(4.25)
with the number of zones $N_{\text{zones}}$, the number of counters $N_{\text{count}}$, the time period of the matrix $T$, and the duration $H$ of the time interval with which measurements are given.

If each (off-diagonal, see Section 4.4.2) O-D matrix cell is treated as an independent parameter, or if K-factors were to be used, the number of required parameters would be:

$$N_{\text{naive}} = \frac{(N_{\text{zones}})^2 - N_{\text{zones}} + N_{\text{counts}}}{\delta} T$$

(4.26)

If we relate the parameters to the (coarser) external time periods $K$ instead of the internal ones $k$, we get:

$$N_{\text{classic}} = \frac{(N_{\text{zones}})^2 - N_{\text{zones}} + N_{\text{counts}}}{\Delta} T$$

(4.27)

Denote $\frac{N_{\text{naive}}}{N_{\text{parametrised}}}$ as $\text{Ratio1}$, and $\frac{N_{\text{classic}}}{N_{\text{parametrised}}}$ as $\text{Ratio2}$. To illustrate the importance of the parametrisation, we take $\delta = 6$ sec., $\kappa = H = 600$ sec (10 min), $N_{\text{counts}} = 3$ and tabulate $N_{\text{naive}}, N_{\text{classic}}, N_{\text{parametrised}}, \text{Ratio1}$, and $\text{Ratio2}$ as shown in the Figure 4.6.

<table>
<thead>
<tr>
<th>$N_{\text{zones}}$</th>
<th>Naive</th>
<th>Nclassic</th>
<th>Nparametrised</th>
<th>Ratio1</th>
<th>Ratio2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>6000</td>
<td>60</td>
<td>84</td>
<td>71</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>10800</td>
<td>108</td>
<td>108</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>18000</td>
<td>180</td>
<td>132</td>
<td>136</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>27600</td>
<td>276</td>
<td>156</td>
<td>177</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>39600</td>
<td>396</td>
<td>180</td>
<td>220</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>54000</td>
<td>540</td>
<td>204</td>
<td>265</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>70800</td>
<td>708</td>
<td>228</td>
<td>311</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>90000</td>
<td>900</td>
<td>252</td>
<td>357</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>11160</td>
<td>1116</td>
<td>276</td>
<td>404</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>13560</td>
<td>1356</td>
<td>300</td>
<td>452</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>16200</td>
<td>1620</td>
<td>324</td>
<td>500</td>
<td>5</td>
</tr>
<tr>
<td>13</td>
<td>19080</td>
<td>1908</td>
<td>348</td>
<td>548</td>
<td>5</td>
</tr>
<tr>
<td>14</td>
<td>22200</td>
<td>2220</td>
<td>372</td>
<td>597</td>
<td>6</td>
</tr>
<tr>
<td>15</td>
<td>25560</td>
<td>2556</td>
<td>396</td>
<td>645</td>
<td>6</td>
</tr>
<tr>
<td>20</td>
<td>45960</td>
<td>4596</td>
<td>516</td>
<td>891</td>
<td>9</td>
</tr>
<tr>
<td>25</td>
<td>72360</td>
<td>7236</td>
<td>636</td>
<td>1138</td>
<td>11</td>
</tr>
<tr>
<td>100</td>
<td>1188360</td>
<td>118836</td>
<td>2436</td>
<td>4878</td>
<td>49</td>
</tr>
<tr>
<td>200</td>
<td>4776360</td>
<td>477636</td>
<td>4836</td>
<td>9877</td>
<td>99</td>
</tr>
<tr>
<td>300</td>
<td>10764360</td>
<td>1076436</td>
<td>7236</td>
<td>14876</td>
<td>149</td>
</tr>
<tr>
<td>400</td>
<td>19152360</td>
<td>1915236</td>
<td>9636</td>
<td>19876</td>
<td>199</td>
</tr>
<tr>
<td>500</td>
<td>29940360</td>
<td>2994036</td>
<td>12036</td>
<td>24876</td>
<td>249</td>
</tr>
<tr>
<td>1000</td>
<td>119880360</td>
<td>11988036</td>
<td>24036</td>
<td>49875</td>
<td>499</td>
</tr>
</tbody>
</table>

Figure 4.6: Number of O-D matrix cells and parameters needed for parametric model

It is clear that the number of free parameters in the naive parametrisation increases rapidly with the dimension of the O-D matrix, and requires approximately $\frac{HN_{\text{zones}}}{2h}$ times as many parameters than the parametrisation proposed. This difference (a factor of 404 for $N_{\text{zones}} = 10$ and a factor of 4878 for $N_{\text{zones}} = 100$) could spell the difference
between a feasible and an infeasible problem. The difference between the proposed parametrisation and one that uses one free parameter for all O-D cells relating to a single O-D relationship for time interval $H$ is much smaller (ranging from a factor of 4 for $N_{zones} = 10$, to a factor of 49 for $N_{zones} = 100$) but is still significant when the estimation is to be applied to real-world cases.

**Remarks on the parametrisation**

This parametrisation has the following advantages:

1. it can take advantage of any destination-choice model available to provide the apriori matrix $T_h^{*s.0}$

2. it avoids the need to incorporate a full-blown disaggregate destination and mode choice model in the estimation, which would otherwise be necessary in order to reconcile distribution with the L.O.S. resulting from the assignment

3. it is better able to pick up geographic anomalies than a destination-choice model because traffic counts and cordon counts directly affect the estimation result instead of through a spatial distribution function

4. it automatically gives the matrix that is both closest to the apriori matrix (with respect to the entropy norm) and minimises the discrepancy with the observed quantities. This means that it uses the distribution model to provide default values for any part of the matrix for which no observations are available.

**4.4.3 Final objective function after parametrisation and further simplification**

Substituting the parametrisation of eqn. 4.24 into eqn. 4.21 and removing the term dealing with the apriori matrix (this is now implicit through the parametrisation) gives us the following unconstrained least-squares problem for the first level problem:

$$z(\Lambda) = z(\Lambda)$$
$$= \sum_{lk} (s_{lk} - S_{lk}(\Lambda)) \left( \sigma_{ij}^{S,obs} \right)^{-1} (s_{lk} - S_{lk}(\Lambda))^T$$

The objective function in eqn. 4.28 may have multiple minima. Whether this is the case depends on the problem instance. The minimisation is now over the parameter vector $\Lambda$ and gives the matrix estimate indirectly:

$$\hat{\Lambda} = \min_{\Lambda} z(\Lambda)$$
$$\hat{T}(\Lambda) = T(\hat{\Lambda})$$
In view of the nonlinearity of 4.28 induced by the factor $T(\Lambda)$, a direct solution is not possible, and an iterative solution method must be used. The methods reviewed selected in chapter 8 all require a vector of derivatives $\nabla_A z(\Lambda)$ of the objective function $z(\Lambda)$ in eqn. 4.28. In view of the form of $z(\Lambda)$ and the symmetry of $S$ we have (see Mardia et al. (1979), appendix A):

$$\nabla_A z(\Lambda) = 2W^{-1} (s_{ik} - S_{ik}(\Lambda))^T \nabla_A S_{ik}(\Lambda)$$

The difficult part is the calculation of $\nabla_A [S_{ik}(\Lambda)]$ as the dynamic link-path incidence matrix $\Delta$ and the route- and departure time choice fractions $P$ depend on the value of the dynamic O-D matrix $T(\Lambda)$, so that in principle the calculation of $\nabla_A z(\Lambda)$ involves\taking the derivative of the entire dynamic equilibrium traffic assignment procedure as described in chapter 6. This reflects the bilevel nature of the dynamic O-D matrix estimation problem, as will be reviewed in chapter 5. It turns out however (see chapter 8) that heuristics exist for the solution of eqn. 4.28 that assume that the assignment is relatively unchanged by changes in the dynamic O-D matrix and which use the following approximate descent directions:

$$\nabla_A z(\Lambda) \approx 2W^{-1} (s_{ik} - S_{ik}(\Lambda))^T \nabla_A T(\Lambda)$$

### 4.5 Summary

In chapter 3 a behavioural framework was proposed for dynamic O-D matrix estimation, which already anticipated that the O-D matrix would be parameterised in order to remove the overparametrisation of the problem.

In this chapter we extended a statistical approach for static O-D matrix estimation to the dynamic case by assuming that we can model the travel demand for each time-interval separately. Starting from the ML formulation of the O-D matrix estimation problem, existing literature shows how to derive an approximate least-squares type of objective function of the classical type. This objective function has the disadvantages that it is a function of the complete O-D matrix to be estimated and therefore is very over-parameterised, and needs to have its elements constrained to be positive. Both the over-parametrisation and the need for constraints can be removed through a parametrisation of the O-D matrix in terms of a classical multiplicative model involving the O-D matrix and any observable properties of the transportation system that can be expressed as linear sums of O-D flows. With this substitution, the objective function is parameterised in terms of a parameter vector $\Lambda$ which has as many elements as there are data sources; the parametrisation can only change the O-D matrix in those cells for which observations are available; for all other cells the parametrisation leaves the apriori matrix unchanged.

This gives the objective function from the statistical point of view. Minimisation of this objective function can then be considered with minimal reference to the nature of the objective function or the statistical structure of the problem, as set out in chapter 5 below.
Chapter 5

Dynamic O-D matrix estimation as an optimisation problem

5.1 Introduction

In chapter 3 we presented a structure for behavioural O-D matrix estimation while in chapter 4 we defined our objective function based on the statistical properties of the estimation problem that result from the distributional assumptions made.

In doing this we did not consider the effect of changes in the matrix on the routing or on the propagation of traffic flows through the network (as discussed in chapter 6 on dynamic traffic assignment). In this chapter we will present the mathematical structure of the matrix estimation problem, highlighting the interdependence of the dynamic O-D matrix and the dynamic traffic flows. As in the static case, in the structure of the resulting problem mathematical problem we can recognize existing mathematical problems such as the bi-level problem and more specifically the Mathematical Program with Equilibrium Constraints (MPEC). A considerable amount of theory exists for these problems, which is applicable to our case.

If we denote the matrix estimation process itself by $\mathcal{E}$ in the sense that $\hat{T}_h^{rs} = \mathcal{E}(\hat{T}_h^{rs,0}, S_{lk})$ with $\hat{T}_h^{rs}$ the matrix estimate, $\hat{T}_h^{rs,0}$ an apriori matrix and $S_{lk}$ the observations, then for consistency we can express the matrix estimation problem as a fixed-point problem: determine $\hat{T}_h^{rs*}$ such that $\hat{T}_h^{rs*} = \mathcal{E}(\hat{T}_h^{rs*}, S_{lk})$. This is a very general formulation, but unfortunately one that is less suitable from the point of view of algorithm development. Moreover in this representation the partial dependency of the assignment fractions of the synthetic group totals on the matrix itself is not represented in an insightful way. A less general formulation, but one which is more insightful allowing an easier expression of this relationship is the bi-level formulation, in which the adjustment of the matrix and the ensuing reaction of the dynamic traffic assignment are modelled as two different (but linked) optimisation problems.
5.2 Mathematical formulation of the dynamic matrix estimation problem with deterministic route choice

We will formulate the dynamic matrix estimation problem as an analogue to the static case. We seek an inelastic (trip generation is not influenced by the level of service between zones) dynamic O-D matrix $T_{h}^{rs}$ ($r, s$ being the origin and destination, and $h$ denoting a time-slice) whose cells $t_{h}^{rs}$ take on a fixed value during each time interval $h$, using observed flows and an apriori matrix. We will use $\Xi$ to denote the length of the time period under study, and time instances with $h \in H$. We have a network $G$ consisting of arcs $a \in A$ and nodes $n \in N : G = (A, N)$. A subset of the nodes are zones $z \in Z$; $Z \subset A$, zones produce and attract trips.

5.2.1 The dynamic assignment map

If we use a Deterministic Dynamic User Equilibrium (DUE-DET) assignment, then we get a dynamic assignment map $\Phi_{T_{h}^{rs}}$ that links the time-dependent vector of traffic link-inflows $U_{ak}$ to the dynamic O-D matrix $T_{h}^{rs}$ as in:

$$ U_{ak} = \Phi_{T_{h}^{rs}} \circ T_{h}^{rs} $$

(5.1)

Link inflows are used as the definition of dynamic traffic flow on a link. The assignment map $\Phi_{T_{h}^{rs}}$ (with elements $\phi_{aphk}^{rs}$) combines the following ingredients:

- the link-path incidence matrix $\Delta_{ap}^{rs}$ that indicates which links are on which paths;
- the link travel times $\tau_{ak}$
- the route fractions $\Phi_{rph}^{rs}$ that specify the distribution over equal cost routes when setting out for destination $s$ from origin $r$ during time interval $h$.

The link-path incidence matrix and the travel times together define the Dynamic Network Loading (DNL) map $\delta_{aphk}^{rs}$ which determines which proportion of the trips that started during period $h$ on path $p$ from $r$ to $s$ entered link $a$ during time period $k$. The DNL map links the vector $\mathcal{U}(k)$ of dynamic link inflows (with components $u_{a}(k)$) to the route flows $f_{aphk}^{rs}$ as in:

$$ u_{a}(k) = \sum_{rs} \sum_{p} \sum_{h} f_{aphk}^{rs} \delta_{aphk}^{rs} $$

(5.2)

The assignment map links the O-D flows $T_{h}^{rs}$ to the link flows as:

$$ u_{a}(k) = \sum_{rs} \sum_{p} \sum_{h} T_{h}^{rs} \phi_{aphk}^{rs} $$

(5.3)
The assignment map summarises the allocation of travel demand and infrastructure supply, but until this point no mention was made of any equilibrium between the two. The assignment map that defines a Dynamic User Equilibrium is a special one that satisfies certain conditions (for a more complete account, see chapter 6) which we will state here in terms of variational inequalities. We will use the notation \( \langle \cdot, \cdot \rangle \) to denote an inner product.

The assignment map is defined so that the resulting flows satisfy the VIP defined by the inner product (see Section 6.5):

\[
\langle C^*_h \left( U^*_{ak} \left( T^*_h \right) \right), (U_{ak} - U^*_{ak}) \rangle \geq 0 \quad \forall U_{ak} \in \Omega
\]

(5.4)

with \( U^*_{ak} \left( T^*_h \right) \) the dynamic equilibrium flow, \( C^*_h \left( U^*_{ak} \left( T^*_h \right) \right) \) the cost vector associated with equilibrium flow, and \( \Omega \) the set of all permissible traffic flow patterns (assumed fixed and given).

Undeterminedness of the assignment mapping

Unfortunately the assignment map is not uniquely defined for the deterministic assignment, because whilst the equilibrium link flow vector \( U^*_{ak} \) is uniquely defined by eqn. 5.4, DUE-DET assignment allows the pathflows to be changed at will within the constraints imposed by eqn. 5.2. In practice one might simply settle for the route fractions \( \mathcal{P}^*_{ph} \) returned by the solution algorithm used, but this is an arbitrary choice. This problem can be avoided by using a DUE-STOCH assignment as presented in Section 5.3.

5.2.2 The objective function

The objective function for the O-D matrix estimation problem \( z \left( T^*_h \right) \) is the one specified in eqn. 4.28 in Section 4.4.3, and is related to the flows and the parameter vector \( \Lambda_{lk} \) as:

\[
z \left( \Lambda_{lk} \right) = z \left( U_{ak} \left( T^*_{h} \left( T^{rs,0}_{h}, \Lambda_{lk} \right) \right), \bar{U}_{ak} \right)
\]

(5.5)

where \( U_{ak} \left( T^*_h \right) \) is the vector of traffic flows obtained by dynamic assignment of the O-D matrix \( T^*_h \), \( U_{ak} \) are the observed traffic flow volumes, and \( T^{rs,0}_{h} \) is an apriori matrix which is considered to be fixed. Note that the vector of free parameters \( \Lambda_{lk} \) is nested two levels deep:

\[
T^*_h : \Lambda_{lk} \rightarrow T^*_h \left( T^{rs,0}_{h}, \Lambda_{lk} \right)
\]

(5.6)

and

\[
T^*_h : \mathbb{R}^{|L| \times K} \rightarrow \mathbb{R}^{|Z| \times |Z| \times K}
\]

(5.7)

and

\[
z : \left( U_{ak} \left( T^*_h \right), \bar{U}_{ak} \right) \rightarrow z \left( U_{ak} \left( T^*_h \right), \bar{U}_{ak} \right)
\]

(5.8)

\[
z : \mathbb{R}^{|Z| \times |Z| \times K} \rightarrow \mathbb{R}
\]

(5.9)
Note also that the parameter vector $\Lambda_{lk}$ of dimension $|L| \times K$ is used to parameterise a dynamic O-D matrix of dimension $|Z| \times |Z| \times K$, where $|L| \approx 2|Z| + \text{const}$, where typically $2|Z| + \text{const} \ll |Z| \times |Z|$ (see also Section 4.4.2). The difference is made up by the apriori matrix $T_{lk}^{\text{pid}}$, which highlights the amount of information contained in it.

5.2.3 Formulation of the dynamic O-D matrix estimation problem

With the notation of this chapter, the problem of determining the matrix $T^*_i$ given an apriori matrix $T^{r_o,0}$ and observed traffic flows $\bar{U}_{ak}$ (OD-ESTIM-DET) can be stated as the following bi-level optimisation program.

\[
\begin{align*}
\hat{T}_{nk}^* &= \arg\min_{\Lambda_{lk}} z(T^*_k(\Lambda_{lk})) \\
\text{s.t.} \quad \langle C_h^*(\mathcal{U}_{ak}^* (T_{nk}^*)), (\mathcal{U}_{ak} - \mathcal{U}_{ak}^*) \rangle &\geq 0 \quad \forall \mathcal{U}_{ak} \in \Omega(T_i^{*}) \quad (5.10b)
\end{align*}
\]

This bi-level problem has a special structure in that the lower level problem represents an equilibrium condition; this type of problem is known as an MPEC or Generalised Bi-Level Problem (GBLP).

The relevance of this observation is twofold. First of all it highlights the modular structure of the problem: it is possible to change the nature of the equilibrium conditions without affecting the structure of the O-D matrix estimation problem; we will take advantage of this fact in sections 5.3 and 5.4 by extending the assignment to include stochastic route choice and departure-time adjustment. Secondly, it highlights the connection with game-theory.

The difference with the ordinary GBLP as it emerges in the formulation of the static O-D matrix estimation problem) is that in the above problem, the feasible set $\Omega$ of the flows depends on the unknown parameter $T^*_i$ as shown by the notation $\Omega(T_i^{*})$. As a result, problem, (5.10b) is not a VI problem, but is known as a quasi-VI problem (see Bielemer (2001)).

The relationship between the upper and the lower problems is illustrated in Figure 5.1.
Objective function
\[ T_k = \arg \min_{T_k} Z(T_k) \]

\[ v_{ak}(T_k) \]

Constraints
\[ c'v_i + v_i - v_i' \geq 0 \quad \forall v_i \]

Dynamic user equilibrium assignment

Leader

\[ T_k \]

\[ v_{ak}(T_k) \]

Followers

5.2.4 The MPEC formulation and game theory

The relevance of identifying our bi-level problem as an MPEC is that it highlight a formal link with game theory (imposing the equilibrium constraint is equivalent to constraining the lower level to be the solution of a Nash game). The equivalence of this problem to a Stackelberg game is noted in Yang et al. (1994), and Luo et al. (1997).

Any MPEC can be interpreted as a game: the maximisation problem is the top-level player, and the equilibrium constraints result from a n-player Nash game. Suppose the player at the top level knows the exact cost function and with it the precise response of the bottom level to any change made at the toplevel. This player could then use this response function to eliminate the response of the bottom level player from his own cost function, and would be left with a cost function solely in terms of his own control variables. This function could then be maximised without further reference to the reaction of the bottom level players as their actions are accounted for anyway. This is known as a Stackelberg game (see Basar and Olsder (1999), section 4.6).

In the event that the toplevel player knows nothing about the response surface of the bottom level players, we have a Cournot game (see Basar and Olsder (1999), section 4.6, example 4.5). In this case a way to look for a mutually stable solution would be to "play the game", i.e. let each player in turn optimise its objective function conditional on the previous action of the other player. In general both games may have different solutions, as is noted e.g. in Bell and Allsop (1998). By definition the payoff of the toplevel player in the Stackelberg game (in our case the objective function), cannot be worse than in the Cournot game, so that the Stackelberg game should be expected to yield the best estimator.

Various formulations of the dynamic O-D matrix estimation can then by characterised by the amount of information on the response of the bottom level problem that is...
available to the top-level player. However, in view of the complicated structure of the bottom level problem, the large number of variables involved and its the nonlinearity, it is very difficult (if not computationally intractable) to eliminate the response of the bottom level problem from the top-level objective function altogether and obtain a true Stackelberg game. It is feasible however to present the top-level problem with a linear approximation of the response of the bottom level problem through a first-order Taylor approximation, and the resulting game will be one of imperfect information. One might in principle present the toplevel game with a quadratic approximation of the bottom level problem, which would require the calculation of the Hessian of the bottom level problem. We have not investigated this, concentrating instead on the linear approximation.

The question then can be reduced to the determination of derivatives of the response surface of the dynamic assignment problem. Even this modest objective runs into computational complications, so that in Section 6.7.2 we will turn to truncated derivatives of the response surface of the bottom level problem.

From a game-theoretic point of view this means that one can view the Cournot and Stackelberg games as two extremes of a continuum of games, in which all games with imperfect information (including those with truncated derivatives) can be fitted. This explains why solution algorithms using approximate derivatives (see chapter 8) can make sense: they solve a slightly different MPEC which is closer to or further from the true Stackelberg game.

The MPEC character is also fundamental in the sensitivity analysis of the problem (see Tobin and Friesz (1988), Patriksson (1994), and Patriksson and Rockafellar (2001)) that points the way towards calculation of the derivatives of the lower level problem. Such derivatives can be used to find a suitable descent direction for the MPEC, as proposed in e.g. Clegg and Smith (2001) and Luo et al. (1997).

5.3 The dynamic matrix estimation problem with stochastic route choice

In the previous section the dynamic matrix estimation problem with deterministic route choice was presented as an MPEC, with a deterministic dynamic user-equilibrium assignment as the equilibrium constraint. As a result of the separation of the estimation problem in an upper and a lower level, stochastic route-choice can be accommodated by substituting a Stochastic Dynamic User Equilibrium assignment (DUE-STOCH) for the bottom-level problem.

If we use a DUE-STOCH assignment, then we also get a dynamic assignment map $\Phi_T$ that links the time-dependent vector of traffic link-inflows $\ell_{nk}$ to the dynamic O-D matrix $T^*_k$ as in Section 5.2.1. The main difference is in the determination of the matrix of route choice probabilities $P_{rs}^p(h)$.

Both the objective function and the dynamic assignment map are identical to the one specified in Section 5.2.3, but the equilibrium conditions that give rise to the
dynamic assignment map are slightly different. In the formulation with departure time adjustment, the VIP is:

\[ \langle G_p^s (h), (F_p^s (h) - F_p^{rs} (h)) \rangle \geq 0 \quad \forall F_p^{rs} \in \Omega \]  
\[ \text{s.t.} \]  
\[ f_p^s (h) = t_k^s p_p^s (h) \]  
\[ g_p^s (h) = [f_p^s (h) - t_k^s p_p^{rs} (k)] \frac{\partial \eta_p^s (k)}{\partial t_k^s} \]  
\[ \frac{\partial \eta_p^s (k)}{\partial t_k^s} > 0 \]  
(5.11a)

The formulation of the O-D estimation problem OD-ESTIM-STOCH then becomes:

\[ \hat{F}_h^s = \arg \min_{\hat{A}_{ls}} z (T_k^s (A_{ls})) \]  
\[ \text{s.t.} \]  
\[ \langle G_p^s (k), (F_p^s (k) - F_p^{rs} (k)) \rangle \geq 0 \quad \forall F_p^{rs} \in \Omega \]  
(5.12a)

5.4 Formulation of the dynamic matrix estimation problem with stochastic route choice and stochastic departure time adjustment

In the previous section the dynamic matrix estimation problem was presented as an MPEC, with dynamic user-equilibrium assignment as the equilibrium constraint. By enriching the structure of the equilibrium constraint to incorporate departure time adjustment one obtains a slightly different but behaviourally interesting problem of dynamic O-D matrix estimation with departure time choice.

Both the objective function and the stochastic dynamic assignment map are identical to the one specified in Section 5.2.3, but the bottom-level problem that gives rise to the dynamic assignment map is slightly different. In the formulation with departure time adjustment, the VIP is defined as in Section 6.8.

The equilibrium conditions are:

\[ \langle G_p^s (k), (F_p^s (k) - F_p^{rs} (k)) \rangle + \langle J_p^s (k), (T_k^s - T_k^{rs}) \rangle \geq 0 \quad \forall F_p^{rs}, T_k^s \in \Omega \]  
\[ \text{s.t.} \]  
\[ f_p^s (k) = t_k^s p_p^s (k) \]  
\[ g_p^s (k) = [f_p^s (k) - t_k^s p_p^{rs} (k)] \frac{\partial \eta_p^s (k)}{\partial t_k^s} \]  
\[ \frac{\partial \eta_p^s (k)}{\partial t_k^s} > 0 \]  
\[ j_p^s (k) = [t_k^s - t_k^{rs0} p_k^s (k)] \frac{\partial \xi_p^s (k)}{\partial t_k^s} \]  
\[ \frac{\partial \eta_p^s (k)}{\partial t_k^s} > 0, \frac{\partial \xi_p^s (k)}{\partial t_k^s} > 0 \]  
(5.13a)
The formulation of the O-D estimation problem OD-ESTIM-SDTAC then becomes:

\[
\hat{T}^{rs}_{ht} = \arg\min_{\Lambda_{ik}} z(T^{rs}_{k} (\Lambda_{ik})) \\
\text{s.t.} \\
\langle G^{rs}_{p} (k), (F^{rs}_{p} (k) - F^{rs*}_{p} (k)) \rangle + \langle J^{rs}_{p} (k), (T^{rs}_{k} - T^{rs*}_{k}) \rangle \geq 0 \quad \forall F^{rs}_{ph}, T^{rs}_{h} \in \Omega
\]  

(5.14a)

(5.14b)

5.5 Summary

In this chapter we took the objective function from chapter 4, which represents the O-D matrix estimation from a statistical point of view, and considered its optimisation as a mathematical problem. We noted that (as in the static case) the O-D matrix estimation problem has a bilevel form. The O-D matrix estimation problem can be cast as an MPEC (Mathematical Program with Equilibrium Constraints), with the optimisation problem itself forming the upper level and the equilibrium traffic assignment forming the lower level. This has three implications. Firstly, the modularisation of the problem ensures that it is possible to use a variety of dynamic assignment techniques without affecting the structure of the O-D matrix estimation problem. Secondly, the modularisation highlights a link with game theory that clarifies why approximate solution methods may be acceptable. Thirdly, the bilevel nature of the problem means that O-D matrix estimation is not a straightforward minimisation problem except in very restricted cases where the user equilibrium assignment plays no role.

From this point of view it is clear that the lower level problem, the user equilibrium assignment, plays an essential and intricate role which we will elucidate in chapter 6 below. As noted in Section 5.2.1, due to the indeterminateness of the assignment mapping of the DUE-DET assignment, the estimation problem OD-ESTIM-DET has some arbitrary aspects, which are not present in the stochastic version OD-ESTIM-STOCH or OD-ESTIM-SDTAC. Therefore we will restrict ourselves to the stochastic dynamic assignments DUE-STOCH and DUE-SDTAC.

The actual calculation of the solution to the MPEC problem using both the lower and the upper level is considered in chapter 8.
Chapter 6

Dynamic user equilibrium assignment

6.1 Introduction

In this chapter we will address the issue of calculating the dynamic assignment map that results from Dynamic User Equilibrium (DUE) conditions with respect to joint route choice and departure time choice. This map is crucial for the dynamic O-D matrix estimation problem (stated in chapter 3, Section 3.4 and chapter 5, Section 5.2, p.72) because it gives the relationship between the unobservable dynamic O-D matrix and the observable quantities related to the dynamic link flows (see p.72). This map is calculated by the Dynamic Traffic Assignment (DTA) procedure.

The difference between static and dynamic O-D matrix estimation lies in the difference between the behavioural and mathematical aspects of the dynamic assignment map and its static counterpart, and in the additional degree of freedom presented by departure time choice (or departure time adjustment). The behavioural aspect is that in addition to the route choice provided by the DTA, the specification in Section 3.4 requires a user equilibrium with respect to both route choice and departure time choice, which will be described in Section 6.8. The joint user equilibrium involving both route choice and departure time choice will be formulated in Section 6.8. Whereas the static assignment simply consists of the product of the route choice probability matrix and the link-path incidence matrix, a dynamic assignment map also specifies when certain path-flows will contribute to which link on the selected path. This naturally affects the dynamic link costs, thereby the dynamic route costs, and therefore the dynamic route choice probabilities. This calculation is carried out by the Dynamic Network Loading procedure, and profoundly complicates the calculation of the assignment map. Last but not least, there is the question of existence, uniqueness and differentiability of the dynamic assignment map.

For clarity we have listed the acronyms for User Equilibrium models used in this chapter in table 6.1. The definition of the network can be found in Section 2.2.1.

The SUE-DET assignment corresponds to the classical static assignment as covered in Patriksson (1994). The SUE-STOCH assignment corresponds to static assignment
Table 6.1: List of acronyms for User Equilibrium assignments

with route choice as discussed in Patriksson (1994), Sheffi (1985), and Daganzo (1977).

An overview of deterministic dynamic user equilibrium (DUE-DET) dynamic assignment models and their development can be found in Bliemer (2001). Unfortunately deterministic assignments don’t incorporate the behavioural aspects specified in the framework of chapter 3, and especially route choice in Section 3.8. Therefore we need to use Stochastic Dynamic User Equilibrium (DUE-STOCH) assignment formulated as a VIP problem, for which we have selected the implementation described in He (1997) and Chabini and He (1998). In our context we use the term "stochastic" to indicate the presence of an error term in the user’s valuation of route cost, which leads to route choice probabilities that give the expected value of the traffic flows. The error term covers many things from perception errors, taste variation, use of proxy variables (see Ben-Akiva and Lerman (1989)) to subjective choice sets (see Bovy and Stern (1990)).

The expected value of the flows does not in itself contain any stochastic component. The usual models give analytic expressions for the choice probabilities and hence the expected values of the flows. In those cases where we cannot obtain analytic expressions for the route choice probabilities, we would have to approximate the expectations using simulation, which then would lead to stochastic components in the resulting flows in the sense that two different model runs would produce slightly different realisations of the stochastic variables involved.

We note that the dynamic assignment mapping forms the primary link between the (unobservable) O-D matrix and the (observable) dynamic group observations $S_{ik}$, and is therefore of importance in the determination of the properties of the objective function $z$, which is minimised in the determination of the O-D-t matrix estimate.

Next we present a Variational Inequality (VI) formulation of the DUE-STOCH and combined DTAC-STOCH-DUE-STOCH methods selected for use, and show how it relates to the DTA methods discussed e.g. in Bliemer (2001). Then we will specify the calculation of assignment fractions from the assignment map given by the DTA procedure in He (1997).

Finally we will turn towards the question of sensitivity and stability of the DUE-STOCH and DTAC-STOCH-DUE-STOCH methods with respect to perturbation of the O-D matrix, and show that under reasonable conditions the assignment result depends in a continuous way on perturbed input parameters. This is important later on when we will consider the properties of the solution of the O-D matrix estimation problem formulated as a bilevel problem, and the convergence of algorithms for the O-D matrix estimation problem in chapter 8.
6.2 Literature survey on User Equilibrium assignment models

A list of analytic dynamic assignment models that appear in the literature is presented in table 6.2. The DUE-DET assignment is covered in Ran and Boyce (1996), Chen (1999), and Bliemer (2001). The DUE-STOCH assignment is covered in Ran and Boyce (1996) and Chen (1999). We note that DUE-DET assignments do not incorporate route choice behaviour nor departure time choice. The DTC-STOCH equilibrium model is a model in which users are free to choose their departure time by 30-60 minutes and therefore to some extent to (re)schedule their activities. The DTAC-STOCH model is one in which users have a preferred departure time and cannot change their activity patterns, but try to minimise their schedule delay by making small adjustments to their departure time.

A shortlist of sources dealing with the formulation of dynamic user equilibria as Variational Inequality Problems (VIP’s) or Nonlinear Complementarity Problems (NCP’s) is given in table 6.2. The headings of the table include: Disc/Cont (Discrete-time / Continuous time), Determ/Stoch (Deterministic or Stochastic route choice), Path/Link (Path-based assignment / Link-based assignment), Publication contains proof of: VIP Eqn. (Equivalence between the VIP formulation used and the user-equilibrium conditions), Erlist (Existence of a solution to the VIP), Uniq. (uniqueness of the solution), Ct (Continuity of the solution w.r.t. e.g. changes in the O-D matrix, Diff (Differentiability of the solution w.r.t. changes in the O-D flows), and FIFO (assignment respects the FIFO conditions).

<table>
<thead>
<tr>
<th>Nr</th>
<th>Publication</th>
<th>Disc Cont.</th>
<th>Determ Stoch</th>
<th>Path/Link</th>
<th>Publication contains proof of:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ran and Boyce (1996)</td>
<td>C</td>
<td>D,S</td>
<td>P</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>He (1997)</td>
<td>D</td>
<td>S</td>
<td>P</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>Zhu and Marcotte (2000)</td>
<td>C</td>
<td>D</td>
<td>P</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>Wie et al. (2002),</td>
<td>D</td>
<td>D</td>
<td>L</td>
<td>NCP +</td>
</tr>
<tr>
<td>8</td>
<td>Huang and Lam (2002)</td>
<td>D</td>
<td>D</td>
<td>P</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 6.2: Literature on the dynamic VIP formulations of UE

1 Uniqueness of path flows is proved for path flows aggregated over all destinations.

As noted in Chen (1999), for discrete-time path-based VIP's the question of existence and uniqueness of solutions is covered by the uniqueness and existence theorems found in Nagurney (1993), which rely on the continuity and monotonicity of the cost functions.
involved. As shown in Bliemer (2001), there is a difference between path-based and link-based formulations in that the former are true VI problems but the latter are only quasi-VI problems, as the solution-space depends on the traffic propagation characteristics. We will use path-based VI formulations only. References to stochastic dynamic user-equilibrium (DUE-STOCH) assignments can be found in Ran and Boyce (1996) and He (1997) on the one hand, and Chen (1999) on the other. The approaches of Ran and Boyce (1996) and He (1997) work with completely general choice models and are therefore the most promising. The approach taken in Chen (1999) extends the cost function with an entropy term related to the choice model, which leads to a MultiNomial Logit (MNL, see Ben-Akiva and Lerman (1989)) model of route choice, which (see chapter 7) is inappropriate in a network context with path overlap.

The recent article of Wie et al. (2002), proves the existence of equilibrium in a link-based VIP formulation. The (static) stochastic user equilibrium (SUE-STOCH) problem has been formulated by Fisk (1980) as a minimisation problem. As the SUE-STOCH can be written as an optimisation problem only when the Jacobian of the link cost functions is positive definite, this formulation is not directly applicable to the dynamic (DUE-STOCH) case.

All authors show the equivalence between the equilibrium conditions sought and the VIP formulation but few consider uniqueness; only Smith (1993) and Huang and Lam (2002) deal with the question of continuity of the solutions, and none with their differentiability. For our purpose the question of uniqueness is important as the well-definedness of the objective function of the O-D matrix estimation problem depends on it.

Existence proofs

The existence proofs of the VI formulation for DUE-DET and DUE-STOCH rely on the application of Brouwer's fixed-point theorem to an equivalent fixed-point formulation (see Nagurney (1993)). Application of Brouwer's theorem requires the path flow cost to be continuous in the path flows. These costs however depend on the inflow of traffic on a link during each time period, and therefore on the continuity of flow propagation as specified by the dynamic network loading model. We note that the O-D demand takes on fixed values during each time period, and is therefore discontinuous in time. For this reason the continuity of the link flows and therefore the path costs with respect to the path flows is not immediately obvious. For existence proofs, we refer to Bliemer (2001).

Sensitivity analysis and derivatives of the solution

To the best of our knowledge, no results are currently available on the sensitivity or derivatives of solutions to dynamic traffic assignments. Fortunately the VI formulation for the static case shares some essential properties with those of the dynamic case (which is also formulated as a VI problem) so that we can re-use or adapt the results from the static case.

A shortlist of the literature on derivatives of the UE solution in the static case is given in table 6.3, which lists the publications and whether they refer to deterministic or
stochastic assignment ("Det/Stoch"), and whether the publication contains proof of the VIP properties (equivalence with the equilibrium conditions ("Eqn."), existence ("Exist"), uniqueness ("Uniq"), continuity ("Ct"), and differentiability ("Diff"), and whether the publications pays attention to the existence ("Exist") and formulation ("Formul") of the derivatives of the solution.

<table>
<thead>
<tr>
<th>Nr</th>
<th>Publication</th>
<th>Det/ Stoch</th>
<th>Proof of VIP properties</th>
<th>Solution derives</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bell and Ilda (1997)</td>
<td>D/S</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>Tobin and Friesz (1988)</td>
<td>D</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>Patriksson and Rockafeller (2002)</td>
<td>D/S</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>Patriksson (2002)</td>
<td>D/S</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Nugurney (1993)</td>
<td>D</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>Dafermos and Nagurney (1984)</td>
<td>S</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.3: Literature on sensitivity analysis of VIP formulations of static UE

The sensitivity and stability of DUE-STOCH assignments with asymmetric Jacobian and elastic demand were investigated by Dafermos and Nagurney (1984), who establish that under certain conditions the path costs between an O-D pair cannot decrease when the demand between that O-D pair increases, and that travel costs between O-D pair 1 cannot decrease when the demand between O-D pair 2 increases. An important theorem by Tobin and Friesz (1988) gives a formula for the calculation of the sensitivity (read derivatives) of an assignment map formulated as a VIP w.r.t. the demand, the cost functions, or any other parametrisation. Later Bell and Ilda (1997) pointed out that the validity of the formula given in Tobin and Friesz (1988) depends on the network topology, and may be invalid in certain cases. Recent work by Patriksson and Rockafeller (2002) criticises the work of Tobin and Friesz (1988) and clarifies the issue. It establishes a mathematical framework from which it follows that the assignment map for deterministic assignments is not necessarily differentiable (in the classical sense), but satisfies the weaker property of directional differentiability (which is sufficient to allow the construction of minimisation algorithms). For our specific problem formulation (which uses stochastic user equilibrium exclusively) the validity of the Tobin and Friesz (1988) theorem is assured.

6.3 Distinguishing between route choice equilibrium and departure time equilibrium

The structure of the behavioural dynamic O-D matrix estimation problem, as set out in Section 3.4, Figure 3.6 specifies two behavioural choice models: the DTAC-STOCH model for combined route choice and departure time adjustment, which contains the DUE-STOCH model for dynamic assignment with stochastic route choice as a sub-model. The relationship between the two models is highlighted in Figure 6.1.
Shown in the top-left box is the Stochastic Departure Time Adjustment Choice (DTAC-STOCH) model. Input to this model is a dynamic O-D matrix $T^{\ast\ast}_{h \mid h} \subseteq T^{\ast\ast,0}_{h \mid h} \subseteq \Lambda$ with preferred departure times, which is transformed into the matrix $\hat{T}^{\ast\ast}_{h}$ through departure time adjustment, as specified by the time adjustment probabilities $P^{\ast\ast}_{h \mid h}(C^{\ast\ast}, \Theta)$ that depend on the equilibrium travel costs $C^{\ast\ast}$ resulting from the DUE-STOCH assignment. The DUE-STOCH assignment takes the matrix $\hat{T}^{\ast}_{h}$ (shown as a dotted arrow) and exports the equilibrium O-D travel costs $C^{\ast\ast}$ (shown as a dotted arrow). This is achieved by determining a consistent set of path flows $f^{\ast}_{p}(h)$, link flows $v_{a}(h)$, link travel times $\tau_{a}(h)$, path costs $C^{\ast\ast}_{p}(h)$, and route choice proportions $P^{\ast}_{p}(C^{\ast\ast}_{p}(h), \Phi)$.

The Stochastic Dynamic User Equilibrium problem in the bottom right-hand box is a DUE-STOCH assignment where demand is (temporarily) rigid (i.e. the entire matrix $\hat{T}^{\ast}_{h}$ is unaffected by the DUE-STOCH assignment) and route choice is determined. In order to obtain a joint equilibrium, equilibrium conditions must be satisfied simultaneously for the DUE-STOCH model and for the departure time adjustment model.

### 6.4 Dynamic Equilibrium Conditions

The purpose of Dynamic Traffic Assignment is to determine equilibrium traffic flow patterns in time and space. In this Section we will state the conditions for stochastic user equilibrium. With respect to Figure 3.6 p.42, we can identify two types of choice for users: route choice and departure time adjustment. By appealing to the STEN concept, we can conceptually deal with both types of choice as route choice through an extended network.

The literature distinguishes variations of network user equilibria, which can be related to the properties of the users that give rise to such equilibria. Network equilibria are examples of to the game-theoretic equilibria referred to in Section 2.4.1, and derive their
properties from those of the players involved (i.e. network users), especially with respect to their informedness about the actions of other players (fully informed players versus partly-informed players, players with perception errors), their preferences (the player's payoff function), and their behaviour (all players are assumed to be uncompromising in maximising their payoff, which consists of making a minimum-cost journey through the transport network).

As we have assumed a single-userclass assignment, all users are assumed to affect the cost (i.e. travel time increase) of the link they use to other players in the same way. Travel costs are assumed to be continuous monotonic functions of the number of users. Again because of the single-userclass assumption, users' tastes in evaluating these costs are identical, although users' perceptions may differ in a random way.

Various specific definitions and characterisations of network traffic equilibria exist (see e.g. Bliemer (2001)), which however can be seen as specific instances of games by choosing appropriate payoff functions.

6.4.1 Equilibrium conditions for users with random taste variation and perception errors: DUE-STOCH equilibrium

For users whose route cost valuation is subject to a random error term (either because we don't know their individual taste, their cost perception is flawed, their knowledge of the least-cost routes is flawed, or because they have information that is unavailable to the modeler), their route choice behavior is assumed equivalent to the following route choice condition:

For each O-D pair \((r, s)\) for any departure time \(h\), the perceived travel costs of all chosen paths are equal and minimal.

This corresponds to a Stochastic Dynamic User Equilibrium. A user's perceived travel cost, denoted as \(\overline{c}_p^r(h)\), is a random variable with a certain distribution that can be thought of as the true travel cost \(c_p^r(h)\) plus a random error term with zero expectation:

\[
\overline{c}_p^r(h) = c_p^r(h) + \varepsilon^r(h), \quad E(\varepsilon^r(h)) = 0
\]

The user equilibrium conditions can be written in mathematical form as a nonlinear complementarity problem (see Bazaara et al. (1993), and Patriksson (1994) and the references therein), with \(f_{p}^{rs}(h)\) the equilibrium path flows, \(\overline{c}_p^{rs}(h)\) the equilibrium route costs, and \(\pi^{rs}(h)\) the minimum of all perceived travel costs in equilibrium:

\[
\begin{align*}
\overline{c}_p^{rs}(h) - \pi^{rs}(h) & \geq 0 \quad \forall (r, s), \forall p \in P_{rs}, \forall h \in H \quad (6.15a) \\
f_{p}^{rs}(h) \left[ \overline{c}_p^{rs}(h) - \pi^{rs}(h) \right] & = 0 \quad \forall (r, s), \forall p \in P_{rs}, \forall h \in H \quad (6.15b) \\
f_{p}^{rs}(h) & \geq 0 \quad \forall (r, s), \forall p \in P_{rs}, \forall h \in H \quad (6.15c)
\end{align*}
\]

The probability that path \(p\) is chosen by a stochastic user among the set of available paths \(P_{rs}\) is equal to the probability that the cost \(\overline{c}_p^{rs}(h)\) of path \(p\) is perceived as
the minimum cost $\pi^{rs*}(h)$; this leads to the probabilistic choice models presented in chapter 7. If we assume that all users are homogeneous (i.e., their perceived path travel times are independent, identically distributed random variables), then for each O-D pair $(r, s)$, the flow on path $p$ is equal to product of the total O-D flow and the probability that path $p$ is chosen by users: $f_p^{rs}(h) = t_h^{rs} p_p^{rs}(h)$. As this property also holds for the equilibrium solution $f_p^{rs*}(h)$, the route choice condition can be expressed as:

$$f_p^{rs*}(h) - t_h^{rs} p_p^{rs*}(h) = 0, \quad \forall (r, s), \forall p \in P_{rs}, \forall h \in H$$

(6.16)

Although the user equilibrium conditions in eqns. 6.15 (after adding flow-propagation equations) constitute a nonlinear complementarity problem which is a solvable mathematical program in its own right, it is preferable to use the VIP formulation for DUE assignment problems because of its convenience in proofs and its direct relationship to solution algorithms. In the following sections we will give the VIP formulation for the DUE-STOCH assignment. For an account of the DUE-DET assignment, see Chen (1999) or Bliemer (2001).

6.5 VIP formulation of DUE-STOCH assignment

In this section we will give the formulation of the DUE-STOCH assignment as a VIP. The literature contains two formulations, one based on perturbed cost, the other incorporating consequence of the equilibrium situation into the VIP itself.

The formulation based on perturbed cost, which is presented in e.g. Chen (1999), has the advantage of being based on a direct extension of the Wardrop principle to perceived instead of real cost. The drawback of this formulation is that it needs further work before it is freed of the assumption that the perception errors are I.I.D. Gumbel distributed (which neglects correlation through route overlap). Examples of this for the static case are given in Bekhor et al. (2002), which however rely on the existence of potential functions of which the costs are gradients. Such gradients do not exist for the dynamic case.

We have selected a different, more general formulation based on incorporating the fact that under equilibrium conditions the path flow probabilities follow from arbitrary route-choice models, which is presented in Section 6.5.1, and gives total flexibility regarding the specification of the route-choice model. Possible model specifications can be found in chapter 7. The disadvantage is that the link with the Wardrop equilibrium principle is more difficult to see.

We will present both methods, but we will use the approach described in Section 6.5.1 for practical computations.

6.5.1 Characterisation by equilibrium conditions

Under equilibrium conditions, route flows follow the formulation in eqn. 6.17, which is essentially the one put forward in Ran and Boyce (1996), and was adapted by He (1997). This formulation is more general in the sense that it can accommodate any route choice model.
Equilibrium conditions

The equilibrium conditions are:

\[ f_p^{rs} (h) = t_h^{rs} p_p^{rs} (k) \quad (6.17) \]

Sheffi (1985) provides the link with stochastic route costs for cost minimising users. Denoting the stochastic route cost as \( c_{ph}^{rs} = c_{ph}^{rs} + \varepsilon_{ph}^{rs} \), with \( c_{ph}^{rs} \) the objective cost and \( \varepsilon_{ph}^{rs} \) a stochastic error term, the route choice probabilities are defined as

\[ p_{ph}^{rs} = P \left[ \varepsilon_{ph}^{rs} \leq \varepsilon_{qh}^{rs} \right] \quad \forall q \in \mathcal{P}^{rs} \quad (6.18) \]

a satisfaction function is defined as:

\[ \bar{c}_h^{rs} = E \left[ \min_{p \in \mathcal{P}^{rs}} c_{ph}^{rs} \right] \quad (6.19) \]

with

\[ p_{ph}^{rs} = \frac{\partial \bar{c}_h^{rs}}{\partial c_{ph}^{rs}} \quad (6.20) \]

On this base, we can define a set of route pseudo-cost functions as:

\[ G_p^{rs} (h) = \left[ f_p^{rs} (h) - t_h^{rs} p_{ph}^{rs} (C_p^{rs} (h)) \right] \frac{\partial c_{ph}^{rs}}{\partial h} \]

This pseudo-cost function penalises deviations of the actual route flows \( f_p^{rs} (k) \) from the ones that would be expected on basis of the route choice probabilities \( P_{ph}^{rs} (C_p^{rs} (h)) \), which depend on the route costs \( C_p^{rs} (h) \), which in turn depend on the route flows. In the equilibrium solution, we have \( G_p^{rs} (h) = 0 \), so that \( f_p^{rs} (h) = t_h^{rs} p_{ph}^{rs} \) or \( \frac{\partial c_{ph}^{rs}}{\partial h} = 0 \) (or both). If we also assume that the route costs vary *monotonically* with the demand:

\[ \frac{\partial c_{ph}^{rs}}{\partial h} > 0 \quad (6.21) \]

then the condition \( G_p^{rs} (h) = 0 \) ensures that \( f_p^{rs} (h) = t_h^{rs} p_{ph}^{rs} \), which implies equilibrium.

VIP formulation

The following VIP specifies equilibrium:

\[ \langle G_p^{rs} (h), (F_p^{rs} (h) - F_p^{rs*} (h)) \rangle \geq 0 \quad \forall F_p^{rs} (h) \in \Omega \quad (6.22) \]

s.t. \[ g_p^{rs} (h) = \left[ f_p^{rs} (h) - t_h^{rs} p_{ph}^{rs} (h) \right] \frac{\partial \eta_{ph}^{rs} (h)}{\partial h^{rs}} \]

ensures equilibrium, because the pseudo-cost functions will all be driven to zero in the optimum.
As the amount of traffic is finite, the set of all permissible path flows is bounded, and is a polytope (see Patriksson (2002)). Let

\[ \mathcal{F} = \{ f \mid T = BF, F \geq 0 \} \]  

be the set of feasible path flows.

Noting that we have a formulation with stochastic route choice and enumerated routes, the RUM-based route choice model ensures that every path has a non-zero flow. Therefore any solution to the VIP will be in the interior (and not on the boundary) of the set of permissible path flows \( F \), and the function \( G \) will take on the value 0 at the equilibrium point (see Nagurney (1993), corollary 1.3 (provided one exists)).

### 6.5.2 Path costs for the VIP formulation

The paths costs \( c^p s (h) \) in the STEN (see Figure 6.2) consist of three terms: cost of departure time shift, travel cost, and schedule-delay cost.

![Figure 6.2: Cost components associated with STEN links](image)

The total path costs are given in eqn. 6.24:

\[ c^p s (h) = \beta DT c^p DT (h) + \beta TT c^p TT, p (h) + \beta AT c^p AT (h + \tau^p s (h)) \]  

(6.24)
with:

- $c_{DT}^{\beta} (h)$: the departure scheduling costs associated with the departure time $h$ and $\beta_{DT}$ the weight factor for departure time

- $c_{T_{T,p}}^{\beta} (h)$: the travel cost along path $p$ when starting during time interval $h$ with $\beta_{TT}$ incurred on the physical network

- $c_{AT}^{\beta} (h + \tau_{p}^{rs} (h))$: the arrival schedule delay costs of arriving at time $h + \tau_{p}^{rs} (h)$ with $\beta_{AT}$ the weight factor for schedule delay

- $\tau_{T_{T,p}}^{rs} (h)$: the travel time along path $p$ when starting during time interval $h$ with $\beta_{TT}$

The departure time costs, the travel time disutility and the schedule delays costs and the models for choosing between joint departure time and path alternatives $(h, p)$ are set out in chapter 7.

We will now examine the path cost components one by one and inject assumptions where needed to obtain the desired result.

Path travel time The path travel time on path $p$ from $r$ to $s$ starting during time interval $h$: $\tau_{p}^{rs} (h)$ is the sum of the link travel times $\tau_{a} (h)$. Defining the travel time on path $p$ starting during period $k$ from $r$ to $s$ until link $a$ is reached as the with $\tau_{p}^{rs} (h)$, and denoting the precursor of link $a$ on path $p$ by $a^{-}$, we have the following recursive definition of $\tau_{p}^{rs} (h)$:

$$\tau_{p}^{rs} (h) = \tau_{pa^{-}}^{rs} (h) + \tau_{a^{-}} (\tau_{pa^{-}}^{rs} (h))$$

which gives the total path travel time as the travel time up to the last link on the path $(a)$ plus the travel time on link $a$ at the time of arrival:

$$\tau_{p}^{rs} (h) = \tau_{pa}^{rs} (h) + \tau_{a} (\tau_{pa}^{rs} (h))$$ \hspace{1cm} (6.25)

Link travel time on link $a$ for traffic entering this link during period $k$ depends through a function $D_{a} (X_{ak})$ on the amount of traffic $X_{ak} = X_{a} (k)$ on the link at the time:

$$\tau_{ak} = \tau_{a} (k) = D_{a} (X_{a} (k))$$ \hspace{1cm} (6.26)

In order to have a path travel time function that is monotone w.r.t. the amount of traffic that selects the path, we assume that the link travel time function $D_{a} (X_{ak})$ is continuously differentiable, and strictly monotone. The strict monotony might be questioned in the case of very light traffic, but from an engineering point of view there is no objection to having strictly monotonously increasing link cost functions whose derivative is kept close to zero for the uncongested part. In this way we can satisfy the formal mathematical requirements of strict monotonicity while still retaining engineering plausibility. Therefore we have the

**Assumption** $\frac{dD_{a}(X_{ak})}{dX_{ak}} > 0 \quad \forall X_{ak}$
**Schedule delay costs** The schedule delay costs depend on the arrival time at the destination, which is the sum of the departure time and travel time between O-D pairs. If we assume a network equilibrium, the O-D travel times are equal for all used paths. Although we have associated the path schedule delay costs with the links connecting the STEN with the physical destination, the link cost does not depend on the flow over that particular link, but rather on the entire flow vector which causes the delays in the network. From a VI point of view this is legitimate, as the link costs are again a function of the flow vector, albeit a non-separable one.

**Departure time scheduling costs** The departure time scheduling costs are associated with the connected network linking the preferred departure times \( r'(h') \) with the realised departure times \( r(h) \). These costs depend neither on the flow through the artificial links between \( r'(h') \) and \( r(h) \) nor on the travel through the network, but solely on the difference between the preferred departure time and the departure time that gives rise to the minimum path least cost in terms of travel time and schedule delay (i.e. \( \beta_{TT}T_{TT,p}^{r_{TT}}(h) + \beta_{AT}C_{AT}^{r_{AT}}(h + \tau_{p}^{r}(h)) \)) that is associated with the choice of departure time \( h \).

**Path cost properties**

Certain properties of the cost functions are crucial for the existence and uniqueness of the solution of VIP’s. These properties are:

- continuity of path costs w.r.t. path flows
- differentiability of path costs w.r.t. path flows
- monotonicity of path costs w.r.t. path flows

We can therefore expect these properties of path cost functions for the DTA (flows, travel times), to have the same influence on the existence of derivatives of the flows generated by the DTA w.r.t. the cells of the dynamic O-D matrix.

We note that since we are working with a discrete-time version of the DTA, we need to verify the conditions for each time period separately, and that the discrete nature of the DTA does not in itself contradict the continuity and differentiability. We also note that our derivation in this respect needs to be supplemented by assumptions, which we justify by considering that from a transportation engineering point of view we may expect path travel times to vary more smoothly with the path flows and since we are using a stochastic assignment, therefore also with the O-D demand.

### 6.6 Dynamic Network Loading

There is need to consider the Dynamic Network Loading (DNL) function (see Figure 6.3), for two reasons. Firstly, the DNL it affects the smoothness, monotonicity of the
cost functions and the traffic flow, therefore the properties of the DUE-STOCH solution and hence the objective function of the O-D matrix estimation problem. For example the DNL governs the convexity of the set of permissible link-flows. Secondly, for assignment fractions (needed to disentangle group observations such as traffic counts) to be extracted from the DTA the dynamic link path incidence matrix must be extracted from the DNL.

On closer inspection, obtaining a fast, correct and realistic DNL procedure turns out to be a complex undertaking (see e.g. Chabini and Kachani (2001)).

![Influence diagram for the discrete-time DTA algorithm](image)

**Figure 6.3:** Influence diagram for the discrete-time DTA algorithm

The relationship between the dynamic O-D matrix \( T^*_k \), the route choice probabilities \( P^r_{pk} \), the path inflows \( f^r_{pk} \) that result from these, the link flows \( u_{ak} \), the link travel times \( \tau_{ak} \), and the path travel times \( \tau^*_k \), and the dynamic link-path incidence matrix \( \delta^r_{pk} \) that results from the discrete-time dynamic assignment is shown in Figure 6.3.

The function of the DTA algorithm is to find a self-consistent set of path flows and path travel times given a relationship between link traffic load and link travel times. The function of the DNL algorithm is to determine a self-consistent set of values for the dynamic link-path incidence matrix \( \delta^r_{pk} \), the link travel time, and the link inflows.

### 6.6.1 The discrete-time Dynamic Network Loading model

We will use the discrete-time Dynamic Network Loading (DNL) model proposed in Chabini (2001). It can be seen as a discretised version of the continuous-time DNL.

Where in the static case link flows are simply combinations of path flows, in the dynamic case we have time-dependent link flows which depend on paths flows at very specific time.
instances. Moreover those precise time instances depend on the travel times through the network, and hence on the link flows and the path flows. As a matter of interest, the DNL is the central factor that distinguishes deterministic dynamic (DUE-DET) from deterministic static (SUE-DET) equilibrium assignment (see also Bliemer (2001)).

Another aspect of the DNL is that it determines the continuity, monotonicity and differentiability (or lack thereof) of the route- and link cost functions. This is important, as the existence of solutions to the SUE-DET and DUE-DET assignments hinges on the continuity of the cost functions. Furthermore the differentiability of the flows and travel times in and near equilibrium depend on the differentiability of the link- and path cost functions.

The DNL adopted in the dynamic O-D matrix estimation software was taken from Chabini (2001) and He (1997), and is reproduced here in order to illustrate the DNL algorithm itself, its shortcomings for our purposes, and the way in which these shortcomings can be overcome.

The DNL uses discrete timesteps of length $\delta$, with

$$\delta = \frac{\Delta}{M}, \quad M \in \mathbb{Z}^{+}$$

with $\Delta$ the minimum free-flow time of all links in the network:

$$\Delta = \min_{a \in A} \tau_{a}^{ff}$$

The relationship between $\delta$ and $\Delta$ is shown in Figure 6.4.

$$\delta \leq \min_{a \in A} \tau_{a}^{ff}$$

The time-dependent link inflows $u_{ap}^{s}(t)$ are discretised:

$$u_{ap}^{s}(t) = u_{ap}^{s}(k\delta) \quad \forall t \in [k\delta, (k+1)\delta] \quad \forall k \in \mathbb{Z}, \forall (r,s), \forall p \in P_{rs}, \forall a \in A$$

The amount of traffic on a link at time instant $k\delta$: $X_{ap}^{s}(k\delta)$ is the difference between the cumulative inflow and the cumulative outflow at that time, given by:
\[ X_{ap}^r(k\delta) = U_{ap}^r(k\delta) - V_{ap}^r(k\delta) \quad (6.30) \]

The inflow rate on a link is the path flow for the first link on a path, and the outflow of its upstream neighbor for all other links:

\[ u_{ap}^r(k\delta) = \begin{cases} 
 f_p^r(k\delta) & \text{if } a \text{ is first link on path } p \\
 v_{ap}^r(k\delta) & \text{if } a \text{ follows directly after } a^p 
\end{cases} \quad \forall k, \forall (r, s), \forall p \in P_{rs} \quad (6.31) \]

The cumulative outflow during period \( k \) is made up of the product of inflow rates and interval durations of all time periods \( l \) that exit no later than period \( k \). This means that the time of entry plus the travel time at the time of entry does not exceed \( k \):

\[ l\delta + \tau_a(l\delta) \leq (k - 1)\delta \]

This leads to:

\[ V_{ap}^r(k\delta) = \sum_{j \in J_a(k)} u_{ap}^r(j\delta) \delta \quad \forall k, \forall (r, s), \forall p \in P_{rs}, \forall a \in A \quad (6.32) \]

With the index set \( J_a(k) \): defined as:

\[ J_a(k) = \{ l | 0 \leq l\delta + \tau_a(l\delta) \leq (k - 1)\delta \} \quad (6.33) \]

The link travel time during time period \( k\delta \) is given by \( \tau_a(k\delta) \), which is related to the total traffic flow on the link \( X_a(k\delta) \) through a link-performance function defined by \( D_a(X_a) \).

\[ \tau_a(k\delta) = D_a(X_a(k\delta)) \quad (6.34) \]

The total amount of traffic on link \( a \) is the sum of the path-flow related amounts of traffic:

\[ X_a(k\delta) = \sum_{r \in P} X_{ap}^r(k\delta) \quad (6.35) \]

\[ U_{ap}^r(0) = 0, \quad V_{ap}^r(0) = 0, \quad X_{ap}^r(0) = 0 \quad \forall (r, s), \forall p \in P_{rs}, \forall a \in A \quad (6.36) \]

The link outflow rate, formally the derivative of the cumulative link outflow rate, is:

\[ v_{ap}^r(k\delta) = \frac{V_{ap}^r(k\delta) - V_{ap}^r((k-1)\delta)}{\delta} \quad \forall k, \forall (r, s), \forall p \in P_{rs}, \forall a \in A \quad (6.37) \]

Similarly for the cumulative inflow:

\[ U_{ap}^r(k\delta) = \sum_{j \in 0}^k u_{ap}^r(j\delta) \delta \quad (6.38) \]
6.6.2 Algorithm for discrete-time dynamic network loading

The algorithm used for network loading is called the C-load (for "Chronological load") algorithm given in Chabini and He (1998).

Note that the algorithm is a one-to-one implementation of the model defined in (6.27) - (6.38), and that no dynamic assignment fractions are calculated. Also the algorithm does not calculate the derivatives of the flow with respect to O-D-t matrix cells required for the optimisation of the upper level.

Influence diagram for the traffic state variables

The influence diagram for the DNL algorithm is shown in Figure 6.5.

![Influence diagram for the C-load discrete-time DNL algorithm](image)

Figure 6.5: Influence diagram for the C-load discrete-time DNL algorithm

The network loading is driven by the path flows \( f_{pk}^{rs} \) that are generated by the DTA. When combined with the time-dynamic link-path incidence matrix \( \delta_{apk}^{rs} \) the dynamic path-specific link inflows linked \( u_{apk}^{rs} \) to that particular path flow on link \( a \) result.

The path-specific link inflows are then used to determine path-specific cumulative link outflows \( V_{apk}^{rs} \). Taking the difference quotient of the path-specific cumulative link outflows results in the path-specific link inflow rates \( u_{apk}^{rs} \) for the downstream link, which in turn are used to calculate path-specific cumulative link inflows \( V_{apk}^{rs} \). The current path-specific traffic load on a link \( X_{apk}^{rs} \) is then calculated as the difference between the cumulative path-specific link inflow and the cumulative path-specific link outflow. By summing these over all paths and O-D pairs the total traffic load \( X_{ak} \) on a link is obtained, which is used to calculate the current link travel time \( \tau_{ak} \). Finally the link-travel-times are used to determine the time-dynamic link-path incidence matrix.

The C-load algorithm solves the interdependencies in the DNL problem in a single pass.
6.6.3 Dynamic assignment fractions

As will be shown in Section 6.7.2, assignment fractions are needed for the analytical calculation of derivatives of the DTA.

The DTA itself can be used to calculate these assignment fractions as shown below. By combining (6.37) and (6.32) we get the following relationship for link \(a\):

\[
v_{ap}^{r_s} (k \delta) = \sum_{j \in J_a (k)} u_{ap}^{r_s} (j \delta) \delta - \sum_{j \in J_a (k-1)} u_{ap}^{r_s} (j \delta) \delta
\]

(6.39)

\[
= \sum_{j \in J_a (k) - J_a (k-1)} u_{ap}^{r_s} (j \delta) \delta
\]

(6.40)

\[
= \sum_{j \in J_a^* (k)} u_{ap}^{r_s} (j \delta) \delta
\]

with

\[
J_a^* (k) = \{ j | (k-1) \delta < j \delta + \tau_a (j) \leq k \delta \}
\]

(6.41)

From (6.31) we can work back along the links of the path by using the outflow of the upstream link as the inflow of the current link, and apply 6.39 to calculate outflow from inflow. Eventually we will encounter the beginning of the path, see that the path-specific linkflows for this DTA procedure can also be written as a weighted sum of path entry flows, which in turn are directly related to the route choice fractions and the O-D flows:

\[
u_{ap}^{r_s} (k \delta) = \sum_{h \in J_h} \varphi_{aphk}^{r_s} T_{ph}^{r_s}
\]

(6.42)

\[
= \sum_{h \in J_h} \delta_{aphk}^{r_s} T_{ph}^{r_s}
\]

(6.43)

with \(J_h\) the set of all cells that "touch" link \(a\) during period \(k\).

Note that in the model formulation in (6.27) - (6.38) the composition of the link exit flow \(u_{ap}^{r_s} (k \delta)\) in terms of O-D-\(t\) matrix cells is not specified. In other words, dynamic assignment fractions \(\varphi_{aphk}^{r_s}\) are not explicitly calculated. Indeed this is to be avoided in view of the high dimensionality of the array \(\varphi_{aphk}^{r_s}\). However it suffices that the assignment fractions can be reconstructed from the loaded network that results after the DTA algorithm has finished because by then all link travel times are known. The assignment fractions can be calculated as needed in order to derive the required derivatives.

Given that it needs to be carried out only after a solution to the DNL problem has been obtained and after the solution to the VI problem governing the route choice part has stabilised, we decided to redo part of the calculations of the C-load algorithm and to separate the calculation of assignment fractions from the C-load algorithm.
6.7 Derivatives of traffic flow with respect to dynamic O-D flows

As noted in chapter 5, the dynamic O-D matrix estimation problem needs certain input from its lower level problem (the DTA) regarding the reaction of the lower level to changes in the settings of the variables controlled by the upper level. This is difficult to realise, but one of the proxies for this information is represented by the derivatives of link flows w.r.t. dynamic O-D flows. The reason is that the derivatives can be used to check for optimality, and to help find descent directions for the O-D matrix estimation problem.

The general objective is to determine the derivatives of link flow with respect to matrix cells: \( \frac{\partial o_d}{\partial u_h} \). In view of the fact that we deal with stochastic assignments so that the differentiability issues are largely mitigated through the influence of the (fully differentiable) route choice, we will generally assume that the derivatives exist despite the possible pitfalls of this approach highlighted in Patriksson and Rockafeller (2002).

6.7.1 Calculating the derivatives

In this section we present a way of calculating the derivatives of link flow in a dynamic traffic assignment setting with respect to changes in the dynamic Origin-Destination matrix. It turns out that our specific case (stochastic equilibrium assignment with enumerated paths) is one in which the derivatives can be calculated using the result of Tobin and Friesz (1988), that cost functions are continuous and differentiable.

In the following we will assume that the DNL has the properties needed to ensure that the cost functions are differentiable.

The sensitivity analysis of Tobin and Friesz

The derivatives we need to calculate follow from the sensitivity analysis presented in Tobin (1986) and Tobin and Friesz (1988). The sensitivity results in Tobin (1986) hold for general VIP’s. and are therefore applicable to our problem.

As noted in Bell and Ilida (1997), the analysis in Tobin and Friesz (1988) uses that the Jacobian of the cost functions w.r.t. path flows is related to the one based on link flows:

\[
\nabla^2 C(F) = \Delta^T \nabla^2 C(\mathbf{u}) \Delta
\]

(6.44)

with \( F \) the path flow vector, \( \Delta \) the link-path incidence matrix, and \( \mathbf{u} \) the link flow vector.

If \( \mathcal{F}^* = \{ F | \mathbf{u}^* = \Delta F, T = BF, F \geq 0 \} \) is the set of path flows that reproduce the (unique) equilibrium path flows pattern \( F^* \), with \( B \) the O-D pair - path incidence matrix. For stochastic assignments the path flows are uniquely determined: \( \mathcal{F}^* = \{ F^* \} \), but for deterministic assignments \( \mathcal{F}^* \) generally has more then one element.

The well-posedness of the derivative hinges on there not being two different sets of paths flows that give rise to the same set of link flows. In general this is not the case,
and the analysis shows that any minimal set of path flows (i.e., after elimination of path flows that are linearly dependent) gives the required sensitivities.

As noted in Bell and Iida (1997), the analysis presented in Tobin and Friesz (1988) hinges on the invertibility of a matrix of the form $\Delta_u \nabla C(\Omega) \Delta_u$ with $\Delta_u$ a matrix of path flows that does not contain dependent path flow vectors, and $\nabla C(\Omega)$ the gradient of the cost functions. For large networks and for deterministic assignments, this matrix is not invertible. For stochastic assignments however the invertibility of the matrix is guaranteed.

An analysis of the existence and calculation of derivatives with a view to obtaining a solution algorithm for MPEC’s is presented in Patriksson and Rockafeller (2002).

### 6.7.2 Direct calculation of derivatives

Calculation of the derivatives of link flows with respect to changes in the matrix cells is complicated by two facts:

1. derivatives of upstream links w.r.t. matrix cells will impact those of downstream links

2. the DTA and DNL feedback loops co-determine the net value of derivatives w.r.t. O-D-t matrix cells $T_{rs}^h$

In this section we will investigate the propagation of the derivatives through the DNL procedure using the influence diagram of the DNL procedure. The influence diagram of the DTA-DNL algorithm is shown in Figure (6.6).

![Influence diagram for DTA-DNL algorithm](image)

Figure 6.6: Influence diagram for DTA-DNL algorithm
Note that the relationship between $v_{ak}^r$ and $\delta_{apk}^r$ and $u_{ak}^r$ is shown as a dotted line, signifying that a "pro-forma" relationship: the way the DNL algorithm calculates $v_{ak}^r$ from $u_{ak}^r$ is through the loop at the bottom.

The derivative of the link flow can be decomposed in terms of derivatives of path-dependent link flows as:

$$\frac{\partial v_a(k)}{\partial T_h^s} = \sum_{rs} \sum_{p} \frac{\partial v_{ap}(k)}{\partial T_h^rs}$$

The derivative of the path-specific link flows can be decomposed into assignment fractions and matrix cells as:

$$\frac{\partial v_{ap}(k)}{\partial T_h^rs} = \sum_{h \in J_h} \frac{\partial (\delta_{apkh}^rT_h^rs)}{\partial T_h^rs} = \sum_{h \in J_h} \frac{\partial (\delta_{apkh}T_{ph}^rT_h^rs)}{\partial T_h^rs}$$

This derivative is:

$$\frac{\partial (\varphi_{apkh}^rP_{ph}^rT_h^rs)}{\partial T_h^rs} = \frac{\partial \delta_{apkh}^rP_{ph}^rT_h^rs}{\partial T_h^rs} + \delta_{apkh}^r \frac{\partial (P_{ph}^rT_h^rs)}{\partial T_h^rs}$$

$$\frac{\partial \delta_{apkh}^rP_{ph}^rT_h^rs}{\partial T_h^rs} + \delta_{apkh}^r \left[ \frac{P_{ph}^rT_h^rs}{\partial T_h^rs} + T_h^rs \frac{\partial P_{ph}^r}{\partial T_h^rs} \right]$$

$$\frac{\partial \delta_{apkh}^rP_{ph}^rT_h^rs}{\partial T_h^rs} + \delta_{apkh}^r \frac{\partial P_{ph}^r}{\partial T_h^rs} + \frac{\partial T_h^rs}{\partial T_h^rs} \delta_{apkh}^rP_{ph}^rT_h^rs$$

$$\varphi_{apkh}^r + \frac{\partial \delta_{apkh}^rP_{ph}^rT_h^rs}{\partial T_h^rs} + \frac{\partial P_{ph}^r}{\partial T_h^rs} \delta_{apkh}^rT_h^rs$$

This expression for the derivatives could be used, but has some drawbacks. Firstly the last two terms are cumbersome to calculate, and require considerable amounts of storage. Secondly the dynamic link-path incidence matrix and the route choice probabilities depend on the O-D matrix cells through feedback loops, i.e. indirectly. In principle the derivatives of the travel times with respect to demand involve the reaction of the system (travel time, as represented by the DNL procedure) and the reaction of the users (represented by the DTA problem, specifically its route choice component). In either case the causal chain that affects $\frac{\partial v_{ap}(k)}{\partial T_h^rs}$ can be read off from Figure 6.6.

According to the theory of bi-level problems the derivatives $\frac{\partial v_{ap}(k)}{\partial T_h^rs}$ that do not take account of the feedback mechanisms represented by the DNL and DTA are meaningful, and can be represented as selecting a game-theoretic formulation of the bi-level problem known as a Cournot game. When the feedback loops (and the corresponding fixed-point
problems) are taken into consideration, the resulting derivatives correspond to a game-theoretic formulation of the bi-level problem known as a Stackelberg game (see Basar and Olsder (1999)).

Therefore we will approximate the derivatives using only the first term, neglecting the last two terms.

### 6.7.3 Approximations through truncated derivatives

We will now list the approximations.

**Approximation 1**

A first approximation of the derivative 6.47 can be found by neglecting the terms $\frac{\partial \bar{r}_{ph}}{\partial T_h} P_{ph} T_h + \frac{\partial \bar{r}_{ph}}{\partial \delta_{ph}} \delta_{ph} T_h$ in eqn. 6.47 representing the reactions of the assignment and route choice to the flows, and work with the product of the dynamic link-path incidence matrix and the route choice fractions $\delta_{ph} P_{ph}$. This would correspond to a dynamic proportional assignment. The corresponding causal diagram is shown in Figure (6.7):

![Causal diagram](image)

*Figure 6.7: Influence diagram when the DNL and the DTA are not considered

The derivative then becomes:
\[ \frac{\partial \nu_a(k\delta)}{\partial T_h^{rs}} = \sum_{rs} \sum_p \frac{\partial \nu_{ap}^{rs}(k\delta)}{\partial T_h^{rs}} \]  
(6.48)

\[ = \sum_{rs} \sum_p \sum_{h' \in J_h} \frac{\partial \nu_{aphk}^{rs} \cdot p_{ph}^{rs} T_h^{rs}}{\partial T_h^{rs}} \]  
(6.49)

\[ = \sum_{rs} \sum_p \sum_{h \in J_h} \nu_{aphk}^{rs} \]  
(6.50)

and the derivatives are essentially proportional to the assignment fractions.

This approximation of the derivatives was implemented.

Approximation 2

As a second approximation (which we have not yet implemented) one can incorporate the reaction of the DNL to changes in the matrix, while neglecting route-changes. The upper level problem would then be aware of reactions of the DNL, but not about re-routing behaviour.

The influence diagram in Figure (6.8), shows that only the straight-line influence of \( T_h^{rs} \) on \( \nu_{ak} \) and the DNL feedback loop are taken into account.

---

Figure 6.8: Influence diagram when only the effect of the DNL is taken into consideration
Note that both the straight-line influence and the effect of the DNL affect the outflow through the \( u_{apk}^{rs} \). The derivative then is calculated as:

\[
\frac{\partial u_{apk}^{rs} (k\delta)}{\partial T_{h}^{rs}} = \sum_{rs} \sum_{p} \sum_{j \in J} \delta_{appk}^{rs} P_{ph}^{rs} + \frac{\partial \delta_{appk}^{rs} \delta_{appk}^{rs} T_{h}^{rs}}{\partial T_{h}^{rs}} \left[ \delta_{appk}^{rs} P_{ph}^{rs} + P_{ph}^{rs} T_{h}^{rs} \right]
\]

The first term was calculated above. The second term can be calculated as:

\[
\frac{\partial r_{p}^{rs} (h)}{\partial T_{h}^{rs}} = \frac{\partial r_{ak} \partial X_{ak} \partial X_{apk}^{rs}}{\partial X_{ak} \partial X_{apk}^{rs}} \frac{\partial X_{apk}^{rs}}{\partial T_{h}^{rs}} \left( r_{rs} \sum_{p} \frac{\partial X_{apk}^{rs}}{\partial T_{h}^{rs}} \frac{\partial X_{apk}^{rs}}{\partial T_{h}^{rs}} \right)
\]

\[
= D' (X_{ak}) \cdot \frac{\partial X_{apk}^{rs}}{\partial T_{h}^{rs}} \left( r_{rs} \sum_{p} \frac{\partial X_{apk}^{rs}}{\partial T_{h}^{rs}} \frac{\partial X_{apk}^{rs}}{\partial T_{h}^{rs}} \right)
\]

\[
= D' (X_{ak}) \cdot \frac{\partial X_{apk}^{rs}}{\partial T_{h}^{rs}} \left( r_{rs} \sum_{p} \frac{\partial X_{apk}^{rs}}{\partial T_{h}^{rs}} \frac{\partial X_{apk}^{rs}}{\partial T_{h}^{rs}} \right) \left[ V_{apk}^{rs} - U_{apk}^{rs} \right]
\]

Combined this gives:

\[
\frac{\partial u_{apk}^{rs} (k\delta) \partial r_{p}^{rs} (h)}{\partial T_{h}^{rs}} = P_{apk}^{rs} (h\delta) + \sum_{a \in \text{path}} \left( \frac{\partial r_{ak} \partial X_{ak} \partial X_{apk}^{rs}}{\partial X_{ak} \partial X_{apk}^{rs}} \frac{\partial X_{apk}^{rs}}{\partial T_{h}^{rs}} \right)
\]

\[
= \sum_{rs} \sum_{p} \sum_{j \in J} \varphi_{apkh}^{rs} (j\delta) \left[ P_{apk}^{rs} (h\delta) + \sum_{a \in \text{path}} \left( \frac{\partial r_{ak} \partial X_{ak} \partial X_{apk}^{rs}}{\partial X_{ak} \partial X_{apk}^{rs}} \frac{\partial X_{apk}^{rs}}{\partial T_{h}^{rs}} \right) \right]
\]

This approximation of the derivatives has not been implemented.

**Approximation 3**

As a third approximation (also not implemented) one could incorporate the effect of re-routing. The influence diagram is shown in Figure (6.9).
Figure 6.9: Influence diagram when the effects of both DNL and DTA are accounted for

In this situation both feedback loops are taken into consideration. Note that we now have the straight-line influence, and the effect through the DNL that affect the outflow through the \( t_{apk}^{rs} \), and the effect through the DTA, which affects the outflow through the intermediary of \( f_{pk}^{rs} \), which involves the route choice since \( f_{pk}^{rs} = P_{pk}^{rs}T_h^{rs} \).

**The path choice derivative** The derivative of the path choice is:

\[
\frac{dP_{p}^{rs}(h)}{dT_{k'}^{rs}} = \frac{dP_{p}^{rs}(c_{p}^{rs}(h))}{dT_{k'}^{rs}} = \frac{\partial P_{p}^{rs}}{\partial c_{h}^{rs}} \frac{dc_{h}^{rs}(h)}{dT_{k'}^{rs}}
\]

\[
= \frac{\partial P_{p}^{rs}}{\partial c_{h}^{rs}} \frac{\partial c_{p}^{rs}(h)}{\partial d_{p}^{rs}(h)} \frac{dr_{p}^{rs}(h)}{dT_{k'}^{rs}}
\]

This means that we need the derivative \( \frac{dr_{p}^{rs}(h)}{dT_{k'}^{rs}} \) of the travel time for each path in the choice set of every single matrix cell \( t_{ap}^{rs} \). As \( \frac{dr_{p}^{rs}(h)}{dT_{k'}^{rs}} = \sum_{a \in p, k \in K} \delta_{apk}^{rs} \frac{dr_{apk}^{rs}(k)}{dT_{k'}^{rs}} \), we need to calculate \( \frac{dr_{apk}^{rs}(k)}{dT_{k'}^{rs}} \) for every link in path \( p \).

This approximation of the derivatives was not yet implemented.
6.8 Combined path choice and departure time choice

In this section we will focus on the extension of the DTA procedure to include departure time adjustment: the combined Departure Time Choice and Dynamic Stochastic User Equilibrium (DTC-DUE-STOCH). In Section 3.8, we noted that we may approximate the combined DTC-DTA problem as an assignment on an extended network (a STEN; see p.50). Unfortunately in this approach it is not clear how to provide convergence proofs, and the articles that refer to the application of this approach: Lindveld and Van Der Zijpp (2000) and Van Der Zijpp and Lindveld (2000)) do not contain such proofs.

Convergence proofs exist in the framework of VIP formulation of traffic assignments, but in order to re-use these proofs for our formulation we need to formulate our DTC-DTA model as a single VI problem, prove existence and uniqueness, and use the results to show that the STEN approach also has a unique solution. We will exclude within-trip re-routing: all travellers remain committed to the path that they select at departure.

The equilibrium conditions for route choice are:

\[ f_p^r (h) = t_p^r t_p^r (h) \]  \hspace{1cm} (6.57)

and those for departure time choice are:

\[ t_h^r = \sum_{h} t_h^r,0 t_h^r \]  \hspace{1cm} (6.58)

The VIP formulation is given in terms of the vectors \( F_p^r (h), G_{ph}^r, J_h^r \) and \( T_h^r \) (with elements \( f_p^r (h), g_p^r, j_h^r \) and \( t_h^r \)):

\[ \langle G_{ph}^r, (F_p^r (h) - F_p^{rs} (h)) \rangle + \langle J_h^r, (T_h^r - T_h^{rs} (h)) \rangle \geq 0 \quad \forall v \in \Omega \]  \hspace{1cm} (6.59)

\[ s.t. \quad g_p^r (h) = \left[ f_p^r (h) - t_p^r t_p^r (h) \right] \frac{\partial \eta_p^r (h)}{\partial T_h^r} \]

\[ j_p^r (h) = \left[ t_h^r - t_h^r,0 (k') t_h^r (k) \right] \frac{\partial \xi_p^r (h)}{\partial T_h^r} \]

\[ \frac{\partial \eta_p^r (h)}{\partial T_h^r} > 0 \]  \hspace{1cm} (6.60)

\[ \frac{\partial \xi_p^r (h)}{\partial T_h^r} > 0 \]  \hspace{1cm} (6.61)
6.8.1 Solution algorithm for the combined path choice and departure time choice assignment

We solved the VIP formulation in eqn. 6.59 using the MSA heuristic in a way that reflects the model structure put forward in Figure 3.6 (pp. 42) as shown in Figure 6.1 and Figure 6.1 on p. 84).

\[ T_i = T_i \]
\[ i = 0 \]
\[ C_i = DSUE(T_i) \]

Departure-time adjustment
\[ T_i' = P_i T_i \]
\[ \lambda = \Delta(i) \]
\[ T_i = \lambda T_i + (1 - \lambda) T_i' \]

\[ |T_i - T_i'| < \varepsilon \]

STOP

Figure 6.10: MSA solution algorithm for the combined DTA / DTC equilibrium problem

Figure 6.10 shows the MSA iteration process: on top is the original O-D-t matrix $T_{h^{*},0}$ enters the loop and is assigned to the variable $T_i$, and $i$ is set to 0. Next the matrix of generalised travel costs $C_i$ is calculated from the current matrix $T_i$ on basis of an DUE-STOCH assignment (see Section 6.5). The matrix of departure time change probabilities $P_i$ is calculated on basis of these costs, and applied to the current matrix $T_i$ to give the adjusted matrix $T_i'$. The mixing factor for this iteration $\lambda_i$ is calculated, and a revised version of the current matrix $T_i$ is calculated from the previous version $T_i$ and the adjusted matrix $T_i'$. The iterative process stops when the revised and adjusted matrices no longer show significant differences.
6.9 Miscellaneous

In this section we make two observations which, although relevant, are complementary to the line of the argument in the previous sections. They concern the role of STEN's, and some aspects of route enumeration for stochastic assignment models.

6.9.1 DUE traffic assignment and STENs

In chapter 3 it was illustrated that an augmented STEN can be used to simultaneously model route and departure time choice. The importance of STEN's is that they can be used to obtain a discrete-time approximation of DTA (see Chen (1999)). The STEN may be thought of as a graph $G$, the structure of which is determined by the link travel times $t$.

If spillback effects are ignored, DUE-STOCH-DTAC-STOCH assignment can be viewed as an SUE-STOCH assignment on a STEN whose link travel times are consistent with the delay that is caused by the assigned flow. Unfortunately however, the STEN itself is determined by the link travel times and therefore a change in travel times results in a different STEN; the STEN is one of the unknowns.

If the STEN $G$ is fixed, the link flows $q$ and link travel times $\tau$ are uniquely implied and can be computed by performing a SUE-STOCH assignment of the travel demand $T^*$ to this STEN using a Frank-Wolfe iterative scheme. During this computation it is important to keep track of the assignment map $\Phi$, that describes the relationship between the dynamic O-D demand and the link flows, because this matrix will be needed during the O-D estimation step. The link travel times again imply the structure of the STEN. Hence $G = G(\tau)$. A solution of the DUE-DET-DTC-STOCH assignment problem is found if a STEN $G$ is found that is consistent with its user optimal link flows. In search for such a solution we use the iterative procedure introduced by Chen (1999) consisting of an inner loop that solves a UE assignment problem and an outer loop in which the STEN is updated. The relevant outputs of this UE assignment step are an auxiliary vector of link travel times ($\tau^*$) and the assignment map (A). As long as the STEN implied by the auxiliary vector ($G(\tau^*)$) differs from the STEN that was used in the latest UE assignment ($G(\tau)$), the travel time vector $\tau$ is changed to a convex combination of $\tau$ and $\tau^*$, and the UE assignment is repeated. A novelty relative to Chen (1999) is the determination of the stepsize $\delta$. Instead of defining $\delta$ as a decreasing function of the iteration number, we choose the minimum value of $\delta$ that implies a change in the STEN. This is a shortcut relative to the original iterative scheme. This shortcut is intended to prevent the procedure from recomputing UE assignments for the same STEN over and over again.

Although we note that a STEN based approach towards dynamic assignment is possible in principle, we feel that it is more suited as a conceptual tool than as a numerical tool and will not use it in actual calculations.
6.9.2 Path enumeration for stochastic assignment models

A property of stochastic route choice models is that in principle (depending on whether the probability distribution has finite or infinite support) they will assign a non-zero flow to any path connecting an origin and a destination. There exist many possible paths which are quite improbable but which formally would have to be processed by a stochastic model. For these reasons we will exclusively work with fixed sets of enumerated routes per O-D pair, say $\mathcal{P}_{rs}$. The set of routes for the whole network would simply be the union of these route sets $\mathcal{R} = \bigcup_{rs} \mathcal{P}_{rs}$. This has three significant consequences.

The first consequence is that apart from restricting the route set, we replace our network $G(A,N)$ with a sub-network $G^*(A^*,N^*)$ with $A^* \subseteq A$, and $N^* \subseteq N$ that contains all origin and destination nodes of $G$, but contains only those arcs and intermediate nodes that are on at least one of the routes in $\mathcal{R}$. The construction of $\mathcal{R}$ can be undertaken along the lines set out in Catalano et al. (2001), .

The second consequence is that whilst the DTA will find a DUE-STOCH solution $f^*$ on $G^*(A^*,N^*)$, this solution need not be an equilibrium solution for $G(A,N)$. A heuristic remedy would be to load the linkflows obtained for $G^*(A^*,N^*)$ onto $G(A,N)$, check if any reasonable paths (otherwise this heuristic will end up with all paths) exist in the loaded network $G(A,N)$ that were previously overlooked, and add them to $\mathcal{R}$, if needed. For a definition of what constitutes a reasonable path, see e.g. Catalano et al. (2001), . Whilst this is not satisfactory from a mathematical point of view, it is makes sense from an engineering point of view as it can be assumed that inspection of the links and nodes that are in $G$ but not in $G^*: G - G^*$, will soon reveal if any important features of the original network $G(A,N)$ remain unused.

The third consequence of path enumeration is that it approach ensures the computational feasibility of the DUE-STOCH assignment while (if done properly) retaining all relevant routes. The case with which this approach can be carried out depends on the network topology: in networks with strong hierarchy of routes (see e.g.Bovy and Stern (1990) and Van Ncs (2002)) this will result in a fairly small set $\mathcal{R}$. Transportation networks with a grid structure (e.g. Sioux Falls) will lead to larger choice sets. In the case where departure time choice is represented through a STEN, this results in a choice set $\mathcal{R}' = |\mathcal{H}| * \mathcal{R}$ with $|\mathcal{H}|$ the number of allowed departure time periods.

The fourth consequence is that there are no unused paths. This seemingly trivial remark will ensure the applicability of the theorem by Tobin and Friesz (see Tobin and Friesz (1988)) on the sensitivity of the equilibrium solution w.r.t. perturbations (read small changes) in the dynamic O-D matrix. existence, overcoming the problem noted in Section 5.4 of Bell and Iida (1997) of rank-deficiency of the link-path incidence matrix.

6.10 Summary

In this chapter we have defined the dynamic user equilibrium assignment that constitutes the lower level of the O-D matrix estimation problem formulated in sections 5.2.3, 5.3, 5.4.
For reasons noted in chapter 5, and to incorporate the behavioural aspects noted in chapter 3, we have focused on the assignments DUE-STOCH and SDTAC. Starting from the equilibrium conditions, and using STEN's as a conceptual framework we have stated VIP formulations for both DUE-STOCH and SDTAC.

In order to provide the assignment map required by the estimation problems OD-ESTIM-STOCH and OD-ESTIM-SDTAC in chapter 5, we have adapted the dynamic network loading part of the DTA model that we adopted.

In order to provide derivatives of the objective function in these optimisation problems, it is necessary to calculate derivatives of the assignment map. Taking advantage of the fact that we are using an analytic DTA with stochastic route choice, we have to calculate analytic derivatives of the assignment map. In order to simplify the calculations, we have truncated the derivatives of the assignment map, and argued that in terms of a game theoretic view of the O-D matrix estimation this is acceptable.

Furthermore this chapter highlights further modularity in the formulation of the O-D matrix estimation problem in the sense that we have adopted a VIP formulation for user equilibrium assignment that can handle arbitrary route choice and departure-time choice models. These models are discussed in chapter 7 below.

Solution methods for the OD-ESTIM-STOCH and OD-ESTIM-SDTAC problems are discussed in chapter 8.
Chapter 7

Behavioural choice models for route and departure time

7.1 Introduction

In chapters 4 to 6 we have developed the specifications of the statistical, optimisation, and dynamic traffic assignment aspects of the model framework proposed in chapter 3, Figure 3.1. In this chapter we will provide details of the behavioural models in our framework. We will discuss the route choice, the departure time choice, and the departure time adjustment models in view of the issues raised in section 3.

Our approach towards dynamic O-D matrix estimation, as summarised in figures 3.5 (p. 41) and 3.6 includes behavioural responses, which are represented through the use of discrete-choice models of (individual user's) behaviour, as represented by the elements for departure time choice, departure time adjustment, and route choice shown in 3.5, and represented by the expressions $P_{t|tr}$ (for departure time choice), $P_{h|t}$ ($C_p, \Theta$) (departure time adjustment), and $P_r^*$ ($C_p, \Phi$) (route choice). In this chapter we will address the behavioural aspects incorporated in our dynamic O-D matrix estimation approach, and explain how they are implemented in our approach. Specifically we will set out the requirements that the model structure presented in chapter 3 imposes on the behavioural models for route choice and departure time adjustment.

After that we will give a brief literature survey of the rationale and workings of the models that fit our needs. We will examine and compare a few such models for departure time choice, departure time adjustment, and route choice, and select a model for route choice. We will also obtain a departure time adjustment model by simplifying and re-estimating a departure time choice model. These models are the ones that are implemented in the software to test the O-D matrix estimation framework.

7.2 Determining appropriate choice models

One of the issues noted in section 1.2.2 is how to model the behavioural aspects identified in section 3.5.
7.2.1 Role of the models

The role of the models is to calculate the preferred departure-times $P_{k|r,s}$ given the origin $r$, the destination $s$, the probabilities for departure time adjustment $P_{k'|h}(C_p, \Theta)$, given the generalised route costs $C_p$, and the route choice probabilities $P_{p|r,s}^*(C_p, \Phi)$ given the generalised route costs $C_p$, and a parameter vector $\Phi$, as suggested in figures 3.5 (pp. 41) and 3.6 (pp.42).

The departure time model converts a static O-D matrix into a dynamic one by specifying a departure time distribution for each O-D pair, and so determine the dynamic apriori matrix. Such models could be based on an analysis of activity choice, or be based on empirical departure time distributions.

The departure time adjustment models update the dynamic O-D matrix on basis of the results of the assignment (travel time and cost of the available routes). The route choice models work within the assignment module and model the trade-off between perceived time and cost on alternative routes.

7.2.2 Model requirements

The models should at least satisfy the requirements of being:

1. a model of rational decision maker in the sense of Ben-Akiva and Lerman (1989) section 3.3, i.e. having consistent and transitive preferences

2. sensitive to type of generalised cost (departure time, travel time, monetary cost, schedule delay etc.) that is affected by the assignment procedure

3. able to work with the modest amount of information available within the context of transportation planning methods

4. sensitive to essential differences between various population types and travel purposes

5. giving reproducible results across iterations in order not to interfere with the convergence of the model system

Requirement no. 1 can be debated at length from a behavioural point of view, but is desirable from an engineering point of view. Requirement no. 2 is essential if the effect of the actual level of service (LOS) given by the assignment on the choice probabilities is to be reflected in the choice probabilities. Requirement no. 3 is needed to ensure that the model can function with the aggregate information available within transportation planning models (which certainly includes items such as geographic zones, time periods, travel time, distance, cost, and may include items such as socio-economic characteristics). Requirement no. 4 is needed to separate population segments whose behaviour is completely different (e.g., children versus working population or pensioners, high or low income groups, constrained versus unconstrained segments), and activities which give rise to different behaviour (work versus leisure). The question of whether traffic
observations can distinguish between population segments is not essential as the segmentation and indeed the entire choice-model specification can be seen as a form of apriori information for which no observations direct are available. Requirement no. 5 seems obvious, but e.g. simulation models may give (slightly) different choice probabilities when repeatedly applied to the same choice problem. This might seem to exclude simulation-based models, but this is not necessarily the case, as one may draw all the random samples in one go, and re-use them for every iteration of the model.

In view of these requirements we have adopted Random Utility discrete choice models of the logit type (see e.g. Ben-Akiva and Lerman (1989)) for use.

7.3 Logit models for route choice

In the application of RUM-based DCM's to route choice the choice set is the set \( P_{rs} \) of enumerated routes \( p \) between the origin \( r \) and destination \( s \), as specified in section 6.9.2.

The problem of correlation of error terms through route-overlap has been recognised, and a number revised DCM's have been proposed to deal with the problem, as reported in Bell and Iida (1997) and Ramming (2002). We will present and motivate our choice of route choice models to implement and use.

The DTA implementation we used is laid out for path-based (as opposed to link-based) discrete-choice models for calculating the route probabilities (see 7 on page 109 for a review). Whilst this approach may rightly be criticised as covering an insufficient number of paths in view of the search for a user equilibrium, empirical work reported in Ben-Akiva et al. (1984), and Bousall et al. (1997), shows that for inter-urban O-D relationships the number of different routes used is fairly small (about 2-10). We note that this probably depends on the network structure, since Ramming (2002) finds that using a k-shortest path algorithm to generate route-sets, even generating 40 route alternatives only resulted in 57% of the observed routes being matched (for dataset that was considered).

We implemented a path-size model in the DTA code (see Ramming (2002), Cascetta (2001)) for route choice based on travel time and travel cost, based on static path overlap (see 7).

7.3.1 The C-logit model

The C-logit model (first proposed in Cascetta et al. (1996), ) gives the following route choice probabilities:

\[
P(i|C_n) = \frac{\exp \mu (V_{in} - CF_{in})}{\sum_{j \in C_n} \exp \mu (V_{jn} - CF_{jn})}
\]

(7.1)

with \( V_{in} \) the systematic part of the utility function for alternative \( i \) and decision maker \( n \), and \( CF_{in} \) the commonality factor. Four forms of the commonality factor are specified in Cascetta et al. (1996), , of which we have selected
\[ CF_{in} = \beta_{CF} \ln \left( \sum_{j \in C_n} \left( \frac{L_{ij}}{\sqrt{L_i L_j}} \right)^\gamma \right) \]  

(7.2)

in which \( \beta_{CF} \) and \( \gamma \) are model parameters, \( L_i \) and \( L_j \) are the lengths of the paths in the choice set, and \( L_{ij} \) is the length of the overlap between paths \( i \) and \( j \). The C-logit model, whilst intuitively attractive has the disadvantage that it is not based on choice theory.

7.3.2 The path-size model

A related choice model that is consistent with discrete choice theory is the path-size model given Ben-Akiva and Bierlaire (1999). The path-size model gives the following route choice probabilities:

\[ P(i|C_n) = \frac{\exp \mu (V_{in} + \ln S_{in})}{\sum_{j \in C_n} \exp \mu (V_{jn} + \ln S_{jn})} \]  

(7.3)

in which \( V_{in} \) the systematic part of the utility function for alternative \( i \) and decision maker \( n \), and \( S_{in} \) is defined as:

\[ S_{in} = \sum_{a \in \Gamma_i} \frac{1}{L_i} \sum_{j \in C_n} \frac{l_i}{\delta_{aj} \frac{L_{C_n}}{L_j}} \]  

(7.4)

with \( a \) a link on path \( \Gamma_i \), \( \delta_{aj} \) an indicator function that is 1 if link \( a \) is on path \( j \), and 0 otherwise, and \( L_{C_n} \) the length of the shortest path in \( C_n \). Although variations on this model are possible (see Ramming (2002)), the model in eqn. (7.3) was adopted for practical reasons.

7.3.3 The Logit-kernel model

The logit-kernel model or "mixed-logit model" (see Bekhor et al. (2002), and Ben-Akiva and Bierlaire (1999)) is a DCM with the following utility function (assuming that the number of elements in the choice set \( |C_n| = J \):

\[ U_p = \beta^T X + \varepsilon_p = \beta^T X + F \xi + u_p = \beta^T X + FT \xi + u_p \]  

(7.5)

with: \( U_p \) the \( J \)-dimensional vector of utilities of the route alternatives for decision maker \( n \), \( \beta \) a \( K \)-dimensional vector of parameters, \( X \) a \( J \times K \) matrix of explanatory variables, \( F \) a \( J \times M \) factor loadings matrix, \( T \) an (unknown) \( M \times M \) lower triangular matrix, \( \xi \) a \( T \) a \( N(0, TT^T) \) distributed random vector, and \( \zeta \) a vector of i.i.d. \( N(0, 1) \) distributed random variables.

The resulting conditional choice probabilities are (if the \( \zeta \) were known):
\begin{equation}
P(p|\varsigma = \varsigma) = \frac{\exp \left( \mu \left( X_p \beta + F_p T \varsigma \right) \right)}{\sum_{q \in C_n} \exp \left( \mu \left( X_q \beta + F_q T \varsigma \right) \right)} \tag{7.6}
\end{equation}

Since (as a rule) the $\varsigma$ are unknown, the probability of choosing route $p$ is:

\begin{equation}
P(p) = \int_{\varsigma} P(p|\varsigma = \varsigma) \, dP(\varsigma) \tag{7.7}
\end{equation}

Ways of calculating the integral in eqn. (7.7) are given in Evans and Schwartz (2000). For high-dimensional $\varsigma$ the method of choice is Monte-Carlo simulation, in which eqn. (7.7) is approximated by repeatedly drawing $\varsigma$ from a $N(0, 1)$ distribution, calculating $P(i|\varsigma)$, and averaging:

\begin{equation}
P(p) \simeq \frac{1}{N} \sum_{l=1}^{N} P(p|\varsigma^l) \tag{7.8}
\end{equation}

Application to the route choice problem can be done as in Bekhor et al. (2002), by assuming:

- link-specific factor to be i.i.d. Normal
- variance proportional to link length
- setting $F$ to be the link-path incidence matrix
- $T$ to be the link variance matrix (diagonal)

In order to avoid problems of reproducibility, a single realisation of the vector $\varsigma$ can be sampled at the start of the O-D matrix estimation run, and re-used throughout. In order to reduce variance, a drawing from a quasi-random sequence is preferred.

**Application of the route choice models to our case**

In applying the route choice models to our problem, we have selected a utility function that contains the length $L_i$ and travel time $\tau_i$ as explanatory variables

\begin{equation}
U_{in} = V_{in} + \xi_{in} = \beta_1 L_i + \beta_2 \tau_i + \xi_{in} \tag{7.9}
\end{equation}

**Specification of the logit-kernel model** In order to apply the Kernel logit model to our problem, we either need to estimate the matrix $F$, or specify its structure apriori. We have opted for specifying $F$ on basis of the overlap matrix $O_{ij}$, defined as

\begin{equation}
O_{ij} = \frac{L_{ij}}{\sqrt{L_i L_j}} \tag{7.10}
\end{equation}
7.3.4 Discussion

In view of the fact that the logit kernel model can reflect the complete route overlap structure of a network through its factor loadings, and in view of the favourable results reported in the simulation studies by Batley et al. (2001), this seems to be the best route choice model. Unfortunately however, this model does not have an analytical expression of its probabilities so that its application in practice requires simulation in order to calculate an approximation of the true route choice probabilities (see eqn. 7.8). Therefore we are prepared to make some concessions in order to have an analytical model.

7.4 Logit models for departure time choice

Departure time choice models have been reported in Vickrey (1969), Small (1987), and have subsequently received attention in Hendrickson and Kocur (1981), Arnott et al. (1990), More general aspects of departure time choice modelling were presented in Van Vuren et al. (1988), and Van Vuren et al. (1999). A recent study which estimated departure time models on data from The Netherlands is reported in ?, which also contains an extensive list of time-of-day choice models. These models typically use variables such as travel time, departure time, arrival time etc. to estimate the preferred departure time.

These models essentially deal with the situation sketched in Figure 7.1.

\[ t_{dep} \rightarrow t \rightarrow t_{PAT} \]

Figure 7.1: Departure time, travel time, and arrival time

Travellers depart at \( h = t_{dep} \), experience a travel time \( \tau_h^{rs} \) that depends on the origin \( r \), the destination \( s \), and the departure time \( h \), and arrive at their destination at \( t_{arr} = h + \tau_h^{rs} \).

7.4.1 Utility function

A classical disutility function based on Vickrey (1969) and proposed in ? is:
\[ r_{hs}^r = \alpha \tau_{hs}^r + \beta \max(0, t_{PAT} - (h + \tau_{hs}^r)) + \gamma \max(0, (h + \tau_{hs}^r) - t_{PAT}) \] (7.11)

Notation

\( \tau_{hs}^r \): travel time between \( r \) and \( s \) when starting at time \( h \)

\( u_{hs}^r \): systematic part of the disutility incurred by setting out from \( r \) to \( s \) at time \( h \)

\( t_{PAT} \): preferred arrival time

This (dis)utility function penalises the travel time \( \tau_{hs}^r \) by a factor \( \alpha \), late arrival (later than \( t_{PAT} \)) by a factor \( \gamma \), and early arrival (before \( t_{PAT} \)) by a factor \( \beta \).

Note that this utility function does not represent any preference for departure time: travellers are assumed to care only for their arrival time regardless of the time at which they have to depart to arrive in time. From a behavioural point of view this is not very plausible, so that we extend the utility function with a term that captures departure time preference. Travellers are assumed to have a Preferred Departure Time \( t_{PDT} \) in addition to their preferred arrival time.

### 7.4.2 Model form

Model forms for RUM-based departure-choice models range from the ordinary MNL model (see Small, K. (1982), Nuzzolo and Russo (1998) for intercity transit) to the OGEV model (Small (1987)) and the ECL model (7).

Strictly speaking the MNL model has the drawback that it does not take account of the fact that adjacent departure time intervals look more "alike" than widely separated intervals, a deficiency which is remedied by the OGEV model and in a more general way by the ECL model, and ultimately Probit models. However, for practical reasons the simple MNL model was selected for implementation in order to demonstrate the principle. The use of more advanced models is a clear candidate for improvement and further research.

### 7.4.3 Application issues

A recurrent issue in choice models transportation is that the models must be applied to a synthetic population, for which not all the explanatory variables are known with the same precision as they were for the population on which the model was estimated. This issue is discussed in some depth in chapter 6 of Ben-Akiva and Lerman (1989).

In our context this relates to all parameters in eqn. 7.11, including the preferred arrival time (\( t_{PAT} \)). For practical reasons we will restrict ourselves to a single segment with a single set of parameters \( \alpha, \beta, \gamma \) and \( t_{PAT} \). In practical application of departure time choice models this simplification will need to be addressed.
7.5 Departure time adjustment modelling

In order to combine route choice and departure time choice, a model is needed that describes the trade-off between schedule delay and other components of the travel disutility such as travel time and cost. For the moment we will assume that such a model is available, and that its parameters have been determined using SP or RP experiments. A classical result is presented in Small, K. (1982) a recent example is given in ?.

In this context we will assume that the VOT and the schedule delay costs are given; these values can be estimated in SP and /or RP experiments.

We propose the following slightly altered version of the departure time choice utility \( v^{rs}_h \) as a basis for departure time adjustment:

\[
v^{rs}_h = v^{rs}(h) = \beta_1 c^{rs}_h + \left( \beta_2 I_{h < t_{PDT}}(h) + \beta_3 I_{h > t_{PDT}}(h) \right) |h - t_{PDT}| + \left( \beta_4 I_{t_{arr} < t_{PAT}}(t_{arr}) + \beta_5 I_{t_{arr} > t_{PAT}}(t_{arr}) \right) |t_{arr} - t_{PAT}|
\]

where:

\( h \): the departure time

\( c^{rs}_h \): the generalised travel cost between \( r \) and \( s \) when starting during period \( h \); \( c^{rs}_h = \tau^{rs}_h + \text{tolls}(h) \), where \( \text{tolls}(h) \) stands for the tolls that we would incur along the path if one were to leave during time interval \( h \).

\( \beta_1 \): the weight factor for travel cost

\( \beta_2, \beta_3 \): the weight factors for departure time schedule delay (departing early or departing late)

\( \beta_4, \beta_5 \): the weight factors for arrival time schedule delay (arriving early or arriving late)

\( t_{arr} \): the arrival time \( t_{arr} = h + \tau^{rs}_h \)

\( I_{h < t_{PDT}}(h) \): an indicator function, defined as: \( I_{h < t_{PDT}}(h) = \begin{cases} 1 & \text{if } h < t_{PDT} \\ 0 & \text{otherwise} \end{cases} \)

Although the overall structure of the utility function in eqn. 7.12 is almost identical to that of eqn. 7.11, two details are added. The first detail is the addition of a disutility term for departing earlier or later than the preferred time:

\[
(\beta_2 I_{h < t_{PDT}}(h) + \beta_3 I_{h > t_{PDT}}(h)) |h - t_{PDT}|
\]

and the second is the use of generalised travel cost \( c^{rs}_h \) instead of travel time \( \tau^{rs}_h \), which allows for the inclusion of time-dependent tolls (see Viti et al. (2003), for an example). This situation is illustrated in Figure 7.2.
Shown on the horizontal axis is the departure time $h = t_{dep}$, on the vertical axis the arrival time $t_{arr}$; the line $t_{dep} = t_{arr}$ that serves as an offset to which the travel time $\tau_h^{rs}$ is to be added to obtain the arrival time. The time-dependent travel time $\tau_h^{rs}$ is indicated using the thick solid line. Travellers departing at their preferred departure time $h_2 = t_{PDT}$ experience travel time $\tau_h^{rs}$ and arrive at $t_{PDT} + \tau_h^{rs}$. We have assumed that there is some tension between preferred departure time and preferred arrival time so that this arrival time is later than the preferred arrival time. Travellers wishing to arrive at their preferred arrival time $t_{PAT}$ need to depart at time $t_{PAT} - s_{tPAT}^{rs}$, with $s_{tPAT}^{rs}$ the inverse of the travel time function, i.e. it takes arrival time as its argument and returns the time at which one would have to depart to arrive at that instant, taking account of the changing travel time.

We note that when we consider travel purpose "work", the preferred arrival times are similar for many people, so that the total travel demand peaks around $t_{PAT} - s_{tPAT}^{rs}$, which in our society leads to congestion and increased travel times. Empirical evidence described in Vellay and De Jong (2001) bears out the assumption that commuters leave earlier than they would have preferred.

As illustrated in Figure 7.3, utility maximising traveller could thus be expected to minimise the expression in eqn. 7.12
Figure 7.3: Minimisation problem for a utility maximising traveller

Equation 7.12 displays a pleasing symmetry, which allows it to capture both situations in which the choice process can be considered to be primarily determined by the arrival time (such as the morning commute), and situations where the choice process can be considered to be primarily determined by the departure time (e.g. the evening commute, and discretionary activities such as sport and other forms of recreation).

7.5.1 Application issues: determining $t_{PDT}$ and $t_{PAT}$

In application we face the same issues as noted in section 7.4.3, i.e. the values of the parameters $\beta, \cdots, \beta_5$, $t_{PDT}$ and $t_{PAT}$.

We will simplify matters by making the following assumptions:

1. there is only one set of values for $\beta, \cdots, \beta_5$ (i.e. we assume the population to be homogeneous and allow no taste variation)

2. we have an apriori matrix $T_h^{rs}$ that is given by a departure time choice model, i.e. for every O-D cell: $t_{PDT} = h$

3. we will assume that the $PAT$ that is implicit in the apriori matrix is related to the $PDT$ through the free-flow travel time as follows: $t_{PAT} = t_{PDT} + \tau^{rs,ff}$

4. we will assume that the choice process is determined by the arrival time,
that we have observations of the true arrival time, and that these values are clustered around the value of $t_{P\cdot AT}$:

With these strong simplifications the schedule-delay term $|t_{arr} - P\cdot AT|$ becomes:

$$|t_{arr} - t_{P\cdot AT}| = |h + \tau_{h}^{rs} - (h + \tau_{ff}^{rs})|$$  \hspace{1cm} (7.16)

$$= |\tau_{h}^{rs} - \tau_{ff}^{rs}|$$  \hspace{1cm} (7.17)

$$v_{h}^{rs} = v^{rs}(h) = \beta_{1} c_{h}^{rs}$$  \hspace{1cm} (7.18)

$$+ (\beta_{2} I_{h < t_{P\cdot DT}}(h) + \beta_{3} I_{h > t_{P\cdot DT}}(h)) |h - t_{P\cdot DT}|$$  \hspace{1cm} (7.19)

$$+ (\beta_{4} I_{t_{arr} < t_{P\cdot AT}}(t_{arr}) + \beta_{5} I_{t_{arr} > t_{P\cdot AT}}(t_{arr})) |\tau_{h}^{rs} - \tau_{ff}^{rs}|$$  \hspace{1cm} (7.20)

The utility of a departure time shift from departure time period $h$ to departure period $h'$ then becomes:

$$v_{h'}^{rs} = v_{h}^{rs} - v_{h'}^{rs}$$  \hspace{1cm} (7.21)

$$= \beta_{1} (c_{h'}^{rs} - c_{h}^{rs})$$  \hspace{1cm} (7.22)

$$+ (\beta_{2} I_{h < t_{P\cdot DT}}(h) + \beta_{3} I_{h > t_{P\cdot DT}}(h)) |h' - h|$$  \hspace{1cm} (7.23)

$$+ (\beta_{4} I_{t_{arr} < t_{P\cdot AT}}(t_{arr}) + \beta_{5} I_{t_{arr} > t_{P\cdot AT}}(t_{arr})) |\tau_{h'}^{rs} - \tau_{h}^{rs}|$$  \hspace{1cm} (7.24)

which includes the extra term $(\beta_{2} I_{h < t_{P\cdot DT}}(h) + \beta_{3} I_{h > t_{P\cdot DT}}(h)) |h' - h|$ as compared to the case where only arrival time is considered.

### 7.6 Summary

We have selected choice models which are completely analytical and very tractable. For the route choice models we have implemented the C-logit model, the path-size model, and the Logit-Kernel model; for practical reasons we will restrict ourselves to the C-logit model. departure time choice is strongly related to the choice of a daily activity pattern, but this is external to our model framework. The justification for keeping it external is that home-work travel causes most of the travel demand during the peak hours, and the corresponding arrival times are mostly not discretionary, and therefore not very much influenced by congestion. Departure time adjustment is the process whereby individual travellers determine their optimal trade-off between travel-time, late arrival, early arrival and early departure, is therefore directly affected by congestion, and is endogenous to our models. We have used the Small-Vickrey formulation of departure time adjustment. The dynamic apriori matrix incorporates both destination-choice and departure time choice (both exogenous to our models) and is formed by adding an (observed or modelled) departure time profile to a given static O-D matrix.

In this chapter we discussed the behavioural choice models that link the raw demand to path flows and departure time choice. These models form a bridge between the mathematical aspects of the framework for dynamic O-D matrix estimation and an attempt
to describe the reality of actual travellers’ decisions in the models. From a transportation point of view they are crucial, whereas from an optimisation point of view they are invisible (being mere parts of the calculation of the SUE-DET assignment). We can now go back to the formulation of the dynamic O-D matrix estimation problem as an MPEC problem recognised in chapter 5, whose lower level was discussed in chapter 6, and focus our attention on ways of solving the MPEC problem in chapter 8 below.
Chapter 8

Solution methods

8.1 Introduction

As shown in chapter 5 the dynamic O-D matrix estimation problem can be formulated as a bi-level programming problem, more precisely an MPEC. Unfortunately these are difficult to solve exactly, because a step in the descent direction of the top-level problem does not have to be a descent direction of the MPEC because of the reaction of the second-level problem (in our case the DTA). This is a general feature of bi-level programming problems. A number of exact algorithms have been proposed as well as some heuristics. In this chapter we will present and compare these and select one for implementation.

8.2 Solution algorithms

In view of the nonlinearities, an iterative solution procedure must be used, which produces a sequence of matrices \( \{ T_k^{rs,n} \} \) that converges to the final matrix estimate \( \bar{T}_k^{rs} \). This can be formally expressed as

\[
T_k^{rs,n+1} = F \left( T_k^{rs,n}, T_k^{rs,0}, S_{lk}, \bar{S}_{lk} \right) = F \left( T_k^{rs,n}, T_k^{rs,0}, S_{lk}, \left( DTA \left( T_k^{rs,n} \right) \right), \bar{S}_{lk} \right) \quad (8.1)
\]

Where the map \( F \) constructs the new matrix estimate on basis of the current group totals, the current matrix estimate \( T_k^{rs,n} \), the apriori matrix \( T_k^{rs,0} \), and the observed group totals \( S_{lk} \), a revised estimate \( T_k^{rs,n+1} \) of the O-D matrix, as shown in Figure 8.1.
Figure 8.1: Iterative method for O-D matrix estimation

The method starts off with an initialisation step in which the dynamic apriori matrix is assigned to the current matrix estimate: $T^{s,n}_{k} = T^{s,0}_{k}$. Then the algorithm, enters the iteration step and assigns the current matrix to the network using a Dynamic Traffic Assignment (DTA). On basis of this assignment result, all group totals $S_{ik}$ that depend on assignment fractions, link flows, or travel times can be calculated. The objective is to calculate a fixed point $T^{r,s,\infty}_{k}$ of this map: $T^{r,s,\infty}_{k} = F(T^{r,s,\infty}_{k})$. The bi-level nature of the O-D matrix estimation problem is reflected in the map $F$, which can be decomposed in a map $F$ that reflects the toplevel problem: $T^{r,s,n+1}_{k} = F(T^{r,s,n}_{k}, T^{r,s,0}_{k}, S_{ik}, \bar{S}_{ik})$ and a map $DTA$ that reflects the bottom level:

$$F(T^{r,s,\infty}_{k}, S_{ik}(DTA(T^{r,s,\infty}_{k})))$$  \hspace{1cm} (8.2)

All but the Sequential Quadratic Problem (SQP) algorithms we found in the literature (Luo et al. (1997), Patriksson and Rockafeller (2002), Clegg and Smith (2001), and Maher et al. (2001)) use a repeated line-search method shown in Figure 8.2. The SQP approach generates a sequence of quadratic optimisation problems whose solutions converge to that of the MPEC.

Shown on top is the initial solution pair $(\rho^{0}, f^{0})$ where $\rho^{0}$ signifies the free parameter of the MPEC, and $f^{0}$ the corresponding flow pattern. Next a descent direction $(\delta\rho, \delta f)$ is sought, simultaneously in the parameter $\rho$ and the flow $f$. Once the descent direction is known there results only a line-search for the optimum from the current solution pair $(\rho, f)$, i.e. the determination of an optimal stepsize $\gamma$ to find the optimum $(\rho + \gamma \delta \rho, f + \gamma \delta f)$ along the ray formed by the current solution and the descent direction. Convergence can be measured e.g. by a criterion on the stepsize $\gamma$. 

Figure 8.2: The generic repeated line-search algorithm

All but the SQP algorithms differ only in the way in which they calculate the descent direction \((\delta \rho, \delta f)\) and how they determine the step size \(\gamma\). We note here the importance of determining a joint descent direction \((\delta \rho, \delta f)\) instead of only a descent direction in \(\delta \rho\) for computational efficiency, because in the latter case we would change \(\rho\) and then have to calculate the corresponding value of \(f\) for each new value of \(\rho\). As the calculation of \(f(\rho)\) equals solving a SUE-DET, DUE-DET, SUE-STOCH, or DUE-STOCH assignment, it is essential to avoid this.

We have listed the main algorithms in table 8.1.

<table>
<thead>
<tr>
<th>Nr</th>
<th>Reference</th>
<th>Descent direction</th>
<th>Solution guaranteed?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yang (1995)</td>
<td>(\delta \rho)</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>Maher et al. (2001)</td>
<td>((\Delta \rho, \Delta f))</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>Clegg and Smith (2001)</td>
<td>(\delta \rho)</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>Patriksson and Rockafeller (2002)</td>
<td>((\delta \rho, \delta f))</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>Luo et al. (1997), (PIPA)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>Luo et al. (1997), (Implicit Programming based method)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>Luo et al. (1997), (SQP)</td>
<td></td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 8.1: Comparison of MPEC solution methods

The first algorithm is the one reported in Yang (1995), which iterates the minimisation of the objective function and the calculation of the response of the lower level and can take a long time to converge. This algorithm was improved in Maher et al. (2001) (presented in section 8.2.2), who calculate an exact descent direction for the upper
level problem and calculate a linear approximation of the response of the lower level problem to a step in the descent direction of the upper level to effectively yield a joint descent direction for line-search. This algorithm was seen to work well in a practical setting in the sense that it converged to a credible fixed point.

The first algorithm for which a convergence proof could be provided is presented in Clegg and Smith (2001). This method attempts to simultaneously solve the upper and lower level problems, and is based on the idea that a descent direction for the upper level problem is safe to use if one can prove that along this direction the objective function for upper level problem descends steeper than the lower level problem can counteract by changing the solution. Such a direction can be obtained by projecting the descent direction of the upper level problem on the set of all problems for which the lower level problem is not worsened. The descent direction is calculated in a way that gives precedence to solving the equilibrium when the system is far from equilibrium, and gives precedence to the solution of the upper level problem when the lower level problem is "sufficiently close" to equilibrium.

Three types of solution algorithms are presented in Luo et al. (1997), (algorithms 5,6,7 in table 8.1), of which no. 5 and 6 assume a constraint or an LCP as the second-level problem. Algorithm 7 allows a VI problem at the second level, which has to be directionally differentiable. The directional differentiability of the (static) assignment problem was proved in Patriksson (2002), so that this algorithm could in principle be used.

The most recent method was presented in Patriksson and Rockafellar (2002), which results in a joint descent direction \((\delta \rho, \delta f)\) that results from solving a quadratic minimisation program. From a theoretical point of view this algorithm seems best as a convergence proof could be given, and the algorithm holds out the promise of a quick practical implementation, e.g. by using Maher's heuristic to obtain a rough solution, and to refine this solution using the more complicated algorithm when improvement seems to level off. Unfortunately this algorithm works in a generalisation of the space of feasible network flows; feasible network flows are obtained by projection of the abstract flows onto the space of feasible flows which requires further work to translate into a working implementation.

8.2.1 Descent algorithms using approximate derivatives

If one applies the line-search procedures of the previous section with an approximate descent direction, one can still obtain improvement of the MPEC. With truncated-gradient algorithms we use approximate derivatives to find an approximate descent direction. The line-search procedures work as before. As noted in section 6.7.2, we use approximate derivatives that result in approximate descent directions. The properties of a solution method that uses approximate (why e.g. it might lead to a solution at all) can be understood from a game-theoretic point of view of the problem.
A game-theoretic view

Any MPEC can be interpreted as a game: the maximisation problem is the top-level player, and the equilibrium constraints result from a n-player Nash game. Suppose the player at the top level knows the exact cost function and with it the precise response of the bottom level to any change made at the toplevel. This player could then use this response function to eliminate the response of the bottom level player from his own cost function, and would be left with a cost function solely in terms of his own control variables. This function could then be maximised without further reference to the reaction of the bottom level players as their actions are accounted for anyway. In the optimum for the toplevel, the reactions of the bottom level players could be calculated if of interest. This is known as a Stackelberg game (see Basar and Olsder (1999), section 4.6).

In the event that the toplevel player know nothing about the response surface of the bottom level players, we have a Cournot game (see Basar and Olsder (1999), section 4.6, example 4.5). In this case a way to look for a mutually stable solution would be to "play the game", i.e. let each player in turn optimise its objective function conditional on the previous actions of the other player.

Obtaining correct derivatives of the traffic flow w.r.t. the O-D matrices is a way of giving the toplevel problem an approximation of the shape of the response surface of the second level problem, and the resulting game will then be one of imperfect information.

8.2.2 Maher's heuristic

It turns out that in the case of O-D matrix estimation even without derivatives one can obtain approximate but useful solutions using a heuristic, as described in Maher et al. (2001) and applied in Zhang et al. (2002). The essence of Maher's heuristic is the derivation of a (possibly sub-optimal) descent direction together with a heuristic to speed up the line search once a descent direction is decided on.

Where Maher et al. (2001) and Zhang et al. (2002), apply the heuristic to an objective function based on link flow targets, we propose to apply Maher's algorithm to the more comprehensive objective function in eqn.

A flowchart for the heuristic using the more general objective function is shown in Figure 8.3.
Figure 8.3: Flow diagram of Maher’s heuristic for O-D matrix estimation adapted to the dynamic situation

An approximate search direction is calculated by determining an improved solution of the upper level (the dynamic O-D matrix) while completely ignoring the effects of the lower level (the assignment). Next the assignment is calculated for this O-D matrix, giving a pair \((T, x)\). Next an approximate line-search is carried out between the original solution and the new solution, whereby the assignments corresponding to the intermediate matrices are approximated by corresponding linear combinations of the assignments at the end points. This results in large computational savings, at the cost of guaranteed convergence.

### 8.3 Choice of a solution algorithm

We have selected the heuristic specified in Maher et al. (2001) for implementation for practical reasons. It is faster than the heuristic algorithm proposed in Yang et
al. (1994), yet very easy to implement. We also note that for our use of stochastic assignment Maher’s assumption that the response of the second level can be linearised is likely to be reasonable, since with stochastic assignment the path flows depend in a differentiable way on the path costs.

The problem of determining the O-D matrix given the assignment fractions was solved using the software described in Lawrence et al. (1997).

The two exact algorithms have practical drawbacks. The algorithm specified in Clegg and Smith (2001) has theoretical appeal, but is reported to incur a heavy computational penalty as a result of the switch between the optimisation of the upper and lower level problems. Whilst for theoretical reasons the algorithm formulated in Patriksson and Rockafeller (2002) is preferable and this algorithm holds much computational promise, its implementation is left as future work in view of the added requirement to implement an efficient method to project the abstract flows of that algorithm onto real network flows. Therefore the simple iterative method presented Yang (1995) and the heuristic presented in Maher et al. (2001) were used.

8.4 Summary

In chapter 5 we formulated the dynamic O-D matrix estimation problem as an MPEC. The specification of the upper level objective function was set out in chapter 4, the properties of the bottom level problem of the MPEC (the SUE-DET assignment) was discussed in chapter 6. In this chapter we have listed a number of solution algorithms for it, ranging from exact to approximate. After comparison of the available algorithms, a heuristic specified in Maher et al. (2001) was selected for implementation. This heuristic was applied to a simple test case and a simple case study, as reported in chapter 10.
Chapter 9

Implementation and verification of the model

9.1 Introduction

In previous chapters a model framework was presented for dynamic O-D matrix estimation, which has been implemented in software. Both the model and the software are complex, and some verification of the correctness of both is needed. As a complete validation of all constituent models would at least require model-runs with estimated models on a realistic population, observed dynamic O-D matrices, with transportation networks to match, such a validation is a subject of future work.

Instead the models and software will be verified to behave in accordance to engineering judgement of what they should do, which will be much more limited a full-scale validation.

The verification of the model and the software was carried out according to the setup shown in table 9.1.

<table>
<thead>
<tr>
<th>Verification step</th>
<th>Required ingredients</th>
<th>Verification of software</th>
<th>Verification of model properties</th>
<th>Empirical test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>artificial network, artificial data</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>real network, artificial data</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>real network, real data</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Table 9.1: Setup for model verification

9.1.1 Verification: aspects considered, criteria, norm

The aspect considered is the ability of the software to recover a mangled O-D matrix when a consistent set of observations related to the true O-D matrix is available. The criteria are the agreement between the O-D matrix cells of the estimated and the true O-D matrix (i.e. the one that resulted in the observations) and the agreement between
the observed group totals and the group totals estimated from the assignment on a synthetic network. The norm is that the software: 1) does not change the true solution if it is entered, 2) is capable of correcting an apriori O-D-t matrix to agree with the group observations.

A second step of the verification is to run the model system on a real transportation network with artificial (but life-like) data to show that the model system does not show counter-intuitive sensitivities when fed with plausible inputs. These steps would in principle verify the correctness of the model system and its implementation.

Further verification would entail an empirical test in which an actual dynamic O-D matrix is estimated for a real network with real data. As this will inevitably involve a fair amount of model calibration (especially the link-delay functions used in the DNL algorithm as they determine the travel-times) this will be left for future work.

9.2 The software

The software architecture loosely follows the model architecture shown in Figure 3.6, and can be represented as in Figure 9.1.

Figure 9.1: Toplevel structure of the software implementation

Shown on top are the observations $S(k)_{obs}$, the initial parameter vector $\Lambda$, and the apriori matrix $T_0$. Together with the parameter vector $\Lambda$, the apriori matrix
is the input of the matrix model, which results in an actual matrix T. This matrix serves as input into the combined departure time adjustment and assignment model and gives rise to estimated versions of the observations Ski. The Lambda parameters are updated to obtain a match between observed and estimated group observations. Unless the process is converged, the new Lambda vector is used in a fresh iteration.

The DTC-DTA problem is solved using MSA.

9.3 Verification of the software

In order to test the implementation of the solution algorithm for the dynamic O-D matrix estimation problem, the model was applied to a synthetic network. In this section we will describe the test and the test results.

9.4 The test-case

The test-case consists of the following basic inputs:

- a network with specification of the zone centroids and reasonable link-performance functions
- a set of paths between each O-D pair
- the travel demand in the form of an apriori O-D-t matrix
- a reasonable but simple specification of the route choice and departure time choice models
- a set of reasonable parameter values for the route choice and departure time choice models

The following sets of artificial group observations are needed for the verification:

- group observations derived directly from the assignment of the apriori matrix in order to show that the model system does not mess up a perfect solution matrix
- non-randomly perturbed group observations in order to show that the model system can correct simple errors in the apriori matrix
- randomly perturbed group observations in order to check if the model system will be able to reproduce the original matrix
9.4.1 The network

![Diagram of the network](image)

Figure 9.2: Simple test-network

The network shown in Figure 9.2 contains three zone centroids (nodes 1, 2, and 5), of which centroids 1 and 2 exclusively function as origins and centroid 5 only as destination. The centroids are connected to the network through centroid connectors (1-10), (2-20), and (50-5).

The route set shown in Table 9.2 contains the following nine routes:

<table>
<thead>
<tr>
<th>O-D pair</th>
<th>Route nr.</th>
<th>Node sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 → 5</td>
<td>1</td>
<td>1-10-30-50-5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1-10-20-30-50-5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1-10-30-40-50-5</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1-10-20-30-40-50-5</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1-10-20-50-5</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1-10-50-5</td>
</tr>
<tr>
<td>2 → 5</td>
<td>7</td>
<td>2-20-5-5</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>2-20-30-50-5</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>2-20-30-40-5-5</td>
</tr>
</tbody>
</table>

Table 9.2: Link flows

The O-D flows carried by the individual links are shown in Table 9.3

<table>
<thead>
<tr>
<th>Link</th>
<th>O-D flows</th>
<th>Average flow share</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-20</td>
<td>1 → 5</td>
<td>45%</td>
</tr>
<tr>
<td>10-30</td>
<td>1 → 5</td>
<td>27%</td>
</tr>
<tr>
<td>10-50</td>
<td>1 → 5</td>
<td>27%</td>
</tr>
<tr>
<td>20-30</td>
<td>1 → 5, 2 → 5</td>
<td>7%</td>
</tr>
<tr>
<td>20-50</td>
<td>1 → 5, 2 → 5</td>
<td>11%</td>
</tr>
<tr>
<td>30-40</td>
<td>1 → 5, 2 → 5</td>
<td>7%</td>
</tr>
<tr>
<td>30-50</td>
<td>1 → 5, 2 → 5</td>
<td>38%</td>
</tr>
<tr>
<td>40-50</td>
<td>1 → 5, 2 → 5</td>
<td>19%</td>
</tr>
</tbody>
</table>

Table 9.3: O-D flows
9.4.2 The apriori demand matrix

The demand is given by the following apriori O-D matrix 9.4:

\[
T_h^r \begin{array}{c|c|c}
 & \text{Destination} & \\
 & 5 & \text{Sum} \\
\hline 
\text{Origins} & 1 & 200 & 200 \\
 & 2 & 3 & 3 \\
\hline 
\text{Sum} & & 203 & 203 \\
\end{array}
\]

Table 9.4: The O-D matrix and its margin totals

The matrix cells were chosen to be different in size so that they would show up in different locations of the scatterplot of O-D cells.

The demand enters the network during four time periods according to the following departure time profile:

<table>
<thead>
<tr>
<th>Percentage departure</th>
<th>O1</th>
<th>O2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time period</td>
<td>10%</td>
<td>40%</td>
</tr>
<tr>
<td>1</td>
<td>30%</td>
<td>50%</td>
</tr>
<tr>
<td>2</td>
<td>50%</td>
<td>10%</td>
</tr>
<tr>
<td>3</td>
<td>10%</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 9.5: Table Caption

This test case contains two difficulties. The first difficulty is characteristic for O-D matrix estimation in general in that the link flows on all links except (20-40) consist of an unobserved mix of O-D flows. The second difficulty is the recovery of the departure time profiles for origins O1 and O2.

There will be four versions of the apriori demand matrix:

1. perfect match with the group observations
2. apriori too high
   (a) apriori too low

9.5 The observations

We will assume the following group observations:

- zonal productions
- zonal attractions (which can be interpreted as flow observations on the centroid connectors (1-10), (2-10), (30,3), and (40-4))
- traffic counts on links (10-30), (10-50), and (20-50).
9.6 Estimation results with only route choice

The estimation results shown below were obtained without departure time choice, using only production and attraction as observations. The estimation results were compared numerically, of which we reproduce only the scatterplots of apriori and estimated O-D cells versus the true ones.

9.6.1 Apriori agrees with group observations

Shown in Figure 9.3 are scatterplots of the estimated group observations versus the observed, before and after the estimation.

![Scatterplots](image)

Figure 9.3: Estimation results with perfect apriori

The model clearly does not change the (correct) apriori matrix, which shows that the model recognises that the matrix fits.

9.6.2 Apriori 20% higher than group observations

Next the true matrix scaled up by 20% was used as apriori matrix and fitted. Shown in Figure 9.4 are scatterplots of the estimated group observations versus the observed, before and after the estimation.
The model clearly succeeds in bringing the matrix back in line with the original.

### 9.6.3 Apriori 20% lower than group observations

Next the true matrix scaled down by 20% was used as apriori matrix and fitted. Shown in Figure 9.5 are scatterplots of the estimated group observations versus the observed, before and after the estimation.

![Graphs showing estimation results when apriori is too low](image)

Figure 9.5: Estimation results when apriori is too low

### 9.7 Estimation results with simultaneous choice of route and departure time

The estimation results shown below were obtained with simultaneous choice of route and departure time, using only production and attraction as observations. Again the estimation results were compared numerically to the true values, and scatterplots are
reproduced of the cells of the apriori matrix and the estimated O-D matrix versus the true ones.

9.7.1 Apriori agrees with group observations

Shown in Figure 9.6 are scatterplots of the estimated group observations versus the observed, before and after the estimation.

![Figure 9.6: Estimation results with perfect apriori](image)

The model clearly does not change the (correct) apriori matrix, which shows that the model recognises that the matrix fits.

9.7.2 Apriori 20% higher than group observations

Next the true matrix scaled up by 20% was used as apriori matrix and fitted. Shown in Figure 9.7 are scatterplots of the estimated group observations versus the observed, before and after the estimation.

![Figure 9.7: Estimation results when apriori is too high](image)
The model clearly succeeds in bringing the matrix back in line with the original.

9.7.3 Apriori 20% lower than group observations

Next the true matrix scaled down by 20% was used as apriori matrix and fitted. Shown in Figure 9.8 are scatterplots of the estimated group observations versus the observed, before and after the estimation.

![Apriori cells vs true cells and Estimated cells vs true cells](image)

Figure 9.8: Estimation results when apriori is too low

9.8 Conclusions

The software was able to correct the deformed apriori matrices and recover the matrix that matches the observations, both with a DTA that has only route choice and with a DTA that has both route choice and departure time choice. In both cases the estimator was able to correct the deformed apriori matrices and recover the matrix that matches the observations.

The estimation results with the pure route-choice DTA show that the O-D estimator is materially correct. The estimation results with simultaneous choice of route and departure time show that an O-D estimator with simultaneous route- and departure time choice is feasible in principle.
Chapter 10

Case study: Kruithuisweg

10.1 Introduction

In the preceding chapter results were presented on estimation runs with synthetic networks and artificial data, in order to verify the sanity of the software implementation. As the counts needed for an estimation on empirical data were unfortunately not yet available at the time of writing, the case-study was also carried out on synthetic data. The model will be applied to a (modest) real-world network with artificial data (so that the performance of the estimator can be checked), and the apriori matrix will be varied in order to check the effect on the estimation results. The advantage of using a real-world network, with demand that has the right order of magnitude, is that it is easier to assess the estimation results than in a purely hypothetical case.

10.2 Description of the study area

The study area is an urban arterial road in Delft, known as the "Kruithuisweg", which is shown in Figure 10.1. It is about 3 km in length, and connects the A4 motorway to the east of Delft with the A13 motorway to the west of Delft. The A13 motorway forms the main North-South connection in the West of The Netherlands, connecting Amsterdam in the North and The Hague in the center with the port of Rotterdam. The city of Delft lies between The Hague in the North and Rotterdam in the South.

A schematic close-up of the Kruithuisweg is shown in 10.2.
Figure 10.1: The Kruithuisweg study area

Figure 10.2: Intersections on the Kruithuisweg

The transportation network and the zone centroids used are shown in Figure 10.3.
The zone centroids are listed in table 10.1.

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Centroid description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A-13 motorway</td>
</tr>
<tr>
<td>2</td>
<td>TU Delft quarter</td>
</tr>
<tr>
<td>3</td>
<td>Hydraulics laboratory, TUD buildings</td>
</tr>
<tr>
<td>4</td>
<td>Industrial zone &quot;De Nieuwe Haven&quot;</td>
</tr>
<tr>
<td>5</td>
<td>Industrial zone &quot;Delft South&quot;</td>
</tr>
<tr>
<td>6</td>
<td>Residential area Voorhof</td>
</tr>
<tr>
<td>7</td>
<td>Residential area Tanthof-Oost</td>
</tr>
<tr>
<td>8</td>
<td>Residential area Buitenhol</td>
</tr>
<tr>
<td>9</td>
<td>Residential area Tanthof-West</td>
</tr>
<tr>
<td>10</td>
<td>Residential area Buitenhol</td>
</tr>
<tr>
<td>11</td>
<td>Residential area Tanthof-West</td>
</tr>
<tr>
<td>12</td>
<td>Urban arterial to Delft and Rijswijk</td>
</tr>
<tr>
<td>13</td>
<td>A4 motorway</td>
</tr>
</tbody>
</table>

Table 10.1: Zone centroid names

10.2.1 Dynamic observations

The group observations used consisted of (synthetic) directional dynamic counts on all centroid connectors (shown in figure 10.3). Synthetic counts were obtained through dynamic assignment of an O-D matrix to the network, after which the dynamic counts were determined from the assigned link flows. Counts of inflow and outflow on centroid connectors function as observations of dynamic zonal traffic production and attraction.
10.2.2 Dynamic demand

A hypothetical (static) O-D matrix for the Kruthuisweg is shown in Figure 10.4.

<table>
<thead>
<tr>
<th>Origin</th>
<th>Destinatie</th>
<th>A13</th>
<th>TNO/TU</th>
<th>WL</th>
<th>Schie</th>
<th>Doet Z</th>
<th>Voozdop</th>
<th>Tarthof O</th>
<th>Butendief</th>
<th>Tarthof W</th>
<th>LVN N</th>
<th>LVN Z</th>
<th>Provenc W</th>
<th>AM</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A13</td>
<td></td>
<td>1 1</td>
<td>500</td>
<td>500</td>
<td>10</td>
<td>10</td>
<td>100</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>1710</td>
</tr>
<tr>
<td>TNO/TU</td>
<td></td>
<td>2 2</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>60</td>
</tr>
<tr>
<td>WL</td>
<td></td>
<td>3 3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>60</td>
</tr>
<tr>
<td>Schie</td>
<td></td>
<td>4 4</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>20</td>
<td>20</td>
<td>122</td>
</tr>
<tr>
<td>Doet Z</td>
<td></td>
<td>5 5</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>20</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>20</td>
<td>20</td>
<td>122</td>
</tr>
<tr>
<td>Voozdop</td>
<td></td>
<td>6 6</td>
<td>100</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>180</td>
</tr>
<tr>
<td>Tarthof O</td>
<td></td>
<td>7 7</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>180</td>
</tr>
<tr>
<td>Butendief</td>
<td></td>
<td>8 8</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>180</td>
</tr>
<tr>
<td>Tarthof W</td>
<td></td>
<td>9 9</td>
<td>25</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>LVN N</td>
<td></td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>LVN Z</td>
<td></td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>Provenc W</td>
<td></td>
<td>12</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>170</td>
</tr>
<tr>
<td>AM</td>
<td></td>
<td>13</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>71</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>327</td>
<td>407</td>
<td>537</td>
<td>72</td>
<td>57</td>
<td>147</td>
<td>57</td>
<td>57</td>
<td>57</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>2997</td>
</tr>
</tbody>
</table>

Figure 10.4: Hypothetical (static) 2-hour O-D matrix for the Kruthuisweg

The static O-D matrix can be visualised through a static traffic assignment. As there is no route-choice in the network an all-or-nothing assignment is sufficient. The assignment result is shown in Figure 10.5.

![Visualisation of the hypothetical static O-D matrix](image)

Figure 10.5: Visualisation of the hypothetical static O-D matrix

A dynamic O-D matrix was obtained from the static O-D matrix shown in figure Figure 10.4 by applying the departure time profile shown in Figure 10.6 to each of its cells.
Figure 10.6: Hypothetical departure-time profile per O-D matrix cell

Shown on the horizontal axis is the time period, shown on the vertical axis is the proportion of the demand that leaves during that time period.

10.3 Model description

10.3.1 Modelling of intersections in analytic dynamic traffic assignment

Most of the intersections on the Kruithuisweg arc signalised and equipped with loop-detectors. The municipality of Delft has brought the loops on-line, and Delft University of Technology is collecting this data within the framework of the Regiolab project. As noted in section 5.2.1, the travel-time along a path directly influences the assignment map, and therefore the set of cells associated with an observation. It is therefore essential to have reasonable estimates for travel time in order to be able to match flow observations to the correct dynamic O-D flows.

Since a significant part of the travel time on the Kruithuisweg is taken up by the waiting for traffic lights, these traffic lights should be modelled in the DTA. Within the framework of the DTA used, two approaches are possible: one approach is to represent the intersection by links whose capacity is set to zero when the traffic light is red (as in a simulation), which was used in Lindveld and Catalano (2002). The disadvantage of this approach is the increased storage requirements and lower performance of the DNL, which weigh heavily in an O-D matrix estimation setting. Another approach is to leave the capacity fixed and to model the average delay as a function of the crossing flows (with specialised link-performance functions as reported in Taale and Van Zuylen (2001)). For practical reasons it was decided to adopt the latter approach.
Akcelik’s delay time function

The delay time at intersections was modelled using a standard formulation found in the literature: Akcelik’s function for time-dependent delay (see Roupail (1992)):

$$\tau = \begin{cases} \frac{c(1-\frac{q}{c})^2}{2(1-\frac{q}{c})} & \text{if } x < 1 \\ \frac{(c-g)}{2} & \text{if } x \geq 1 \end{cases} + \frac{Q_0}{C}$$  \hspace{1cm} (10.1)$$

with:

- $x$: the degree of saturation; $x = \frac{3}{c} = \frac{3}{T}$, and $q$ the arrival flow rate
- $c$: the cycle length [sec]
- $g$: the effective green time
- $S$: queue outflow rate
- $Q_o$: the average queue length
- $C$: the capacity rate [veh/sec]

The average queue length is modelled as:

$$Q_o = \begin{cases} \frac{CT}{4} \left( x-1 + \sqrt{(x-1)^2 + \frac{12(x-x_0)}{CT}} \right) & \text{if } x > x_0 \\ 0 & \text{if otherwise} \end{cases}$$  \hspace{1cm} (10.2)$$

with:

- $T$: the cycle time [sec]
- $x_0$: the threshold level above which queues are expected to arise

$$x_0 = 0.67 + \frac{Sg}{600}$$  \hspace{1cm} (10.3)$$

Determination of the green time $g$

The green time $g$ on the intersection is in practice determined by the traffic controller for that intersection; on the Kruithuisweg these controllers react to the amount of traffic that approaches the intersection both from the lateral and the transversal directions. This leads to a classical problem in traffic assignment: non-separability of the link-cost functions. Whilst the VI formulation used for the DTA should be capable of handling this sort of problem, it is a complication whose detailed exploration does not materially add to the investigation of the problem at hand. It turns out that the green time can be approximated by a linear function of both the lateral and transversal flow based on microsimulation experiments reported in Miska (2003).
10.3.2 Link-performance functions

For all links the Smulders speed-density function (see Bliemer (2001)) was used, which is specified as follows

\[ v(\rho) = \begin{cases} 
 v^{ff} \left(1 - \frac{\rho}{\rho^{crit}}\right) & \text{if } 0 \leq \rho \leq \rho^{crit} \\
 v^{ff} \rho^{crit} \left(1 - \frac{1}{\rho^{jam}}\right) & \text{if } \rho^{crit} < \rho \leq \rho^{jam}
\end{cases} \quad (10.4) \]

with: \( v(\rho) \) the traffic speed as a function of traffic density on a network link, \( v^{ff} \) the free-flow traffic speed, \( \rho \) the traffic density on a network link, \( \rho^{crit} \) the "critical density" at which traffic changes from free-flowing to congested, and \( \rho^{jam} \) the maximum traffic density that is physically possible.

For links that end in an intersection, we have two options:

- use the Smulders function and ignore the error in travel time caused by the queuing delay
- use the Akcelik function.

In order to carry out the estimation on real data, the second option would have been necessary in order to get the link travel times right. For the synthetic data used in the numerical experiments however, this complication was left for further work and the pure Smulders function was used for all links.

10.3.3 Route choice

Route choice was not present in this network.

10.4 Experimental setup

As in chapter 9, four numerical experiments were carried out to test the estimator. These experiments differ only in the choice of the apriori matrix, as listed below:

- 1. apriori matrix equal to the true matrix
- 2. apriori matrix equal to the true matrix scaled down by a factor of 0.8
- 3. apriori matrix equal to the true matrix scaled up by a factor of 1.2
- 4. apriori matrix a randomly perturbed version of the true matrix

The first experiment is used to check if the estimator can recognise that it has the correct solution, and does not disturb this. The second and third experiments are used to verify that the estimator can scale an apriori matrix that has the correct structure up or down to match observations. The fourth experiment is used to investigate if the
estimator is able to return the right solution when the apriori matrix no longer has the correct structure in addition to being scaled.

The main difference between the experiments carried out here and those in chapter 9 is the size of the estimation problem: the number of parameters to be estimated is now: 13 zones x 12 time periods x 2 = 312, and the number of O-D matrix cells is 13 x 13 x 12 = 2028.

10.5 Estimation results

The estimation results shown below were obtained without departure time choice, using only production and attraction as observations. The estimation results were compared numerically, of which we reproduce the scatterplots of apriori and estimated O-D cells versus the true ones, scatterplots of estimated group observations versus true group observations, and a graph of the decrease of the objective function as the estimation progresses.

As in the previous chapter four estimation runs were carried out: one in which the apriori matrix was identical to the matrix used to obtain the observations, one in which the apriori matrix was scaled 20% up, one in which it was scaled 20% down, and one in which it was randomly perturbed.

10.5.1 Comments on computational aspects

The estimation results shown below were obtained using (truncated) analytical derivatives. Attempts to estimate O-D matrices for the Kruthuisweg case using numerical derivatives led to computation times of about 31 hours (on a 3 Ghz. pentium-4 Pc.); with analytical derivatives the computation time dropped to approximately 1 hour. This clearly illustrates the importance of using derivative information. Memory requirements (when all data was kept in-core) were about 500 Mb. for 13 O-D pairs and 12 time intervals for observations. Therefore scaling the estimator up to larger networks will present a computational challenge, in which the classical trade-off of memory versus speed appears: the memory-consuming assignment fractions may be written out to disk, and read whenever an evaluation of the objective function is needed. This should reduce the core memory requirement at the cost of increased computation time.

10.5.2 Apriori agrees with group observations

Shown in Figures 10.7, Fig_apri_exact_grpcm are scatterplots of the estimated group observations versus the observed, before and after estimation. Shown in figure 10.9 are the values of the objective function.
Figure 10.7: Estimation results with perfect apriori (O-D matrix cells)

Figure 10.8: Estimation results with perfect apriori (Group observations)

Figure 10.9: Objective function z by iteration number

The model clearly does not change the (correct) apriori matrix, except in minute ways which can be attributed to roundoff error in the software, and the limited number of decimals used in the specification of the apriori matrix. This shows that the estimator
recognises that the O-D matrix fits the observations, which constitutes a check on both the software and the estimator.

10.5.3 Apriori 20% higher than group observations

In this test, the true matrix scaled up by 20% was used as apriori matrix. Shown in Figure 10.10 are scatterplots of the estimated group observations versus the observed, before and after the estimation.

![Apriori cells versus true cells](image1)

![Estimated cells versus true cells](image2)

Figure 10.10: Estimation results with inflated apriori (O-D matrix cell values)

![Group values (before estimation)](image3)

![Group values (after estimation)](image4)

Figure 10.11: Estimation results with inflated apriori (group observations)
The model clearly succeeds in bringing the matrix back in line with the true values. The decrease of the objective function shown in Figure 10.12 shows that the estimator can quickly correct for large and obvious discrepancies, but that it is more difficult to remove smaller discrepancies. Note that the quality of the fit between true and estimated O-D matrix cells and the one between observed and estimated group counts is almost the same, which is probably caused by the fact that the structure of the apriori matrix is identical to that of the true O-D matrix.

10.5.4 Apriori 20% lower than group observations

In this test, the true matrix scaled down by 20% was used as apriori matrix. Shown in Figure are scatterplots of the estimated group observations versus the observed, before and after the estimation.

Figure 10.13: Estimation results with deflated apriori matrix (O-D matrix cells)
The estimation results in Figures 10.13 and 10.14 show that the estimator corrected the deflated apriori matrix. A number of large group observations and cell values remains uncorrected after 10 iterations. This is thought to be due to the weight matrix used (the inverse of the variance-covariance matrix; see eqn. 4.28), which essentially penalises the absolute error in the group observations and not its square, so that the estimator may get stuck in a local minimum. Note that as in section 10.5.3 the quality fit between true and estimated O-D matrix cells and between observed and estimated group counts is almost the same.
10.5.5 Randomised apriori matrix

![Graph 1](image1.png)

**Figure 10.16:** Estimation results with randomly perturbed apriori matrix (O-D matrix cells)

![Graph 2](image2.png)

**Figure 10.17:** Estimation results with a randomly perturbed apriori matrix (group observations)

![Graph 3](image3.png)

**Figure 10.18:** Objective function $z$ by iteration number
As shown in Figures 10.16 and 10.17 the estimator corrected the perturbed O-D matrix. The resulting fit of the group values is better than that for the O-D matrix cells, which can be attributed to the under-determinedness of the O-D matrix estimation problem, whereas the parameter estimation problem for the group observations is not under-determined. In view of the multiplicative model used (see eqn. 4.24) it is hypothesized that this is due to the random nature of the perturbations, which change the structure of the O-D matrix in ways the uniform deflation or inflation used in sections 10.5.3 does not.

The decrease of the objective function shown in Figure 10.18 again shows that the estimator is able to quickly correct for large discrepancies, but that it has more difficulties finding the best parameter estimate among the local minima.

10.5.6 Conclusions

The estimator was able to determine when it could no longer improve the O-D matrix, and was able to correct scaled-down and perturbed apriori matrices and was able to find a matrix that matches the observations. This shows that the O-D estimator is materially correct. It was also shown that the estimator can be applied to a modest but real network.

The estimation results show that it is possible for the estimator to be more successful at reconciling the differences between observed and estimated group counts than at reproducing the reproducing the true O-D matrix. This is seen as a consequence of the nature of the O-D matrix estimation problem (which is underdetermined), as opposed to the problem of determining a parameter vector to minimise the discrepancy between estimated and observed group counts (which is not underdetermined).
Chapter 11

Extensions and applications

11.1 Introduction

The model framework for O-D-t matrix estimation as introduced and verified in the preceding chapters is an example of an MPEC being applied to a transportation system, i.e. optimisation of a scalar objective function on traffic flows of a transportation system subject to the constraint that the underlying transportation system is in user-equilibrium. Many topical problems in transportation modelling can be cast as MPEC problems, e.g. optimal toll settings, optimal traffic control measures, optimal link-performance estimation, optimal parameters for route choice algorithms, and calibration of multi-modal route choice models. The behavioural models for route choice and departure time choice used in this thesis present great scope for improvement both from a theoretical and from a practical point of view. The numerical solution algorithms used in this thesis were generally selected for their simplicity and ease of use rather than for rigour or numerical efficiency. Whilst this is defensible in view of the large uncertainties associated with O-D matrix estimation in practice, they leave room for improvement. Finally, although pains were taken to use datastructures and algorithms capable of being scaled, further work is required in order to handle even medium-scale application (say in the order of 400 zones and 100 time periods).

Therefore four clear directions of further research can be distinguished:

- refinement of the solution algorithms
- improvement of the choice models used
- application extension of the framework developed here to other MPEC's
- scaling the software to handle medium to large-scale applications.

These will be discussed in subsequent sections.
11.2 Refinement of the solution algorithms

First of all, the solution methods presented in this thesis are to some extent heuristic, and can be improved and refined. The first thing that comes to mind is the solution of the MPEC problem (now carried out using Maher's heuristic). This solution method can be vastly improved from a theoretical viewpoint (e.g. it can be replaced by a method for which a convergence proof is available), as shown in Pariotsson and Rockafeller (2002) for the static case. A challenging (but clear-cut and widely applicable) research direction would be to extend their work to the dynamic case.

Next, a method of checking for and dealing with local minima is needed in view of the complexity of the optimisation problem and the absence of guarantees for unique minima.

It would also be useful if the estimator would be able produce confidence intervals for its estimates.

In addition, the solution of the VI problems that serve to define the equilibrium conditions, currently done using the MSA method, can benefit from the more rigorous (and more complicated) solution methods such as the projected gradient method presented in Nagurney (1993) and Chen (1999).

11.3 Improvement of the choice models used

The behavioural element was introduced into the framework through RUM-based discrete choice models for route choice and for departure time choice. Both the route choice models and the departure time choice models used in this thesis can be improved, both from a theoretical and from a practical point of view. From a theoretical point of view the use of more advanced choice models that are better able to capture the covariance structure of the problem begs for attention. More fundamentally, how to best capture human route choice behaviour in networks using choice models is still an active area of research.

That this is not just an academic pursuit is shown by the recent work of Ramming (2002) which shows that e.g. the basic problem of choice-set generation in car networks is far from solved. In addition, to the best of our knowledge, no completely satisfactory analytic route choice model for car networks exists to date.

11.4 Extension of the model to more general cases

The model framework that was operationalised in this thesis to solve the MPEC formulation of the dynamic O-D matrix estimation problem may be applied to other transportation planning problems formulated as MPEC's with relatively small modifications. In this section we will identify a number of extensions and possible applications of the model structure used.
11.4.1 Extension to multiple userclasses

We first note that the current framework can be extended to multiple user-classes and that it can incorporate both elastic and inelastic demand. In chapter 1 we claimed that the framework can be extended to take account of the supply-demand interaction of multiple user-classes. In terms of the framework this means a) the inclusion of a dynamic multi-userclass equilibrium assignment, as described in Blicmer (2001) and b) adaptation of the objective function to deal with two matrices simultaneously.

With respect to the estimation, the observations used should be able to distinguish between userclasses because otherwise the only information on the userclasses would come from the apriori matrix and from the assignment mapping as different userclasses would have different route choice probabilities. We note that although the current generation traffic measurement equipment can often distinguish between vehicle types (passenger cars, vans, lorries, busses), but cannot distinguish between userclasses on other grounds than vehicle characteristics, this may change with the advent of electronic vehicle identification tags.

11.4.2 Extension to a multimodal setting

One of the (currently unaddressed) problems in the research on multimodal transport modelling through supernetworks (see Carlier et al. (2002) and Catalano et al. (2001), for a description) is the lack of calibration tools for the transportation models that can be built using multimodal route choice models.

We expect that some form of calibration will prove necessary to account for geographic components in the utility function of multimodal routes. A specific example is the feeling of unsafety at certain train or metro stations, which is a geographic effect. It is therefore unlikely that a multimodal route choice model estimated on a survey that lacks the relevant geographic aspects will be able to capture such effects. Application of the estimated model in a network context will then lead to discrepancies between observed and predicted mode shares that vary by geographic location. This problem can be addressed by adding special, geographically based taste parameters, to e.g. the mode constant and calibrating these constants using the framework described in this thesis.

We also note that the current framework can be applied to estimate O-D matrices in multimodal transport on supernetworks of multimodal transportation analysis on supernetworks) as the mathematical structure is identical.
Figure 11.1: Adaptation of the model structure to incorporate multimodal transportation on supernetworks.

The model structure of chapter 3 can be adapted as shown in Figure 11.1; the difference consists of a more elaborate route choice model that gives path choice probabilities $P^r_{rs}(C_p, \Phi)$ that depend on path cost and a parameter vector $\Phi$. If the supernetwork contains subnetworks (e.g. for a particular mode, a particular route, or e.g. a bus line) then the O-D matrix for that subnetwork may be found by aggregating the path-flows that make use of this subnetwork at the entry- and exit point of this subnetwork. Such matrices are known as E-E (Entry-Exit) matrices.

11.4.3 Extension to elastic demand

In the exposition of the behavioural model to be used in dynamic O-D matrix estimation we have implicitly taken the view that total travel demand is inelastic with respect to destination (no destination choice allowed) and with respect to the total amount of traffic generated ("rigid demand").

In applications the amount of traffic generated is sometimes made flexible by extending the physical network with a "shadow route" between each O-D pair with an appropriate route cost, so that the equilibrium route cost through the real network determines the amount of traffic that will choose the physical network, and hence will make a real trip. As the amount of traffic on the network now depends on route choice, this is equivalent to estimating an O-D matrix that uses a particular mode, and we have reduced the problem to one that fits into the framework used.

From a theoretical point of view the inclusion of elastic demand merely adds one extra term to the VIP formulation of traffic equilibrium. From this perspective the only
thing that would have to change in the current software framework is to replace the combined route choice and departure time choice dynamic assignment with an elastic-demand route choice and departure time choice dynamic assignment, as proposed in Chen (1999).

**Extension to determine the uncongested O-D-t matrix.**

The O-D matrix estimator presented in this thesis produces an estimate of the dynamic O-D matrix that takes account of departure time rescheduling. Of considerable interest in planning studies is the O-D matrix that would have resulted if congestion were absent. It turns out that the O-D matrix estimation framework presented in this thesis allows one to estimate this matrix by reversing the departure time adjustment process on the O-D matrix estimate.

This procedure starts from the O-D-t matrix that would result if all travellers would experience only free-flow travel times on the entire network, and models their actual departure times as a re-routing within the STEN from their preferred departure time to their actual departure time. Such a matrix can be calculated using an iterative procedure, as proposed in Lindveld and Van Der Zijpp (2000) and demonstrated in Van Der Zijpp and Lindveld (2000). This process is modelled in so-called DUO-D& R (Dynamic User optimum-Departure time and Route) assignments. The input for such assignments consists of an O-D matrix with preferred departure times under uncongested conditions, and a choice model for the actual departure time given the expected travel times. In this context we will assume that the VOT and the schedule delay costs are given; these values can be estimated from SP and/or RP experiments. Another approach is the one presented in Teekamp et al. (2002),.

**Extension to dynamic O-D matrix estimation in the presence of traffic control**

It is likely that some sort of traffic control measures will be present on the transportation network under consideration. Such measures typically affect the travel time or the route cost or both, and are often responsive to the existing traffic situation as reflected in the traffic volumes. The impact of a given set of traffic control measures can be incorporated by enriching the network with special links that reflect the effect of the traffic control, or it can be modelled explicitly if the traffic control strategy is known. In the latter case the effect of traffic control can be represented as in 11.2.
To the O-D matrix estimation, the only difference would be a different DTA module that generates a different assignment map and different derivatives of link flows with respect to O-D matrix elements.

11.4.4 Estimation of dispersion parameters

Liu and Fricker (1996) and Yang et al. (2001), estimate dispersion parameter along with an O-D matrix in the static case. From a mathematical point of view this introduces one extra parameter, which can estimated along with the multitude of parameters that define the O-D matrix. Extension of this idea to the dynamic case is straightforward when viewed within the MPEC framework: as with the static case we have a single additional parameter to estimate, subject to the assumption of dynamic user equilibrium.

11.4.5 Estimation of optimal tolls and optimal traffic control

The problem of optimal traffic control, and the optimal toll problem which results when the control measures are restricted to tolling, are also examples of MPEC’s in that the determination of an optimal toll or optimal signal settings can be formulated as the minimisation of some aspect of the tolls (e.g. minimum total toll) subject to a user equilibrium for the traffic flows that result from those tolls. By replacing the module that minimises the objective function for O-D matrix estimation with one that optimises for signal settings or tolling, one could re-use the software framework for O-D matrix estimation to solve those problems. This would also constitute a more rigorous solution approach to the tolling application presented in Viti et al. (2003), .
Calculation of optimal traffic control settings can also be described as an MPEC in here one minimises an aspect of the signal settings (e.g. total delay) subject to a user equilibrium for the traffic flows that result from those settings. The interaction between the (upper level) traffic control and the (lower level) user choice, and its potential effect on the optimum control is illustrated in Bell and Allsop (1998), and the existence of multiple equilibria is shown in Taale and Van Zuylen (2003).

11.5 Summary

The solution methods used in this thesis can clearly be improved, which constitutes an avenue of further research. The models used to account for behavioural aspects (specifically route choice in car networks and departure time choice) are still in a state of flux and can certainly be improved. The formulation of the O-D-t matrix estimation problem as an MPEC, the model system proposed in this thesis to account for behavioural aspects, is an instance of a rich class of transportation planning and traffic control problems. Therefore the software framework prepared to solve the resulting MPEC can be applied to many other problems with relatively minor modifications. Last but not least, for dynamic O-D matrix estimation to be practically useful it must be capable of being applied up to real-life networks. Scaling up the software implementation to handle realistic networks would require a non-trivial amount of work in software development and in the search for efficient solution methods for large scale MPEC's.
Chapter 12

Conclusions

12.1 Introduction

In this thesis we have developed a new dynamic O-D matrix estimator that is founded on a behavioural basis. This estimator is able to produce a statistically sound time-dependent OD trip table estimate given flow observations and an apriori trip table, such that the estimated dynamic OD flows are consistent with dynamic assignment predictions of network flows based on the estimated dynamic OD table.

The set of unknown OD flows are calculated by means of a model parameterisations, the parameters of which are estimated using an optimization approach which ensures maximum agreement of resulting OD flows with observations and a priori information under conditions of dynamic equilibrium in the related network flows.

Apart from the dynamics included in the adopted exogenous information (an apriori dynamic OD matrix and dynamic traffic flow observations) the essential factors in dynamising the OD-table are a dynamic traffic assignment model and a flow-dependent departure time adjustment model.

Key elements in the developed estimator are therefore:

- an exogenous a priori dynamic OD-table;
- observations of groups of OD-flows and groups of route flows;
- a parametrisations of the endogenous OD-table to be estimated;
- a dynamic user-equilibrium assignment model;
- a departure time adjustment model
- a parameter optimization procedure.

In this chapter we will present the following:

- a brief outline of the problem at hand and its characteristics;
• a short summary of the work done in developing the estimator;
• an overview of essential new scientific achievements and contributions
• an outline of possible generalisations and extensions of the work done
• an evaluation of the results achieved in the thesis research.

12.2 The problem and its characteristics

The problem addressed in this is the estimation of a dynamic OD-trip table representing the interzonal travel demand as a function of time. In general, OD-t flows as such are not directly observable but may be estimated using other observations such as e.g. traffic counts and behavioural information embedded in travel models. Key element is how to determine the time-dependent pattern in the OD-flows such that consistency exists with observations and with assigned dynamic flows in the network.

This problem is characterised by the following issues:

1. In general, the available observations do not uniquely define the unknown OD-t table so that the estimation problem is highly underspecified.

2. Observations always come with observational errors and deviations from representative days which lead to a statistical estimation problem rather than a ‘fitting’ problem.

3. The system to which the observations as well as the target OD-t matrix refer is characterized by various forms of demand-supply interaction by which travel choices (destination, route, departure and arrival times) continuously are adapting to changing levels of service in the network.

4. The target OD-t matrix is characterized by a relatively stable underlying spatial and temporal structure that should be preserved to a sufficient degree in the estimation.

5. The available observations for estimation somehow have to be related to the targeted OD-t flows. This requires a highly advanced dynamic network assignment model from which a coupling of OD-flows and route flows to observations can be derived.

6. Key to the problem is the adequate consideration of the time-dependent travel behaviour of network users such as departure time choice, arrival time choice, and the like for which adequate behavioural models have to be used.

7. The many demand-supply interdependencies in the system make the estimation problem highly complex for which no standard optimization procedures are available.
12.3 Summary of the work done

We will now list the main elements of the work done to address these problem characteristics.

12.3.1 Behavioural framework of the estimator

In the thesis, a modular framework has been developed for the OD-t matrix estimation (see Chapter 7). At the heart of this framework is a multiplicative model which is used to parameterise the unknown matrix. The multiplicative model operates on an exogenous dynamic apriori matrix. The parameters of this model are estimated by optimizing an objective function that expresses the agreement between observations and predictions on the one hand and between a priori and estimated matrix on the other.

The estimated dynamic matrix is subject to adaptation through a departure time adjustment model, the choice probabilities of which depend on time-dependent travel costs which follow from a dynamic traffic assignment of the current matrix estimate. At the same time the assignment produces the assignment map which links flow observations to the estimated O-D-t matrix cells. In this way deviations between observed traffic flows and those expected from assignment of the current O-D-t matrix estimate are related to matrix cells, which in serves as input to the objective function. Part of the assignment procedure is a probabilistic route choice model the parameters of which are exogenous to the estimator.

12.3.2 Parameterisations to solve the underspecification

Through the use of a model parametrisation of the unknown O-D-t matrix, the underspecified O-D-t matrix estimation problem is reduced to a properly specified parameter estimation problem. The parameterisations of the unknown matrix consists of an exogenous apriori matrix and a parameter vector which are linked through a multiplicative model. This multiplicative models yields the OD-t matrix that is closest to the a priori matrix while exactly satisfying the observational and other constraints. By changing the constraints into objectives, and adjusting the parameters of the multiplicative model, the target matrix is obtained that is closest to the a priori matrix while maximizing the objectives. In this approach, a crucial role is played by the dynamic apriori matrix, which is assumed to be based on destination and departure time choice models.

12.3.3 Statistical estimation

In order to account for the uncertainties, errors and mutual contradictions in the observations a tractable distribution-free statistical estimation procedure has been developed based on Generalized Least Squares estimators. The chosen parameterisation, the multiplicative model mentioned above, ensures positivity of the OD-t matrix so that additional constraints in the estimation are superfluous.
12.3.4 Solution approach and use of derivatives

In view of the interdependence between the upper and the lower level the O-D matrix estimation is carried out iteratively. Given an O-D-t matrix estimate the equilibrium traffic flows are calculated, along with first-order derivatives of the equilibrium flows with respect to changes in the matrix. The flows corresponding to this equilibrium, along with their derivatives are used in a minimisation step that minimises the discrepancy between the counts and the equilibrium flows and between the apriori matrix and the O-D-t matrix estimate.

In this minimisation step, a descent direction for upper level minimisation problem is found using a first-order approximation of the derivatives of the lower level, as given by the assignment map.

12.3.5 Development and implementation of a DTA procedure

A DTA procedure was developed that determines a Stochastic Dynamic User Equilibrium, with respect to simultaneous choice of route and departure time, which also calculates and store the assignment map, in an existing Dynamic Network Loading model was used. As the derivative information is used as a proxy for the true response surface of the lower level game, the resulting game corresponds to a cross between a Stackelberg game and a Cournot game. In addition, the quality of the derivative information determines whether the resulting game is closer to a Cournot game or to a Stackelberg game, and hence the quality of the estimate.

12.4 Contributions of this thesis

A detailed list of contributions of this thesis is given in sections 1.4.1 and 1.4.2, of which we will review the main points here.

First of all we have developed a consistent behavioural framework is developed for dynamic O-D matrix estimation. This framework incorporates models for route choice and departure-time choice, or more specifically, departure time adjustment. Departure time adjustment is the most marked additional behavioural component which in dynamic O-D matrix estimation compared to static O-D matrix estimation. Departure-time adjustment can be viewed as a special type of route choice in an extended network: the Space Time Extended Network.

We have formulated the dynamic O-D matrix estimation problem as an optimisation problem constrained by a dynamic traffic equilibrium. In the dynamic case this leads to a general class of problems known as MPEC’s, with the equilibrium conditions given in terms of a Variational Inequality.

A known problem in O-D matrix estimation in general is the strong under-determinedness of the estimation problem: an O-D matrix typically has more unknown cells than independent observations are available, and many O-D matrices exist that will for a given set of observations. Various methods have been proposed in the literature to resolve
this problem. In this thesis we extend an existing parametrisation (involving an apriori matrix and all available observations) to the dynamic case, and show that it requires 50 to 100 times fewer independent parameters than the approach in which all O-D matrix cells are considered as free parameters. The trade-off is that somewhat more emphasis is placed on the quality of the apriori matrix.

We have adapted existing DTA software to fit the needs of the O-D matrix estimation in several ways. First of all the we altered the software to explicitly calculate and store the dynamic assignment map. This assignment map is needed to disentangle aggregate flow observations (such as traffic counts) into their constituent O-D matrix cells. The assignment map also serves as an approximation of the derivative of traffic flow with respect to the O-D-t matrix cells, which is needed to determine a suitable descent direction for the O-D matrix estimation in the top level optimisation problem. Secondly we extended the DTA software with stochastic route choice models and departure time adjustment models in such a way that these models can be replaced by more sophisticated models.

We integrated the DTA module with software for solving the MPEC, and demonstrated the resulting program on test networks and a small case study.

12.5 Avenues for further work

A number of extensions and application of the behavioural framework developed in this thesis have been listed in Chapter 11, of which we mention improved solution methods, confidence intervals for the estimates produced, extension of the software to handle larger problem instances, refinement of the choice models used, estimation of dispersion parameters, and application of the framework to other MPEC problems such as pricing and traffic control.

12.6 Evaluation

12.6.1 Verification

On basis of this framework, an O-D-t matrix estimator was implemented in software and tested on a synthetic network. In the problem instances examined, the O-D matrix estimator was able to retrieve dynamic O-D matrices that were deliberately deformed. On basis of these tests it was concluded that the O-D-t matrix estimator works properly.

12.6.2 Practical significance

Inasmuch as this thesis demonstrates a consistent behavioural framework for O-D-t matrix estimation on general networks, it is of relevance for the development of O-D-t matrix estimators for a full-size (say at least 400 zone) network. Further work is needed to implement the estimator for larger networks, but this could be restricted to careful software engineering.
12.6.3 Advantages of the approach used

The advantages of the approach towards O-D-t matrix estimation followed in this thesis are:

- the estimator always gives an answer, whose quality gracefully degrades as the amount of available observational data declines; if no observational data is available at all, the apriori matrix is returned. This provides much flexibility in the amount of observational data needed for estimation (provided we are prepared to trust the structure present in the apriori matrix).

- due to the use of generalised flow observations, the estimator can be extended to work with e.g. triplength distributions, observed partial matrices, and cordon counts.

- the estimator returns an O-D-t matrix that is consistent with the modelled demand-supply interaction mechanism, and best fits available observations.

- the DTA model used provides a framework into which any desired route-choice model and departure time adjustment model may be plugged in, thus allowing more sophisticated behavioural model to be used as they become available.

- due to the modular structure of the DTA, the DNL may be extended to incorporate queuing without any consequences for the O-D matrix estimator.

- due to the generality of the DTA model, the estimator may be:

  - applied to multi-userclass situations
  - applied to other modalities, such as public transport, or inland shipping
  - applied to multi-modal situations
  - extended to estimate elastic-demand matrices
  - used to incorporate specialist statistical O-D-t matrix estimators such as the one developed by Van Der Zijpp (1996), thereby operationalising these to a general network context.

It is clear that there is scope for follow-up research.

12.6.4 Reflection on the assumptions made

The framework on which the O-D matrix is based makes two crucial assumptions:

- the equilibrium hypothesis leads to useful results
- the apriori matrix is not too far wrong in ways that cannot be corrected by the estimation
The equilibrium hypothesis

The O-D-t matrix estimator is largely tied to the equilibrium concept, without considering e.g. day-to-day dynamics. We believe that the structure of the daily traffic pattern is solid enough that the equilibrium hypothesis, whilst strictly speaking false, is close enough to reality to provide valuable insights into the transportation system. In addition the equilibrium hypothesis seems to provide the best avenue for forecasting future developments in a consistent way.

Quality of the apriori matrix

In view of the parametrisation used, the estimator is vulnerable to errors in the apriori matrix on which no observational evidence is available to the estimator. The main issue would probably be the structure of the O-D matrix. In this respect it is necessary to extend some trust to the apriori matrix. We feel that carefully built behavioural models for destination choice and departure time choice are an acceptable source for apriori matrices.
BIBLIOGRAPHY


Dial, R.B. (1971) A probabilistic multipath traffic assignment algorithm which obviates path enumeration. Transportation Research 5, 83-111


time estimation. Estimators of travel time for road networks. eds. Bovy, P.H.L., Thijs, 

time estimators on the Amsterdam orbital motorway: Evaluation results from the 
DACCORD-project, Delft University of Technology. Transportation Planning and 
Traffic Engineering section, Research Report VK2205.335, Delft.

Evaluating travel time estimators on the Padua-Venice corridor: Evaluation results from 
the DACCORD-project. Delft University of Technology. Faculty CiTG. Transportation 

travel time estimators and predictors. Transportation Research Record 1719, 45-53.


network. Transportation Research 30A, 287-305.

Lundgren, J., Peterson, A. (2002) An heuristic for the bilevel origin destination ma-
trix estimation problem. Proceedings of the 13th Mini-Euro conference and the 9th 
Meeting of the EURO working group on Transportation, Bari 2002, pp. 647-651.

user-optimal assignment problem. Transportation Research 36B, 421-443.

Cambridge University Press.

Maher, M. (1983) Inferences on trip matrices from observations on link volumes: a 
Bayesian statistical approach. Transportation Research 17B, 435-447.

matrix estimation and traffic control problems with stochastic user equilibrium link 
flows. Transportation Research 35B, 23-40

sher, D.A. and Button, K.J.. Pergamon.

Guidelines from an Application to Paris. Transportation Research Record 1752, 76-
38.


Academic press.

Restrospective. Unpublished manuscript presented at the International Association


Bibliography


Wardrop (1952) Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers*. Part II, 325-278


Appendix A

Calculation of dynamic assignment fractions

In this appendix we detail the calculation of dynamic assignment fractions which the DTA algorithm used does not provide. Assignment fractions can be defined in two ways: one can base the definition on the time period $k$ during which traffic that departed during a certain time period $h$ is present on a certain link $a$ (as is usual in static assignments), or one can base the definition on when it enters or exits link $a$. We will use the latter definition in order to maximise consistency with ordinary traffic counts, which are based on traffic flowing past a detector rather than being present on a link. In addition this results in a more sharply defined criterion, as traffic entering link $a$ may remain on that link during several simulation time intervals.

The relationship between the link outflows and the matrix cells is given by eqn (6.42), with the assignment fractions $\varphi_{aphk}^a$ defined as:

\[
\varphi_{aphk}^a = \begin{cases} 
0 & \text{if the pathflow from period } h \text{ does not reach link } a \text{ during period } k \\
\varphi & \text{if the pathflow from period } h \text{ has already left link } a \text{ during period } k \\
1 & \text{if the pathflow from period } h \text{ partially exits link } a \text{ during period } k \\
\varphi (0, 1) & \text{if the entire pathflow from period } h \text{ exits link } a \text{ during period } k 
\end{cases}
\]

(A.1)

Defining:

<table>
<thead>
<tr>
<th>The time instant</th>
<th>As the time at which:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{a,h}^{entry} = h\delta + \tau_{pa}(h\delta)$</td>
<td>traffic starting at time $h$ enters link $a$</td>
</tr>
<tr>
<td>$t_{a,h}^{exit} = t_{a,h}^{entry} + \tau_{a}(t_{a,h}^{entry})$</td>
<td>traffic starting at time $h$ exits link $a$</td>
</tr>
<tr>
<td>$t_{a,(h+1)}^{entry} = (h + 1)\delta + \tau_{pa}((h + 1)\delta)$</td>
<td>traffic starting at time $(h + 1)$ enters link $a$</td>
</tr>
<tr>
<td>$t_{a,(h+1)}^{exit} = t_{a,(h+1)}^{entry} + \tau_{a}(t_{a,(h+1)}^{entry})$</td>
<td>traffic starting at time $(h + 1)$ exits link $a$</td>
</tr>
</tbody>
</table>

we get:

$\Delta t_a = t_{a,(h+1)}^{exit} - t_{a,h}^{entry}$ the time at which traffic from the O-D cell starting between $h$ and $(h + 1)$ is on link $a$. 
The following six possibilities exist:

Figure A.1: Six possibilities for incidence of traffic departing during time interval $h$ on link $a$ during time interval $k$

we have the following expressions for the assignment fractions:

\[ \varphi_{aphk}^g = \begin{cases} 
0 & \text{if } \left( \frac{t_{exit}^a - t_{exit}^{a(h+1)}}{\Delta t_a} \leq k\delta \right) \ \ (I) \\
0 & \text{if } \left( \frac{t_{exit}^a - t_{exit}^{a(h+1)}}{\Delta t_a} \geq (k+1)\delta \right) \ \ (II) \\
\left( \frac{t_{exit}^a - t_{exit}^{a(h+1)}}{\Delta t_a} - k\delta \right) & \text{if } \left( \frac{t_{exit}^a - t_{exit}^{a(h+1)}}{\Delta t_a} < k\delta \right) \\
\left( \frac{t_{exit}^a - t_{exit}^{a(h+1)}}{\Delta t_a} - k\delta \right) & \text{if } \left( \frac{t_{exit}^a - t_{exit}^{a(h+1)}}{\Delta t_a} < k\delta \right) \\
\left( \frac{t_{exit}^a - t_{exit}^{a(h+1)}}{\Delta t_a} - k\delta \right) & \text{if } \left( \frac{t_{exit}^a - t_{exit}^{a(h+1)}}{\Delta t_a} < k\delta \right) \\
\left( \frac{t_{exit}^a - t_{exit}^{a(h+1)}}{\Delta t_a} - k\delta \right) & \text{if } \left( \frac{t_{exit}^a - t_{exit}^{a(h+1)}}{\Delta t_a} < k\delta \right) \\
1 & \text{if } \left( \frac{t_{exit}^a - t_{exit}^{a(h+1)}}{\Delta t_a} < k\delta \right) \ \ (VI) 
\end{cases} \]
We note that in this definition of assignment fractions we implicitly split O-D flows, which can be criticized on the basis that the implementation of the DTA used in this thesis does not. The counter-argument is that this way of defining the assignment fractions is more logically consistent than any of the alternatives, which would consist of assigning the entire O-D-t cell to one period or another. If one adopts this practice however, the O-D-t matrix estimation problem as a whole will no longer have a fixed-point solution, but may result in an oscillating solution. In combination with the difficulties of obtaining numerical derivatives this effectively obscures the question of whether a fixed-point has been reached or not, thereby making it more difficult to test and debug the software implementation of the estimator. For this reason it was decided to adopt the above "hairsplitting" definition of the assignment fractions, and to implement this definition.

Algorithm for calculation of the assignment fractions

The idea of the algorithm is simply to track every O-D-t cell through all the links of all paths that it can use, and to calculate its arrival and departure times for each link on each path.

Step 0: (Carry out DTA)
Complete DTA algorithm

Step 1: (Calculate assignment fractions)
1 : for i do
   1.1: for (r, s) do
      for p ∈ Prs do
while \( a \in p \) do

\( [\tau_{in}, \tau_{out}] = \text{Calculate\_arrival\_and\_departure\_times}(r,s,p,a); \)
Calculate\_assignment\_fractions using A.2;
Store\_assignment\_fractions\_in\_cells;
\( a = a \rightarrow \text{nextlink} \)
endwhile \( a \in p \)
endfor \( p \in P_{rs} \)
endfor \( (r,s) \)
endfor \( i \)

determine path travel times from the start and end of H to each link a on the path

![Diagram](image)

**Figure A.2: Calculation of assignment fractions**

\( ([\tau_{pa,in}^{rs}, \tau_{pa,out}^{rs}] = \text{Calculate\_arrival\_and\_departure\_times}(r,s,p,a)) \)

if \( a \) is first link of \( p \) then { 
  for \( i = 0 \) to NSubIntervals { 
    \( \tau_{in}(i) = 0; \)
    \( \tau_{out}(i) = 0; \)
  }
}
else { 
  // link \( a \) is not the first link of the path
for i=0 to NSubIntervals {
    \( \tau_{in}(i) = \tau_{out}(i); \)
}
for i=0 to NSubIntervals {
    \( k_{in} = \text{Round} \left( \frac{\tau_{in}(i)}{\delta} \right); \)
    \( t_{out} = \tau_{in}(k_{in}) + \tau_{a}(k_{in}); \)
    \( k_{out} = \text{Round} \left( \frac{t_{out}}{\delta} \right); \)
    \( \tau_{out}(k_{out}) = t_{out}; \)
}

(Calculate_assignment_fractions(r,s,p,a,\tau_{in},\tau_{out}))
if a is first link of p then {
    for i=0 to NSubIntervals {
        \( \tau_{in}(i) = 0; \)
        \( \tau_{out}(i) = 0; \)
    }
}
else {
    // link a is not the first link of the path
    for i=0 to NSubIntervals {
        \( \tau_{in}(i) = \tau_{out}(i); \)
    }
    for i=0 to NSubIntervals {
        \( t_{out} = \tau_{in}(k_{in}) + \tau_{a}(k_{in}); \)
        \( \tau_{out}(k_{out}) = t_{out}; \)
    }
}

Appendix C: Adaptation of the C-load algorithm to calculate derivatives

The C-load algorithm works as follows:

Algorithm 5 C-load

Step 0: (initialisation)
\( \Delta = \min_{a} D_{a}(0) \)
\( M = \max \left( \frac{\Delta}{\delta}, 1 \right) \) is the number of \( \delta \) intervals within a \( \Delta \) interval.
\( X_{ap}^{rs}(0) = U_{ap}^{rs}(0) = V_{ap}^{rs}(0) = 0 \)
\( \frac{X_{ap}^{rs}(0)}{\delta} = \frac{U_{ap}^{rs}(0)}{\delta} = \frac{V_{ap}^{rs}(0)}{\delta} = 0 \)
\( i = -1 \)
1: for \( i = 0 \) to \( N \) do

Step 1: (solve the system of equations within the \( i \)th interval)

1.1: for \( a \in A \) and \( p \) passing through \( a \) do

for \( k = iM \) to \([ (i+1)M - 1 ] \) do

\[
V_{ap}^{rs} (k\delta) = \sum_{j \in \{ j \mid 0 \leq j \delta - \tau_a (j\delta) \leq (k-1)\delta \}} v_{ap}^{rs} (j\delta) \delta
\]

\[
\frac{\partial V_{ap}^{rs} (k\delta)}{\partial T_p^{rs}} = \sum_{j \in \{ j \mid 0 \leq j \delta - \tau_a (j\delta) \leq (k-1)\delta \}} \frac{\partial v_{ap}^{rs} (j\delta)}{\partial T_p^{rs}} \delta
\]

\[
\frac{\partial V_{ap}^{rs} (k\delta)}{\partial T_a^{rs}} = \sum_{j \in \{ j \mid 0 \leq j \delta - \tau_a (j\delta) \leq (k-1)\delta \}} \frac{\partial v_{ap}^{rs} (j\delta)}{\partial T_a^{rs}} \delta
\]

\[
v_r^s (k\delta) = \sum_{\tau ap} v_{ap}^{rs} (k\delta)
\]

\[
v_a (k\delta) = \sum_{\tau ap} v_{ap}^{rs} (k\delta)
\]

\[
\frac{\partial v_r^s (k\delta)}{\partial T_a^{rs}} = \sum_{\tau ap} \frac{\partial v_{ap}^{rs} (k\delta)}{\partial T_a^{rs}}
\]

endfor

endfor

1.2: for \( a \in A \) and \( p \) passing through \( a \) do

for \( k = iM \) to \([ (i+1)M - 1 ] \) do

\[
u_{ap}^{rs} (k\delta) = \begin{cases} f_p^{rs} (k\delta) & \text{if } a \text{ is first link on path } p \\ v_{ap}^{rs} (k\delta) & \forall k, \forall (r, s), \forall p \in \mathcal{P}_{rs} \end{cases}
\]

\[
u_{ap}^{rs} (k\delta) = \begin{cases} f_p^{rs} (k\delta) & \text{if } a \text{ follows directly after } a' \\ v_{ap}^{rs} (k\delta) & \forall k, \forall (r, s), \forall p \in \mathcal{P}_{rs} \end{cases}
\]

\[
\frac{\partial v_{ap}^{rs} (k\delta)}{\partial T_p^{rs}} = \begin{cases} f_p^{rs} (k) & \text{if } a \text{ is first link on path } p \\ v_{ap}^{rs} (k\delta) & \forall k, \forall (r, s), \forall p \in \mathcal{P}_{rs} \end{cases}
\]

\[
\frac{\partial v_{ap}^{rs} (k\delta)}{\partial T_a^{rs}} = \begin{cases} v_{ap}^{rs} (k\delta) & \text{if } a \text{ follows directly after } a' \\ v_{ap}^{rs} (k\delta) & \forall k, \forall (r, s), \forall p \in \mathcal{P}_{rs} \end{cases}
\]

\[
U_{ap}^{rs} (k\delta) = \sum_{j=0}^{k} v_{ap}^{rs} (j\delta) \delta
\]

\[
X_{ap}^{rs} (k\delta) = U_{ap}^{rs} (k\delta) - V_{ap}^{rs} (k\delta)
\]

\[
\tau_a (k\delta) = D_a (X_{ap}^{rs} (k\delta))
\]

endfor

endfor

endfor

Step 2: (Stopping criterion)

If the network is empty

then \( T_\infty = i \Delta \); stop

else: \( i = i + 1 \); goto step 1

Note that the algorithm is a one-to-one implementation of the model defined in (6.27) - (6.38), and that no dynamic assignment fractions are calculated. Also the algorithm does not calculate any derivatives of the flow with respect to O-D-t matrix cells.

Illustration of the calculation  The expression for \( \varphi_{apKH}^{rs} \) can be interpreted as in Figure A.3. Set out on the horizontal axis is the time, on the vertical axis the distance travelled. As we assume the FIFO property holds, trajectories starting and ending at the beginning and end of small intervals do not intersect.
The amount of traffic exiting link $l_1$ during period $K$ that entered during period $H$ is shown shaded. Note that not all traffic that entered the link during period $H$ exits during period $K$: only the traffic entering during time periods $k_3$, $k_4$, $k_5$ exits in period $K$, in periods $k_{11}$, $k_{12}$, $k_{13}$ respectively. The trajectories exiting $l_1$ during period $K$ enter the link during $H$. Therefore the assignment fraction

$$\varphi_{a,p}^{r,s} = \frac{v_{ap}^{r,s}(k_{11}) + v_{ap}^{r,s}(k_{12}) + v_{ap}^{r,s}(k_{13})}{T_{k_{11}}^{r,s} + T_{k_{12}}^{r,s} + T_{k_{13}}^{r,s} + T_{k_{14}}^{r,s}}$$

(A.4)

The trajectories are piecewise linear, and traffic is moved across trajectory boundaries at the end of every timestep. In the calculation of the assignment fractions it is necessary to account for the fact that the intersection of the trajectories at the boundaries of time periods when traffic enters the network, with links in general do not so on gridpoints. In general a mixture of O-D-t cells will enter and exit links during any small time period. This is illustrated in Figure (A.4), which shows the trajectories of flow packets that entered the network and exit link $l_1$ during time periods $k_{11} - k_{14}$.

The time instant that a flow packet that entered the network during time period $h$ leaves link $a$ is denoted as: $t_{a,h}^{r,s}$

As can be seen from the figure, the proportion of the link flow entering the network during time period $k_3$ is equal to: $\varphi_{1hk} = \frac{t_{1,k_3}^{r,s} - t_{1,k_2}^{r,s}}{t_{1,k_3}^{r,s} - t_{1,k_2}^{r,s}}$
Figure A.4: Calculation of assignment fractions: close-up

The flow exiting link 1 during time period $k_{12}$ is equal to: $\varphi_{1k_4k_{12}} * f_{k_3} + \varphi_{1k_4k_{12}} * f_{k_4}$ or equivalently: $\varphi_{1k_3k_{12}} P_{r_p} T_{r_p} + \varphi_{1k_4k_{12}} P_{r_p} P_{r_p}$. This gives an explicit expression for the composition of the outflow of link 1 in terms of the O-D-t matrix cells.
Appendix B

Calculation of dynamic assignment fractions for attraction

As noted in section , page , in the dynamic case group the assignment fractions related to observations for zonal attraction are affected by the assignment in ways that are absent from the classical static O-D matrix estimation. The reason is that the travel-time between origin and destination depends on the network state, so that when traffic from a certain O-D-t cell arrives at a certain destination during time interval $k$, the time interval $h$ during which it departed from the origin is given by

$$h = t - \tau^{rs} (h)$$  \hspace{1cm} (B.1)

which constitutes a fixed-point problem. Fortunately this fixed-point problem is easily solved when the solution to the DTA problem is at hand: one simply starts off with an initial estimate $h_0 = t - \tau^{rs}_{\text{min}}$ with $\tau^{rs}_{\text{min}}$ the minimum travel-time between $r$ and $s$. The estimate $h_0$ constitutes the latest possible time interval during which traffic could have departed to arrive during time interval $k$. The true departure time interval $h$ can then be found by searching backwards from $h_0$ until eqn. B.1 is satisfied.
Appendix C

Derivatives of the matrix w.r.t. to parameters

C.1 Derivative of the matrix with respect to parameter $\Lambda$

The derivative of matrix cells with respect to the parameter vector $\Lambda$ now involves the assignment fractions. For implementation it is practical to express this derivative in terms of the derivative of the logarithm of $T_h^{rs}$:

$$\ln T_h^{rs} = \ln T_h^{rs,0} + \sum_{l=1}^{L} \lambda_{lk} \cdot f^l P_{lk}^{rs}$$

The relationship between the two derivatives is:

$$\frac{\partial T_h^{rs}}{\partial \lambda_{lk}} = T_h^{rs} \frac{\partial \ln T_h^{rs}}{\partial \lambda_{lk}}$$

and

$$\frac{\partial \ln T_h^{rs}}{\partial \lambda_{lk}} = f^l P_{lk}^{rs}_{hy}$$

so that:

$$\frac{\partial T_h^{rs}}{\partial \lambda_{lk}} = T_h^{rs} f^l P_{lk}^{rs}$$

C.2 Calculation of derivatives $\frac{\partial S_{lk}}{\partial \lambda_{mn}}$

The calculation of $\frac{\partial S_{lk}}{\partial \lambda_{mn}}$ can be split into two parts: the special case where $mn = lk$, and the general case where $mn \neq lk$. 
The group observation $S_{lk}$ is:

$$S_{lk} = \sum_{(r,s,h) \in I_{lk}} f^l_{\varphi_{lthk}} T^s_h$$

$$\frac{\partial S_{lk}}{\partial \lambda_{mn}} = \frac{\partial}{\partial \lambda_{mn}} \sum_{(r,s,h) \in I_{lk}} f^l_{\varphi_{lthk}} T^s_h = \sum_{(r,s,h) \in I_{lk} \cap I_{mn}} \left( f^l_{\varphi_{lthk}} T^s_h \left( \Lambda \right) + f^l_{\varphi_{lthk}} \frac{\partial T^s_h}{\partial \lambda_{mn}} \right)$$

If we approximate this by assuming that the assignment fractions $\varphi^s_{lthk}$ are not affected by a change in the matrix, we get:

$$\frac{\partial S_{lk}}{\partial \lambda_{mn}} = \sum_{(r,s,h) \in I_{lk} \cap I_{mn}} f^l_{\varphi_{lthk}} \frac{\partial T^s_h}{\partial \lambda_{mn}}$$

Substituting the expression for $\frac{\partial T^s_h(\Lambda)}{\partial \lambda_{mn}}$ gives:

$$\frac{\partial S_{lk}}{\partial \lambda_{lk}} = \sum_{(r,s,h) \in I_{lk} \cap I_{mn}} f^l_{\varphi_{lthk}} T^s_h f^m_{\varphi_{mn}}$$

**Special case:** $mn = lk$

In the special case that $mn = lk$ (i.e. we evaluate the effect of a change in a parameter on its own observation group), we obtain:

$$\frac{\partial S_{lk}}{\partial \lambda_{lk}} = \sum_{(r,s,h) \in I_{lk}} f^l_{\varphi_{lthk}}^2 T^s_h$$

**Structure of $I_{lk} \cap I_{mn}$**

In general $I_{lk} \cap I_{mn}$ has to be calculated by actually comparing the lists $I_{lk}$ and $I_{mn}$, and taking the intersection. For special observations such as zonal production and attraction the structure of $I_{lk} \cap I_{mn}$ is simplified and can be calculated in advance.

**Structure of $\frac{\partial S_{lk}}{\partial \lambda_{mn}}$**

When we restrict ourselves to the observation groups production (I), attraction (II) and traffic counts (III), the following cases can be distinguished:

<table>
<thead>
<tr>
<th>$l$</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>II</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>III</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>
Depending on the case, and representing $S_{ik}$ and $\lambda_{ik}$ as vectors, the structure of $\frac{\partial S_{ik}}{\partial \lambda_{ik}}$ can often be simplified and calculated in advance:

\[
\begin{pmatrix}
(production) & (attraction) & (counts) \\
1 & \emptyset & 1 \ldots 1 & ? \ldots ? \\
\emptyset & 1 & 1 \ldots 1 & ? \ldots ? \\
1 \ldots 1 & 1 & \emptyset & ? \ldots ? \\
\emptyset & \emptyset & \emptyset & ? \ldots ? \\
1 \ldots 1 & \emptyset & 1 & ? \ldots ? \\
? \ldots ? & ? \ldots ? & ? \ldots ? & 1 \ldots ?
\end{pmatrix}
\]
Index

Commonality Factor, 111

DDUE, 80
DTA, 12, 79
DTC-DSUE, 103
DUE, 72
  Dynamic Traffic Assignment, 79
GLS, 56
LOS, 110
MPEC, 12
O-D, 1
Poisson distribution, 60
SDTAC, 84
SDUE, 80
STEN, 31
SUE
  Stochastic User Equilibrium, 82
User Equilibrium, 36
VI, 80
Summary

In this thesis we have developed a new dynamic O-D matrix estimator that is founded on a behavioural basis. This estimator is able to produce a statistically sound time-dependent OD trip table estimate given flow observations and an apriori trip table, such that the estimated dynamic OD flows are consistent with dynamic assignment predictions of network flows based on the estimated dynamic OD table.

The set of unknown OD flows are calculated by means of a model parameterisations, the parameters of which are estimated using an optimization approach which ensures maximum agreement of resulting OD flows with observations and a priori information under conditions of dynamic equilibrium in the related network flows.

Apart from the dynamics included in the adopted exogenous information (an apriori dynamic OD matrix and dynamic traffic flow observations) the essential factors in dynamising the OD-table are a dynamic traffic assignment model and a flow-dependent departure time adjustment model.

Key elements in the developed estimator are:

- an exogenous a priori dynamic OD-table;
- observations of groups of OD-flows and groups of route flows;
- a parameterisations of the endogenous OD-table to be estimated;
- a dynamic user-equilibrium assignment model;
- a departure time adjustment model;
- a parameter optimization procedure.
This problem is characterised by the following issues:

1. In general, the available observations do not uniquely define the unknown OD-t table so that the estimation problem is highly underspecified.

2. Observations always come with observational errors and deviations from representative days which lead to a statistical estimation problem rather than a 'fitting' problem.

3. The system to which the observations as well as the target OD-t matrix refer is characterized by various forms of demand–supply interaction by which travel choices (destination, route, departure and arrival times) continuously are adapting to changing levels of service in the network.

4. The target OD-t matrix is characterized by a relatively stable underlying spatial and temporal structure that should be preserved to a sufficient degree in the estimation.

5. The available observations for estimation somehow have to be related to the targeted OD-t flows. This requires a highly advanced dynamic network assignment model from which a coupling of OD-flows and route flows to observations can be derived.

6. Key to the problem is the adequate consideration of the time-dependent travel behaviour of network users such as departure time choice, arrival time choice, and the like for which adequate behavioural models have to be used.

7. The many demand-supply interdependencies in the system make the estimation problem highly complex for which no standard optimization procedures are available.

These issues was addressed within a behavioural framework, in which the O-D matrix estimation proper is related to the response of transportation system to changes in the O-D matrix. The result is a flexible O-D matrix estimator, which has been operationalised in software for modest problem instances, and whose proper functioning has been verified on synthetic data. The estimator itself is considered to be capable of being scaled up to practical problem sizes, and of being adapted to other types of optimisation problems in transportation that use dynamic stochastic user equilibrium as a constraint.
Samenvatting

In dit proefschrift is een nieuwe schatter ontwikkeld voor dynamische Herkomst-Bestemmings matrices, gestoeld op een gedragsmatige basis. Deze schatter is in staat om een statistisch correcte tijdsafhankelijke H-B matrix te schatten op basis van verkeersstroom waarnemingen en een apriori matrix, en wel zodanig dat de geschatte dynamische H-B stromen consistent zijn met de stromen die voorspeld worden door dynamische toedeling van de geschatte H-B matrix.

De H-B stromen worden berekend door middel van een modelmatige parametrisatie van de onbekende H-B matrix, waarvan de parameters geschat worden door middel van een optimalisatieprocedure. Deze optimalisatieprocedure zorgt voor maximale overeenstemming zowel tussen de uit de H-B matrix voortvloeiende verkeersstromen en de waargenomen stromen alsook tussen de geschatte H-B matrix en de apriori matrix.

Kernpunten van de ontwikkelde schatter zijn:

- een (exogene) dynamische apriori H-B matrix;
- waarnemingen van groepen H-B-stromen en groepen route stromen;
- een parametrisatie van de (endogene) te schatten H-B matrix;
- een model voor dynamisch gebruikersevenwicht in het verkeerssysteem;
- a model voor vertrektijdstipaanpassing;
- een parameter optimalistie procedure.
Het schattingsprobleem kan als volgt gekarakteriseerd worden:

1. In het algemeen leggen de waarnemingen de onbekende H-B matrix niet vast, zodat het schattingsprobleem sterk ondergespecificeerd is.

2. Waarnemingen zijn altijd behept met diverse fouten en afwijkingen van de representatieve gemiddelde dag, hetgeen leidt tot een statistisch schattingsprobleem

3. Het verkeersstelsel waarop de waarnemingen en de apriori matrix betrekking hebben wordt gekenmerkt door verscheidene vraag-aanbod interacties, waarbij keuzegedrag (keuze van bestemming, route, vertrektijdstip, en aankomsttijd) continu aan de heersende omstandigheden in het netwerk worden aangepast.

4. De apriori matrix wordt gekenmerkt door relatief stabiele onderliggende ruimtelijke en temporele structuren, die in de schatting in voldoende mate behouden moeten blijven.

5. De beschikbare waarnemingen zijn niet direct te herleiden tot individuele matrixcellen, en moeten daaraan gerelateerd worden door middel van additionele informatie. Deze informatie wordt verkregen uit een geavanceerd dynamisch netwerk toedelingsmodel.

6. Een essentiële kwestie is de adequate modellering van het tijdafhankelijke reisgedrag van de reizigers in het verkeersstelsel, zoals routekeuze, vertrektijdstipkeuze, en keuze van aankomsttijd.

7. De diverse vraag-aanbod interacties in het systeem leiden tot een bijzonder gecompliceerd schattingsprobleem waarvoor geen standaardsoftware beschikbaar is.

Het multiplicatieve model maakt het mogelijk om reeds bekende informatie over de ruimtelijke en temporele structuren (zoals genoemd in punt (4)) als vertrekpunt in te brengen in de schatting, en door middel van statistische schatting van de parameter vector de nodige aanpassingen te verrichten. De schatting kan omgaan met de (2) genoemde meetfouten, en is uitvoerbaar omdat de parameertensor zodanig wordt opgezet dat met iedere parameter een onafhankelijke waarneming correspondeert. Het onder (7) genoemde schattingsprobleem wordt iteratief opgelost, waarbij gebruik gemaakt wordt van de afgeleiden van het verkeers(toedelings)model dat als randvoorwaarde dient voor de schatting.

Het resultaat is een H-B matrixschatter met een zeer flexibele databehoefte die in software geoperationeleiseerd voor bescheiden netwerken, en waarvan de correcte werking geverifieerd is op synthetische data. De schatter zelf wordt gezien als opschalbaar naar probleemgroottes zoals die in de praktijk voorkomen, en wordt tevens gezien als voldoende flexibel om te kunnen worden toegepast op andere optimalisatieproblemen in de vervoersmodellering die een stochastisch dynamisch gebruikersevenwicht als randvoorwaarde stellen.
About the author

Charles Lindveld was born in 1961 in Eindhoven. In 1980 he started his study in Mathematics at the University of Leiden in The Netherlands, where he graduated in 1988 in Applied Mathematics with Numerical Mathematics as main subject, and Information Science, Mathematical Statistics, and Operations Research techniques as subsidiary subjects. After his graduation he fulfilled his military obligations.

In 1989 he joined Hague Consulting Group (HCG), a consultancy specialising in transportation planning, as an Analyst and later as a Consultant. During his stay at HCG, he worked on various projects, such as the port of the National Model System from mainframe to micro-computer, the implementation of a transportation demand model for greater Stockholm, the START-1 and START-2 projects on real-time traffic monitoring and prediction, the estimation of a demand model for the National Railroads, the DRIVE-II project "DYNA" on traffic modelling and prediction for inter-urban motorways, and the estimation of a static O-D matrix for the National Model System.

In 1998 he joined the Transportation Planning and Traffic Engineering section of the Faculty of Civil Engineering and Geosciences of Delft University of Technology in The Netherlands. Here he conducted his Ph.D. research on dynamic O-D matrix estimation, and worked approximately fifty percent of the time on various third-party projects (such as the evaluation of the results of the fourth framework project DACCORD, the inland navigation project BVMS, the Libertel project), and contributed to the teaching activities of the section.
TRAIL Thesis Series

A series of The Netherlands TRAIL Research School for theses on transport, infrastructure and logistics.


