5-2015

Modeling Transportation Planning Applications via Path Flow Estimator

Seungkyu Ryu

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Civil and Environmental Engineering Commons

Recommended Citation
https://digitalcommons.usu.edu/etd/4225

This Dissertation is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact dylan.burns@usu.edu.
MODELING TRANSPORTATION PLANNING APPLICATIONS

VIA PATH FLOW ESTIMATOR

By

Seungkyu Ryu

A dissertation submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Civil and Environmental Engineering

Approved:

__________________________________________  ______________________________________
Anthony Chen                  Ziqi. Song
Major Professor              Committee Member

Gilberto E. Urroz           Haitao Wang
Committee Member            Committee Member

Yong Seog Kim               Mark R. McLellan
Committee Member            Vice President for Research and
                             Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah
2015
ABSTRACT

Modeling Transportation Planning Applications via Path Flow Estimator

by

Seungkyu Ryu, Doctor of Philosophy
Utah State University, 2015

Major Professor: Dr. Anthony Chen
Department: Civil and Environmental Engineering

The Path Flow Estimator (PFE) concept was originally developed to estimate path flows (hence origin–destination flows) and link flows for a whole road network (given some counts at selected roads). It is now further developed as an alternative for modeling different transportation planning applications: (1) a bicycle network analysis tool for non-motorized transportation planning, (2) a multi-class traffic assignment model for freight planning, and (3) a simplified travel demand forecasting framework for small community planning.

The first application of the redeveloped PFE is to develop a two-stage bicycle traffic assignment model for estimating/predicting bicycle volumes on a transportation network. The first stage considers key criteria (e.g., distance related attributes, safety related attributes, air quality related attributes, etc.) to generate a set of non-dominated (or efficient) paths, while the second stage adopts several traffic assignment methods to determine the flow allocations to the network. This two-stage approach can be used as a stand-alone bicycle traffic assignment to the transportation network given a bicycle origin-destination (O-D) matrix. The second application aims to enhance the realism of traffic assignment
models for freight planning by incorporating different modeling considerations into the multi-class traffic assignment problem. These modeling considerations involve developing both model formulation and customized solution algorithm, which in turn involve asymmetric interactions among different vehicle types (i.e., cars versus trucks), a path-size logit (PSL) model (for accounting random perceptions of network conditions with explicit consideration of route overlapping), and various traffic restrictions imposed either individually or together to multiple vehicle types in a transportation network. In the third application, a simplified planning framework is developed to perform planning applications in small communities where limited planning resources hinder the development and application of a full four-step model. Two versions (i.e., base year and future year) of the PFE are proposed to address the specific transportation planning issues and needs of small communities.

These new PFE developments for planning applications are tested with different realistic transportation networks. The results suggest that the new PFE applications proposed in this dissertation provide an alternative to the traditional four-step travel demand forecasting model that can be used as a stand-alone application with better modeling capability and fewer resources.

(221 pages)
PUBLIC ABSTRACT

Modeling Transportation Planning Applications via Path Flow Estimator

by

Seungkyu Ryu, Doctor of Philosophy
Utah State University, 2015

Major Professor: Dr. Anthony Chen
Department: Civil and Environmental Engineering

The current practice for modeling in the field of transportation planning is through a four-step travel demand forecasting procedure (i.e., trip generation, trip distribution, mode choice, and traffic assignment); the practice is commonly referred to as the four-step model. Although such a modeling approach has become standard practice, it is deficient in several areas. Specifically, (1) it lacks capability for modeling non-motorized modes such as bicycles, (2) it is inadequate for modeling multiple vehicle types sharing the same roadway space, and (3) it is difficult to apply to small communities with limited resources. This dissertation recognizes these deficiencies and responds to them through the development of alternative transportation planning applications via the path flow estimator (PFE). The PFE was originally developed as a one-stage network observer capable of estimating path flows and path travel times using only traffic counts from a subset of network links. In this dissertation, the PFE is used to develop the following three transportation planning applications for addressing the three deficiencies of the four-step model: (1) a bicycle network analysis tool for non-motorized transportation planning, (2)
a multi-class traffic assignment model for freight planning, and (3) a simplified travel demand forecasting model for small community planning.

The first application develops a two-stage bicycle traffic assignment model for estimating/predicting bicycle volumes on a transportation network. The first stage considers two key criteria (e.g., distance related attributes and safety related attributes) to generate a set of non-dominated (or efficient) paths, while the second stage determines the flow allocation to the set of efficient paths. In stage one, a bi-objective shortest path problem based on the two key attributes is developed to generate the efficient paths. In stage two, several traffic assignment methods are adopted to determine the flow allocations in a network. In addition, the two-stage bicycle traffic assignment model is further extended to consider multiple user classes and multiple criteria. This two-stage approach can be used as a stand-alone bicycle traffic assignment to the transportation network given a bicycle origin-destination (O-D) matrix.

The second application aims to enhance the realism of traffic assignment models for freight planning by incorporating different modeling considerations into the multi-class traffic assignment problem. These modeling considerations involve developing both model formulation and customized solution algorithm that in turn, involve asymmetric interactions among different vehicle types (i.e., cars versus trucks), a path-size logit (PSL) model (for accounting random perceptions of network conditions with explicit consideration of route overlapping), and various traffic restrictions imposed either individually or together to multiple vehicle types in a transportation network. Specifically, a variational inequality (VI) approach is used to formulate the stochastic multi-class traffic assignment problem with asymmetric vehicle interactions, route overlapping, and traffic
restraints. In addition, a customized path-based solution algorithm, consisting of an iterative balancing scheme, a self-regulated averaging line search scheme, and a column generation scheme, is developed for solving the stochastic multi-class traffic assignment problem.

The third application develops an alternative planning tool to model and forecast network traffic for planning applications in small communities, where resources debilitate the development and applications of the conventional four-step travel demand forecasting model. Two versions of the PFE are developed to address the specific transportation planning issues and needs of small communities: a base year PFE for estimating the current network traffic conditions using field data and planning data if available and a future PFE for predicting future network traffic conditions using forecast planning data and the calibrated origin-destination trip table as constraints. Solution algorithms are also developed to solve the two PFE models; they are integrated into a GIS-based software tool called Visual PFE, which is streamlined for planning applications in small communities.

To show proof of concept, these new PFE developments for planning applications are tested with different realistic transportation networks. The results suggest that the new PFE applications developed in this dissertation provide an alternative to the traditional four-step travel demand forecasting model that can be used as a stand-alone application with better modeling capability and fewer resources.

Seungkyu Ryu
ACKNOWLEDGMENTS

My sincerest appreciation goes to my supervisor, Professor Anthony Chen, for his continuous encouragement, guidance, and support during the entire period of my studies at Utah State University. Without his advice and patience, I would have been unable to complete this dissertation. I also wish to express my sincere thanks to the committee members, Dr. Ziqi. Song, Dr. Gilberto E. Urroz, Dr. Yong Seog Kim, and Dr. Haitao Wang, for their advice. Finally, a very special mention goes to my parents; I would not have been able to complete my studies without their love and continuous support.

Seungkyu Ryu
3.1 Introduction .................................................................................................................. 27
3.2 Two-stage bicycle traffic assignment procedure ....................................................... 29
   3.2.1 Two key cyclist route choice criteria .................................................................. 30
   3.2.2 Stage one: Bi-objective shortest path procedure .............................................. 33
   3.2.3 Stage two: Bicycle traffic assignment methods ................................................ 35
   3.2.4 Extension to multi-class and multi-criteria assignment .................................... 39
3.3 Numerical results .......................................................................................................... 41
   3.3.1 A simple network ............................................................................................... 41
   3.3.2 Winnipeg network ............................................................................................. 44
3.4 Conclusions .................................................................................................................. 59
REFERENCES ....................................................................................................................... 60

4. MODELING ASYMMETRIC VEHICLE INTERACTIONS, ROUTE OVERLAPPING, AND TRAFFIC RERAINTS IN MULTI-CLASS TRAFFIC ASSIGNMENT PROBLEMS ................................................................. 64

Abstract ............................................................................................................................. 64
4.1 Introduction .................................................................................................................. 64
   4.1.1 Relevant studies ................................................................................................. 65
   4.1.2 Objective and contribution ................................................................................ 67
4.2 Modeling considerations ............................................................................................... 68
   4.2.1 Vehicle interactions ........................................................................................... 68
   4.2.2 Route overlapping ............................................................................................. 71
   4.2.3 Traffic restraints ............................................................................................... 74
4.3 Multi-class traffic assignment formulation ................................................................... 76
4.4 Solution algorithm ....................................................................................................... 81
   4.4.1 Direction finding ............................................................................................... 81
   4.4.2 Line search ........................................................................................................ 83
   4.4.3 Column generation ........................................................................................... 85
   4.4.4 Overall solution procedure .............................................................................. 87
4.5 Numerical results ........................................................................................................ 91
   4.5.1 Description of the experiments ......................................................................... 92
   4.5.2 Experiment 1: Convergence characteristics and sensitivity analysis ............. 94
   4.5.3 Experiment 2: Three issues in the multi-class traffic assignment problem ....... 98
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Applications of PFE</td>
<td>11</td>
</tr>
<tr>
<td>2.2 Summary of various outputs at different spatial levels derived from path flows</td>
<td>24</td>
</tr>
<tr>
<td>3.1 Traffic assignment methods</td>
<td>36</td>
</tr>
<tr>
<td>3.2 Estimated route distance and route BLOS and the corresponding generated non-dominated routes</td>
<td>42</td>
</tr>
<tr>
<td>3.3 Comparison of assigned flows using five assignment methods</td>
<td>44</td>
</tr>
<tr>
<td>3.4 Generated total demand for each class</td>
<td>52</td>
</tr>
<tr>
<td>3.5 Summary of criteria used for the utility function of each user class in the two scenarios</td>
<td>52</td>
</tr>
<tr>
<td>3.6 Average traveled distance, BLOS and CO for each user class and all user classes</td>
<td>55</td>
</tr>
<tr>
<td>4.1 Urban street level of service</td>
<td>93</td>
</tr>
<tr>
<td>4.2 Assigned V/C ratios on 14 bridges without and with imposing capacity constraints</td>
<td>110</td>
</tr>
<tr>
<td>5.1 Percentage error for daily traffic volumes by facility type</td>
<td>126</td>
</tr>
<tr>
<td>5.2 Example of trip production and trip attraction estimation using ITE trip rates</td>
<td>129</td>
</tr>
<tr>
<td>5.3 Urban street level of service</td>
<td>155</td>
</tr>
<tr>
<td>5.4 LOS analysis for Main Street</td>
<td>156</td>
</tr>
<tr>
<td>5.5 Example of using ITE trip rates to determine zonal production and attraction for aggregated TAZ 320</td>
<td>161</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Pictorial representation of the conventional four-step model</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Organization of the dissertation</td>
<td>7</td>
</tr>
<tr>
<td>2.1</td>
<td>Pictorial representation of the conventional four-step model</td>
<td>19</td>
</tr>
<tr>
<td>2.2</td>
<td>Illustration of using path flows to derive useful information at different spatial levels</td>
<td>23</td>
</tr>
<tr>
<td>2.3</td>
<td>Path-link incidence matrix</td>
<td>23</td>
</tr>
<tr>
<td>3.1</td>
<td>Graphical illustration of the travel time function with vehicle interactions</td>
<td>30</td>
</tr>
<tr>
<td>3.2</td>
<td>Two-phase procedure for generating non-dominant routes</td>
<td>34</td>
</tr>
<tr>
<td>3.3</td>
<td>Illustration of route choice probability using the DAA method</td>
<td>38</td>
</tr>
<tr>
<td>3.4</td>
<td>Four types of cyclists in Portland (Geller, 2006)</td>
<td>40</td>
</tr>
<tr>
<td>3.5</td>
<td>Test network and link characteristics</td>
<td>41</td>
</tr>
<tr>
<td>3.6</td>
<td>Winnipeg network with bike lanes</td>
<td>45</td>
</tr>
<tr>
<td>3.7</td>
<td>Characteristics of the Winnipeg network</td>
<td>46</td>
</tr>
<tr>
<td>3.8</td>
<td>Generated routes analysis based on route distance and route BLOS</td>
<td>48</td>
</tr>
<tr>
<td>3.9</td>
<td>Link flow patterns of three assignment methods</td>
<td>49</td>
</tr>
<tr>
<td>3.10</td>
<td>Route flow distribution in terms of route distance and route BLOS</td>
<td>50</td>
</tr>
<tr>
<td>3.11</td>
<td>Link flow pattern of each user class for both scenarios</td>
<td>54</td>
</tr>
<tr>
<td>3.12</td>
<td>Effect of multi-class and multi-criteria considerations on route choice probabilities</td>
<td>58</td>
</tr>
<tr>
<td>4.1</td>
<td>Graphical illustration of the travel time function with vehicle interactions</td>
<td>70</td>
</tr>
<tr>
<td>4.2</td>
<td>Illustration of the PSL model in resolving the route overlapping problem</td>
<td>74</td>
</tr>
</tbody>
</table>
4.3 Three types of vehicle restrictions .......................................................... 76

4.4 Comparison of memory requirements between individual class and universal path sets ................................................................. 86

4.5 Flowchart of generating a universal path set for a given O-D pair ............. 87

4.6 Flowchart of the customized solution algorithm ........................................ 88

4.7 Flowchart of the customized solution algorithm ........................................ 93

4.8 Computational comparisons between MSA and SRA schemes .................. 95

4.9 Sensitivity of the two parameters in the SRA line search scheme .............. 96

4.10 Computational efforts under different dispersion parameter values, PCE factors, percentages of trucks, and travel time weight parameters ......................... 97

4.11 Flow allocation comparison between truck pre-loading and simultaneous assignment methods ................................................................. 100

4.12 Congestion level and number of used paths by vehicle class for the truck pre-loading and simultaneous assignment methods ......................... 101

4.13 Flow allocation comparison between MNL and PSL SUE models ............. 103

4.14 Path flow distributions as a function of PS value for MNL and PSL SUE models 104

4.15 Flow allocation comparison between without and with vehicle restrictions .... 106

4.16 Bridge locations in Winnipeg network .................................................... 107

4.17 Convergence characteristics with link capacity constraints ..................... 108

5.1 The PFE-based transportation planning framework .................................. 124

5.2 Network representation of an intersection ............................................... 127

5.3 Zonal production and attraction flows from socio-economic and land use data ... 129

5.4 Base year PFE formulation ........................................................................ 132

5.5 Derivation of future target O-D trips ...................................................... 135

5.6 Future year PFE formulation .................................................................... 137
5.7 Flowchart of the PFE solution algorithm ......................................................... 138
5.8 St. Helena network and locations of intersection turning movement counts ....... 148
5.9 Comparison between observed and estimated flows ....................................... 151
5.10 Base year flow pattern estimates ................................................................. 151
5.11 Future year zonal production and attraction flows ....................................... 153
5.12 Forecast link flows and V/C ratios for two scenarios .................................... 154
5.13 Traffic counts and aggregated TAZ locations in the Eureka network ............ 158
5.14 Procedure for converting land use data to zonal production and attraction trips... 159
5.15 Land use and socioeconomic data in terms of employment and dwelling units
    for years 2005 and 2030 .................................................................................. 160
5.16 Estimated zonal production and attraction flows for years 2005 and 2030 ....... 161
5.17 Estimated zonal production and attraction flows for the base year ............... 163
5.18 Percentage of O-D pairs by the absolute O-D flow difference for the base year .. 164
5.19 Link flow comparison for the base year ....................................................... 165
5.20 Zonal production and attraction flow comparison for the future year .............. 167
5.21 Percentage of O-D pairs by the absolute O-D flow difference for the future year 168
5.22 Forecast link flows and V/C ratios for the future year .................................... 169
5.23 Percentage of links by the forecast V/C ratio for the future year .................... 169
CHAPTER 1

INTRODUCTION

1.1 General background

Transportation is critical to the social, environmental, and economic health of every city. Federal regulations in the United States require urbanized areas with populations of over 50,000 residents to have a Metropolitan Planning Organization (MPO) responsible for transportation planning (Meyer and Miller, 2000). One of the major functions of a MPO is to develop an appropriate analytical program to evaluate transportation alternatives and to support metropolitan decision-making scaled to the size and complexity of the region and to the nature of its transportation issues and the realistically available options.

The current practice for modeling in the field of transportation planning is through a four-step travel demand forecasting procedure. It consists of trip generation (travel choice), trip distribution (destination choice), modal split (mode choice) and traffic assignment (route choice) in a top-down sequential process (Ortuzar and Willumsen, 2001). The outputs of one step serve as the inputs of the next step. This practice is commonly referred to as the four-step model. These steps are depicted in Figure 1.1 and are briefly described below.

![Figure 1.1 Pictorial representation of the conventional four-step model](image-url)
Trip generation (or travel choice) is the first step in the conventional four-step travel demand forecasting process. It predicts the number of trips originating in or destined for a particular traffic analysis zone (TAZ). Once the trip productions and attractions for each zone are computed, the trips can be distributed among the zones using a trip distribution model. The trip distribution model is essentially a destination choice model that generates a trip table using the trip ends produced from the trip generation models and the network attributes (e.g., interzonal travel times). Modal split (or mode choice) models determine how trips will be divided among the available modes. Finally, the last step of the conventional four-step model is traffic assignment (or route choice). This step involves assigning the mode-specific O-D trip tables to a transportation network.

Although such a modeling approach has become standard practice, the conventional four-step model has been criticized for its inherent weakness, the lack of a single unifying rationale that would explain or legitimize all aspects of demand. However, that is not its only weakness; the four step model also suffers from inconsistent consideration of travel times and congestion effects in various steps of the procedure (see Boyce, 2002; Oppenheim, 1995; Garret and Wachs, 1996; McNally, 2000). In addition, the conventional four-step model is deficient in several areas. Specifically, (1) it lacks capability for modeling non-motorized modes such as bicycles, (2) it is inadequate for modeling multiple vehicle types sharing the same roadway space, and (3) it is difficult to apply to small communities with limited resources.
1.2 Problem descriptions

Non-motorized transportation planning

Non-motorized modes such as bicycles constitute an important part of a community’s transportation system, and they are also vital to the success of transit-oriented-developments (TODs). Yet, they are often ignored in transportation planning and travel demand forecasting modeling. At best, non-motorized modes are treated as a byproduct in the planning process. In addition, many cities have begun to invest and promote cycling as a healthy, environmentally friendly, and economical alternative mode of travel to the motorized vehicles (especially private motorized vehicles). However, the current practice in modeling bicycle trips in a network is inadequate. Only a few research efforts focus on network analysis for bicycle trips (e.g., Klobucar and Fricker, 2007; Broach et al., 2011; Mekuria et al., 2012). These methods provide an initial effort to develop a traffic assignment method for bicycle trips, but they are too simplistic (i.e., simply based on all-or-nothing (AON) assignment method using a single attractiveness measure (e.g., distance, safety, or a composite measure of safety multiplied by distance). However, compared to the route choice models for private motorized vehicles, route choice behavior for cyclists is much more complex as there many influential factors affecting their route choice decisions.

Freight Planning

Traffic assignment is an essential and fundamental step in the transportation planning and management process. Given constant travel demands between each origin-destination (O-D) pair (i.e., travelers), and travel cost functions for each link of the network (i.e., transportation network), the traffic assignment problem is to determine the traffic flow
pattern as well as network performance measures (e.g., total system travel time, vehicle miles of travel, vehicle hours of travel, fuel consumption and emission, etc.). In practice, most traffic assignment models are single user class and contain three main modeling assumptions. Firstly, most traffic assignment models employ a separability assumption on the link travel time function (i.e., no interactions). Secondly, most models assume deterministic user equilibrium (DUE) or stochastic user equilibrium (SUE) without accounting for route overlapping. Lastly, most models lack side constraints to describe the limited supply of certain scarce resources (e.g., link capacities) in a network which are shared by multiple vehicle types (e.g., passenger cars and multiple truck types). Similarly, these models may also neglect to limit certain classes of vehicles (e.g., trucks) on underpasses due to height restriction, bridges due to weight restriction, and prohibited lanes due to lane restriction. Modeling the asymmetric interactions among different vehicle types (e.g., the impact of trucks on the travel times of cars is typically higher than the impact of cars on the travel times on trucks) has the potential to improve the realism of traffic assignment models (i.e., better capture the travel times required by different vehicle types to traverse a link). In fact, inadequate considerations of vehicle interactions in the traffic assignment models could result in inaccurate travel time estimates, which could in turn affect flow allocations to routes and network equilibrium process. As truck traffic continues to grow as a result of increasing freight shipments transported by trucks, there is an increasing interest to model multiple vehicle classes. In particular, there would be increasing demand to address the impacts of truck traffic on congestion, infrastructure deterioration, safety, and environmental concerns in many urban cities. In addition, it is important to recognize that there are several different sizes of trucks whose maneuver and
operational characteristics as well as traffic restrictions are quite different. These characteristics emphasize the need to develop a multi-class traffic assignment model that considers the above modeling issues.

**Small community planning**

The current practice in modeling network traffic is through a four-step travel demand forecasting model that requires travel surveys as input and specialized technical staff to develop and estimate the travel demand itself. Although such a modeling approach has been used in practice in major urban areas, Yan (1998) noted that many smaller communities usually do not have the resources to conduct travel surveys, nor to employ technical staff for model development and maintenance. Without data from a travel survey in the study area, trip generation rates of various land use zones are often “borrowed” from such published data as *Trip Generation* of the Institute of Transportation Engineers (ITE) or reports of travel surveys performed in other areas. The unavailability of data on Trip Length Frequency Distribution (TLFD) of local travelers often forces modelers to skip the calibration of trip distribution models. Instead, calibration and validation of the overall model are often carried out by altering the friction factors and adding k-factors (i.e., an empirically zone-to-zone adjustment factor, which takes into account the effects on travel patterns of defined social and economic linkages not otherwise incorporated into the model). The k-factors are added in a trial-and-error fashion to the trip distribution model such that the results of traffic assignment would match traffic counts on selective screenlines and critical links. The calibration process is usually a lengthy process and the resultant models often contain many factors that do not have the necessary behavioral foundation established from travel surveys. Schutz (2000) suggests that development of
innovative methodologies is urgent and necessary for communities who are unable to meet the planning requirements.

1.3 Research objectives

The objectives of this dissertation are to develop alternative transportation planning applications via the path flow estimator (PFE) to address the deficiencies of the four-step model. The PFE was originally developed as a one-stage network observer capable of estimating path flows and path travel times using only traffic counts from a subset of network links. Specifically, the objectives are to:

- **Objective 1:** Develop a bicycle network analysis tool for non-motorized transportation planning application.
- **Objective 2:** Develop a multi-class traffic assignment model for freight planning application.
- **Objective 3:** Develop a simplified travel demand forecasting framework for small community planning application.
- **Objective 4:** Examine the model development and solution algorithm of different transportation planning applications to gain insights for further development.
- **Objective 5:** Conduct case studies on real networks to demonstrate how these new PFE developments can be used in practice.

1.4 Organization

This dissertation consists of six chapters. Fig. 1.2 provides a flowchart of the dissertation organization.
In Chapter 2, the concept of PFE is reviewed. The review includes mathematical programming formulation, solution properties, solution algorithm, and applications of PFE for transportation planning and operation purposes.

In Chapter 3, a bicycle network analysis tool for non-motorized planning application is presented. It consists of two stages: a bi-objective shortest path procedure for generating efficient paths and a bicycle traffic assignment procedure for determining flow allocations to the bicycle network. In addition, extension to multiple user classes and multiple criteria is discussed.

In Chapter 4, a multi-class traffic assignment for freight transportation planning application is presented. Different modeling considerations for improving the
realism of traffic assignment model for freight planning are discussed along with model formulation and customized solution algorithm.

- In Chapter 5, a simplified travel demand forecasting framework for small community planning is presented. Two versions (i.e., base year and future year) of the PFE are provided to address the specific transportation planning issues and needs of small communities. Solution algorithms, integrated into a GIS-based software tool called Visual PFE, are also presented for solving the two PFE models.

- In Chapter 6, concluding remarks and future research directions are summarized.

REFERENCES


CHAPTER 2
LITERATURE REVIEW

This chapter reviews the concept of the path flow estimator (PFE) by summarizing its mathematical programming formulation, solution properties, solution algorithms, and potential applications for transportation planning and operation purposes.

2.1 Concept of path flow estimator

The concept of the Path Flow Estimator (PFE) was first proposed by Sherali et al. (1994) as a linear program (LP) based on the Wardrop (1952) user equilibrium principle. This approach assumed that route choice is under the deterministic user equilibrium principle (DUE), thus path flow pattern is not unique. In addition, the approach requires all link counts in the network to be observed. With these deficiencies (not unique path flow pattern and requiring all observed counts), the approach is less practical. Bell and Shield (1995) and Bell et al. (1997) extended the original LP PFE method to the nonlinear PFE, which is based on the stochastic user equilibrium (SUE) principle, and only required partial link counts in the network to be observed for the estimation process. This nonlinear PFE method is more practical and gives unique path flows. This uniqueness is different from the under-specified problem of the O-D estimation from traffic counts, which is known to have multiple solutions since the number of observations (link counts) is generally less than the number of variables (O-D pairs). The multiplicity of the solutions (O-D estimates) should be quantified by some quality measures [e.g., the maximal possible relative error (MPRE) of Yang et al. (1991), the total demand scale of Bierlaire (2002)].
The core component of the nonlinear PFE method is based on the logit-based route choice model (Fisk, 1980) that interacts with link cost functions to produce a logit-based SUE traffic flow pattern conforming with the traffic counts and other side constraints if available (i.e., intersection turning movement flows, target O-D flows, production flows, attraction flows, etc.). Due to the flexibility of aggregating path flows at different spatial levels, the nonlinear PFE method allows us to perform not only traffic assignment if O-D demand is available but also O-D estimation with partial observed link counts. The theoretical advantage of the nonlinear PFE method is the single-level convex programming formulation with side constraints. Since the objective function is strictly convex with respect to the decision variables (path flows) and the constraints are all linear (equality and inequality) equations, the optimization is guaranteed to yield unique path flows that can be used to derive other useful information at different spatial levels. Due to the unique features discussed above, the PFE has been applied to different transportation applications shown in Table 2.1.

Table 2.1: Applications of PFE

<table>
<thead>
<tr>
<th>Reference</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bell and Shield (1995)</td>
<td>General path flow estimation</td>
</tr>
<tr>
<td>Bell et al. (1997)</td>
<td></td>
</tr>
<tr>
<td>Bell and Grosso (1998)</td>
<td></td>
</tr>
<tr>
<td>Bell and Lam (1996)</td>
<td>Time-dependent multi-class O-D estimation</td>
</tr>
<tr>
<td>Bell and Cassir (1998)</td>
<td>Multi-modal O-D estimation</td>
</tr>
<tr>
<td>Chen et al. (2005, 2012a)</td>
<td>Quality assessment of estimated O-D demand</td>
</tr>
<tr>
<td>Chootinan and Chen (2011)</td>
<td></td>
</tr>
<tr>
<td>Chen et al. (2009, 2010)</td>
<td>Norm approximation of traffic counts inconsistency in O-D estimation</td>
</tr>
<tr>
<td>Chen et al. (2012b)</td>
<td>Intersection turning movement estimation</td>
</tr>
<tr>
<td>Lee et al. (2006)</td>
<td>Planning application for small MPO</td>
</tr>
<tr>
<td>Jansuwan et al. (2012)</td>
<td></td>
</tr>
</tbody>
</table>
2.2 Mathematical formulation

The nonlinear PFE model is similar to the formulation of the logit-based SUE problem proposed by Fisk (1980), except for the constraint set. Notation is provided first for convenience, followed by mathematical programming formulation, optimality conditions, and solution algorithm.

2.2.1 Notation

Consider a transportation network $G = (N, A)$, where $N$ is the set of nodes and $A$ is the set of links on a given network. We define the following list of notations.

Sets

$N$ : set of nodes

$R$ : set of origins, $R \subseteq N$

$S$ : set of destinations, $S \subseteq N$

$RS$ : set of origin-destination (O-D) pairs

$\overline{RS}$ : set of prior origin-destination (O-D) pairs

$K_{rs}$ : set of routes connecting O-D pair $rs$

$\overline{A}$ : set of observed links, $\overline{A} \subseteq A$

$U$ : set unobserved links, $U \subseteq A$

$A$ : set all network links, $A = \overline{A} \cup U$

Parameters and inputs

$\theta$ : dispersion parameter in the logit model

$v_a$ : observed traffic volume on link $a$

$C_a$ : capacity of link $a$
$z_{rs}$ : priori travel demand for O-D pair $rs$

$\varepsilon_a$ : measurement error allowed for traffic count on link $a$

$\varepsilon_{rs}$ : measurement error allowed for target O-D pair $rs$

Variables

$x_a$ : flow on link $a$

$t_a(\cdot)$ : travel time on link $a$

$\delta_{ka}^{rs}$ : route-link indicator, 1 if link a is on route $k$ between O-D pair $rs$ and 0 otherwise

$f_k^{rs}$ : flow on route $k$ connecting O-D pair $rs$

$\epsilon_k^{rs}$ : cost on route $k$ connecting O-D pair $rs$

$q_{rs}$ : O-D flow between O-D pair $rs$

2.2.2 Mathematical formulation

The formulation of PFE is given as follows:

Minimize: 

$$\frac{1}{\Theta} \sum_{rs \in RS} \sum_{k \in K_{rs}} f_k^{rs} (\ln f_k^{rs} - 1) + \sum_{a \in A} \int_0^{t_a} t_a(w)dw$$  \hspace{1cm} (2.1)

subject to:

$$(1 - \varepsilon_a) \cdot v_a \leq x_a \leq (1 + \varepsilon_a) \cdot v_a, \quad \forall a \in A$$  \hspace{1cm} (2.2)

$v_a \leq C_a, \quad \forall a \in U$  \hspace{1cm} (2.3)

$$(1 - \varepsilon_{rs}) \cdot z_{rs} \leq q_{rs} \leq (1 + \varepsilon_{rs}) \cdot z_{rs}, \quad \forall rs \in RS$$  \hspace{1cm} (2.4)

$f_k^{rs} \geq 0, \quad \forall k \in K_{rs}, rs \in RS$  \hspace{1cm} (2.5)

where

$$x_a = \sum_{rs \in RS} \sum_{k \in K_{rs}} f_k^{rs} \delta_{ka}^{rs}, \quad \forall a \in A$$  \hspace{1cm} (2.6)
\[ q_{rs} = \sum_{k \in K_{rs}} f_{rs}^k, \quad \forall rs \in RS \quad (2.7) \]

As opposed to the logit-based SUE model, the PFE finds path flows that minimize the SUE objective function in Eq. (2.1) while simultaneously reproducing traffic counts on all observed links in Eq. (2.2) and prior travel demands of certain O-D pairs in Eq. (2.4) within some predefined error bounds. For the unobserved links, as indicated by Eq. (2.3), link capacity is set as the upper bound to prevent the estimated path flows from producing unrealistically high estimated link flows. This objective function, nonetheless, yields the logit-based probability expression. Error bounds (\( \varepsilon_u \) and \( \varepsilon_r \)) are introduced into Eqs. (2.2) and (2.4) as user-defined inputs to account for measurement errors of traffic counts and the confidence level associated with prior O-D demands, respectively. More reliable information will constrain the estimation (either link flows or O-D flows) to be within a smaller tolerance, whereas less reliable information will allow for a larger deviation. The introduction of the user-defined error bounds in the interval constraints enhances the flexibility of PFE by allowing the user to incorporate local knowledge about the network conditions to the estimation process. Eq. (2.5) constrains the path flows to be non-negative, while Eqs. (2.6) and (2.7) are definitional constraints to obtain link flows and O-D flows from the path flow solution. Overall, the PFE formulation differs from the logit-based SUE model by explicitly considering the interval constraints (2.2) and (2.4) and capacity constraint (2.3). The first-order conditions (equivalency) for the minimization program can now be obtained as follows.
2.2.3 Optimality conditions

The Lagrangian function of the above PFE formulation and its first partial
derivatives with respect to the path-flow variables can be expressed as follows.

\[
L(f,u^+,u^-,d,o^+,o^-) = Z + \sum_{a \in A} u_a^- \left( v_a (1 - \varepsilon_a) - \sum_{r \in \mathcal{R}} \sum_{k \in K_{r\in}} f_k^r \delta_k^{rs} \right) + \\
\sum_{a \in A} u_a^+ \left( v_a (1 + \varepsilon_a) - \sum_{r \in \mathcal{R}} \sum_{k \in K_{r\in}} f_k^r \delta_k^{rs} \right) + \sum_{a \in U} d_a \left( C_a - \sum_{r \in \mathcal{R}} \sum_{k \in K_{r\in}} f_k^r \delta_k^{rs} \right) + \\
\sum_{r \in \mathcal{R}} o_{rs}^- \left( z_{rs} (1 - \varepsilon_{rs}) - \sum_{k \in K_{rs}} f_k^r \right) + \sum_{r \in \mathcal{R}} o_{rs}^+ \left( z_{rs} (1 + \varepsilon_{rs}) - \sum_{k \in K_{rs}} f_k^r \right)
\]

where \( u_a^-, u_a^+, d_d, o_{rs}^-, \) and \( o_{rs}^+ \) are the dual variables of constraints (2.2), (2.3), and (2.4), respectively. The values of \( u_a^+, d_d \), and \( o_{rs}^+ \) are restricted to be non-positive, while the value of \( u_a^- \) and \( o_{rs}^- \) must be nonnegative; \( u_a^- \) and \( u_a^+ \) can be viewed as the corrections in the link cost function, which bring the estimated path flows into agreement with the observed link volumes; similarly, \( o_{rs}^- \) and \( o_{rs}^+ \) can be interpreted as corrections to the O-D travel times that can be used to steer the estimated path flow pattern to within the O-D interval constraints specified by Eq. (2.4). These dual variables are zero if the estimated link flows and O-D flows are within an acceptable range defined by the measurement error bound, and non-zero if they are binding at one of the limits. \( d_d \) is related to the link queuing delay when the estimated link flow reaches its capacity (Bell and Iida, 1997).

\[
f_k^r \frac{\partial L(f,u^+,u^-,d,o^+,o^-)}{\partial f_k^r} = 0, \ \forall \ k \in K_{r\in}, \ rs \in \mathcal{R} \\
u_a^- \frac{\partial L(f,u^+,u^-,d,o^+,o^-)}{\partial u_a^-} = 0, \ \forall \ a \in \mathcal{A} \\
u_a^+ \frac{\partial L(f,u^+,u^-,d,o^+,o^-)}{\partial u_a^+} = 0, \ \forall \ a \in \mathcal{A}
\]
\[
\frac{\partial L(f, u^+, u^-, d, o^+, o^-)}{\partial d_a} = 0, \ \forall a \in U
\]
\[
\frac{\partial L(f, u^+, u^-, d, o^+, o^-)}{\partial o_{rs}^-} = 0, \ \forall \ rs \in \overline{RS}
\]
\[
\frac{\partial L(f, u^+, u^-, d, o^+, o^-)}{\partial o_{rs}^+} = 0, \ \forall \ rs \in \overline{RS}
\]

For \(f_k^{rs} \frac{\partial L(f, u^+, u^-, d, o^+, o^-)}{\partial f_k^{rs}} = 0, \ \forall k \in K_{rs}, rs \in RS:\)

\[
f_k^{rs} \left( \frac{1}{\theta} \ln f_k^{rs} + \sum_{a \in \Lambda} \alpha_a (x_a) \delta_{ka}^r - \sum_{a \in \Lambda} u_a^- \delta_{ka}^r - \sum_{a \in \Lambda} u_a^+ \delta_{ka}^r - \sum_{a \in U} d_a \delta_{ka}^r - o_{rs}^- - o_{rs}^+ \right) = 0, \ \forall k \in K_{rs}, rs \in RS
\]

(2.10)

Since always \(f_k^{rs} > 0, \ \forall k \in K_{rs}, rs \in RS,\) in the logit-based SUE model,

\[
\left( \frac{1}{\theta} \ln f_k^{rs} + \sum_{a \in \Lambda} \alpha_a (x_a) \delta_{ka}^r - \sum_{a \in \Lambda} u_a^- \delta_{ka}^r - \sum_{a \in \Lambda} u_a^+ \delta_{ka}^r - \sum_{a \in U} d_a \delta_{ka}^r - o_{rs}^- - o_{rs}^+ \right) = 0
\]

(2.11)

For \(u_a^- \frac{\partial L(f, u^+, u^-, d, o^+, o^-)}{\partial u_a^-} = 0, \ \forall a \in \overline{\Lambda}:\)

\[
\begin{align*}
  v_a (1 - \varepsilon_a) - \sum_{r \in RS} \sum_{k \in K_{rs}} f_k^{rs} \delta_{ka}^r < 0 & \Rightarrow v_a (1 - \varepsilon_a) < x_a, \ \text{if} \ u_a^- = 0 \\
  v_a (1 - \varepsilon_a) - \sum_{r \in RS} \sum_{k \in K_{rs}} f_k^{rs} \delta_{ka}^r = 0 & \Rightarrow v_a (1 - \varepsilon_a) = x_a, \ \text{if} \ u_a^- > 0
\end{align*}
\]

(2.12)

For \(u_a^+ \frac{\partial L(f, u^+, u^-, d, o^+, o^-)}{\partial u_a^+} = 0, \ \forall a \in \overline{\Lambda}:\)

\[
\begin{align*}
  v_a (1 + \varepsilon_a) - \sum_{r \in RS} \sum_{k \in K_{rs}} f_k^{rs} \delta_{ka}^r > 0 & \Rightarrow v_a (1 + \varepsilon_a) > x_a, \ \text{if} \ u_a^+ = 0 \\
  v_a (1 + \varepsilon_a) - \sum_{r \in RS} \sum_{k \in K_{rs}} f_k^{rs} \delta_{ka}^r = 0 & \Rightarrow v_a (1 + \varepsilon_a) = x_a, \ \text{if} \ u_a^+ < 0
\end{align*}
\]

(2.13)
For $d_a \frac{\partial L(f, u^+, u^-, d, o^+, o^-)}{\partial d_a} = 0$, $\forall a \in U$:

$$C_a - \sum_{rs \in RS} \sum_{k \in K_a} f_k^{rs} \delta_{ka} > 0 \Rightarrow C_a > x_a^+ \quad \text{if} \quad d_a = 0$$

$$C_a - \sum_{rs \in RS} \sum_{k \in K_a} f_k^{rs} \delta_{ka} = 0 \Rightarrow C_a = x_a^+ \quad \text{if} \quad d_a < 0$$

(2.14)

For $\partial L(f, u^+, u^-, d, o^+, o^-)$ $\partial o_{rs}^- = 0$, $\forall rs \in \overline{RS}$:

$$z_{rs} \left(1 - \epsilon_{rs}^-ight) - \sum_{k \in K_a} f_k^{rs} < 0 \Rightarrow z_{rs} \left(1 - \epsilon_{rs}^-ight) < q_{rs}^-, \quad \text{if} \quad o_{rs}^- = 0$$

$$z_{rs} \left(1 - \epsilon_{rs}^-ight) - \sum_{k \in K_a} f_k^{rs} = 0 \Rightarrow z_{rs} \left(1 - \epsilon_{rs}^-ight) = q_{rs}^-, \quad \text{if} \quad o_{rs}^- > 0$$

(2.15)

For $\partial L(f, u^+, u^-, d, o^+, o^-)$ $\partial o_{rs}^+ = 0$, $\forall rs \in \overline{RS}$:

$$z_{rs} \left(1 + \epsilon_{rs}^+ight) - \sum_{k \in K_a} f_k^{rs} > 0 \Rightarrow z_{rs} \left(1 + \epsilon_{rs}^+ight) > q_{rs}^+, \quad \text{if} \quad o_{rs}^+ = 0$$

$$z_{rs} \left(1 + \epsilon_{rs}^+ight) - \sum_{k \in K_a} f_k^{rs} = 0 \Rightarrow z_{rs} \left(1 + \epsilon_{rs}^+ight) = q_{rs}^+, \quad \text{if} \quad o_{rs}^+ < 0$$

(2.16)

Let the generalized route cost $\tilde{c}_k^{rs} = c_k^{rs} - \sum_{a \in A} \left(u_a^- + u_a^+ight) \delta_{ka} - \sum_{a \in U} d_a \delta_{ka} - (o_{rs}^- + o_{rs}^+)$, where

$$c_k^{rs} = \sum_{a \in A} t_a \left(x_a^+ \right) \delta_{ka} ) \right).$$

Rearrange the Eq. (2.11) and obtain:

$$f_k^{rs} = \exp \left(-\theta \cdot \tilde{c}_k^{rs} \right), \forall k \in K_{rs}, rs \in RS$$

(2.17)

Hence, the route choice probability function can be expressed as follows.

$$P_k^{rs} = \frac{f_k^{rs}}{\sum_{k \in K_{rs}} f_k^{rs}} = \frac{\exp \left(-\theta \cdot \tilde{c}_k^{rs} \right)}{\sum_{k \in K_{rs}} \exp \left(-\theta \cdot \tilde{c}_k^{rs} \right)}, \forall k \in K_{rs}, rs \in RS$$

(2.18)
Similar to the logit-based SUE model, path flows from PFE can be derived analytically as a function of path costs and dual variables associated with constraints (2.2), (2.3), and (2.4) as follows.

\[ f_k^{rs} = \exp\left(\theta \left( -c_k^{rs} + \sum_{a \in A} \left( u_a^- + u_a^+ \right) \delta_{ka}^{rs} + \sum_{a \in U} d_a \delta_{ka}^{rs} + o_{rs}^+ + o_{rs}^- \right) \right) \quad \forall k \in K_{rs}, rs \in RS \]  

(2.19)

We proceed to the second-order condition to show the uniqueness of path-flow solution. Differentiating equation (2.11) by another path flow variable gives the following:

\[ \frac{\partial^2 L}{\partial f_k^{rs} \partial f_l^{od}} = \begin{cases} \frac{\partial c_k^{rs}}{\partial f_k^{rs}} + \frac{1}{\theta f_k^{rs}} & \text{if } f_k^{rs} = f_l^{od}, \forall k \in K_{rs}, l \in K_{rs}, rs \in RS, od \in RS \vspace{0.2cm} \\ 0 & \text{otherwise} \end{cases} \]  

(2.20)

Eq. (2.21) indicates all diagonal elements with positive values (i.e., \( \frac{\partial c_k^{rs}}{\partial f_k^{rs}} > 0, \theta > 0 \), and \( f_k^{rs} > 0 \)) and all off-diagonals with zeros respect to O-D pair \( rs \) in the Hessian matrix. In other words,

\[ \frac{\partial^2 L}{\partial f_k^{rs} \partial f_l^{rs}} = \begin{bmatrix} \frac{\partial c_1^{rs}}{\partial f_1^{rs}} + \frac{1}{\theta f_1^{rs}} & 0 & \cdots & 0 \\ 0 & \frac{\partial c_2^{rs}}{\partial f_2^{rs}} + \frac{1}{\theta f_2^{rs}} & 0 & \cdots \\ \vdots & \vdots & \ddots & \frac{\partial c_k^{rs}}{\partial f_k^{rs}} + \frac{1}{\theta f_k^{rs}} \\ 0 & 0 & \cdots & \frac{\partial c_r^{rs}}{\partial f_r^{rs}} + \frac{1}{\theta f_r^{rs}} \end{bmatrix} \]  

(2.21)

Since the diagonal elements of the block matrix with respect to O-D pair \( rs \) are equal to \( \frac{\partial c_k^{rs}}{\partial f_k^{rs}} + \frac{1}{\theta f_k^{rs}} \), the matrix \( \nabla^2 f \) is positive definite for all O-D pair \( rs \). Hence, objective function (2.1) is strictly convex with respect to path flows; therefore, the path-flow solution is unique.
2.3 Solution algorithm

The solution procedure for solving PFE is depicted in Fig. 2.1. It consists of three main modules: (1) iterative balancing scheme, (2) column (or path) generation, and (3) output derivation from path flows. The basic idea of the iterative balancing scheme is to sequentially scale the path flows to fulfill one constraint at a time by adjusting the dual variables. Once the scheme converges, the path flows can be analytically determined. A column generation is included in the solution procedure to avoid path enumeration for a general transportation network. Finally, an output derivation procedure is used to derive information at different spatial levels using the path-flow solution from PFE (e.g., link flows, turning movement flows for all intersections, production flows, attraction flows, O-D flows, screen-line flows, and total demand). Fig. 2.1 depicts an overall flowchart of solution algorithm for solving PFE.

![Flowchart of Solution Algorithm](image)

Fig. 2.1 Pictorial representation of the conventional four-step model
Iterative Balancing Scheme

The iterative balancing scheme can be summarized as follows.

Step 1. *Initialization*
1.1 Set \( n = 0 \),
1.2 Set primal variables: \( x_a^n \) and \( q_{rs}^n = 0 \),
1.3 Set dual variables: \( (u_a^+)^n \), \( (u_a^-)^n \), \( (d_a)^n \), \( (o_{rs}^-)^n \), and \( (o_{rs}^+)^n = 0 \).

Step 2. *Compute Dual and Primal Variables*
2.1 Set \( n = n + 1 \)
2.2 Update dual variables
   - For each measured link \( (a \in A) \), update the dual variables
     \[
     (u_a^+) = \min \left\{ 0, (u_a^+)^{n-1} + \frac{1}{\theta} \ln \left( \frac{(1 + \varepsilon_a) \cdot \nu_a}{x_a^n} \right) \right\}
     \]
     \[
     (u_a^-) = \max \left\{ 0, (u_a^-)^{n-1} + \frac{1}{\theta} \ln \left( \frac{(1 - \varepsilon_a) \cdot \nu_a}{x_a^n} \right) \right\}
     \]
   - For each unmeasured link \( (a \in U) \), update the dual variables
     \[
     (d_a) = \min \left\{ 0, (d_a)^{n-1} + \frac{1}{\theta} \ln \left( \frac{C_a}{x_a^n} \right) \right\}
     \]
   - For each target O-D flow \( (rs \in \overline{RS}) \), update the dual variables
     \[
     (o_{rs}^+) = \min \left\{ 0, (o_{rs}^+)^{n-1} + \frac{1}{\theta} \ln \left( \frac{(1 + \varepsilon_{rs}) \cdot z_{rs}}{q_{rs}^n} \right) \right\}
     \]
     \[
     (o_{rs}^-) = \max \left\{ 0, (o_{rs}^-)^{n-1} + \frac{1}{\theta} \ln \left( \frac{(1 - \varepsilon_{rs}) \cdot z_{rs}}{q_{rs}^n} \right) \right\}
     \]
2.3 Compute primal variables
   - Path flows
     \[
     (f_k) = \exp \left\{ \theta \cdot \left( (c_k) + \sum_{a \in X} \left[ (u_a^+) + (u_a^-) \right] \delta_{ka} + \sum_{a \in U} (d_a) \delta_{ka} + (o_{rs}^+) + (o_{rs}^-) \right) \right\}
     \]
- Link flows
\[ x_a^n = \sum_{rs \in RS} \sum_{k \in K_{rs}} (f_k^{rs})^n \delta_{ka}^{rs}, \forall a \in A \]

- O-D flows
\[ q_{rs}^n = \sum_{k \in K_{rs}} (f_k^{rs})^n, \forall rs \in RS \]

Step 3. Convergence Test

If \( \eta_0 \leq \text{Max} \left\{ \left| (u_a^+)^n - (u_a^-)^{n-1} \right|, \left| (u_a^-)^n - (u_a^+)^{n-1} \right|, \left| (d_a)^n - (d_a)^{n-1} \right| \right\} < \eta, \]

where \( \eta_0 \) is a convergence tolerance (e.g., \( 10^{-6} \)) and \( \eta \) is the upper limit of change in dual variables, then set all parameters of the next iteration equal to those of the current iteration, and go to step 2.

If \( \text{Max} \left\{ \left| (u_a^+)^n - (u_a^-)^{n-1} \right|, \left| (u_a^-)^n - (u_a^+)^{n-1} \right|, \left| (d_a)^n - (d_a)^{n-1} \right| \right\} \geq \eta, \) then set all parameters of the next iteration equal to those of the current iteration, set \( \text{outer} (n) = \text{outer}(n+1) \), and terminate.

In the above procedure, we just provide the adjustment equations for different types of constraints (e.g., observed links, unobserved links, and target O-D flows). The detailed derivations of the adjustment equations can be found in Chen et al. (2009, 2010), and convergence of the iterative balancing scheme is discussed in Bell et al. (1997) and Bell and Iida (1997).
Column Generation

The above iterative balancing scheme assumes that a working path set is given. For large networks, it is not practical to enumerate a working path set in advance since the number of possible paths grows exponentially with respect to network size. To circumvent path enumeration, a column (or path) generation procedure can be augmented to the iterative balancing scheme. Basically, the algorithm introduces an outer loop (or iteration) to iteratively generate paths to the working path set as needed to replicate the observed interval constraints (e.g., link counts, selected prior O-D flows, etc.), and to account for the capacity restraints for the unobserved links as well as the congestion effects, while the iterative balancing scheme iteratively adjusts the primal variables (e.g., path flows, link flows, and O-D flows) and the dual variables in the inner loop for a given working path set from the outer loop. Note that the working path set is generated by a column generation scheme (or a shortest path algorithm) using the generalized link costs, which is based on not only the link costs but also the dual variables from the active side constraints. The dual variables force the column generation scheme to generate paths that satisfy the side constraints. For additional discussions on the issue of using the generalized link costs to generate paths, refer to Bell et al. (1997) and Chen et al. (2009, 2010).

Output

The unique path-flow solution from PFE makes it possible to derive useful information at different spatial levels. As an illustration, consider a simple network shown in Fig. 2.2 with six nodes, seven directed links, two origins, two destinations, and four O-D pairs. There are six paths serving the four O-D pairs. The path-link incidence indicator
is given in Fig. 2.3. Table 2.2 provides a summary of various outputs at different spatial levels derived from path flows.

Fig. 2.2 Illustration of using path flows to derive useful information at different spatial levels

*District level O-D 1 (O-D 1-3 and 2-4)

Fig. 2.3 Path-link incidence matrix
Table 2.2: Summary of various outputs at different spatial levels derived from path flows

<table>
<thead>
<tr>
<th>Output</th>
<th>Equation</th>
<th>Flow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total demand</td>
<td>$T = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} f^{rs}_{k}$</td>
<td>$f_1^{13} + f_2^{13} + f_3^{14} + f_4^{23} + f_5^{24} + f_6^{24}$</td>
</tr>
<tr>
<td>Production 1</td>
<td>$P_1 = \sum_{s \in \text{Sc}} \sum_{k \in \text{K}} f^{rs}_{k}$</td>
<td>$f_1^{13} + f_2^{14} + f_3^{14}$</td>
</tr>
<tr>
<td>Production 2</td>
<td>$P_2 = \sum_{s \in \text{Sc}} \sum_{k \in \text{K}} f^{rs}_{k}$</td>
<td>$f_4^{23} + f_5^{24} + f_6^{24}$</td>
</tr>
<tr>
<td>Attraction 1</td>
<td>$A_3 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} f^{rs}_{k}$</td>
<td>$f_1^{13} + f_2^{14} + f_4^{23}$</td>
</tr>
<tr>
<td>Attraction 2</td>
<td>$A_4 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} f^{rs}_{k}$</td>
<td>$f_3^{14} + f_4^{24} + f_6^{24}$</td>
</tr>
<tr>
<td>O-D (1-3)</td>
<td>$q_{13} = \sum_{k \in \text{K}} f^{rs}_{k}$</td>
<td>$f_1^{13} + f_2^{13}$</td>
</tr>
<tr>
<td>O-D (1-4)</td>
<td>$q_{14} = \sum_{k \in \text{K}} f^{rs}_{k}$</td>
<td>$f_3^{14}$</td>
</tr>
<tr>
<td>O-D (2-3)</td>
<td>$q_{23} = \sum_{k \in \text{K}} f^{rs}_{k}$</td>
<td>$f_4^{23}$</td>
</tr>
<tr>
<td>O-D (2-4)</td>
<td>$q_{24} = \sum_{k \in \text{K}} f^{rs}_{k}$</td>
<td>$f_5^{24} + f_6^{24}$</td>
</tr>
<tr>
<td>District-level O-D 1</td>
<td>$W_1 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} \sum_{a \in \text{Cl}} f^{rs}<em>{k} \delta^{rs}</em>{ka} \delta^{rs}_{lb}$</td>
<td>$f_1^{13} + f_2^{13} + f_5^{24} + f_6^{24}$</td>
</tr>
<tr>
<td>Screenline 1</td>
<td>$s_1 = \sum_{ae} \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} f^{rs}<em>{k} \delta^{rs}</em>{ka}$</td>
<td>$f_1^{13} + f_2^{13} + f_3^{14} + f_4^{23} + f_5^{24} + f_6^{24}$</td>
</tr>
<tr>
<td>Intersection turning movement 1</td>
<td>$t_1 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} \sum_{a \in \text{IN}} \sum_{b \in \text{OUT}} f^{rs}<em>{k} \delta^{rs}</em>{ka} \delta^{rs}_{lb}$</td>
<td>$f_1^{13} + f_3^{44}$</td>
</tr>
<tr>
<td>Intersection turning movement 2</td>
<td>$t_2 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} \sum_{a \in \text{IN}} \sum_{b \in \text{OUT}} f^{rs}<em>{k} \delta^{rs}</em>{ka} \delta^{rs}_{lb}$</td>
<td>$f_2^{13} + f_3^{24} + f_4^{24}$</td>
</tr>
<tr>
<td>Link 1</td>
<td>$x_1 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} f^{rs}<em>{k} \delta^{rs}</em>{ka}$</td>
<td>$f_1^{13}$</td>
</tr>
<tr>
<td>Link 2</td>
<td>$x_2 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} f^{rs}<em>{k} \delta^{rs}</em>{ka}$</td>
<td>$f_2^{14} + f_3^{13}$</td>
</tr>
<tr>
<td>Link 3</td>
<td>$x_3 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} f^{rs}<em>{k} \delta^{rs}</em>{ka}$</td>
<td>$f_6^{24}$</td>
</tr>
<tr>
<td>Link 4</td>
<td>$x_4 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} f^{rs}<em>{k} \delta^{rs}</em>{ka}$</td>
<td>$f_4^{23} + f_5^{24}$</td>
</tr>
<tr>
<td>Link 5</td>
<td>$x_5 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} f^{rs}<em>{k} \delta^{rs}</em>{ka}$</td>
<td>$f_2^{13} + f_3^{14} + f_4^{23} + f_5^{24}$</td>
</tr>
<tr>
<td>Link 6</td>
<td>$x_6 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} f^{rs}<em>{k} \delta^{rs}</em>{ka}$</td>
<td>$f_2^{13} + f_4^{23}$</td>
</tr>
<tr>
<td>Link 7</td>
<td>$x_7 = \sum_{rs \in \text{RS}} \sum_{k \in \text{K}} f^{rs}<em>{k} \delta^{rs}</em>{ka}$</td>
<td>$f_3^{14} + f_5^{24}$</td>
</tr>
</tbody>
</table>
REFERENCES


CHAPTER 3
A TWO-STAGE BICYCLE TRAFFIC ASSIGNMENT MODEL

Abstract

Many cities are observing a rise in cycling as an alternative mode of travel to conventional private motorized vehicles because of health, environmental, and economical benefits. However, this change in modal share is not reflected in current transportation planning and travel demand forecasting modeling processes. The current practices to model bicycle trips in a network are too simplistic; they are based only on the all-or-nothing assignment method, which uses single attributes. Thus, the purpose of this chapter is to develop a two-stage traffic assignment model by considering the key factors (or criteria) in cyclists’ route choice behavior. The first stage considers two key criteria (e.g., distance related attributes and safety related attributes) to generate a set of non-dominated (or efficient) paths, while the second stage determines the flow allocation to the set of efficient paths. In stage one, a bi-objective shortest path problem based on the two key attributes is developed to generate the efficient paths. In stage two, several traffic assignment methods are adopted to determine the flow allocations in a network. In addition, the multi-class, multi-criteria bicycle traffic assignment model is further extended by considering multi-user classes. Numerical experiments are conducted to demonstrate the two-stage approach for bicycle traffic assignment.

3.1 Introduction

Non-motorized modes such as bicycles constitute an important part of a community’s transportation system, and they are also vital to the success of transit-
oriented-developments (TODs). They were, however, often ignored in transportation planning and travel demand forecasting modeling, or at best treated as a byproduct in the planning process. In addition, many cities have begun to invest and to promote cycling as a healthy, environmentally friendly, and economical alternative mode of travel to the motorized vehicles (especially private motorized vehicles). However, the current practice in modeling bicycle trips in a network is inadequate. Only a few research efforts focus on network analysis for bicycle trips (e.g., Klobucar and Fricker, 2006, 2007; Broach et al., 2011; Mekuria et al., 2012). These methods provide an initial effort to develop a traffic assignment method for bicycle trips, but they are too simplistic. It is simply based on all-or-nothing (AON) assignment method that uses single attributes such as distance, safety, or a composite measure of safety multiplied by distance. Compared to the route choice models for private motorized vehicles, route choice behavior for cyclists is much more complex as there many influential factors affecting their route choice decisions. Many empirical studies on bicycle route choice analysis indicate that cyclists choose routes based on a number of criteria (e.g., distance, number of intersections, road grade, bike facility, safety, etc.). Stinson and Bhat (2003), Hunt and Abraham (2007), and Broach et al. (2011) found cyclists are concerned with travel distance or time when making route choice decisions, while Hopkinson and Wardman (1996), Akar and Clifton (2009), Dill and Carr (2003), and Winters et al. (2011) indicated that safety played an important role in selecting a suitable route for cyclists. Sener et al. (2009) also found that the travel distance/time and safety were important factors in cyclists’ route choices. Mekuria et al. (2012) suggested that stress is an important factor in the bicycle trip-making behavior. Using GPS tracking data, Hood et al. (2011) developed a path-size logit model (Ben-Akiva and Birelaire, 1999)
as a cyclist route choice model and performed the bicycle traffic assignment on a pre-enumerated path set generated by the doubly stochastic method (Bovy and Fiorenzo Catalano, 2007). Menghini et al. (2010) also adopted the a path-size logit model for traffic assignment model on pre-generated path set by breadth-first search link elimination approach.

The purpose of this chapter is to develop a two-stage traffic assignment model by considering two key factors (or criteria) in cyclists’ route choice behavior. The first stage considers two key criteria (distance related attribute and safety related attribute) to generate a set of non-dominated (or efficient) paths, while the second stage determines the flow allocation to the set of efficient paths. In stage one, a bi-objective shortest path problem based on the two key attributes is developed to generate the efficient paths. In stage two, several traffic assignment methods are adopted to determine the flow allocations in a network. Numerical experiments are conducted to demonstrate the two-stage approach for bicycle traffic assignment.

The remainder of this chapter is organized as follows. After the introduction, the two-stage bicycle traffic assignment procedure is presented in section 3.2, followed by two numerical experiments to demonstrate the features and applicability of the proposed two-stage procedure in section 3.3, and some concluding remarks in section 3.4.

3.2 Two-stage bicycle traffic assignment procedure

This section describes the proposed two-stage procedure for bicycle traffic assignment as shown in Fig. 3.1. In stage one, two key criteria, namely route distance and route level of service, are used in a bi-objective shortest path algorithm to generate a set of non-dominated (or efficient) paths. In stage two, several traffic assignment methods will
be used to determine the flow allocations to the efficient paths generated in stage one to obtain the complete bicycle flows on the network. The following subsections describe the two key cyclist route choice criteria, the bi-objective shortest path algorithm, and several traffic assignment methods for flow allocations to the efficient paths.

Fig. 3.1 Graphical illustration of the travel time function with vehicle interactions

3.2.1 Two key cyclist route choice criteria

Due to the diverse influential factors in the bicycle route choice model, using the conventional single objective as the sole criterion for determining the route choice decisions as in the private motorized vehicles [i.e., the Wardrop (1952) user equilibrium model based on flow-dependent travel times] may not be adequate in modeling cyclist route choice behavior (Menghini et al., 2010; Kang and Fricker, 2013). In this research, we adopt two key criteria (route distance related attributes and route safety related attributes) to capture the most important factors in affecting cyclist route choice behavior.
Route distance

Route distance is a composite measure of not only the sum of link distances along the route, but also the delays at signalized intersections that the route passes through. For bicycle trips, intersection delays have been shown to be a deterrence to cyclists’ route choice behavior. Since link length and intersection delay measure different qualities (length in kilometers and time in seconds, respectively), we convert delay to an equivalent distance unit with an appropriate conversion factor. The factor we used is as follows:

\[
    d_k^{rs} = \sum_{a \in A} l_a \delta_{ka}^{rs} + \sum_{a \in IN_i, b \in OUT_i} cf_{it}^{rs} d_i^{rs} \delta_{ka}^{rs} \delta_{kb}^{rs}, \quad rs \in RS, \ k \in K_{rs}
\]

where \( d_k^{rs} \) is the distance on route \( k \) connecting O-D pair \( rs \); \( l_a \) is the length on link \( a \); \( \delta_{ka}^{rs} \) is the route-link indicator; 1 if link \( a \) is on route \( k \) between O-D pair \( rs \) and 0; \( cf_{it}^{rs} \) is the conversion factor for turning movement \( t \) at intersection \( i \); \( d_i \) is the delay of turning movement \( t \) at intersection \( i \); \( A \) is the set of links; \( IN_i \) and \( OUT_i \) are the sets of links terminating into and originating out of intersection \( i \); \( RS \) is the set of O-D pairs; and \( K_{rs} \) is the set of routes connecting O-D pair \( rs \). The route distance in Eq. (3.1) can be computed by summing the link distances (first term) and the intersection delays (second term) caused by turning movement from link \( a \) to link \( b \) of intersection \( i \) that comprise of that route.

Route bicycle level of service (BLOS)

There are numerous measures for assessing the safety aspect of bicycle facilities or the suitability for bicycle travel. Lowry et al. (2012) provided a recent review of thirteen methods used in the literature. All methods attempt to provide a score of the perceived safety of bicycle facilities by using a linear regression with variables that represent conditions of the roadway and environment that affect a cyclists’ comfort level. For this
research, we adopt the bicycle level of service (BLOS) developed by the Highway Capacity Manual (HCM, 2010) as a surrogate measure to account for different attributes contributing to the safety of bicycle routes. The BLOS measure is considered as the state-of-the-art method, and has been adopted by many cities in the United States as a guide for bicycle facility design. However, other bicycle safety measures could also be used in our proposed framework for modeling cyclists’ route choice behavior. The route BLOS measure described in Eq. (2) is a composite measure based on the average segment bicycle score on a route \((\text{ABSeg})\), average intersection bicycle score on a route \((\text{ABInt})\), and average number of unsignalized conflicts/driveways per mile on a route \((\text{Cflt})\) as follows:

\[
BLOS_k = 0.200 \cdot (\text{ABSeg}_k) + 0.030 \cdot \left(\exp(\text{ABInt}_k)\right) + 0.050 \cdot (\text{Cflt}) + 1.40,
\]

\(\forall rs \in \text{RS}, \ k \in \text{K}_{rs}\)

where \(BLOS_k\) is the bicycle level of service on route \(k\) connecting O-D pair \(rs\); \(\text{ABSeg}_k\) is the length weighted average segment bicycle score on route \(k\) connecting O-D pair \(rs\)

\[
(\text{ABSeg}_k) = \left(\sum_{a \in \text{A}} l_a \cdot \text{Bseg}_a \cdot \delta_{ka}^{rs}\right) \div \left(\sum_{a \in \text{A}} l_a \cdot \delta_{ka}^{rs}\right);
\]

\(\text{ABInt}_k\) is the average intersection bicycle score on route \(k\) connecting O-D pair \(rs\)

\[
(\text{ABInt}_k) = \frac{\sum_{i \in \text{IN}, a \in \text{OUT}} \text{IntBLOS}_k \delta_{ka}^{rs} \delta_{kb}^{rs}}{N_k};
\]

\(\text{Cflt}_k\) is the number of unsignalized conflicts per mile; \(N_k\) is the number of intersections on route \(k\).

Note that the segment and intersection bicycle scores \((\text{Bseg}_a\) and \(\text{IntBLOS}_k\)) provided in Eqs. (3.3) and (3.4) are calibrated based the volume and speed of motorized vehicles, width configuration of bicycle facilities, pavement conditions, number of intersections, etc. The derived BLOS score is a relative measurement without score unit to
evaluate the comfortableness on cycling route. The details of the BLOS development can be found in NCHRP Report 616 (Dowling et al., 2008).

\[ BSeg_a = 0.507 \ln \left( \frac{V_a}{4 \cdot PHF_a \cdot La_a} \right) + 0.199 Fs_a \left( 1 + 10.38 \cdot HV_a \right)^2 + 7.066 \left( \frac{1}{PC_a} \right)^2 - 0.005(We_a)^2 + 0.76 \]  

(3.3)

\[ IntBLOS_i = -0.2144 \cdot Wt_i + 0.0153 \cdot CD_i + 0.0066 \left( \frac{Vol15_i}{L_i} \right) + 4.1324 \]  

(3.4)

where PHF is the peak hour factor of link \( a \); \( HV_a \) is the proportion of heavy vehicles of link \( a \) (in motorized vehicle volume); \( We_a \) is the average effective width on outside through lane of link \( a \) (ft); \( Fs_a \) is the effective speed factor on link \( a \); \( La_a \) is the total number of directional through lanes on link \( a \); \( V_a \) is the directional motorized vehicle volume on link \( a \) (vph); \( PC_a \) is the FHWA’s five point pavement surface condition rating on link \( a \); \( Wt_i \) is the width of outside through lane plus paved shoulder (including bike lane where present) of intersection \( i \); \( CD_i \) is the crossing distance, the width of the side street (including auxiliary lanes and median) of intersection \( i \); \( Vol15_i \) is the volume of directional traffic during a 15-minute period of intersection \( i \); and \( L_i \) is the total number of directional through lanes of intersection \( i \).

3.2.2 Stage one: Bi-objective shortest path procedure

Solving the bi-objective shortest path problem is similar to solving any multi-objective optimization problem as there may not exist a single optimal solution that dominates all other solutions in all objectives. Hence, solving multi-objective problems requires generating a set of non-dominated (or Pareto) solutions. Bi-objective shortest path
problem belongs to a class of NP-hard problems (Serafini, 1986). Several solution procedures have been developed to solve this complex problem. These include the label correcting approach (Skriver and Andersen, 2000), the label setting approach (Tung and Chew, 1992), the ranking method (Climaco and Martins, 1982), and the two-phase method (Ulungu and Teghem, 1995). However, handling non-additive route cost structure (e.g., route BLOS) may not be easy in these methods. In this paper, we adopt the two-phase procedure used in Ehrgott et al. (2012) to solve the bi-objective shortest problem with non-additive route cost structure. The overall two-phase procedure is described in Fig. 3.2. In the first phase, it uses the distance-related attributes (i.e., link distance and intersection delay) to generate a set of realistic routes without exceeding the maximum allowable bound. In the second phase, the corresponding safety-related attributes are computed for each route in the set to determine the efficient routes according to the two key criteria, route distance and route BLOS.

![Fig. 3.2 Two-phase procedure for generating non-dominate routes](image-url)
3.2.3 Stage two: Bicycle traffic assignment methods

The conventional bi-objective traffic assignment model was introduced by Dial (1979) for the multi-class traffic assignment problem. Dial (1979, 1996, 1997) adopted a linear value of time (VOT) function to convert travel time to an equivalent monetary unit, while Gabriel and Bernstein (1997) introduced a nonlinear VOT function for the non-additive traffic equilibrium problem. Nagurney (2000), Nagurney et al. (2001, 2002), and Nagurney and Dong (2002) proposed variable weights for the multi-criteria traffic assignment problem by assuming a linear generalized cost function for combining the criteria with variable weights. Recently, Raith et al. (2014) introduced four traffic assignment methods for solving the multi-objective traffic assignment problem. In this paper, we adopt these traffic assignment methods for flow allocations to the efficient paths identified in stage 1. In addition, we include a path-size logit (PSL) traffic assignment method, which has been widely adopted in the literature as a multi-path traffic assignment method. Table 3.1 provides a summary of the traffic assignment methods for flow allocations in stage 2.

Equal share assignment (ESA) method

The ESA method evenly assigns the O-D demand to all efficient routes as follows.

\[
f_k^{rs} = \frac{q_{rs}}{|K_{rs}|}
\]  

(3.5)

where \( f_k^{rs} \) is the flow on route \( k \) connecting O-D pair \( rs \); \( q_{rs} \) is the demand between O-D pair \( rs \), and \( |K_{rs}| \) is the number of routes in O-D pair \( rs \). Hence, each efficient route in O-D pair \( rs \) has an equal share of the O-D demand.
Table 3.1: Traffic assignment methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal share assignment (ESA)</td>
<td>O-D demand is split evenly between all efficient routes</td>
<td>Easy to implement</td>
<td>Assignment results not dependent on the objective values</td>
</tr>
<tr>
<td>Travel distance per benefit of BLOS assignment (TBA)</td>
<td>O-D demand is assigned according to the distribution of travel distance per unit of better BLOS compared to the shortest distance route</td>
<td>Enable the assignment of flow to non-supported routes (non-convex points) and non-extreme supported routes</td>
<td>Sensitive to the assumed distribution</td>
</tr>
<tr>
<td>Reference point assignment (RPA)</td>
<td>The route attractiveness is determined by the Euclidean distance to the reference point, and the probability is determined by the route attractiveness relative to the attractiveness of other routes</td>
<td>Easy and intuitive in modifying the shares of demand assigned to each efficient route</td>
<td>Sensitive to the reference point and potential bias with different objective scales</td>
</tr>
<tr>
<td>Dominated area assignment (DAA)</td>
<td>Shares of demand are assigned to the efficient routes based on the part of the objective space dominated by the corresponding route attribute point</td>
<td>Consider the attributes of the non-dominated routes</td>
<td>Sensitive to the maximum objective values (extreme supported routes)</td>
</tr>
<tr>
<td>Path-size logit assignment (PSLA)</td>
<td>O-D demand is assigned based on the combined utilities of two objectives</td>
<td>Account for the total route cost values and an economic interpretation</td>
<td>Require detailed survey data to calibrate the parameters</td>
</tr>
</tbody>
</table>

Travel distance per benefit of BLOS assignment (TBA) method

The TBA method assigns the O-D demand according to the distribution of travel distance per benefit of BLOS relative to the shortest distance route. The slopes ($\rho$) between the shortest distance route and other efficient routes in the Pareto set represent the travel
distance per benefit of BLOS. With the computed slopes, $\rho$, route choice probabilities can be obtained from a predetermined distribution function as follows.

- Compute $\rho_{rs}^k$ between shortest distance route $k_{rs}$ and other efficient route $k$ in O-D pair $rs$
- Compute route choice probability with $\rho_{rs}^k$
  
  $\Pr[k_{rs}] = Pr[0 \leq \rho_{rs}^k < \rho_{rs}^{k+1}] = \int_{0}^{\rho_{rs}^{k+1}} f(\rho) \ d\rho$
  
  $\Pr[k_{rs}] = Pr[\rho_{rs}^k \leq \rho_{rs}^k < \rho_{rs}^{k+1}] = \int_{\rho_{rs}^k}^{\rho_{rs}^{k+1}} f(\rho) \ d\rho$
  
  $\Pr[K_{rs}] = 1 - \sum_{k} \Pr[k_{rs}]$

Reference point assignment (RPA) method

The RPA method assigns the O-D demand based on the route attractiveness. The route attractiveness is determined by the Euclidean distance ($\varepsilon$) to the reference point (i.e., a virtual or an ideal point), and the route choice probability is determined by the computed route attractiveness. Raith et al. (2014) introduced the following three different probability functions.

$$P_{rs}^k = \frac{\sum_{l=1}^{l=k} \varepsilon_{rs}^{l} - \varepsilon_{rs}^{k}}{(|K_{rs}| - 1) \sum_{l=1}^{l=k} \varepsilon_{rs}^{l}} \quad (3.6)$$

$$P_{rs}^k = \frac{\sum_{l=1}^{l=k} (\varepsilon_{rs}^{l})^2 - (\varepsilon_{rs}^{k})^2}{(|K_{rs}| - 1) \sum_{l=1}^{l=k} (\varepsilon_{rs}^{l})^2} \quad (3.7)$$

$$P_{rs}^k = \frac{\prod_{l=k}^{l=k} \varepsilon_{rs}^{l}}{\sum_{l=1}^{l=k} \prod_{l=1}^{l=k} \varepsilon_{rs}^{l}} \quad (3.8)$$
The first function given in Eq. (3.6) is a sum-based approach. The target $\varepsilon_k^{\alpha}$ is extracted from the sum of all routes $\sum_{l=1}^{K} \varepsilon_l^{\alpha}$, and the probability can be determined by dividing the total Euclidean distance of all routes weighted by the number of routes minus the target route. Alternatively, the probability can be computed based on the squares sum as shown in Eq. (3.7). Finally, the product approach is introduced in Eq. (3.8).

**Dominated area assignment (DAA) method**

The DAA method assigns the O-D demand to the probability obtained from the share space computed with both objective values. See Fig. 3.3 for an illustration of the share space and route choice probability.

\[
\text{Share space on route } k = \frac{\sum_{n=1}^{K} S_{kn} / n + S_{k0}}{\sum_{l=1}^{K} \sum_{n=1}^{K} S_{ln} + \sum_{l=1}^{K} S_{l0}}
\]

**Example for route 1**

\[
P_1^{\alpha} = \frac{(S_{11} + S_{12} / 2 + S_{13} / 3 + S_{10})}{30.0} = 0.11
\]

Fig. 3.3 Illustration of route choice probability using the DAA method

**Path-size logit assignment (PSLA) method**

The PSLA method assigns the O-D demand based on the combined utilities of two objectives via the path size logit choice function. The multinomial logit (MNL) model is a
widely used route choice model under the random utility principle. However, it is well known that the major drawbacks in applying the MNL model to the route choice problem are inability to account for overlapping (or correlation) among routes. Ben-Akiva and Bierlaire (1999) proposed the path size logit model (PSL) as an alternative to solve the overlapping problem in MNL. The closed-form probability of PSL is expressed as follows.

\[ P_{krs} = \frac{PS_{krs} \cdot \exp \left( U_{krs} \right)}{\sum_{j=1}^{n} PS_{jrs} \cdot \exp \left( U_{jrs} \right)}, \quad \forall \ k \in K, \ rs \in RS \]  

(3.9)

where, \( U_{krs} = -\left( \left( d_{krs}^{\alpha} \cdot (BLOS_{krs}^{\beta}) \right) \right), \quad \forall \ k \in K, \ rs \in RS \)

\[ PS_{k} = \sum_{\alpha \in a} \left( \frac{L_{rs}^{\alpha}}{L_{krs}^{\alpha}} \right) \sum_{\lambda \in \lambda_{K}} \delta_{\lambda}, \quad \forall \ k \in K, \ rs \in RS \]

\( L_{rs}^{\alpha} \) : path length on path \( k \) between origin \( r \) and destination \( s \); \( l_{s} \) : the length of link \( a \).

### 3.2.4 Extension to multi-class and multi-criteria assignment

In this section, we extend the single class model to multi-class model. Based on the Portland study (Geller, 2006), it has been suggested that there are four types of transportation cyclists as indicated in Fig. 3.4. These four types of cyclists are: (1) strong and fearless, (2) enthused and confident, (3) interested but concerned, and (4) no way no how. Strong and fearless cyclists represent less than 1%; they are the rare daily commuters who “will ride regardless of roadway conditions.” Enthused and confident cyclists represent 7% and are semi-regular cyclists who are “comfortable sharing the roadway with automotive traffic, but they prefer to do so operating on their own facilities” (e.g., bicycle lanes and bicycle boulevards). Interested but concerned cyclists, who represent the 60% of the cyclist, are irregular cyclists who are “curious about cycling” but are concerned with riding a bicycle. Lastly, no way no how travelers represent 33% and are simply “not
interested in bicycling at all, for reasons of topography, inability, or simply a complete and utter lack of interest”. The study also noted that “the separation between these four broad groups is not generally clear-cut”. However, this classification with percentage to each user class serves as a good foundation to develop a multi-class version of the multi-criteria bicycle traffic assignment model. For the extended model, route pollution related attributes are added as follows.

![Fig. 3.4 Four types of cyclists in Portland (Geller, 2006)](image)

**Route pollution**

We consider carbon monoxide (CO) as an important indicator for the level of atmospheric pollution. Other pollutants can be modeled in the similar manner. In this research, the route pollution is computed as follows.

\[
CO_k^a = \sum_{a \in A} g_a \cdot \delta_{rs}^{k} \quad \forall k \in K, rs \in RS
\]

where \( g_a \) is the amount of CO pollution in grams per hour (g/h) on link (segment) \( a \). To estimate the amount of CO pollution, we adopt the nonlinear macroscopic model of Wallace et al. (1998).

\[
g_a(\overline{v}_a) = 0.2038 \cdot t_a(\overline{v}_a) \cdot \exp \left( \frac{0.7962 \cdot l_a}{t_a(\overline{v}_a)} \right), \quad a \in A
\]

where \( \overline{v}_a \) is the motorized vehicle volume on link \( a \); \( t_a(\overline{v}_a) \) is the link travel time by measured in minutes; and \( l_a \) is the link length measured in kilometers.
3.3 Numerical results

To demonstrate the proposed two-stage bicycle traffic assignment procedure, two networks are adopted in the numerical experiments. First, a simple network is used to illustrate the features of the different traffic assignment methods. Then, a real network is employed to demonstrate the applicability of the two-stage procedure for single class and multi-class model.

3.3.1 A simple network

The network shown in Fig. 3.5 is used to illustrate the features of different traffic assignment methods for bicycle trips. To simplify the analysis, we assume both objectives (i.e., distance and BLOS) are obtained from a prior analysis. In the left panel, the numbers in parenthesis next to each link number are the link distance and link BLOS, while the turning delay and intersection BLOS are provided in the right panel. The travel demand from node 1 to node 5 is 10.

![Test Network and link characteristics](image)

Fig. 3.5 Test network and link characteristics
Using the link characteristics in Fig. 3.5, the route distance and route BLOS can be computed as shown in Table 3.2. In this experiment, there are four non-dominated routes. Route 1 is the shortest distance route, route 4 has the least BLOS score (i.e., a lower BLOS score means a higher level of service), while route 5 and route 6 are non-dominated routes between the two extremes (i.e., have route distance and route BLOS between the shortest distance route and the least BLOS route).

Table 3.2: Estimated route distance and route BLOS and the corresponding generated non-dominated routes

<table>
<thead>
<tr>
<th>Route #</th>
<th>Link member</th>
<th>Route distance</th>
<th>Route BLOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-5</td>
<td>6.00</td>
<td>2.33</td>
</tr>
<tr>
<td>2</td>
<td>1-4-10</td>
<td>10.30</td>
<td>2.34</td>
</tr>
<tr>
<td>3</td>
<td>1-4-9-7</td>
<td>12.50</td>
<td>2.29</td>
</tr>
<tr>
<td>4</td>
<td>3-8-5</td>
<td>12.50</td>
<td>1.88</td>
</tr>
<tr>
<td>5</td>
<td>3-10</td>
<td>6.80</td>
<td>2.17</td>
</tr>
<tr>
<td>6</td>
<td>3-9-7</td>
<td>9.00</td>
<td>1.98</td>
</tr>
<tr>
<td>7</td>
<td>2-7</td>
<td>7.00</td>
<td>2.23</td>
</tr>
<tr>
<td>8</td>
<td>2-6-10</td>
<td>8.80</td>
<td>2.33</td>
</tr>
<tr>
<td>9</td>
<td>2-6-8-5</td>
<td>12.50</td>
<td>2.12</td>
</tr>
</tbody>
</table>

Using these generated non-dominated routes, we perform the following bicycle traffic assignment methods:

- ESA: Uniformly allocate the O-D demand to the four efficient routes
- TBA: Gamma distribution with $\alpha:2.00; \beta:2.97$ (i.e., the assumed parameters give the following $\Pr[\rho \leq 10.0] = 85\%$).
- RPA: Route distance and route BLOS for the reference point are 5.0 and 1.90, respectively.
- DAA: Maximum distance is 15.0 and maximum route BLOS is 5.0
• PSLA: $\alpha: 0.862; \beta: 0.117$ [assumed parameters obtained from Kang and Fricker (2013)]

Table 3.3 provides a comparison of assigned flows using the five traffic assignment methods for bicycle trips. From the table, we can observe that all methods assign more flows to the shortest distance route (route 1) with the exception of the ESA and DAA methods. As mentioned, the ESA method assigns equal amount of flows to all four efficient routes regardless of the objective values on the routes, while the DAA method assigns flows according to the share space of the route, which is sensitive to the maximum objective values of the extreme supported routes. The TBA and RPA methods allocate flows to the efficient routes using the objective values (i.e., route distance and route BLOS) in different ways. In the TBA method, the O-D demands are assigned according to the distribution of travel distance per unit of better BLOS compared to the shortest distance route. It enables the assignment of flow to non-supported routes (non-convex points) and non-extreme supported routes. As for the RPA method, it uses three probability functions given in Eqs. (3.6) to (3.8) to assign the O-D demand based on the route attractiveness determined by the Euclidean distance to the reference point. It is intuitive and easy to modify the shares of demand assigned to each efficient route. However, both methods do not explicitly consider cyclists’ actual route choice behavior (i.e., no calibration). The PSLA method, on the other hand, requires additional survey and parameter calibration to fit the cyclists’ choice to the two key criteria. In this study, we adopt the parameter values from Kang and Fricker (2013). From Table 3.3, it seems that both TBA and RPA using Eq. (3.8) can produce assignment results similar to that of the PSLA model. In summary, the ESA and DAA methods, albeit simple, are not suitable for modeling cyclists’ route choice behavior.
since it either does not consider the objective values or is sensitive to the maximum values when assigning flows to the efficient routes. The TBA, RPA using Eq. (3.8), and PSLA methods seem to produce flow patterns that not only account for the objective values but also reflect cyclists’ route choice behavior.

Table 3.3: Comparison of assigned flows using five assignment methods

<table>
<thead>
<tr>
<th>Route #</th>
<th>Route dist.</th>
<th>Route BLOS</th>
<th>ESA</th>
<th>TBA</th>
<th>RPA Eq. (3.6)</th>
<th>RPA Eq. (3.7)</th>
<th>RPA Eq. (3.8)</th>
<th>DAA</th>
<th>PSLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.00</td>
<td>2.33</td>
<td>2.50</td>
<td>4.99</td>
<td>3.08</td>
<td>3.28</td>
<td>4.96</td>
<td>3.78</td>
<td>6.04</td>
</tr>
<tr>
<td>5</td>
<td>6.80</td>
<td>2.17</td>
<td>2.50</td>
<td>2.87</td>
<td>2.91</td>
<td>3.19</td>
<td>2.97</td>
<td>0.70</td>
<td>3.06</td>
</tr>
<tr>
<td>6</td>
<td>9.00</td>
<td>1.98</td>
<td>2.50</td>
<td>1.70</td>
<td>2.41</td>
<td>2.64</td>
<td>1.35</td>
<td>0.87</td>
<td>0.84</td>
</tr>
<tr>
<td>4</td>
<td>12.50</td>
<td>1.88</td>
<td>2.50</td>
<td>0.44</td>
<td>1.60</td>
<td>0.89</td>
<td>0.72</td>
<td>4.64</td>
<td>0.06</td>
</tr>
</tbody>
</table>

3.3.2 Winnipeg network

In this section, the two-stage approach is applied to a real network in the City of Winnipeg, Canada. The Winnipeg network, shown in Fig. 3.6, consists of 154 zones, 1,067 nodes, 2,555 links (1,943 links without centroid connectors), and 4,345 O-D pairs for motorized vehicles. The network structure, O-D trip table for motorized vehicles, and link performance parameters are from the Emme/4 software (INRO Consultants, 2013). The bicycle network is assembled based on the information obtained from the City of Winnipeg (2013a). Among the 2,555 links, 541 links include bike paths or bike lanes. Using the 2006 census data (City of Winnipeg, 2013b), the bicycle O-D demand is created based on the gravity model with the gamma function. To create the multi-class bicycle O-D trip tables, the bicycle O-D demand is segmented into the three user classes mentioned in Fig. 3.4 (i.e., strong and fearless cyclists, enthused and confident cyclists, and interested but concerned cyclists).
Characteristics of the Winnipeg Network

Fig. 3.7 shows the characteristics of the Winnipeg network that are used to compute the three key route choice criteria described in Section 3.2.1 and Section 3.2.4. Fig. 3.7(a) depicts the link length distribution that is used for computing the route choice criteria; Fig. 3.7(b) and Fig. 3.7(c) plot the motorized volume and speed distributions obtained from the multi-class motorized vehicle traffic assignment results provided by the Emme/4 software (INRO Consultants, 2013); Fig. 3.7(d) and Fig. 3.7(e) show the computed bicycle segment and intersection LOS distributions based on Eqs. (3.3) and (3.4) from HCM (2010); and Fig. 3.7(f) plots the link CO distribution based on the nonlinear macroscopic model of Wallace et al. (1998) given in Eq. (3.11).
Fig. 3.7 Characteristics of the Winnipeg network
A segment with a high motorized vehicle volume typically gives a higher BLOS value, while a lower BLOS value is typically for links with a larger effective width on outside through lane. In addition, a segment with a high motorized vehicle volume typically yields a large value of CO due to the congestion effect. These characteristics of the Winnipeg network serve as the input factors for calculating the three route criteria: route distance, route BLOS, and route pollution.

**Stage One: Bicycle BLOS analysis and route generation results**

Fig. 3.8 shows the generated route results in the Winnipeg network. To compute the BLOS measures in Eqs. (3.3) and (3.4), traffic conditions (e.g., motorized vehicle volumes) and space availability (e.g., lane width) are obtained from the multi-class traffic assignment results provided by Emme/4 software and Google Earth, respectively. A segment with a high motorized vehicle volume typically gives a higher BLOS value, while a lower BLOS value is typically for links with a larger outside lane width. After evaluating the BLOS measures, the two-phase bi-objective shortest path procedure is performed to generate the efficient routes in terms of route distance and route BLOS for each O-D pair in the Winnipeg network [see Fig. 3.8(a) and Fig. 3.8(b)]. In total, there are 58,846 efficient routes. Longer distance O-D pairs typically have more efficient routes, while shorter distance O-D pairs have less efficient routes. As for the route distribution in terms of BLOS, most routes are between 2.5 and 4.0, which correspond to BLOS of B, C, and D. Fig. 3.8(c) provides an illustration of the efficient routes for O-D pair (1-8). Route 1 has the shortest distance (1.11 km) and the worst BLOS (3.50), while route 5 has the best BLOS (2.6) and longest distance (1.58 km).
Stage two (1): Bicycle traffic assignment results for single class

Using the generated efficient routes in the first stage, we perform three bicycle traffic assignment methods: TBA, RPA with Eq. (3.8), and PSLA with the following assumed parameters:

- **TBA**: Gamma distribution with $\alpha = 1.50; \beta = 0.32$ (i.e., the assumed parameters give the following $Pr[\rho \leq 1.0 \text{(km per BLOS)}] = 90\%$).

- **RPA**: Route distance and route BLOS for the reference point are $\min\{d^{(r)}_k\}$ and $\min\{BLOS^{(r)}_k\}$.

Fig. 3.8 Generated routes analysis based on route distance and route BLOS
- PSLA: Parameters for the utility function are $\alpha = 0.862; \beta = 0.117$ (Kang and Fricker; 2013).

Figs. 3.9(a), 3.9(b), and 3.9(c) depict the link flow patterns of TBA, RPA, and PSLA, while Fig. 3.9(d) compares the link flow distributions of the three assignment methods.

Visually, the three link flow patterns look similar. The main differences from Fig. 3.9(d) are that TBA and RPA assign a higher percentage of links to low flow values (i.e.,
0 to 10 units), while PSLA assigns a higher percentage of links to medium flow values (i.e., 10 to 50 units). For the high flow values (i.e., 50 to 100+), the three assignment methods identify similar number and location of links in the network as shown by the red color coded links in Figs. 3.9(a), 3.9(b), and 3.9(c).

In terms of the flow distributions allocated by route distance and route BLOS, Fig. 3.10 shows the results of the three assignment methods. The TBA method tends to assign more flows to the shorter distance routes (0 to 3 km) with a higher value of BLOS or a lower level of safety (3 to 4.5+), while both RPA and PSLA methods seem to assign similar percentage of flows by route distance with some variations by route BLOS. The aggregate measures, total traveled distance (TTD) and average traveled distance (ATD), total traveled BLOS (TTB) and average traveled BLOS (ATB), are also computed for the three traffic assignment methods (see the bottom of Fig. 3.10).

<table>
<thead>
<tr>
<th></th>
<th>TBA</th>
<th>RPA</th>
<th>PSLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTD</td>
<td>18,607</td>
<td>20,142</td>
<td>20,161</td>
</tr>
<tr>
<td>ATD</td>
<td>3.34</td>
<td>3.61</td>
<td>3.62</td>
</tr>
</tbody>
</table>

(a) Route distance (km)

<table>
<thead>
<tr>
<th></th>
<th>TBA</th>
<th>RPA</th>
<th>PSLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTB</td>
<td>21,054</td>
<td>20,299</td>
<td>20,626</td>
</tr>
<tr>
<td>ATB</td>
<td>3.78</td>
<td>3.64</td>
<td>3.70</td>
</tr>
</tbody>
</table>

(b) Route BLOS

TTD: Total Traveled Distance; ATD: Average Traveled Distance
TTB: Total Traveled BLOS; ATB: Average Traveled BLOS

Fig. 3.10 Route flow distribution in terms of route distance and route BLOS
Similar to route flow distribution, the TBA method has the TTD and ATD and the highest TTB and ATB. On the other hand, both RPA and PSLA methods have similar TTD (20,142 km and 20,161 km) and ATD (3.61 km and 3.62 km), but the RPA method assigns a slightly lower TTB and ATB than that of the PSLA method (20,299 and 20,626 for TTB and 3.64 and 3.70 for ATB). Overall, the results of the Winnipeg network demonstrate the applicability of the two-stage bicycle traffic assignment procedure with the flexibility of using traffic assignment methods.

Stage two (2): Bicycle traffic assignment results for multi-class

To demonstrate the multi-class, multi-criteria bicycle traffic assignment problem, three classes of cyclists are adopted to develop the numerical experiments for examining the effects of multiple user classes and multiple criteria on the bicycle traffic assignment results. The three cyclist classes are as follows: the “strong and fearless” cyclist class (who compose of less than 1% of the population), the “enthused and confident” cyclist class (7% of the population), and the “interested but concerned” cyclist class. The “no way no how” cyclist class, who comprise 33% of the population, are not included in the numerical experiments because this user class does not consider cyclist as a potential mode.

Based on the obtained O-D demand in single class model, the O-D demand for each class are generated using the proportion in four type of cyclist in Fig. 3.4. The “no way no how” cyclist class, who comprise 33% of the population, is not included in the numerical experiments because this user class does not consider cyclist as a potential mode. Table 3.4 shows the generated total demand for each user class and total demand for bicycle trips.
Table 3.4: Generated total demand for each class

<table>
<thead>
<tr>
<th>Class #</th>
<th>Type</th>
<th>Proportion</th>
<th>Total demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Strong and Fearless</td>
<td>1.5%</td>
<td>82.0</td>
</tr>
<tr>
<td>2</td>
<td>Enthused and Confident</td>
<td>10.3%</td>
<td>573.9</td>
</tr>
<tr>
<td>3</td>
<td>Interest but Concerned</td>
<td>88.2%</td>
<td>4919.1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>100.0%</td>
<td>5575.0</td>
</tr>
</tbody>
</table>

Two scenarios are set up to examine the effects of using different number of criteria as the utility function on the multi-class bicycle traffic assignment model. Table 3.5 provides a summary of the two scenarios. Scenario 1 assumes all user classes adopt two criteria for the utility function, but the two criteria are different for each user class. On the other hand, Scenario 2 assumes the following: class 1, the strong and fearless cyclist class, is only concerned about route distance; class 2, the enthused and confident cyclist class, uses both route distance and route BLOS; and class 3, the interested but concerned cyclist class, adopts all three criteria (route distance, route BLOS, and route CO) for route choice decisions.

Table 3.5: Summary of criteria used for the utility function of each user class in the two scenarios

<table>
<thead>
<tr>
<th></th>
<th>Scenario 1</th>
<th>Scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class 1</td>
<td>Class 2</td>
</tr>
<tr>
<td>Route Distance</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Route BLOS</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Route CO</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
Using the efficient routes generated from the first stage for all user classes, we perform the customized path-based algorithm for assigning the multi-class bicycle O-D trip tables to the network according to the PSL stochastic loading method. In this research, the following parameters are assumed for the utility function in Eq. (3.9): $\alpha = 0.862$, $\beta = 0.117$; (these two values are obtained from Kang and Fricker, 2013), and $\gamma = 0.05$ (this value is assumed). Fig. 3.11 depicts the link flow pattern of each user class for both scenarios. Note that the magnitude of the link flow is color coded and represented by the thickness of the line. For the link flow pattern of class 1 in scenario 1, the total number of efficient routes (using the criteria route distance and route pollution) is 10,695. Conversely, in scenario 2 (which uses route distance as the sole criterion), the total number of efficient routes is 7,368. Therefore the link flow patterns between the two scenarios are quite different since different numbers and route utilities are being used to assign the O-D demand of class 1. On the other hand, users in class 2 of both scenarios use the same two objectives (route distance and route BLOS) to compute the route utilities, and consequently yield the same link flow pattern.

As for class 3, the two scenarios adopt different objectives (i.e., route BLOS and route pollution for scenarios 1 and all three route criteria for scenario 2) and generate different numbers of efficient routes. However, the resulting link flow patterns are similar visually as this class has the largest amount of O-D trips (88% of total demand or 4,919 trips out of 5,575 trips) compared to 656 trips or less than 12% in class 2 and class 3 (see Table 3.4).
<table>
<thead>
<tr>
<th>Class</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1" alt="Map 1" /></td>
<td><img src="image2" alt="Map 2" /></td>
</tr>
<tr>
<td>2</td>
<td><img src="image3" alt="Map 3" /></td>
<td><img src="image4" alt="Map 4" /></td>
</tr>
<tr>
<td>3</td>
<td><img src="image5" alt="Map 5" /></td>
<td><img src="image6" alt="Map 6" /></td>
</tr>
</tbody>
</table>

Fig. 3.11 Link flow pattern of each user class for both scenarios
For the aggregate network measures, Table 3.6 provides the average traveled distance, the average traveled BLOS, and the average traveled CO for each user class computed according the following equations:

Average traveled distance: \( ATD^m = \sum_{rs \in R} \sum_{k \in K^m} d^r_k f^r_{mk} \sum_{rs \in R} \sum_{k \in K^m} f^r_{mk}, \forall \ m \in M \) (3.12)

Average traveled BLOS: \( ATB^m = \sum_{rs \in R} \sum_{k \in K^m} BLOS^r_k f^r_{mk} \sum_{rs \in R} \sum_{k \in K^m} f^r_{mk}, \forall \ m \in M \) (3.13)

Averaged traveled CO: \( ATC^m = \sum_{rs \in R} \sum_{k \in K^m} CO^r_k f^r_{mk} \sum_{rs \in R} \sum_{k \in K^m} f^r_{mk}, \forall \ m \in M \) (3.14)

where \( f^r_{mk} \) is the flow on route \( k \) connecting O-D pair \( rs \) of class \( m \); and \( M \) is the set of class.

Table 3.6: Average traveled distance, BLOS and CO for each user class and all user classes

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Route distance (km/h)</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>Route distance (km/h)</td>
<td>4.808</td>
<td>5.129</td>
<td>5.423</td>
<td>5.384</td>
</tr>
<tr>
<td></td>
<td>Route BLOS</td>
<td>3.863</td>
<td>3.612</td>
<td>3.566</td>
<td>3.575</td>
</tr>
<tr>
<td></td>
<td>Route CO (g/h)</td>
<td>3.999</td>
<td>4.221</td>
<td>4.335</td>
<td>4.318</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>Route distance (km/h)</td>
<td>4.787</td>
<td>5.129</td>
<td>5.136</td>
<td>5.130</td>
</tr>
<tr>
<td></td>
<td>Route BLOS</td>
<td>3.847</td>
<td>3.612</td>
<td>3.639</td>
<td>3.640</td>
</tr>
<tr>
<td></td>
<td>Route CO (g/h)</td>
<td>4.022</td>
<td>4.221</td>
<td><strong>4.198</strong></td>
<td>4.198</td>
</tr>
</tbody>
</table>

*bold and red fonts indicate the criteria used for the specific user class*

A cursory glance at Table 3.6 would reveal several obvious patterns in aggregate network measures. Firstly, the table shows that route distance seems to have a higher impact when comparing the two scenarios (e.g., class 1 and class 3). This is particularly obvious in scenario 2: class 1, which uses route distance as its only criterion, has the lowest average traveled distance among the three user classes and in both scenarios. Secondly, the table shows a positive correlated relationship between route distance and route CO.
Minimizing route distance implicitly reduces the value of route CO [see Eqs. (3.10) and (3.11)]. A closer inspection of Table 3.6 would reveal the effects of using multiple criteria in the calculation of the aggregate network measures. The effect can be readily observed in scenario 1 by examining the values for route BLOS and route CO for class 2 and 3. Since class 2 focuses on minimizing route distance and route BLOS while class 3 focuses on minimizing route BLOS and route CO, we might expect that classes 2 and 3 would have lower values for their respective criteria of focus. However, Table 3.6 shows that Class 3’s value for route CO is higher than Class 2’s value for route CO even though route CO was minimized in Class 3 and not in Class 2. These unexpected results may be attributed to the multi-criteria aspect of the calculation process. It is likely that when more than one criteria affected by the weighting between the two assigned criteria. In the case of class 3’s relatively high route CO value, route BLOS was given more weight than route CO during the calculation process. Thus, we can conclude that the flow patterns are sensitive to the different combinations of criteria.

For the disaggregate analysis, we examine the effect of multi-class and multi-criteria considerations on the route choice probabilities. Specifically, user classes considered in the analysis include single class and multiple classes. For the single user class, two different utility functions are used for comparison; the first utility function uses two criteria (route distance and route BLOS), and the second utility function uses three criteria (route distance, route BLOS, and route pollution). For the multiple user classes, we continue to use the set up in the two scenarios [scenario 1 (S1) and scenario 2 (S2)]. For demonstration purposes, we use O-D pairs (5-2) and (43-4) to respectively represent a short O-D pair and a long O-D pair in the Winnipeg network. Fig. 3.12 shows three major
efficient routes for each O-D pair and the route choice probabilities. For both O-D pairs, route 1 is the shortest-distance route among three efficient routes, while the other two are efficient routes (but these routes do not necessarily have the best value in the other two criteria). For the short O-D pair (5-2), Fig. 3.12(a) shows that the single user class with a bi-criteria utility function assigns a higher probability for all three routes compared to those of the single user class with a three criteria utility function and both scenarios of the multiple user classes. The reason is that the number of efficient routes generated for the short O-D pair using the single user class with bi-criteria utility function is much less compared to the other cases. Therefore, it assigns a higher probability to these efficient routes. Fig. 3.12(b) shows that there is less disparity in the assigned probabilities to the three major efficient routes for the long O-D pair (43-3) compared to the short O-D pair (5-2). Also, scenario 2 assigns a higher probability to route 1 since it only uses the route distance as the objective for generating efficient routes, while scenario 1 considers both route distance and route pollution. From Fig. 3.12(c), we can observe that the route choice probabilities of each class are significantly different in the multi-class analysis. Cyclists from class 1 travel only on the shortest route in both scenarios. Although scenario 1 considers two objectives, the network generates only one efficient route because both objectives (i.e., route distance and route CO) are highly correlated. There are a few notable differences in route choice probabilities within individual user classes. In the long O-D pair (43-4) analysis, the probabilities between route 1 and 3 for class 2 cyclists in both scenarios differ by 4.4 percentage points. For class 3 cyclists, there is little variance in route choice probability for all three routes in scenario 1. However, in scenario 2, class 3 cyclists
experience greater variance in route choice probability; the probabilities between route 2 and 3 differ by 5.1 percentage points.

<table>
<thead>
<tr>
<th>Route #</th>
<th>Single-class</th>
<th>Multi-class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bi-obj.</td>
<td>Tri-obj.</td>
</tr>
<tr>
<td>1</td>
<td>18.1%</td>
<td>12.1%</td>
</tr>
<tr>
<td>2</td>
<td>17.8%</td>
<td>11.9%</td>
</tr>
<tr>
<td>3</td>
<td>17.3%</td>
<td>11.5%</td>
</tr>
</tbody>
</table>

(a) O-D pair (5-2)

<table>
<thead>
<tr>
<th>Route #</th>
<th>Single-class</th>
<th>Multi-class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bi-obj.</td>
<td>Tri-obj.</td>
</tr>
<tr>
<td>1</td>
<td>27.8%</td>
<td>28.2%</td>
</tr>
<tr>
<td>2</td>
<td>27.9%</td>
<td>28.3%</td>
</tr>
<tr>
<td>3</td>
<td>23.4%</td>
<td>23.2%</td>
</tr>
</tbody>
</table>

(a) O-D pair (43-4)

(c) Route choice probabilities of each class in both scenarios

Fig. 3.12 Effect of multi-class and multi-criteria considerations on route choice probabilities
3.4 Conclusions

In this chapter, we presented a two-stage bicycle traffic assignment model with consideration of cyclists’ route choice behavior. In stage one, we considered two key criteria (e.g., route distance, route BLOS) to generate a set of non-dominated (or efficient) paths using a bi-objective shortest path procedure. In stage two, we adopted five traffic assignment methods [equal share assignment (ESA), travel distance per benefit of BLOS assignment (TBA), reference point assignment (RPA), dominated area assignment (DAA), and path-size logit assignment (PSLA)] for flow allocations to the set of efficient paths identified in stage one. In addition, we also provided a multi-class bicycle traffic assignment model as extension of single class model. From the first case study, we found that the ESA and DAA methods, albeit simple, are not suitable for modeling cyclists’ route choice behavior since these methods either do not consider the objective values or are sensitive to the maximum values when assigning flows to the efficient routes. The TBA, RPA using Eq. (3.8), and PSLA methods appeared to produce flow patterns that not only account for the objective values but also reflect cyclists’ route choice behavior. From the second case study, we found there are strong correlations in terms of flow allocations among the TBA, RPA using Eq. (3.8), and PSLA methods, but the RPA method seems to assign similar flow pattern as the PSLA method. In multi-class assignment, the route choice probabilities are highly sensitive as selecting number of objectives (e.g. two objectives in the scenario 1 and three objectives in the scenario 2) and the flows patterns between signal class model and multi-class model are significantly different as single class model could not consider multi attributes of each class. Overall, the results of the Winnipeg network
demonstrate the applicability of the two-stage bicycle traffic assignment procedure with the flexibility of using traffic assignment methods.

In this research, we chose the HCM’s bicycle level of service (BLOS) as a surrogate measure for modeling cyclists’ perception of safety (or risk) on different bicycle facility types. There are other possibilities for surrogate measures; for example, it would be helpful to consider other surrogate measures, such as the bicycle compatibility index (Harkey et al., 1998) or the route stress (Mekuria et al., 2012) as a substitute for perception of safety, or additional criteria such as route cognition using the concept of space syntax (Raford et al., 2007) from the field of urban planning and design, and examine its impact on efficient route generation and flow allocations to the bicycle network. In addition, more tests should be conducted with different network topologies with different bicycle facilities and travelers’ characteristics. Note that the current two-stage bicycle traffic assignment model did not consider the effect of congestion (i.e., link travel times are independent of bicycle flows). As the number of cyclists increases, it would be necessary to consider flow-dependent link travel times to capture the effect of congestion in the bicycle traffic assignment procedure.

REFERENCES


CHAPTER 4
MODELING ASYMMETRIC VEHICLE INTERACTIONS, ROUTE OVERLAPPING, AND TRAFFIC RESTRAINTS IN MULTI-CLASS TRAFFIC ASSIGNMENT PROBLEMS

Abstract

This research presents the model development and computational algorithmic design for addressing asymmetric vehicle interactions, route overlapping, and traffic restraints in multi-class traffic assignment problems. The variational inequality model and the customized solution algorithm are developed for the multi-class traffic assignment problem with different modeling considerations. Three numerical experiments are conducted to demonstrate the computational performance of the customized solution algorithm and features of the model formulation. The results reveal that the consideration of asymmetric vehicle interactions, route overlapping, and traffic restraints can have a significant impact on the network equilibrium flow allocations, and that the customized solution algorithm is a practical approach for solving the multi-class traffic assignment problem with realistic modeling considerations.

4.1 Introduction

Traffic assignment is an essential and fundamental step in the transportation planning and management process. Given constant travel demands between each origin-destination (O-D) pair (i.e., travelers), and travel cost functions for each link of the network (i.e., transportation network), the traffic assignment problem is to determine the traffic flow pattern as well as network performance measures (e.g., total system travel time, vehicle
miles of travel, vehicle hours of travel, fuel consumption and emission, etc.). In practice, most traffic assignment models are single user class and make a number of modeling assumptions including: separability assumption on the link travel time function (i.e., no interactions), deterministic user equilibrium (DUE) or stochastic user equilibrium (SUE) without accounting for route overlapping, and no side constraints to describe the limited supply of certain scarce resources (e.g., link capacities) in a network which are shared by multiple vehicle types (e.g., passenger cars and multiple truck types) or to limit certain classes of vehicles (e.g., trucks) on underpasses due to height restriction, bridges due to weight restriction, and prohibited lanes due to lane restriction. However, as truck traffic continues to grow as a result of increasing freight shipments transported by trucks, there is an increasing interest to model multiple vehicle classes separately, especially in addressing the impacts of truck traffic on congestion, infrastructure deterioration, safety, and environmental concerns in many urban cities. According to the Bureau of Transportation Statistics (BTS), freight shipments transported by trucks account for 71 percent by value in U.S. dollars and 76 percent by weight in tons of all commodity shipments (BTS, 2014). Therefore, the purpose of this research is to develop the model formulation and the customized solution algorithm for addressing asymmetric vehicle interactions, route overlapping, and traffic restraints in multi-class traffic assignment problems involving multiple types of trucks.

4.1.1 Relevant studies

This section reviews some relevant studies on multi-class traffic assignment problem with emphasis on modeling the asymmetric interactions among multiple vehicle types. The multi-class traffic assignment problem is an important extension of the classical
traffic equilibrium problem (i.e., the Beckmann-type mathematical programming formulation) that aims to improve the realism of traffic assignment models by explicitly modeling multiple user classes in the same transportation network with individual cost functions contributing to its own and possibly other classes cost functions in an individual way (Dafermos, 1972). Applications of the multi-class traffic assignment problem include two types: (a) Modeling market segmentation of travelers’ characteristics such as users with different valuations of travel times (Nagurney, 2000; Nagurney and Dong, 2002; Yang and Huang, 2004; Huang and Li, 2007), different risk-taking behaviors (Lo et al., 2006; Shao et al., 2006; Di et al., 2008; Xu et al., 2014), different knowledge levels of network travel times (Yang, 1998; Huang and Li, 2007), and (b) Modeling different vehicle types (e.g., trucks and passenger cars) sharing the same highway network (Daganzo, 1983; Toint and Wynter, 1996; Mahmassani and Mouskos, 1988; Wu et al., 2006; Noriega and Florian, 2007).

Particularly, modeling the asymmetric interactions among different vehicle types has the potential to improve the realism of traffic assignment models (i.e., better capture the travel times required by different vehicle types to traverse a link). However, inadequate considerations of vehicle interactions in the traffic assignment models could result in inaccurate travel time estimates, which could affect flow allocations to routes and to the network equilibrium process. Interest in modeling multiple vehicle types separately in the traffic assignment problem is growing, especially in urban areas where truck flows are of increasing concern. There are also many practical applications such as evaluation of truck-related highway improvements and traffic operations (e.g., capacity addition in the form of additional lanes, lane access restrictions to certain vehicle classes, traffic controls for trucks,
etc.), assessment of truck emission characteristics, pavement deteriorations by trucks, and evaluation of accident patterns by trucks. In addition, it is important to recognize that there are several different sizes of trucks whose maneuver and operational characteristics as well as traffic restrictions are quite different. These characteristics emphasize the need to develop a multi-class traffic assignment model that considers the above modeling issues.

4.1.2 Objective and contribution

Motivated by the needs to enhance the realism of traffic assignment models for multiple user classes, this chapter develops model formulation and customized solution algorithm that consider asymmetric interactions among different vehicle types, path-size logit model for accounting random perceptions of network conditions with explicit consideration of route overlapping, and various traffic restrictions imposed either individually or together to multiple vehicle types in a transportation network. Specifically, variational inequality model is used to formulate the multi-class traffic assignment problem with different modeling considerations. The customized solution algorithm is developed using the iterative balancing scheme, the self-regulated averaging line search method, and column generation with consideration of vehicle restrictions and physical capacity constraints. The contributions of the paper to the literature are: (i) development of a multi-class traffic assignment model for addressing asymmetric vehicle interactions, random perceptions with route overlapping, and various traffic restraints, and (ii) development of a customized solution algorithm for solving for the multi-class traffic assignment problem with different modeling considerations.

The remainder of this chapter is organized as follows. The modeling issues considered in this research including vehicle interactions, route overlapping, and traffic
restraints are introduced in Section 4.2. Section 4.3 provides the multi-class traffic assignment formulation and some qualitative properties. Section 4.4 presents a customized solution algorithm specifically designed for solving the proposed multi-class traffic assignment problem with different modeling considerations. Section 4.5 presents three numerical experiments to demonstrate the proposed solution algorithm and model formulation. Finally, some concluding remarks are given in Section 4.6.

4.2 Modeling considerations

This section discusses the modeling issues in the multi-class traffic assignment problem, which includes vehicle interactions considered in the travel time functions, route overlapping in the stochastic user equilibrium framework, and various traffic restraints imposed either individually or together to multiple vehicle types in a transportation network.

4.2.1 Vehicle interactions

Vehicle interactions occur when several vehicle types are sharing the same roadway space by interacting with each other on the same transportation network. These interactions typically occur between vehicles with very different maneuver and operational characteristics such as between trucks and passenger cars. These interactions could be further detailed as heavy trucks (3+ axle, 6+ tire, combination unit commercial vehicles), medium trucks (2+ axle, 6+ tire, single unit commercial vehicles), light trucks (2-axle, 4-tire commercial vehicles), and passenger cars (FHWA, 2011).

In the past, the separability assumption on the link travel time function is generally required to develop a mathematical programming formulation for the traffic assignment
problems (Sheffi, 1985; Patriksson, 1994). However, this mathematical convenience may lead to modeling inaccuracies due to the over simplification of not considering vehicle interactions when modeling multi-class traffic assignment problems with multiple vehicle types. Since interactions between vehicles are designed to model flow reactions among various vehicle types sharing the same physical infrastructure, most link travel time functions limit the interaction effects to one link (Mahmassani and Mouskos, 1988; Noriega and Florian, 2007). An example for two vehicle classes (e.g., passenger car and truck) is shown in Fig. 4.1, where each vehicle class has its own travel time and at the same time it contributes to other vehicle class’s travel time in an individual way.

Fig. 4.1(a) provides a typical link travel time function for modeling passenger car and truck interactions with passenger car equivalence (PCE) values and free-flow travel time weights for the two vehicle classes. Fig. 4.1(b) shows the overall travel time functional surfaces for the two modes. Due to the characteristics of trucks, the truck travel time surface always lies above of the travel time surface for passenger cars. This means that for any combination of passenger car and truck flows on a link, the travel time for trucks is always larger than that for passenger cars (i.e., trucks take longer time to traverse the link). To better illustrate the differences of the two travel time surfaces, Fig. 4.1(c) and Fig. 4.1(d) show travel time variations as a function of increasing passenger car and truck flows, respectively. For example, the truck free flow travel time at zero passenger car flow in Fig. 4.1(c) is about 15% higher than the free flow travel time for passenger car (i.e., 11.5 and 10 minutes). As the passenger car flow increase, the travel time variation for both vehicle classes also increases (i.e., at 100 passenger cars, travel time ranges from 11.5 to 15.8 minutes for passenger car and from 13.2 to 18.1 minutes for truck). On the other hand, the
travel time variation as a function of truck flows in Fig. 4.1(d) has a larger range for both passenger cars and trucks. As the truck flow increases to 20, the travel time variation ranges from 10.0 to 15.8 minutes for passenger cars and from 11.5 to 18.1 minutes for trucks, indicating that a small amount of truck traffic can have a significant impact on travel time variation for both vehicle types. Hence, the travel time function can model not only the individual travel times required to traverse a link but also the asymmetric interactions between the two vehicle modes.

\[
T_a^n = 10 \cdot P^n \cdot \left(1 + 0.15 \cdot \left(\frac{x_a^c + PCE^c \cdot x_a^t}{100}\right)^4\right)
\]

<table>
<thead>
<tr>
<th></th>
<th>(P^n)</th>
<th>(PCE^n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger car</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Truck</td>
<td>1.15</td>
<td>2.00</td>
</tr>
</tbody>
</table>

(a) Link travel time function and parameters

(b) Travel time surfaces

(c) Travel time variation as a function of passenger car flows

(d) Travel time variation as a function of truck flows

Fig. 4.1 Graphical illustration of the travel time function with vehicle interactions
For the general case with multiple vehicle types, the following non-separable link travel time function in Eq. (4.1) is adopted for modeling the asymmetric interactions among different modes.

$$t_a^m(\cdot) = p^m t_a^0 \left[ 1 + \alpha_m \left( \sum_{m \in M} PCE^m \cdot x_a^m \right) \beta_a \right]$$

(4.1)

where $t_a^m(\cdot)$ is the travel time on link $a$ of class $m$; $p^m$ is the travel time weight parameter of class $m$; $t_a^0$ is the free-flow travel time of link $a$; $PCE^m$ is the passenger car equivalent factor of class $m$; $x_a^m$ is the flow on link $a$ of class $m$; $x_a = \sum_{m \in M} PCE^m \cdot x_a^m$ is the total flow on link $a$; $C_a$ is the capacity of link $a$; and $\alpha_m$ and $\beta_a$ are the parameters of travel time function on link $a$.

### 4.2.2 Route overlapping

In the stochastic user equilibrium (SUE) problem, route overlapping is one of the major concerns in the route choice models (Prashker and Bekhor, 2004; Chen et al., 2012b). In the literature, there are two modeling approaches for the SUE problem: open-form and closed-form probability expressions depending on the probability distribution used to model the random error term in the random utility function. Open-form probability expression does not have an analytical solution (i.e., cannot solve the intractable multidimensional integrals). Examples include the multinomial probit (MNP) SUE model suggested by Daganzo and Sheffi (1977) using the normal distribution, and the logit kernel SUE model suggested by Bekhor et al. (2002). Thus, the open-form SUE models must be solved by the approximation method (Maher, 1992), the numerical method (Rosa and
Maher, 2002), or the Monte Carlo simulation (Sheffi and Powell, 1982; Meng and Liu, 2012). These methods typically require significant computational efforts.

For the closed-form models, two approaches have been developed in the literature to alleviate the route overlapping problem: (a) modification of the deterministic (or systematic) term of the utility function to account for the overlapping problem while still retaining the single-level tree structure of the route choice model (e.g., C-logit by Cascetta et al. (1996) and Zhou et al. (2012), path-size logit by Ben-Akiva and Bierlaire (1999) and Chen et al. (2012b), and path-size weibit by Kitthamkesorn and Chen (2013), and (b) modification of the random error term of the utility function to capture the similarity effect among routes by allowing a more flexible error structure through a two-level tree structure (e.g., cross-nested logit (CNL) by Bekhor and Prashker (1999), paired combinatorial logit (PCL) by Bekhor and Prashker (1999), Pravinvongvuth and Chen (2005), and Chen et al. (2014), and Ryu et al. (2014), and generalized nested logit (GNL) by Bekhor and Prashker (2001).

In this research, the path-size (PS) factor is adopted to handle the route overlapping problem due to its simplicity and relatively well performance compared to other closed-form models discussed above. The PS factor accounting for different path sizes is determined by the length of links within a path and the relative lengths of paths that share a link as follows.

\[
PS_{km}^{rs} = \sum_{a \in k} \left( \frac{1}{L_{rs}^{km}} \right) \left( \frac{1}{\sum_{l \in K_{rs}^m} \delta_{lm}} \right), \quad \forall \ m \in M, \ k \in K_{rs}^m, \ rs \in RS
\] (4.2)
where $PS_{km}^{rs}$ is the PS factor of path $k$ between origin $r$ and destination $s$ in class $m$; $l_a$ is the length of link $a$; $L_{km}^{rs}$ is the length on path $k$ between origin $r$ and destination $s$ in class $m$; and $\delta_{kam}^{rs}$ is the path-link indicator, 1 if link $a$ is on path $k$ for class $m$ between O-D pair $rs$ and 0 otherwise. Paths with a heavy overlapping with other routes have a smaller PS value, while a larger PS value indicates the paths are more distinct. For other functional forms of the PS factor, see Bovy et al. (2008) and Prato (2009).

With the derived PS value in Eq. (4.2), the PS-logit (PSL) probability for the multi-class SUE problem can be expressed as follows.

$$
P_{km}^{rs} = \frac{PS_{km}^{rs} \cdot e^{-\theta_m c_{km}^{rs}}}{\sum_{l \in K_{rs}^m} PS_{lm}^{rs} \cdot e^{-\theta_m c_{lm}^{rs}}}, \quad \forall \ m \in M, \ k \in K_{rs}^m, rs \in RS
$$

(4.3)

where $P_{km}^{rs}$ is the probability of selecting path $k$ between origin $r$ and destination $s$ in class $m$; $\theta_m$ is the dispersion parameter for class $m$; and $c_{km}^{rs} = \sum_{a \in A} t_a^m(\cdot) \delta_{kam}^{rs}$ is the cost of path $k$ between origin $r$ and destination $s$ in class $m$.

While the PSL model is relatively simple, it has been shown to perform well relative to more complex model forms such as the CNL model (Bekhor et al., 2006; Prato and Bekhor, 2006; Prato and Bekhor, 2007). Fig. 4.2 provides an illustration of how the PSL model resolves the route overlapping problem using the loop-hole network. The route choice probability and PS-value with different degrees of route overlapping are shown in Fig. 4.2(b) and Fig. 4.2(c), respectively.
When there are no route overlaps ($x=0$), the PSL model gives the same choice probability as the MNL model (i.e., independence assumption is fully satisfied). However, when there are route overlaps, the choice probability of the two overlapping routes becomes smaller with an increasing $x$ value in the PSL model (i.e., the green line showing the probabilities of PSL for R2 and R3), which is more reasonable comparing to the results of the MNL model (i.e., the red line showing the probabilities of MNL for all three routes).

4.2.3 Traffic restraints

Traffic restraints described by various types of side constraints has been suggested as a promising approach to improve the realism of traffic assignment models (Larsson and Patriksson, 1999; Chen et al., 2011). In the literature, side constraints have been considered to model queuing and congestion effects (Bell, 1995; Yang and Lam, 1996; Larsson et al., 2004; Chen et al., 2011), link toll pricing (Ferrari, 1995; Yang and Bell, 1997), restraining traffic flows (Ferrari, 1995; Yang and Bell, 1997; Chen et al., 2011; Meng et al., 2014), replicating observed flows (Bell et al., 1997; Chen et al., 2005, 2009, 2010, 2012a;
Chootinan et al., 2005), and restricting vehicle emissions (Ferrari, 1995; Chen et al., 2011; Xu et al., 2015).

All of the above side constraints are applied to the aggregate flow on a link (or a node) without differentiating vehicle types, and can be written as follows.

\[ g_a(x) \leq 0, \quad \forall \ a \in \bar{A}, \tag{4.4} \]

where \( x \) is a link flow vector; \( \bar{A} \) is a subset of links (or nodes) in the network, which involves traffic restraint; and \( g_a(x) \) is a function of link flows and a predetermined threshold value. An example of \( g_a(x) \) is the physical link capacity constraint that restricts the total flow of all vehicle classes to be less than or equal to the capacity on link \( a \).

\[ x_a \leq C_a, \quad \forall \ a \in \bar{A}, \tag{4.5} \]

For other physical and environmental constraints, see Chen et al. (2011) and Xu et al. (2015). However, trucks are often perceived to restrict the flows or safety on certain road segments. The most common reasons for truck flow restriction are improving highway operations and safety hazard including crashes, pavement and structural considerations, and work zone construction restrictions. In general, there are three types of restriction: lane restriction, height restriction, and weight restriction (see Fig. 4.3). For lane restriction on highways, truck flows are typically restricted to the right or outside lane (e.g., curb lane) to improve the efficiency of passenger car travel. For height restriction, trucks are restricted on tunnels or underpasses if the vehicles exceed certain height requirements. For weight restriction, heavy trucks exceeding certain weight limit are typically banned on bridges or elevated roadways. Both height and weight restrictions used to improve safety hazards can be posed as constraints as follows:
\[ g_{am}(x^m) = x^m_a = 0, \quad \forall a \in \hat{A}_m, m \in M, \]  

(4.6)

where \( x^m \) is a link flow vector of class \( m \); \( \hat{A}_m \) is a subset of links (or nodes) of class \( m \) in the network, which involves imposing vehicle restriction constraint of class \( m \) on a subset of links \( \hat{A}_m \). This study considers not only the physical capacity constraints of total flow on a link but also flow restrictions for different types of vehicles, particularly for trucks, to enhance the realism of multi-class traffic assignment models.

![Fig. 4.3 Three types of vehicle restrictions](image)

(a) Lane restriction  
(b) Height restriction  
(c) Weight restriction

4.3 Multi-class traffic assignment formulation

In the section, we provide the model formulation for the multi-class traffic assignment problem with considerations of the asymmetric vehicle interactions, route overlapping, and various traffic restraints discussed in Section 4.2. Some usual assumptions are made: (a) The multi-class O-D demands are given and fixed, and stationary traffic conditions exist in a transportation network; (b) all user classes have imperfect knowledge of network travel times, and make their route choice decisions in a stochastic user equilibrium (SUE) manner; (c) side constraints are active only on a subset of links in the network, but the network, as a whole, has sufficient capacity to accommodate
the rerouting traffic of each O-D pair. Note that assumption (a) can be easily relaxed by considering the multi-class traffic assignment model to have mode choice with fixed total demand or elastic demand for each user class. Assumption (b) can be further extended by considering network uncertainty (e.g., the stochastic reliability-based user equilibrium models by Shao et al. (2006); or the stochastic mean-excess traffic equilibrium model by Chen et al., 2011); user classes in these models are assumed to make probabilistic route choice decisions under both random perception and network uncertainty; the last assumption (c) ensures that a solution of the multi-class traffic assignment problem exists (i.e., the network has alternative paths and sufficient capacity to reroute traffic of each O-D pair).

Let \( \eta_{km}^{rs} \) denote the equivalent cost of the PSL model on path \( k \) between origin \( r \) and destination \( s \) in class \( m \) as follows.

\[
\eta_{km}^{rs} = c_{km}^{rs} + \frac{1}{\theta_{km}} \ln \left( \frac{f_{km}^{rs}}{PS_{km}^{rs}} \right), \quad \forall \ m \in M, \ k \in K_{rs}^{m}, \ rs \in RS
\] (4.7)

Denote \( \eta \) and \( f \) as the vectors of equivalent path costs \( \ldots, \eta_{km}^{rs}, \ldots \)\( ^T \) and path flows \( \ldots, f_{km}^{rs}, \ldots \)\( ^T \), respectively. Then, the multi-class traffic assignment problem with different modeling considerations can be formulated as a variational inequality (VI) problem as follows.

Find a path flow solution \( f^{*} \in \Omega \), such that

\[
\eta(f^{*})^T(f - f^{*}) \geq 0, \ \forall f \in \Omega
\] (4.8)

where \( \Omega \) represents the feasible set defined by Eqs. (4.9)-(4.13).
\[ \sum_{k \in K^m_{rs}} f^m_{km} = q^m_{rs}, \quad \forall m \in M, \quad rs \in RS \tag{4.9} \]

\[ x^m_a = \sum_{rs \in RS} \sum_{k \in K^m_{rs}} f^m_{km} \delta^rs_{ka}, \quad \forall a \in A, \quad m \in M \tag{4.10} \]

\[ x_a = \sum_{m \in M} PCE^m \cdot x^m_a, \quad \forall a \in A \tag{4.11} \]

\[ f^m_{km} \geq 0, \quad \forall m \in M, \quad k \in K^m_{rs}, \quad rs \in RS \tag{4.12} \]

\[ g_l(\cdot) \leq 0, \quad \forall l \in L \tag{4.13} \]

Eq. (4.9) is the travel demand conservation constraint for class \( m \) between O-D pair \( rs \); Eq. (4.10) is a definitional constraint that sums up all path flows of class \( m \) that pass through a given link \( a \); Eq. (4.11) is the total flow on link \( a \) using the PCE factors to convert the different vehicle classes to equivalent passenger car units; Eq. (4.12) is a non-negativity constraint on the path flows; and Eq. (4.13) is the generalized side constraint describing the traffic restraints individually or together to the multiple vehicle types in a transportation network, where the set \( L \) is a subset of links and/or links of class \( m \). Eq. (4.13) can be an individual vehicle restriction constraint on link \( a \) of class \( m \), i.e., Eq. (4.6), or a link capacity constraint restricting the total flow of all vehicle classes to be less than or equal to the capacity on link \( a \), i.e., Eq. (4.5).

It should be mentioned that the above VI model is different from the traditional VI formulation for the asymmetric traffic assignment problem (Smith, 1979; Dafermos, 1980) and the mathematical programming (MP) formulation for the standard traffic assignment model (Beckmann et al., 1956; Sheffi, 1985) by explicitly considering multiple user classes with different random perceptions, route overlapping, asymmetric vehicle interactions, and
side constraints on physical capacity of total flow on a link and flow restrictions for
different types of vehicles, particularly for trucks.

Assume that the feasible set $\Omega$ is non-empty, $g_k(\cdot)$ is a convex and continuously differentiable function, which satisfies an appropriate constraint qualification (Facchinei and Pang, 2003). Let $\lambda_l, \forall l \in L$, denote the Lagrangian multiplier associated with the
generalized side constraint (4.13), and $\theta_m, \forall m \in M, rs \in RS$, denote the Lagrangian multiplier between O-D pair $rs$ of class $m$ associated with the demand conservation constraint in Eq. (4.9), the generalized path cost function $\eta_{km}^{rs}$ can be specified as

$$
\eta_{km}^{rs} = \eta_{km}^{rs} + \sum_{l \in L} \delta_{km}^{rs} \left( \lambda_l \frac{\partial g_l}{\partial x_l} \right), \forall m \in M, k \in K_{rs}^m, rs \in RS
$$

(4.14)

We have the following proposition.

**Proposition 1.** Assume the equivalent path cost function $\eta(f)$ is positive, then the solution of the VI problem (4.8)-(4.13) satisfies the equilibrium conditions.

**Proof.** Since a constraint qualification is assumed to hold for the VI problem (4.8)-(4.13), from the Karush-Kuhn-Tucker (KKT) conditions of variational inequality (Facchinei and Pang, 2003), we have

$$
J_m^{rs} (\eta_{km} (f^*) - \theta_m^{rs} ) = 0, \quad \forall m \in M, k \in K_{rs}^m, rs \in RS
$$

(4.15)

$$
\eta_{km}^{rs} (f^*) - \theta_m^{rs} \geq 0, \quad \forall m \in M, k \in K_{rs}^m, rs \in RS
$$

(4.16)

$$
\lambda_l \geq 0, \quad g_l(\cdot) \leq 0, \quad \text{and} \quad \lambda_l \cdot g_l(\cdot) = 0, \quad \forall l \in L
$$

(4.17)

and Eqs. (4.9) and (4.12).
It is easy to see that the generalized equilibrium conditions are satisfied.

\[ \tilde{\eta}^r_{km}(t^* ) - o^r_m = \begin{cases} 0 & \text{if } (f^*_{km})^s > 0 \\ \geq 0 & \text{if } (f^*_{km})^s = 0 \end{cases}, \quad \forall m \in M, \ k \in K^m_{rs}, \ rs \in RS \quad (4.18) \]

This completes the proof. \( \square \)

The Lagrangian multiplier values, \( \lambda_i, \ \forall l \in L \), can be regarded as the ‘shadow prices’ corresponding to the side constraints. It should be noted that the Lagrangian multipliers are generally not unique despite that the aggregate link-flow pattern is unique. Moreover, the complementary conditions (4.17) of traffic flow restraints imply that if the side constraint is strictly satisfied (i.e., \( g_i(\cdot) < 0 \)), the corresponding Lagrangian multiplier should be equal to zero. On the other hand, if the side constraint is binding (i.e., \( g_i(\cdot) = 0 \)), then the corresponding Lagrangian multiplier should be greater than or equal to zero. For example, if the generalized side constraints are specified in terms of the physical link capacity constraints of link flows in Eq. (4.5), where the second term of the generalized path cost function (4.14) can be interpreted as the link queuing delay (Bell, 1995). Queue will occur when flows reach its capacity corresponding to the active capacity constraint (i.e., \( x_a = C_a \)) and will be zero when flows are strictly below its capacity (i.e. \( x_a < C_a \)).

\[ d_a = \sum_{l \in L} \delta^r_{km} \left( \lambda_i \frac{\hat{g}_i}{\hat{\lambda}_a} \right), \quad \forall a \in \hat{A} \quad (4.19) \]

Similarly, for the individual vehicle restriction constraints on link \( a \) of class \( m \) in Eq. (4.6), where the second term of the generalized path cost function (4.14) can be interpreted a high penalty cost for restricting vehicle class \( m \) from using link \( a \).

\[ u_a^m = \sum_{l \in L} \delta^r_{km} \left( \lambda_i \frac{\hat{g}_i}{\hat{\lambda}_a^m} \right), \quad \forall a \in \hat{A}_m, \ m \in M \quad (4.20) \]
Thanks to the closed-form PSL probability expression in Eq. (4.3), path flows can be derived analytically as a function of path cost and dual variables as follows.

\[ f_{mk}^{rs} = P S_{km}^{rs} \cdot \exp \left( -\theta^{m} \left( c_{km}^{rs} + \sum_{a \in A} d_{a}^{rs} \delta_{kam}^{rs} + \sum_{a \in A_{m}} u_{a}^{m} \delta_{kam}^{rs} - o_{rs}^{m} \right) \right), \quad \forall m \in M, \ k \in K_{rs}^{m}, \ rs \in RS \]  

(4.21)

4.4 Solution algorithm

This section presents a customized solution algorithm specifically designed for solving the proposed multi-class traffic assignment problem with consideration of asymmetric interactions, route overlapping, and various vehicle restrictions. This customized solution algorithm has three main steps: direction finding, line search, and column generation. These three modules of the solution algorithm work together to handle the multi-class traffic assignment, the asymmetric vehicle interactions through the link performance functions, the PSL model to account for route overlapping, and both the aggregate and the individual flow restriction constraints.

4.4.1 Direction finding

Unlike the traditional multi-class user equilibrium (UE) assignment model, which only considers the asymmetric interactions between modes, the proposed model is a multi-class path-size logit (PSL) SUE traffic assignment problem considering asymmetric interactions between modes, route overlapping, and various side constraints. Therefore, the simply all-or-nothing (AON) assignment method typically used in the traditional multi-class UE assignment model without route overlapping consideration and side constraints cannot be used directly to determine the search direction for the proposed model. Thanks to the closed-form PSL probability expression, one can adopt the iterative balancing
scheme to determine the search direction by solving the dual problem (Bell and Iida, 1997).

The basic idea is to scale the primal variables (i.e., $f'_{km}$) to fulfill one constraint at a time by adjusting the dual variables (i.e., $u^m_a$, $d^m_a$ and $o^m_{rs}$) associated with constraints.

For a given iteration with a fixed path set, the iterative balancing scheme is performed using the adjustment equations until the convergence (e.g., insignificant change of the primal and dual variables) for direction finding. The adjustment of the dual variable is used to steer the violated constraint back to satisfactory. Here we provide the adjustment equations for constraints (4.5), (4.6) and (4.9) on the physical link capacity constraint, vehicle restriction constraint, and O-D demand conservation constraint, respectively. The derivations of the adjustment equations are provided in Appendix A.

Handling vehicle restriction constraints

The adjustment factor ($\gamma^m_a$) for the dual variable ($u^m_a$) of the vehicle restriction constraint is given as follows.

$$\gamma^m_a = -\frac{1}{\theta^m} \ln \left( \frac{0}{x^m_a} \right) = \infty, \quad \forall a \in \hat{A}_m, m \in M$$  \hspace{1cm} (4.22)

Note that the theoretical adjustment factor is infinity in order to enforce the strict vehicle restriction constraint. However, for practical implementation, the adjustment factor is set at a sufficiently large value (e.g., $10^{10}$).

Handling link capacity constraints

The adjustment factor ($\pi^m_a$) for the dual variable ($d^m_a$) of the capacity constraint is given as follows.
\[ \sum_{m \in M} PCE^m \cdot \chi^m \cdot \exp \left( \theta^m \left( -\pi_a \right) \right) = C_a \quad (4.23) \]

To find a suitable \( \pi_a \) (i.e., an aggregate link variable) to adjust the individual vehicle class \( m \) sharing the same roadway space, a numerical method is adopted since there is no closed-form adjustment factor for computing \( \pi_a \) (see Appendix A for details).

**Handling O-D demand conservation constraints**

The adjustment factor \( \psi^m_{rs} \) for the dual variable \( \phi^m_{rs} \) of the O-D demand conservation constraint is given as follows.

\[ \psi^m_{rs} = \frac{1}{\theta^m} \ln \left( \frac{q^m_{rs}}{\sum_{k \in K^m_{rs}} f^m_{rs}} \right) \quad \forall m \in M, \ k \in K^m_{rs}, \ rs \in RS \quad (4.24) \]

**4.4.2 Line search**

To avoid evaluating the complex function required in the traditional line search schemes (e.g., bisection, golden section, and Armijo) and the slow convergence of the method of successive averages (MSA) scheme, the self-regulated averaging (SRA) scheme developed by Liu et al. (2009) is adopted to determine a suitable stepsize for updating the solution vector. The SRA scheme updates the stepsize as follows:

\[ \alpha^{(n)} = \frac{1}{\beta^{(n)}} \quad (4.25) \]

\[ \beta^{(n)} = \begin{cases} 
\beta^{(n-1)} + \lambda_1, & \text{if } \| \tilde{f}^{(n)} - f^{(n-1)} \| \geq \| \tilde{f}^{(n-1)} - f^{(n-2)} \| \\
\beta^{(n-1)} + \lambda_2, & \text{otherwise}
\end{cases} \quad (4.26) \]

where \( \lambda_1 > 0 \) and \( 0 < \lambda_2 < 1 \). \( \tilde{f}^{(n)} \) and \( f^{(n-1)} \) are the vector forms of \( \tilde{f}^{\pi_{rs}(n)}_{km} \) and \( f^{\pi_{rs}(n-1)}_{km} \) at two consecutive iterations \( n-1 \) and \( n \), where \( f^{\pi_{rs}(n)}_{km} \) is the auxiliary flow on path \( k \) between O-D
pair $rs$ of class $m$. Similar to the MSA scheme, the following conditions are also satisfied in order to guarantee convergence (Robbins and Monro, 1951; Blum, 1954; Liu et al., 2009).

$$\alpha^{(n)} > 0, \sum_{n=1}^{\infty} \alpha^{(n)} = \infty, \text{ and } \lim_{n \to \infty} \alpha^{(n)} = 0 \left( \text{or } \sum_{n=1}^{\infty} \left( \alpha^{(n)} \right)^2 < \infty \right)$$ (4.27)

We can see that the stepsize sequence from the SRA scheme is still strictly decreasing. However, the decreasing speed of stepsize is different. It is controlled by either $\lambda_1$ or $\lambda_2$ depending on the residual error between two consecutive iterations. When the current residual error is increased compared to the previous iteration (i.e., tends to diverge), parameter $\lambda_1 > 1$ is used to make the stepsize reduction more aggressive. In contrast, when the residual error is decreased (i.e., tends to converge), parameter $0 < \lambda_2 < 1$ is used to make the stepsize reduction more conservative. Hence, the stepsize sequences from the SRA scheme indeed satisfy the above conditions for convergence. Note that the SRA scheme has been successfully adopted in solving different traffic assignment models (Xu and Chen, 2013; Chen et al., 2014; Kitthamkesorn and Chen, 2013) and combined travel demand models (Yang et al., 2013; Yao et al., 2014) with separable link travel time functions (i.e., no interactions) formulated as a convex programming formulation. To our best knowledge, applying the SRA scheme within a customized path-based algorithm for solving the stochastic multi-class traffic assignment problem with non-separable link travel time functions (i.e., asymmetric interactions) formulated as a VI problem has not been attempted in the literature.
4.4.3 Column generation

The iterative balancing scheme for direction finding in section 4.4.1 assumes that a working path set is given. For large networks, this may not be practical to enumerate a working path set in advance since the number of possible paths grows exponentially with respect to network size. To circumvent path enumeration, a column generation (or shortest path algorithm) is embedded into the solution procedure for a general transportation network. Link costs are updated based on the new primal and dual variables [i.e., \( t_a^m = t_a^m(x) + u_a^m + d_a \)]. The shortest path algorithm is then used to generate paths for each O-D pair and each user class accordingly in order to satisfy the constraints (e.g., link capacity constraint for all vehicles and vehicle restriction constraint based on vehicle class). If the generated paths are new, they are added into the current path set. Note that the dual variables force the shortest path algorithm to generate paths that satisfy the link capacity and vehicle restriction constraints. For example, if the link flows are binding on the capacity in a previous iteration, the link cost is increased with adding positive dual variable (e.g., considering queuing delay for the link capacity constraint). For the vehicle restricted links, the link cost has the infinity cost with adding infinity dual variable \( u_a^m \), hence the path including the vehicle restricted link in a class may not generated.

To reduce the intensive memory requirements of storing paths, a universal path set is designed to store paths for all vehicle classes efficiently without the need to separately store paths for each individual vehicle class (i.e., a path can be shared or used by multiple vehicle classes). A binary (true/false) indicator in each vehicle class is used to determine the individual class path set out of the universal path set. This simple scheme can help to reduce the memory requirements by eliminating the storage of redundant paths for each
vehicle class. Fig. 4.4 provides an illustration of the individual class path set and the universal path set and a comparison of the memory requirements for the two path storage methods.

<table>
<thead>
<tr>
<th>Path</th>
<th># of links</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>True</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>True</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>False</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>False</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>True</td>
</tr>
</tbody>
</table>

Required memory:
\[ 4 \cdot (5 \cdot 7) + 2 \cdot (5 \cdot 3) = 170 \text{bytes} \]

*Memory size: 4 byte for a link
2 byte for an indicator*

For a given O-D pair, the algorithm starts with the first vehicle class by determining the shortest path with the updated link costs obtained from the previous iteration. It then compares the shortest path with paths in the current universal path set to determine whether the shortest path is new or not. If the shortest path is a new path, the current universal path set is updated by augmenting this new path \((k=k+1)\), and the indicator for this path is activated. If not, this means the shortest path is already in the current universal path set. The existing path will be flagged as the shortest path. These processing steps are repeated
until the last vehicle class. The overall procedure for generating a universal path set of a given O-D pair is depicted in Fig. 4.5.

Fig. 4.5 Flowchart of generating a universal path set for a given O-D pair

4.4.4 Overall solution procedure

The overall solution procedure for solving the proposed multi-class traffic assignment problem with considerations of asymmetric vehicle interactions, route overlapping, and various traffic restraints is shown in Fig. 4.6 and detailed algorithmic steps are provided as follows.
Step 0. Input.

- Network data: network connectivity, link capacity \( (C_a) \), free-flow link travel time \( (t_a^m)^0 \), capacity constrained links \( (a \in \bar{A}) \), and restricted links \( (a \in \hat{A}_m) \).
- O-D demands: Multi-class O-D demands \( (q_{rs}^m, \forall m \in M, \; rs \in RS) \)
- Parameters: PCE factors \( (PCE_a^m) \); travel time weight factors \( (p_a^m) \), and dispersion parameters \( (\theta^m) \).
Step 1: Initialization.
- Set iteration counter: \( n = 0 \)
- Initialize primal variables: \( (f_{km}^r)^n, (x_{a}^m)^n \) and \( (x_a)^n = 0 \)
- Initialize dual variables: \( (d_a)^n \) and \( (\sigma_{rs}^m)^n = 0 ; u_a^m = \infty \)
- Initialize path set: \( (K_{rs}^m)^n = \emptyset \)

Step 2: Update the generalized link costs with PCE factors, weight factors, and dual variables.
- \( n = n + 1 \)
- Generalized link cost: \( (t_a^m)^n = p_a^m \cdot t_a^m \left( \sum_{m \in M} PCE_a^m \cdot x_a^m \right)^{n-1} + (u_a^m)^{n-1} + (d_a)^{n-1} \)

Step 3: Column generation.
- Determine the shortest path: \( (k_{rs}^m)^n \)
- Update universal path set: \( (K_{rs}^m)^n = (K_{rs}^m)^{n-1} \cup (k_{rs}^m)^n \)
- Update PS values using Eq. (4.1): \( (PS_{km}^r)^n \)

Step 4: Direction finding using iterative balancing scheme.

4.1 Initialization. Set \( \tilde{n} = 0 \), \( (d_a)^{\tilde{n}} \) and \( (\sigma_{rs}^m)^{\tilde{n}} = 0 \), and \( (u_a^m)^{\tilde{n}} = \infty \).
- Compute initial primal variables (path flows, class link flows and total link flows)
  \[
  \left( f_{km}^r \right)^{\tilde{n}} = PS_{km}^r \cdot \exp \left( -\theta^m \cdot \left( c_{km}^r + \sum_{a \in \Lambda} (d_a)^{\tilde{n}} \cdot \delta_{kam} + \sum_{a \in \Lambda_m} (u_a^m)^{\tilde{n}} \cdot \delta_{kam} \cdot (\sigma_{rs}^m)^{\tilde{n}} \right) \right) 
  \]
  \[
  \left( x_{a}^m \right)^{\tilde{n}} = \sum_{r \in RS} \sum_{k \in K_a^m} \left( f_{km}^r \right)^{\tilde{n}} \cdot \delta_{kam} 
  \]
  \[
  \left( x_a \right)^{\tilde{n}} = \sum_{m \in M} PCE_a^m \cdot \left( x_a^m \right)^{\tilde{n}} 
  \]
4.2 Determine adjustment factors.
- \( \pi_a \) and \( \psi_{rs}^m \) with Eq. (4.23) and (4.24), and \( \gamma_a^m = \infty \)

4.3 Update dual variables.

\[
\begin{align*}
(o_{rs}^m)^{\hat{t}+1} &= (o_{rs}^m)^{\hat{t}} + \psi_{rs}^m \\
(d_a)^{\hat{t}+1} &= \text{Max} \left\{ 0, (d_a)^{\hat{t}} + \pi_a \right\} \\
(u_a)^{\hat{t}+1} &= (u_a)^{\hat{t}} + \gamma_a^m = \infty
\end{align*}
\]

4.4 Update primal variables (path flows, class link flows, and total link flows)

\[
\begin{align*}
(f_{km}^c)^{\hat{t}+1} &= PS_{km}^c \exp \left\{ -\theta_m \cdot \left( c_{km}^c + \sum_{a \in A} (d_a)^{\hat{t}+1} \delta_{kam} + \sum_{a \in A_a} (u_a)^{\hat{t}+1} \delta_{kam} - (o_{rs}^m)^{\hat{t}+1} \right) \right\} \\
(x_a^m)^{\hat{t}+1} &= \sum_{rs \in RS} \sum_{k \in K_{rs}^m} (f_{km}^c)^{\hat{t}+1} \delta_{kam} \\
(x_a^m)^{\hat{t}+1} &= \sum_{m \in M} PCE_a^m \cdot (x_a^m)^{\hat{t}+1}
\end{align*}
\]

4.5 Convergence test of iterative balancing scheme.

Compute \( \bar{\xi} = \max \left\{ \| (o_{rs}^m)^{\hat{t}+1} - (o_{rs}^m)^{\hat{t}} \|, \| (d_a)^{\hat{t}+1} - (d_a)^{\hat{t}} \|, \| (u_a)^{\hat{t}+1} - (u_a)^{\hat{t}} \| \right\} \)

If \( \eta_0 \leq \bar{\xi} \leq \eta \), set \( \tilde{n} = \tilde{n} + 1 \) and go to Step 4.2. If \( \bar{\xi} \leq \eta_0 \), terminate. Otherwise stop.

Note that \( \eta \) is the upper limit of the adjustment allowed (e.g., \( 10^6 \) for detecting divergence)

**Step 5:** Line Search using SRA method.
- Compute the norm between previous consecutive flows:
- \( \| \hat{f}(n) - f^{(n-1)} \| \) and \( \| \hat{f}(n-1) - f^{(n-2)} \| \)
- Compute \( \beta^{(n)} : \beta^{(n)} = \begin{cases} 
\beta^{(n-1)} + \lambda_1, & \text{if } \| \hat{f}(n) - f^{(n-1)} \| \geq \| \hat{f}(n-1) - f^{(n-2)} \| \\
\beta^{(n-1)} + \lambda_2, & \text{otherwise}
\end{cases} \)
- Compute step size \( \alpha^{(n)} : \alpha^{(n)} = 1/\beta^{(n)} \)
Step 6: Update solution vector.

- Update path flows: \( f_{km}^{rs} = (f_{km}^{rs})^{n-1} + \alpha^n (f_{km}^{rs})^n - (f_{km}^{rs})^{n-1} \)
- Update class link flows: \( x_a^m = \sum_{r \in R \cdot s \in K_{rs}^m} (f_{km}^{rs})^n \delta_{km}^{rs} \)
- Update total link flows: \( x_a = \sum_{m \in M} PCE_a^m \cdot (x_a^m)^n \)

Step 7: Convergence test.

- If Relative Gap (RG) > \( \epsilon \), set \( n = n + 1 \) and go to Step 2. Otherwise stop and go to step 8.

Step 8: Output.

- Class path flows and travel times: \( f_{km}^{rs} \) and \( c_{km}^{rs} \)
- Class link flows and travel times: \( x_a^m \) and \( t_a^m \)
- Total link flows: \( x_a = \sum_{m \in M} PCE_a^m \cdot x_a^m \)

4.5 Numerical results

In this section, we present three numerical experiments to examine the customized solution algorithm and model formulation developed for the multi-class traffic assignment problem with different modeling considerations. Experiment 1 is designed to examine the effect of line search methods and vehicle restriction constraints on the customized solution algorithm. Experiment 2 focuses on the assignment results of the multi-class traffic assignment model formulation with considerations of vehicle restrictions, vehicle interactions, and route overlapping. Experiment 3 applies the customized solution algorithm and model formulation to explore the impacts of imposing physical capacity constraints on the bridges in a transportation network.
4.5.1 Description of the experiments

In all three experiments, the Winnipeg network shown in Fig. 4.7 is used to examine the solution algorithm and its application of the multi-class traffic assignment problem with different modeling considerations in a real network setting. The network consists of 154 zones, 1,067 nodes, 2,535 links, and 4,345 O-D pairs. The network structure, O-D trip tables, and link performance parameters are from the Emme software (INRO Consultants, 2013). Three classes of vehicle types are considered: passenger cars, medium trucks, and heavy trucks. The total travel demand is 59,333 trips, and the deterministic UE assignment results show that this is a moderately congested network (i.e., the average volume to capacity (V/C) ratio is about 0.7). The aggregated truck demands from the Emme software are assumed to disaggregate by 70% and 30% to obtain the O-D trips for medium trucks and heavy trucks. The final demands and percentages for three vehicle classes are shown in Table 4.1.

The link restrictions by vehicle class are obtained from the City of Winnipeg (2013). The BPR function in Fig. 4.1(a) with link-specific parameters, PCE factors, and travel time weight parameters are obtained from Noriega and Florian (2007) and shown in Table 4.1. The values of the dispersion parameter for each class are also shown in Table 4.1. In this study, we assume the dispersion parameter values of 0.5, 0.75, and 1.0 for passenger cars, medium trucks, and heavy trucks, respectively. These values indicate that approximately 8%, 2%, and 1% of the passenger car, medium truck, and heavy truck drivers will choose the route with a higher cost given a 5-minute difference between two routes in the multinomial logit (MNL) model.
For the SRA line search method, the parameters are set as: $\lambda_1 = 1.90$ and $\lambda_2 = 0.01$.

The customized solution algorithm specifically designed for solving the proposed multi-class traffic assignment problem with various considerations is coded in Intel Visual FORTRAN XE2013 and run on a 3.40GHz processor with 16.00GB of RAM. In three
experiments, the stopping criterion is based on the relative gap of the equivalent
generalized path cost given in Eq. (4.14) as follows.

\[
RG = \frac{1}{|M|} \sum_{m \in M} \sum_{rs \in R, k \in K_{rs}^m} \left( \tilde{\eta}_{km}^{rs} - \tilde{\eta}_{km}^{rs} \right) \cdot f_{km}^{rs}
\]

(4.28)

where \( |M| \) is the number of classes, \( \tilde{\eta}_{km}^{rs} \) is the generalized path cost on path \( k \) between O-D pair \( rs \) of class \( m \), and \( \tilde{\eta}_{km}^{rs} \) is the minimum generalized cost between O-D pair \( rs \) of class \( m \).

4.5.2 Experiment 1: Convergence characteristics and sensitivity analysis

Experiment 1 is designed to examine the convergence characteristics of the
customized path-based algorithm for solving the stochastic multi-class traffic assignment
problem. The computational performance compares two line search schemes: method of
successive averages (MSA) and SRA. Note that both SRA and MSA line search schemes
do not need function evaluations or derivative evaluations. Convergence results in terms
of relative gap in log scale using the customized path-based algorithm with MSA and SRA
schemes are shown in Fig. 4.8(a). As expected, the MSA scheme exhibits a sublinear rate
of convergence and has difficulty reaching the desired level of accuracy (1E-10), while the
SRA scheme exhibits a linear rate of convergence toward the required accuracy. In fact,
the MSA stepsize scheme only reaches an accuracy of 1E-4 at the maximum amount of
CPU time (100 seconds) allowed. Fig. 4.8(b) also plots the stepsize trajectories of the two
line search schemes. One can see the stepsize sequences of both schemes are decreasing
(i.e., satisfying the conditions in Eq. (4.22) to guarantee convergence). However, the
decreasing rate of SRA scheme is more efficient since the next stepsize is determined
according to the residual error between two consecutive iterations and the two control parameters ($\lambda_1$ and $\lambda_2$) in Eqs. (4.25) and (4.26). As can be seen, the SRA stepsizes are larger than those generated by the MSA scheme. This self-regulating feature helps to improve not only the algorithmic convergence for obtaining highly accurate solution, but also the convergence rate of the customized path-based algorithm (i.e., from a sublinear convergence rate due to the MSA scheme to an approximate linear convergence rate using the SRA scheme).

![Fig. 4.8 Computational comparisons between MSA and SRA schemes](image)

In addition, the sensitivity analysis of $\lambda_1$ and $\lambda_2$ in the SRA stepsize scheme is examined in Fig. 4.9. These two parameters adjust the decreasing speed of stepsize. When higher parameter values are adopted, the stepsize is more aggressively decreased. Parameter $\lambda_1$ is active when the current residual error is increased compared to the previous iteration; otherwise, $\lambda_2$ is active. As can be seen, the effect of $\lambda_1$ is quite marginal, while $\lambda_2$ has a significant impact on the computational effort for both CPU time and number of iterations. This result indicates that the flow difference of two consecutive iterations is
marginal. Thus, our recommendation for the parameters selection, at least for this type of highly congested networks, is to choose a small value of $\lambda_2$ (e.g., less than 0.10) and a large value of $\lambda_1$ in order to enhance the convergence.

![Fig. 4.9 Sensitivity of the two parameters in the SRA line search scheme](image)

The above analysis indicates that the customized path-based algorithm with the SRA line search scheme can promise convergence. In the following analyses, we conduct several sensitivity analyses of model parameters on computational performance using the SRA line search scheme. Specifically, the effects of dispersion parameter, PCE factor, percentages of trucks, and travel time weight parameter on the computational performance are displayed in Fig. 4.10. Fig. 4.10(a) examines various combinations of dispersion parameter for passenger car (PC), medium truck (MT), and heavy truck (HT). Note that the dispersion parameter for the base case is 0.75 for PC, 1.0 for MT, and 1.25 for HT. As the dispersion parameter increases, computational effort also increases due to less dispersed traffic and more interactions on path flows of all vehicle classes concentrated on fewer paths.
Fig. 4.10 Computational efforts under different dispersion parameter values, PCE factors, percentages of trucks, and travel time weight parameters

Fig. 4.10(b) and Fig. 4.10(c) examine the impact of PCE factors and percentages of trucks on the computational performance. The PCE factors and percentages of demands for the base case are 1.0 and 92% for PC, 2.5 and 6% for MT, and 3.5 and 2% for HT. As the PCE factors and percentages of trucks increase, CPU times for both cases also increase. This is expected since both cases increase the number of equivalent PC vehicles in the network, which results in higher congestion level and requires more computational efforts.
to equilibrate the path flows of all vehicle classes, particularly for the MT and HT classes. The last case examines the travel time weight parameter with the base case of 1.0 for PC, 1.1 for MT, and 1.15 for HT. As can be seen from Fig. 4.10(d), the computational performance is not affected by the increase of the travel time weight parameter. The reason is that the travel time weight parameter only affects the travel time function of its own vehicle class, unlike the PCE factors which affect not only its own vehicle class but also other vehicle classes; therefore, flow allocations (i.e., adjustments of primal and dual variables) are not sensitive to the changes of the travel time weight parameter. Overall, the sensitivity analyses reveal that the customized path-based algorithm is effectiveness in solving the stochastic multi-class traffic assignment problem with respect to different model parameters despite that some cases require more computational efforts to reach a highly accurate solution.

4.5.3 Experiment 2: Three issues in the multi-class traffic assignment problem

Experiment 2 examines three issues in the multi-class traffic assignment problem: Effect of vehicle interactions, effect of vehicle restrictions, and effect of route overlapping.

4.5.3.1 Effect of vehicle interactions

To examine the effect of vehicle interactions, two multi-class traffic assignment methods are considered: truck pre-loading assignment and simultaneous assignment. The truck pre-loading assignment is a popular multi-class traffic assignment method involving different vehicle types. It preloads the heavy vehicles (i.e., medium and heavy trucks) first and then equilibrates the passenger cars. Hence, the passenger cars are affected by the congestion of the preloaded truck flows, while the trucks are not affected by the congestion
caused by passenger cars. In this method, trucks serve as background traffic; hence, there are no vehicle interactions among the different vehicle types. To account for vehicle interactions, all vehicle types are simultaneously assigned to the network using the non-separable link travel time function in Eq. (4.1).

To facilitate visualization of the assignment results, the vehicle class link flow and total link flow differences between the two assignment methods are depicted in a geographical information system (GIS) map in Fig. 4.11, while Fig. 4.12 reports the congestion level in terms of total travel time and number of used paths for the two assignment methods. Links are coded in color and thickness to highlight the link flow differences (e.g., green color indicates that the truck pre-loading method assigns more flows than that of the simultaneous method, while red color shows the reverse). Fig. 4.11(a), Fig. 4.11(b), and Fig. 4.11(c) show the link flow differences between the two assignment methods for passenger car, medium truck, and heavy truck, respectively, while Fig. 4.11(d) provides the aggregate (or total) link flow differences for all three vehicle types. From these figures, one can visually observe that the link flow patterns between the two assignment methods are quite different, particularly for the individual vehicle types as indicated by Figs. 4.11(a) to 4.11(c) for passenger cars and trucks. The reason is that the truck demands are assigned according to the all-or-nothing (AON) scheme used in the truck pre-loading assignment. Essentially, both medium and heavy trucks are assigned only to one path per O-D pair [see Fig. 4.12(b)]. This may lead to heavy congestion on some links in the background traffic, while the passenger car demands are assigned according to the SUE principle without considering the effect of vehicle interactions. On the other hand, the simultaneous assignment method assigns all three vehicle classes concurrently according
to the SUE principle with explicit considerations of vehicle interactions and different perceptions of network congestion (see the dispersion parameters in Table 4.1).

As a result, the congestion level of the truck pre-loading assignment method is generally higher compared to that of the simultaneous assignment method with vehicle interactions [see Fig. 4.11(a)]. The reason is that the truck flows for both medium and
heavy trucks were loaded without consideration of the congestion contributed by passenger cars (i.e., 92% of the overall network demands). As shown in Fig. 4.11(b), the numbers of used paths for these two truck classes are only one per O-D pair due to the AON assignment used in the truck pre-loading method, while the simultaneous assignment method assigns the truck demands to many more used paths, on the average 8.1 and 8.8 paths per O-D pair, for the medium and heavy trucks, respectively. As for the passenger cars, the average numbers of used paths are similar for the two assignment methods as both accounts for congestion in the equilibration, albeit one considers vehicle interactions while the other does not. Hence, more links with V/C ≥ 1 are assigned under the truck pre-loading method, which leads to higher congestion level for all vehicle classes shown in Fig. 4.11(a).

![Fig. 4.12 Congestion level and number of used paths by vehicle class for the truck pre-loading and simultaneous assignment methods](image)

* the value in parenthesis is the average travel time per vehicle in (a)
* the value in parentheses is the average number of used paths per O-D pair in (b)

4.5.3.2 Effect of route overlapping

As discussed in Section 4.2.2, route overlapping is one of the major concerns in the route choice models used in the SUE problem (Prashker and Bekhor, 2004; Chen et al.,
2012b). The section examines the differences between the multinomial logit (MNL) and path-size logit (PSL) models for the multi-class traffic assignment problem Fig. 4.13 and Fig. 4.14 provide the flow allocation comparison and path flow distributions as a function of PS value between MNL and PSL SUE models, respectively. The flow allocations results in terms of individual vehicle classes and total flow between two models are depicted in GIS maps with coded color and thickness to highlight the link flow differences (e.g., green color indicates that MNL link flows are higher than PSL link flows). Figs. 4.13(a), 4.13(b), and 4.13(c) show the link flow differences between MNL and PSL models for passenger cars, medium trucks, and heavy trucks, respectively, while Fig. 4.13(d) provides the aggregate (or total) link flow differences for all three vehicle types. From these figures, there are visible differences of the assignment results between the MNL and PSL SUE models for each vehicle class as well as the total flows from all three vehicle classes. Particularly, the flow allocations for the passenger car have a significant difference for the two SUE models in the CBD area, where roads are much denser with more overlapping among paths. The MNL SUE model assigns more flows to paths passing through the CBD area as indicated by the links in red (or darker) color, while the PSL SUE model allocates less flows on these paths, because similarity among routes is explicitly considered in the path-size (PS) factor of the PSL route choice model. For the two truck classes, the difference is less significant as the truck demands are relatively less compared to the passenger car demands (8% for trucks and 92% for passenger cars), and vehicle restrictions for trucks are more important in determining the paths for flow assignment.
In terms of path flow allocations between the two SUE models, Fig. 4.14 presents the path flow distributions as a function of the PS factor given in Eq. (4.2). Recall that a smaller PS value implies a higher overlapping among paths, while a larger PS value indicates the paths are more distinct. As can be seen, the MNL model generally assigns a higher percentage of path flows to paths with a lower value of PS to all vehicle classes compared to those of the PSL model. The reason is that the MNL model does not account
for route overlapping, while the PSL model uses the PS factors to adjust the route choice probabilities and flow allocations to the paths with overlap.

![Fig. 4.14 Path flow distributions as a function of PS value for MNL and PSL SUE models](image)

### 4.5.3.3 Effect of vehicle restrictions

This section focuses on improving the realism of multi-class traffic assignment models by considering vehicle restrictions on underpasses due to height, on bridges due to weight, and on certain lanes due to lane restriction. Similar to the effect of vehicle interactions, the vehicle class link flow differences between with and without vehicle restrictions are graphically illustrated using GIS maps shown in Fig. 4.15. For each vehicle class, the link flow differences are separated by non-restricted and restricted links (i.e.,
both medium and heavy truck restrictions) according to those identified in Fig. 4.7. Therefore, there are six figures in total that describe the link flow differences of non-restricted and restricted links. Each pair of figures describe a different vehicle class: Fig. 4.15(a) and Fig. 4.15(b) describe passenger cars; Fig. 4.15(c) and Fig. 4.15(d) describe medium trucks; and Fig. 4.15(e) and Fig. 4.15(f) describe heavy trucks. Note that green color indicates that the assignment method with vehicle restrictions assigns more flows than that of the assignment method without vehicle restrictions, while red color shows the reverse. In addition, the mean absolute error (MAE) is also calculated to quantify the effect of imposing vehicle restrictions on certain links in the network. As can be seen, there are clear differences on the flow allocations between without and with vehicle restrictions for all vehicle classes despite that vehicle restrictions are only applied to the two truck classes. These differences are obvious in the central business district (CBD) area as many truck restricted links are located in this area. Because of these imposed restrictions on medium and heavy trucks, the link flow patterns for the non-restricted links for all vehicle types are also affected as shown in Fig. 4.15(a), Fig. 4.15 (c), and Fig. 4.15 (e). The MAE measure shows that the passenger cars have the highest errors for both non-restricted and restricted links since 92% of the total demand belongs to this vehicle class. As for the medium and heavy trucks, the MAE measures for the restricted links are 14.67 and 8.33, respectively, because the demands for these two truck classes are 6% and 2%. However, the heavy truck restrictions seem to have a higher MAE on the non-restricted links than that of the medium truck restrictions (i.e., 19.91 for heavy trucks versus 16.6 for medium trucks) despite that the demand of heavy trucks is lower.
Fig. 4.15 Flow allocation comparison between without and with vehicle restrictions
4.5.4 Experiment 3: Imposing link capacity constraints

Experiment 3 provides further analysis on the customized solution algorithm and the multi-class traffic assignment model with link capacity constraints imposed on 14 bridges (or 28 directional links) in the Winnipeg network as depicted in Fig. 4.16. In this section, convergence characteristics and flow allocations on bridges (with and without link constraints) will be analyzed. This analysis could be useful in assessing and managing bridge safety in practice.

Fig. 4.16 Bridge locations in Winnipeg network
4.5.4.1 Convergence characteristics with link capacity constraints

As the customized solution algorithm consists of an inner loop (Step 4 for determining the search direction using an iterative balancing scheme) and an outer loop (Step 2 to Step 7), Fig. 4.17 depicts the convergence characteristics for both loops with the center graph showing the overall convergence of the outer loop in terms of CPU time (seconds) and the peripheral graphs showing the convergence results of the inner loop of two dual variables (i.e., bridge B5 WB and bridge B11 NB) in terms of number of inner iterations at three outer iterations (i.e., 1st outer iteration at 0 second, 170th outer iteration at 1,497 seconds, and 347th outer iteration at 3,199 seconds).

Fig. 4.17 Convergence characteristics with link capacity constraints
Unlike experiment 1, this experiment takes more CPU time to adjust the dual variables for the capacity constraints imposed on the 14 bridges [or 28 link capacity constraints in Eq. (4.6)] in the inner loop. From the peripheral graphs in Fig. 4.17, we can observe how the dual variables are adjusted for selected outer iterations. In the first outer iteration (0 second), the dual variables have not stabilized as there are not enough paths in the path set to spread the overloaded links. At the 170th outer iteration, we can observe that the dual variables have stabilized (i.e., satisfying the convergence test of Step 4.5), but the primal variables (path flows) have not satisfied the convergence test of Step 7. Finally, at the 347th outer iteration, both primal and dual variables satisfy the convergence tests of both inner and outer iterations.

4.5.4.2 Effect of capacity constraints on 14 bridges

Table 4.2 provides the assigned V/C ratios of the 14 bridges in both directions without and with imposing the capacity constraints. Without imposing the link capacity constraints, there are 9 overloaded bridges (i.e., assigned link flow exceeds the capacity). The V/C ratios of these 9 bridges range from 1.08 to 1.51, particularly on B7-WB, B10-NB, and B11-NB with a significant amount of truck flows, which may pose a safety hazard. With the link capacity constraints imposed on all bridges, all link flows are within the capacity limits. Some portions of the assigned flows on the overloaded bridges were diverted to other paths using other bridges with spare capacities (e.g., V/C on B1-WB increases from 0.84 to 1.00, reaching its capacity). Hence, the flow pattern on the bridges is generally more dispersed, and more importantly, all assigned flows are within the capacity limit of the bridges.
Table 4.2: Assigned V/C ratios on 14 bridges without and with imposing capacity constraints

<table>
<thead>
<tr>
<th>Bridge ID</th>
<th>Direction</th>
<th>Without capacity constraints</th>
<th></th>
<th>With capacity constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>V/C</td>
<td>PC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PC</td>
</tr>
<tr>
<td>B1</td>
<td>EB</td>
<td>501</td>
<td>56</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>1088</td>
<td>131</td>
<td>195</td>
</tr>
<tr>
<td>B2</td>
<td>EB</td>
<td>559</td>
<td>44</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>455</td>
<td>240</td>
<td>0</td>
</tr>
<tr>
<td>B3</td>
<td>EB</td>
<td>195</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>1517</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B4</td>
<td>EB</td>
<td>403</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>2147</td>
<td>155</td>
<td>0</td>
</tr>
<tr>
<td>B5</td>
<td>EB</td>
<td>832</td>
<td>41</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>2445</td>
<td>297</td>
<td>0</td>
</tr>
<tr>
<td>B6</td>
<td>EB</td>
<td>667</td>
<td>61</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>1301</td>
<td>95</td>
<td>83</td>
</tr>
<tr>
<td>B7</td>
<td>EB</td>
<td>649</td>
<td>61</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>1748</td>
<td>93</td>
<td>116</td>
</tr>
<tr>
<td>B8</td>
<td>EB</td>
<td>481</td>
<td>76</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>272</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>B9</td>
<td>NB</td>
<td>2485</td>
<td>299</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>SB</td>
<td>647</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>B10</td>
<td>NB</td>
<td>2054</td>
<td>211</td>
<td>188</td>
</tr>
<tr>
<td></td>
<td>SB</td>
<td>444</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>B11</td>
<td>NB</td>
<td>1962</td>
<td>81</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>SB</td>
<td>555</td>
<td>37</td>
<td>23</td>
</tr>
<tr>
<td>B12</td>
<td>NB</td>
<td>422</td>
<td>63</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>SB</td>
<td>585</td>
<td>19</td>
<td>33</td>
</tr>
<tr>
<td>B13</td>
<td>NB</td>
<td>3433</td>
<td>273</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>SB</td>
<td>1316</td>
<td>99</td>
<td>0</td>
</tr>
<tr>
<td>B14</td>
<td>NB</td>
<td>650</td>
<td>75</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>SB</td>
<td>995</td>
<td>134</td>
<td>98</td>
</tr>
</tbody>
</table>

Note: **NB**: Northbound, **SB**: Southbound, **EB**: Eastbound and **WB**: Westbound. **PC**: Passenger car, **MT**: Medium truck, **HT**: Heavy truck and **TF**: Total flow in PCE units.

4.6 Conclusions

In this chapter, we presented the model development and the customized solution algorithm for a multi-class traffic assignment model. They will contribute to the literature by including the following array of important modeling features that are necessary for the real-world applications: (a) asymmetric travel time functions to model the vehicle
interaction effect of multiple vehicle types in urban networks, (b) path-size logit model for modeling different random perceptions and accounting for overlapping routes, and (c) both vehicle restriction constraints for individual vehicle class (e.g., lane restriction, height restriction, and weight restriction for trucks) and aggregate physical link capacity constraints to enhance model realism. To show proof of concept, a real network in the City of Winnipeg, Canada was conducted as a case study. Numerical results revealed that: (1) disregarding vehicle interactions could significantly affect the traffic assignment results of multiple vehicle classes, particularly with medium and heavy trucks that have a strong asymmetric interactions with other vehicle classes such as passenger cars, (2) neglecting the route overlapping problem could lead to biased route flow allocations and hence overload links with heavy route overlaps, (3) ignoring vehicle restriction constraints could lead to assignment results not conforming to actual flow restrictions applied in the real world, (4) imposing link capacity constraints could have a significant impact on the flow allocations, particularly for truck flows, which may pose a safety hazard on bridges, and (5) incorporating the SRA stepsize scheme was effective for solving the multi-class traffic assignment problem with different modeling considerations.

It should be pointed out that the multi-class trip tables were assumed available from the Emme software (INRO Consultants, 2013). Particularly, the medium and heavy truck trip tables were assumed to be disaggregated from the truck demands by 70% and 30%, respectively. In practice, the trip tables for passenger car and different truck classes need to be estimated simultaneously due to the asymmetric interactions between auto traffic and truck traffic and different vehicle restrictions applied to different vehicle classes. Hence, future research will focus on developing a multi-class origin-destination trip estimation
method for multiple vehicle classes with multiple data sources. In addition, different physical and environmental side constraints (e.g., Chen et al., 2011) can be considered to model different physical configurations (e.g., lane drop, merging, nodal capacity, etc.), traffic control policies, and environmental restraints (e.g., emission and noise) to improve the realism of multi-class traffic assignment models.

REFERENCES


CHAPTER 5
PATH FLOW ESTIMATOR FOR PLANNING APPLICATIONS IN SMALL COMMUNITIES

Abstract

This research presents an alternative planning framework to model and forecast network traffic for planning applications in small communities, where limited resources debilitate the development and applications of the conventional four-step travel demand forecasting model. The core idea is to use the path flow estimator (PFE) to estimate current and forecast future traffic demand while taking into account of various field and planning data as modeling constraints. Specifically, two versions of PFE are developed: a base year PFE for estimating the current network traffic conditions using field data and planning data, if available, and a future year PFE for predicting future network traffic conditions using forecast planning data and the estimated base year origin-destination trip table as constraints. In the absence of travel survey data, the proposed method uses similar data (traffic counts and land use data) as a four-step model for model development and calibration. Since the Institute of Transportation Engineers (ITE) trip generation rates and Highway Capacity Manual (HCM) are both utilized in the modeling process, the analysis scope and results are consistent with those of common traffic impact studies and other short-range, localized transportation improvement programs. Solution algorithms are also developed to solve the two PFE models and integrated into a GIS-based software called Visual PFE. For proof of concept, two case studies in northern California are performed to
demonstrate how the tool can be used in practice. The first case study is a small community of St. Helena, where the city’s planning department has neither an existing travel demand model nor the budget for developing a full four-step model. The second case study is in the city of Eureka, where there is a four-step model developed for the Humboldt County that can be used for comparison. The results show that the proposed approach is applicable for small communities with limited resources.

5.1 Introduction

Transportation is critical to the social, environmental, and economic health of every metropolitan city. Because of its importance, federal regulations in the United States require each urbanized area over 50,000 in population to have a Metropolitan Planning Organization (MPO) responsible for transportation planning (Meyer and Miller, 2000). Yet, according to the U.S. Census, over 40 percent of all U.S. communities have populations less than 50,000. In California, there are 333 municipalities (out of 535 municipalities) that have a population less than 50,000 (U.S. Census Bureau, 2009). Such small communities usually do not have sufficient resources to conduct travel surveys or embark on model development and maintenance for carrying out various planning functions. Current practice in modeling network traffic is through a four-step travel demand forecasting model (i.e., trip generation, trip distribution, mode choice, and traffic assignment), commonly referred to as the four-step model, that requires travel surveys as input and specialized technical staffs to develop and estimate. Although such a modeling approach has been used in practice in major urban areas, Yan (1998) noted that many smaller communities usually do not have sufficient resources to conduct travel surveys, nor to house technical staffs for model development and maintenance. Without data from a travel survey in the study area,
trip generation rates of various land use zones are often “borrowed” from such published data as Trip Generation of the Institute of Transportation Engineers (ITE, 2008) or reports of travel surveys performed in other areas. The unavailability of data on Trip Length Frequency Distribution (TLFD) of local travelers often forces modelers to skip the calibration of trip distribution models. Instead, calibration and validation of the overall model are often carried out by altering the friction factors and adding k-factors (i.e., an empirically zone-to-zone adjustment factor, which takes into account the effects on travel patterns of defined social and economic linkages not otherwise incorporated into the model), in a trial-and-error fashion, to the trip distribution model such that the results of traffic assignment would match traffic counts on selective screenlines and critical links. The calibration process is usually a lengthy process and the resultant models often contain many factors that do not have the necessary behavioral foundation established from travel surveys. Schutz (2000) suggests that for those communities to meet the planning requirements, development of innovative methodologies is urgent and necessary. Many researchers and practitioners have proposed techniques for modeling networks in small communities. Turnquist and Gur (1979) introduced a method for estimating origin-destination (O-D) trip tables from observed link volumes for evaluation of short-range, sub-regional traffic improvement plans (e.g., quick response traffic impact analysis when improvement to the transportation systems are proposed). They noted that estimating trip tables from traffic counts represents a cost-effective alternative to conventional trip generation and distribution models that depend on expensive, time-consuming surveys and labor-intensive data preparation and analyses. Because many jurisdictions regularly conduct traffic counts on streets and intersections, estimating trip tables from observed
traffic volumes can also significantly reduce the effort and time associated with data collection. However, Turnquist and Gur's study deals exclusively with the estimation of trip table. It does not demonstrate how the method can be used in a modeling process when changes in land use and transportation network are expected. Many of the subsequent studies on the estimation of O-D trip tables continue to focus on the methods and applications in evaluating traffic operation strategies, which usually do not involve regional growth due to land use changes. As a result, planners and modelers dealing with forecasting network traffic involving land use development cannot easily apply the O-D estimation methods.

Path Flow Estimator (PFE), originally developed by Bell and Shield (1995) and further enhanced by Chen et al. (2005, 2009, 2010b) and Chootinan et al. (2005a), is a one-stage network observer capable of estimating path flows and path travel times using only traffic counts from a subset of network links. The core component of PFE is a logit-based route choice model (Fisk, 1980) that interacts with link cost functions to produce a stochastic user equilibrium (SUE) traffic flow pattern conforming with the traffic counts and other side constraints if available (e.g., intersection turning movement flows, target O-D flows, production flows, attraction flows, etc.). The theoretical advantage of the nonlinear PFE is the single-level convex programming formulation with side constraints. Since the objective function is strictly convex with respect to the decision variables (path flows) and the constraints are all linear (equality and inequality) relationships, the optimization is guaranteed to yield unique path flows that can be used to derive other useful information at different spatial levels. Note that the uniqueness of path flows in PFE is different from the underspecified problem of the O-D estimation from traffic counts, which
is known to have multiple solutions since the number of observations (i.e., link counts) is generally less than the number of variables (i.e., O-D pairs). The multiplicity of the solutions should be quantified by some quality measures (e.g., the generalized demand scale of Chen et al., 2012a).

The basic idea of PFE is to find a set of unique path flows that can reproduce the observed link counts as well as other available side information, such as historical O-D patterns or link capacities, to increase the observability of the O-D trip table. The resulting path flows can be used to derive other flows, such as flows on unobserved links, intersection turning movement flows, O-D flows, production flows of a traffic zone, attraction flows of a traffic zone, and total travel demand in the entire network. The flexibility of aggregating path flows at different spatial levels allows us to develop a simplified PFE planning model that makes use of not only existing field data (e.g., traffic counts, intersection turning movement flows, etc.) but also planning data (e.g., socio-economic and land use data converted through the ITE trip generation rates (ITE, 2008) to obtain zonal production and attraction flows) for estimating and forecasting network traffic in small communities in the absence of trip generation models. Using the trip rates to capture the trip-making propensity of a given land use configuration in the study area is a common practice, and provides an economical and reasonable estimate when planning resources are limited.

The main objective of this research is to develop a simplified planning framework that exploits the O-D estimation capability of the path flow estimator (PFE) to perform planning applications in small communities where limited planning resources hinder the development and application of a full four-step model. Two versions (i.e., base year and
future year) of the PFE are proposed to address the specific transportation planning issues and needs of small communities. The base year PFE is used for estimating the O-D trip table using current year field measurement data (and planning data if any) as constraints, and the future year PFE is used for predicting network traffic conditions using future trip production and trip attraction and scaled base year (calibrated) O-D trip table as constraints. The unique features of PFE make the proposed framework less resource intensive than the conventional four-step process, and particularly suited for small communities with limited resources. These two versions for PFE are finally integrated into a GIS-based software tool called Visual PFE for planning applications in small communities.

The rest of this chapter is organized as follows. The tailored PFE models developed for transportation planning in small communities are discussed in Section 5.2. Section 5.3 provides the solution algorithm for solving the two PFE models. Section 5.4 presents two case studies to demonstrate how the Visual PFE tool can be used in practice. Section 5.5 provides some concluding remarks.

5.2 Simplified PFE planning tool

Fig. 5.1 depicts the overall framework of using PFE as a simplified planning tool for estimating the base year O-D trip table and predicting future year network traffic conditions. The major thrust of the proposed approach is that model estimation and forecasting are each accomplished through PFE constrained by various field data and planning data. Specifically, the simplified PFE planning tool includes two modules as depicted in Fig. 5.1: (a) a base year PFE for estimating and calibrating the origin-destination (O-D) trip table using current field data (e.g., traffic counts) and planning data (e.g., historical/target trip table, trip production, trip attraction) as constraints, and (b) a
future year PFE for predicting network traffic conditions, taking into account future trip production and trip attraction and the scaled base year O-D trip table to match future total demand as constraints. When only field data (e.g., traffic counts) are available in the base year for model estimation, modeling network traffic with the base year PFE can be more efficient than with a traditional four-step model because it can produce estimates within an acceptable error bound in one estimation step. When field data are not available in the future, the scaled base year O-D trip table can be used as constraints to preserve the trip-making pattern in the future. These unique features of PFE make it a promising planning tool for modeling and forecasting network traffic for small communities with limited resources.

Fig. 5.1 The PFE-based transportation planning framework
5.2.1 Input data for the base year PFE

Unlike the conventional four-step model which requires extensive surveys for each step (i.e., socio-economic and land use survey, household survey, roadside survey, internal and external survey, on-board transit survey, traffic counts, etc.), PFE just requires some existing field data and planning data that can be obtained in public domains. For details, we provide a description of each type of data and its mathematical expression associated with PFE as follows.

5.2.1.1 Observed traffic counts

Traffic counts on streets and intersections are regularly collected by various government agencies (e.g., State Department of Transportation, County, City, MPO, etc.). These observed traffic counts are collected in different formats: hourly traffic volume, annual average daily traffic (AADT), annual average weekday traffic (AAWT). Using the peak hour traffic factor and directional distribution factor (Roess et al., 2004), it is possible to convert them into a common and usable format for PFE. In addition, PFE does not require the estimated link flows to reproduce the observed traffic counts exactly, but within an acceptable range defined by the error bounds (i.e., upper and lower bounds). More reliable counts will constrain the estimation to be within a smaller tolerance, whereas less reliable counts will allow for a larger deviation. The introduction of the user-defined error bounds in the traffic count constraints enhances the flexibility of PFE by allowing the user to incorporate local knowledge about the network conditions in the estimation process. The observed traffic count on a link is expressed as follows:
where \( v_a \) is the observed traffic volume on link \( a \); \( x_a \) is the estimated flow on link \( a \); \( \epsilon_a \) is the percentage of measurement error allowed for the traffic count on link \( a \); \( \overline{A} \) is the set of network links with measurements. The error bounds are inputs provided by the user. This typically involves knowledge and experience of the user to provide appropriate error bounds that reflect the traffic measurements within the study area. Alternatively, these error bounds can be specified according to the road functional classification (i.e., freeway, principle arterial, minor arterial, and collector) of roadways as suggested by the Federal Highway Administration (FHWA, 1990) and Cambridge Systematics (2010) in Table 5.1.

Table 5.1: Percentage error for daily traffic volumes by facility type

<table>
<thead>
<tr>
<th>Facility Type</th>
<th>FHWA Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freeway</td>
<td>+/- 7%</td>
</tr>
<tr>
<td>Major Arterial</td>
<td>+/- 10%</td>
</tr>
<tr>
<td>Minor Arterial</td>
<td>+/- 15%</td>
</tr>
<tr>
<td>Collector</td>
<td>+/- 25%</td>
</tr>
</tbody>
</table>

*Source: FHWA (1990) and Cambridge Systematics (2010)*

Similarly, turning movement counts at intersections can also be used in the base year PFE. Note that modeling intersection turning movements typically requires network expansion at each intersection in order to represent all turning movements. However, adding nodes and links to the network to model intersection turning movements is an expensive proposition. Consider a single intersection represented as a node (left as shown in Fig. 5.2), to model the intersection turning movements would require adding 3 nodes and 12 links for each intersection (right, Fig. 5.2). For a network with 1,000 nodes and 4,000 links, this would require 4,000 nodes and 16,000 links to fully model all intersection...
turning movements in the network. This is a 4-fold increase in terms of nodes and links. For large-scale networks, this approach is infeasible.

Fig. 5.2 Network representation of an intersection

In this study, we make innovative use of the path-flow solution to derive intersection turning movements without the need to expand the network. For the turning movements at an intersection, the orientations of links connected to intersection are needed so as to determine the individual turning movement (e.g., left, right, or through movement) from the used paths without the need to expand the network for representing turning movements [see Chen et al. (2012b) for the procedure for deriving intersection turning movements using path flows]. The observed intersection movement counts at an intersection can be expressed as follows:

\[(1 - \varepsilon_m^i) \cdot g_m^i \leq t_m^i \leq (1 + \varepsilon_m^i) \cdot g_m^i, \quad \forall m \in M_i, i \in \overline{I} \]  \hspace{1cm} (5.2)

where \(g_m^i\) is the observed traffic volume on turning movement \(m\) at intersection \(i\); \(t_m^i\) is the estimated flow on turning movement \(m\) at intersection \(i\); \(\varepsilon_m^i\) is the percentage of measurement error allowed for turning movement \(m\) from intersection \(i\); \(M_i\) is the set of turning movements at intersection \(i\); and \(\overline{I}\) is the set of intersections with measurements.
Similar to the traffic counts on a link, PFE does not require the estimated intersection movement flows to reproduce the observed intersection turning movement counts exactly, but within an acceptable range defined by the error bounds (i.e., upper and lower bounds). The user can incorporate local knowledge about the network conditions to provide these error bounds.

5.2.1.2 Zonal production and attraction flows

One of the features of PFE is that it can account for not only traffic counts, but also such planning data as zonal productions and attractions as inputs to the O-D estimation problem. Thus, it has the potential to be an alternative to the conventional trip generation and distribution models for the evaluation of short-range and sub-regional transportation plans. Unlike the traditional O-D estimation methods that use only observed traffic counts and prior O-D trips (or target trip table) to update the trip table, the zonal production and attraction constraints can reflect the spatial interactions of the land use development and travel patterns within a community. In the absence of trip generation models, existing socio-economic and land use information in the base year (e.g., population, employment, number of dwelling units, dwelling types, etc.) can be used to estimate zonal productions and zonal attractions as shown in Fig. 5.3. This requires using the ITE trip generation rates (ITE, 2008) to convert the socio-economic and land use information to zonal productions and zonal attractions. The trip generation rates published by ITE are expressed in the unit of vehicle trips for: (1) an average weekday, Saturday and Sunday, (2) weekday morning and evening peak hours of the generator, and (3) the weekday morning and evening peak hours on the adjacent street traffic with the percentage of trip productions and attractions for each zone. Table 5.2 shows an example of applying the ITE trip generation rates. The
entering and exiting trip ends are used to estimate vehicle trip attractions and productions, respectively.

Fig. 5.3 Zonal production and attraction flows from socio-economic and land use data

Table 5.2: Example of trip production and trip attraction estimation using ITE trip rates

<table>
<thead>
<tr>
<th>TAZ</th>
<th>Parcel ID</th>
<th>Land Use</th>
<th>Unit</th>
<th>Number</th>
<th>Trip Rate</th>
<th>Total Trip Ends</th>
<th>Entering (Attraction)</th>
<th>Exiting (Production)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>General office</td>
<td>ksf</td>
<td>215</td>
<td>1.49</td>
<td>320</td>
<td>265</td>
<td>54</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>Single Family (Medium density)</td>
<td>acres</td>
<td>84</td>
<td>2.74</td>
<td>230</td>
<td>78</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Single Family (High density)</td>
<td>acres</td>
<td>23</td>
<td>3.74</td>
<td>86</td>
<td>29</td>
<td>57</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Agriculture</td>
<td>ksf</td>
<td>60</td>
<td>2.01</td>
<td>121</td>
<td>61</td>
<td>60</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>General office</td>
<td>ksf</td>
<td>156</td>
<td>1.49</td>
<td>232</td>
<td>193</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Single Family</td>
<td>acres</td>
<td>3.2</td>
<td>2.74</td>
<td>9</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

ksf = Thousands of Square Feet
Using the trip rates to capture the trip-making propensity of a given land use configuration in the study area is a common practice, and it provides an economical and reasonable estimate when planning resources are limited. In the same manner as the traffic count constraints, the zonal production and zonal attraction constraints can be expressed as follows.

\[(1 - \varepsilon_r) \cdot O_r \leq P_r \leq (1 + \varepsilon_r) \cdot O_r, \forall r \in \bar{R} \quad (5.3)\]
\[(1 - \varepsilon_s) \cdot D_s \leq A_s \leq (1 + \varepsilon_s) \cdot D_s, \forall s \in \bar{S} \quad (5.4)\]

where \(O_r\) and \(D_s\) are the observed trip production of origin \(r\) and observed trip attraction of destination \(s\) obtained by converting land use data via the ITE trip rates; \(P_r\) and \(A_s\) are the estimated trip production of origin \(r\) and estimated trip attraction of destination \(s\); \(\varepsilon_r\) and \(\varepsilon_s\) are the error bounds allowed for trip production of origin \(r\) and trip attraction of destination \(s\); and \(\bar{R}\) and \(\bar{S}\) are the sets of zones with planning data. The introduction of error bounds in Eqs. (5.3) and (5.4) provides the flexibility to have differential land use developments among the different zones.

5.2.1.3 Prior O-D trip tables

It is well known that O-D estimation from traffic counts is a highly underspecified problem (i.e., the number of O-D demands to be estimated is much greater than the number of independent traffic counts). Therefore, there are multiple O-D trip table estimates that can reproduce the same link flows. In order to increase the observability of the O-D estimation problem from traffic counts, target (or outdated) O-D demands are commonly included to preserve the spatial distribution of the O-D demand pattern. Target (or prior) O-D flow in the base year can be expressed as follows:
\[(1 - \varepsilon_{rs}) \cdot z_{rs} \leq q_{rs} \leq (1 + \varepsilon_{rs}) \cdot z_{rs}, \quad \forall rs \in \overline{RS} \quad (5.5)\]

where \(z_{rs}\) is the target O-D flows between origin \(r\) and destination \(s\); \(q_{rs}\) is the estimated O-D flows between origin \(r\) and destination \(s\); \(\varepsilon_{rs}\) is the error bound allowed for the target O-D demands between origin \(r\) and destination \(s\); and \(\overline{RS}\) is the set of target (or prior) O-D pairs. Similar to the zonal production and attraction flows, reliable O-D pairs are constrained within a smaller tolerance, while less reliable O-D pairs can have a larger deviation.

5.2.1.4 Total demand

Total demand is a useful measure to assess the overall traffic demand level in the study area. It can be expressed as follows:

\[(1 - \varepsilon) \cdot F \leq T \leq (1 + \varepsilon) \cdot F \quad (5.6)\]

where \(F\) is the target total demand; \(T\) is the estimated total demand; and \(\varepsilon\) is the error bound allowed for the target total demand.

5.2.2 Base year PFE formulation

The core component of base year PFE is a logit-based route choice model with various side constraints as discussed above. Based on the equivalent mathematical programming formulation given by Fisk (1980), the base year PFE formulation can be formulated as a constrained convex program as shown in Fig. 5.4.
**Objective Function**

Minimize: $Z(f) = \frac{1}{\theta} \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{k}^{rs} \left( \ln f_{k}^{rs} - 1 \right) + \sum_{a \in A} \int_{0}^{x_{a}} t_{a}(\omega) d\omega$ \hspace{1cm} (5.7)

**Constraints:**

### Field Data

Observed traffic counts from Eq. (5.1) \hspace{1cm} (Core data)

Intersection turning movement counts from Eq. (5.2) \hspace{1cm} (Optional data)

### Planning Data

Zonal production flows from Eq. (5.3) \hspace{1cm} (Optional data)

Zonal attraction flows from Eq. (5.4) \hspace{1cm} (Optional data)

Target O-D trip table from Eq. (5.5) \hspace{1cm} (Desirable data)

Total demand from Eq. (5.6) \hspace{1cm} (Optional data)

### Others

$x_{a} \leq C_{a}, \quad \forall a \in U$ \hspace{1cm} (5.8)

$f_{k}^{rs} \geq 0, \quad \forall k \in K_{rs}, \text{ } rs \in RS$ \hspace{1cm} (5.9)

### Definition

$x_{a} = \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{k}^{rs} \delta_{ka}^{rs}, \quad \forall a \in A$ \hspace{1cm} (5.10)

$t_{m}^{i} = \sum_{rs \in RS} \sum_{k \in K_{rs}} \sum_{a \in \Omega_{k}} \sum_{k \in OUT_{k}} f_{k}^{rs} \delta_{ka}^{rs} \delta_{kb}^{rs}, \quad \forall m \in M_{s}, i \in I$ \hspace{1cm} (5.11)

$P_{r} = \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{k}^{rs}, \quad \forall r \in R$ \hspace{1cm} (5.12)

$A_{s} = \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{k}^{rs}, \quad \forall s \in S$ \hspace{1cm} (5.13)

$q_{rs} = \sum_{k \in K_{rs}} f_{k}^{rs}, \quad \forall rs \in RS$ \hspace{1cm} (5.14)

$T = \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{k}^{rs}$ \hspace{1cm} (5.15)

---

Fig. 5.4 Base year PFE formulation
where $\theta$ is the dispersion parameter in the logit model; $f_{rk}$ is the flow on path $k$ connecting O-D pair $rs$; $t_a()$ is the travel time on link $a$; $x_a$ is the estimated traffic volume on link $a$; $\delta_{ka}$ is the path-link indicator, 1 if link $a$ is on path $k$ between O-D pair $rs$ and 0 otherwise; and the rest of the variables are previously defined.

The objective function (5.7) has two terms: an entropy term and a user equilibrium term. The entropy term seeks to spread trips onto multiple paths according to the dispersion parameter, while the user equilibrium term tends to cluster trips on the minimum cost paths. As opposed to the traditional logit-based SUE model, the base year PFE finds path flows that minimize the SUE objective function in Eq. (5.7) while simultaneously reproducing traffic counts on all observed links in Eq. (5.1), turning movement counts on all observed intersections in Eq. (5.2), zonal production and attraction flows of certain origins and destinations in Eqs. (5.3) and (5.4), prior travel demands of certain O-D pairs in Eq. (5.5), and total demand in Eq. (5.6) within some predefined error bounds. These error bounds are essentially confidence levels of the observed data at different spatial levels used to constrain the path flow estimation. For the unobserved links, the estimated flows cannot exceed their respective capacities as indicated by Eq. (5.8). This constraint is incorporated for the same purpose as in the capacitated traffic assignment (Larsson and Patriksson, 1995), which is to prevent producing unrealistically high link flow estimates. Eq. (5.9) constrains the path flows to be non-negative. Eqs. (5.10), (5.11), (5.12), (5.13), (5.14) and (5.15) are definitional constraints that sum up the estimated path flows to obtain the link flows, intersection turning movement flows, zonal production flows, zonal attraction flows, O-D flows, and total demand, respectively. The analytical path flow solution for the base
year PFE formulation can be derived as a function of path costs and dual variables associated with the constraints (see Appendix B for the detailed derivation).

5.2.3 Input data for the future year PFE

Since one cannot observe future traffic counts, the forecasting process for the future network traffic conditions needs to make full use of available planning data (i.e., future socioeconomic and land use data) and the O-D demand pattern estimated in the base year.

5.2.3.1 Future zonal production and attraction flows

Similar to the base year PFE, future socioeconomic and land use data can be used to generate future zonal production and attraction flow using the ITE trip generation rates. The future zonal production and zonal attraction constraints can be expressed as follows.

\[
(1 - \varepsilon_r) \cdot O_r \leq P_r \leq (1 + \varepsilon_r) \cdot O_r, \quad \forall r \in \tilde{R} \tag{5.16}
\]

\[
(1 - \varepsilon_s) \cdot D_s \leq A_s \leq (1 + \varepsilon_s) \cdot D_s, \quad \forall s \in \tilde{S} \tag{5.17}
\]

where \(O_r\) and \(D_s\) are the future zonal production of origin \(r\) and future zonal attraction of destination \(s\) obtained by converting future land use data via the ITE trip rates; \(P_r\) and \(A_s\) are the estimated future zonal production of origin \(r\) and estimated future zonal attraction of destination \(s\); \(\varepsilon_r\) and \(\varepsilon_s\) are the forecasting error bounds for trip production of origin \(r\) and trip attraction of destination \(s\); and \(\tilde{R}\) and \(\tilde{S}\) are the sets of zones with future data. The error bounds in Eqs. (5.16) and (5.17) provide the flexibility to have differential land use developments among the different zones to reflect the spatial interactions of the future land use development and travel patterns within a community.
### 5.2.3.2 Future target O-D trip table

Since future traffic counts are not available, it is necessary to make use of the base year spatial demand distribution as a guide to infer the trip-making pattern in the future year. Due to uneven land use changes, some O-D pairs may experience substantial changes in its demand relative to other O-D pairs, so the scaling over the base year O-D may not apply to all the O-D pairs in the same way. Fig. 5.5 provides a graphical illustration of the steps used to compute future year target O-D trip table (or selected future target O-D pairs).

<table>
<thead>
<tr>
<th>Base Year PFE</th>
<th>Future Year Landuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base Year O-D Trip</td>
<td>Future year Production, Attraction and Total Demand</td>
</tr>
<tr>
<td>$r$</td>
<td>$s$</td>
</tr>
<tr>
<td>$q_{rs}^B$</td>
<td>$P_r^B$</td>
</tr>
<tr>
<td>$A_s^B$</td>
<td>$T^B$</td>
</tr>
<tr>
<td>$r$</td>
<td>$s$</td>
</tr>
<tr>
<td>$O_r^L$</td>
<td>$D_s^L$</td>
</tr>
<tr>
<td>$A_r^L$</td>
<td>$F^L$</td>
</tr>
</tbody>
</table>

$r$ : origin  
$s$ : destination  
$B$ : base year  
$L$ : future year

$$s_{rs} = \frac{P_r^B}{O_r^L}, \quad s_{rs} = \frac{A_s^B}{D_s^L} \quad \text{or} \quad s_{rs} = \frac{T^B}{F^L}, \quad \forall \  rs \in RS$$

$S_{rs}$ : Scaling factor on O-D pair $rs$

$$z_{rs} = q_{rs}^B \cdot s_{rs}, \quad \forall \ rs \in RS$$

$Z_{rs}$ : Future Target O-D Trip on O-D pair $rs$

**Fig. 5.5 Derivation of future target O-D trips**
With the scaling factors, the future target O-D flows can be expressed as follows:

\[(1 - \varepsilon_{rs}) \cdot z_{rs} \leq q_{rs} \leq (1 + \varepsilon_{rs}) \cdot z_{rs}, \quad \forall rs \in \overline{RS}\]  

(5.18)

where \(z_{rs}\) is the future target O-D flows between origin \(r\) and destination \(s\); \(q_{rs}\) is the estimated future O-D flows between origin \(r\) and destination \(s\); \(\varepsilon_{rs}\) is the error bound allowed for the target O-D demands between origin \(r\) and destination \(s\); and \(\overline{RS}\) is the set of target O-D pairs. These future target O-D flow constraints can help to infer the spatial distribution of O-D demands.

5.2.3.3 Future total demand

Similar to the base year, future total demand can be included to constrain the estimated total demand as follows.

\[(1 - \varepsilon) \cdot F \leq T \leq (1 + \varepsilon) \cdot F\]  

(5.19)

where \(F\) is the future target total demand; \(T\) is the estimated future total demand; and \(\varepsilon\) is the error bound allowed for the target total demand.

5.2.4 Future year PFE formulation

The core component of the future year PFE is a logit-based route choice model with only planning data (i.e., future trip production and trip attraction) and the scaled base year O-D trip table to match future total demand as constraints for predicting future network traffic conditions. Similar to the base year PFE, the future year PFE can be formulated as a convex program with various side constraints as shown in Fig. 5.6.
**Objective Function**

Minimize: \( Z(f) = \frac{1}{\theta} \sum_{r \in \mathcal{RS}} \sum_{k \in \mathcal{K}_r} f_{krs}^r \left( \ln f_{krs}^r - 1 \right) + \sum_{a \in \mathcal{A}} \int_{0}^{\infty} t_a(\omega) d\omega \)  

(5.20)

**Constraints:**

**Planning Data**

Zonal production flows from Eq. (5.16)  
(Core data)

Zonal attraction flows from Eq. (5.17)  
(Core data)

Target O-D trip table from Eq. (5.18)  
(Core data)

Total demand from Eq. (5.19)  
(Optional data)

**Others**

\( f_{krs}^r \geq 0, \quad \forall k \in \mathcal{K}_r, rs \in \mathcal{RS} \)  
(5.21)

**Definition**

\( x_a = \sum_{r \in \mathcal{RS}} \sum_{k \in \mathcal{K}_r} f_{krs}^r x_{rsa}, \quad \forall a \in \mathcal{A} \)  
(5.22)

\( P_r = \sum_{a \in \mathcal{S}} \sum_{k \in \mathcal{K}_r} f_{krs}^r, \quad \forall r \in \mathcal{R} \)  
(5.23)

\( A_s = \sum_{r \in \mathcal{R}} \sum_{k \in \mathcal{K}_r} f_{krs}^r, \quad \forall s \in \mathcal{S} \)  
(5.24)

\( q_{rs} = \sum_{k \in \mathcal{K}_r} f_{krs}^r, \quad \forall rs \in \mathcal{RS} \)  
(5.25)

\( T = \sum_{r \in \mathcal{RS}} \sum_{k \in \mathcal{K}_r} f_{krs}^r \)  
(5.26)

Fig. 5.6 Future year PFE formulation

Similar to the base year PFE formulation, the future year PFE formulation finds path flows that minimize objective function (5.20) while conforming to the future zonal production and attraction flows of certain origins and destinations in Eqs. (5.16) and (5.17), future target travel demands of certain O-D pairs in Eq. (5.18), and future total demand in Eq. (5.19) within some predefined error bounds. The analytical path flow solution for the
future PFE formulation can be derived as a function of path costs and dual variables associated with the constraints (see Appendix C for the detailed derivation).

5.3 Solution algorithm

The solution procedure for solving the two versions of PFE is depicted in Fig. 5.7. It consists of three main modules: (1) an iterative balancing scheme, (2) column (or path) generation, and (3) output generation from path flows.

![Flowchart of the PFE solution algorithm]

Fig. 5.7 Flowchart of the PFE solution algorithm

The basic idea of the iterative balancing scheme is to sequentially scale the path flows to fulfill one constraint at a time by adjusting the dual variables. Once the scheme
converges, the path flows can be analytically determined. A column generation is included in the solution procedure to avoid path enumeration for a general transportation network. Finally, an output generation procedure is used to derive information at different spatial levels using the PFE path-flow solution (e.g., link flows, turning movement flows for all intersections, production flows, attraction flows, O-D flows, and total demand).

5.3.1 Iterative balancing scheme

Note that the iterative balancing scheme is demonstrated for the base year PFE with both field data and selected planning data. The same scheme can also be used to solve the future year PFE with only planning data as constraints. The steps are summarized as follows.

**Step 1. Initialization**

1.1 Set inner iteration \( (n) = 0 \)

1.2 Set primal variables: \( x^n, t^n, P^n, A^n, q^n \) and \( T^n = 0 \)

1.3 Set dual variables: \( (u^n), (u^n), (\tau^{+}m)^n, (\rho^{+}r)^n, (\rho^{-}r)^n, (\alpha^{+}s)^n, (\alpha^{-}s)^n, (o^{-}rs)^n, (o^{+}rs)^n, (\psi^{-})^n, (\psi^{+})^n \) and \( (d_a)^n = 0 \)

where \( u^n, u^n, \tau^{+}m, \tau^{-}m, \rho^{+}r, \rho^{-}r, \alpha^{+}s, \alpha^{-}s, o^{-}rs, o^{+}rs, \psi^{-}, \psi^{+} \) and \( d_a \) are the dual variables of constraints (5.1), (5.2), (5.3), (5.4), (5.5), (5.6) and (5.8), respectively.

**Step 2. Compute Dual and Primal Variables**

The values of, \( u^n, \tau^{+}m, \rho^{+}r, \alpha^{+}s, o^{+}rs, \psi^{+} \) and \( d_a \) are restricted to be non-positive, while the values \( u^n, \tau^{-}m, \rho^{-}r, \alpha^{-}s, o^{-}rs \) and \( \psi^{-} \) must be nonnegative. \( u^n, u^n, \tau^{-}m \) and \( \tau^{+}m \) can be viewed as the corrections to the link travel times and intersection turning movement delays,
respectively, by adjusting the estimated path flows to match with the link count and intersection turning movement count constraints specified by Eqs. (5.1) and (5.2). Similarly, $\rho_r^-$, $\rho_r^+$, $\alpha_r^-$, $\alpha_r^+$, $o_{rs}^-$, $o_{rs}^+$, $\psi^-$ and $\psi^+$ are can be viewed as the corrections to the zonal production levels, zonal attraction levels, O-D travel times, and total demand level, respectively. These corrections can be used to steer the estimated path flows to match with the observed zonal production, zonal attraction, target O-D flow, and total demand constraints specified by Eqs. (5.3), (5.4), (5.5), and (5.6). $d_a$ is related to the link queuing delay (Bell et al., 1997) when the estimated link flow reaches its capacity. These dual variables are zero if the estimated primal values (e.g., link flows, intersection turning movement flows, production flows, attraction flows, O-D flows, and total demand) are within an acceptable range defined by the measurement error bounds, non-zero if they are binding at one of the limits, and infinity (or very large positive or negative values) if there exists no solution that can fulfill the constraints (Bell et al., 1997).

2.1 Update dual variables

a. For each measured link ($a \in \tilde{A}$), update the dual variables

$$\left( u_a^+ \right)^n = \text{Min} \left\{ 0, \left( u_a^+ \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{(1 + \epsilon_a) \cdot v_a}{x_a^n} \right) \right\}$$

$$\left( u_a^- \right)^n = \text{Max} \left\{ 0, \left( u_a^- \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{(1 - \epsilon_a) \cdot v_a}{x_a^n} \right) \right\}$$

b. For each unmeasured link ($a \in U$), update the dual variables

$$\left( d_a \right)^n = \text{Min} \left\{ 0, \left( d_a \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{C_a}{x_a^n} \right) \right\}$$
c. For each measured intersection \((i \in I)\), update the dual variables
\[
(t^+_m) = \min \left\{ 0, \left( t^+_m \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{\left(1 + \varepsilon^+_i \right) \cdot g^+_m}{t^+_m} \right) \right\}
\]
\[
(t^-_m) = \max \left\{ 0, \left( t^-_m \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{\left(1 - \varepsilon^+_i \right) \cdot g^-_m}{t^-_m} \right) \right\}
\]

d. For each zonal production flow \((r \in R)\), update the dual variables
\[
(\rho^+_r) = \min \left\{ 0, \left( \rho^+_r \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{\left(1 + \varepsilon_r \right) \cdot O_r}{\rho^+_r} \right) \right\}
\]
\[
(\rho^-_r) = \max \left\{ 0, \left( \rho^-_r \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{\left(1 - \varepsilon_r \right) \cdot O_r}{\rho^-_r} \right) \right\}
\]

e. For each zonal attraction flow \((s \in S)\), update the dual variables
\[
(\alpha^+_s) = \min \left\{ 0, \left( \alpha^+_s \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{\left(1 + \varepsilon_s \right) \cdot D_s}{\alpha^+_s} \right) \right\}
\]
\[
(\alpha^-_s) = \max \left\{ 0, \left( \alpha^-_s \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{\left(1 - \varepsilon_s \right) \cdot D_s}{\alpha^-_s} \right) \right\}
\]

f. For each target O-D flow \((rs \in RS)\), update the dual variables
\[
(o^+_rs) = \min \left\{ 0, \left( o^+_rs \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{\left(1 + \varepsilon_{rs} \right) \cdot z_{rs}}{o^+_rs} \right) \right\}
\]
\[
(o^-_rs) = \max \left\{ 0, \left( o^-_rs \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{\left(1 - \varepsilon_{rs} \right) \cdot z_{rs}}{o^-_rs} \right) \right\}
\]

g. For the total demand, update the dual variables
\[
(\psi^+) = \min \left\{ 0, \left( \psi^+ \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{\left(1 + \varepsilon \right) \cdot F}{T^n} \right) \right\}
\]
\[
(\psi^-) = \max \left\{ 0, \left( \psi^- \right)^{n-1} + \frac{1}{\theta} \ln \left( \frac{\left(1 - \varepsilon \right) \cdot F}{T^n} \right) \right\}
\]
2.2 Compute primal variables: path flows can be derived analytically as a function of path costs and dual variables (see Appendix B and Appendix C for the analytical derivation of path flow solution).

a. Compute path flows

\[
(f_{k}^{rs})^n = \exp \left( \sum_{a \in A} t_a \left( x_a^n \right) \delta_{kd}^{rs} + \sum_{a \in A} \left( u_a^{-} \right)^n + \left( u_a^{+} \right)^n \cdot \delta_{kd}^{rs} + \sum_{a \in U} \left( d_a \right)^n \delta_{kd}^{rs} + \left( \rho_r^{+} \right)^n + \left( \rho_r^{-} \right)^n + \left( \alpha_s^{-} \right)^n + \left( \alpha_s^{+} \right)^n + \left( \alpha_r^{-} \right)^n + \left( \alpha_r^{+} \right)^n + \left( \psi^{-} \right)^n + \left( \psi^{+} \right)^n + \sum_{m \in M_i} \sum_{i \in I} \sum_{a \in A_{IN_i}} \sum_{b \in A_{OUT_i}} \left( \tau_m^{i-} \right)^n + \left( \tau_m^{i+} \right)^n \cdot \delta_{km}^{rs} \delta_{kb}^{rs} \right) \right) \quad \forall k \in K_{rs}, \forall rs \in RS
\]

b. Compute link flows

\[
x_a^n = \sum_{rs \in RS} \sum_{k \in K_{rs}} \left( f_{k}^{rs} \right)^n \delta_{kd}^{rs}, \quad \forall a \in A
\]

c. Compute intersection turning movement flows

\[
t_m^{i,n} = \sum_{rs \in RS} \sum_{k \in K_{rs}} \sum_{a \in A_{IN_i}} \sum_{b \in A_{OUT_i}} \left( f_{k}^{rs} \right)^n \delta_{kd}^{rs} \delta_{kb}^{rs}, \quad \forall m \in M_i, i \in I
\]

d. Compute zonal production and attraction flows

\[
\begin{align*}
P^n_r &= \sum_{rs \in RS} \sum_{k \in K_{rs}} \left( f_{k}^{rs} \right)^n, \quad \forall r \in R \\
A^n_s &= \sum_{rs \in RS} \sum_{k \in K_{rs}} \left( f_{k}^{rs} \right)^n, \quad \forall s \in S
\end{align*}
\]

e. Compute O-D flows

\[
q^n_{rs} = \sum_{k \in K_{rs}} \left( f_{k}^{rs} \right)^n, \quad \forall rs \in RS
\]

f. Compute total demand

\[
T^n = \sum_{rs \in RS} \sum_{k \in K_{rs}} \left( f_{k}^{rs} \right)^n
\]
**Step 3. Convergence and divergence Test**

\[ \xi = \text{Max}\left\{ \left| u^+_a - (u^+_a)^{n-1} \right|, \left| u^-_a - (u^-_a)^{n-1} \right|, \left| d_a^n - (d_a)^{n-1} \right|, \left| r^+_m - (r^+_m)^{n-1} \right|, \left| \alpha^+_s - (\alpha^+_s)^{n-1} \right|, \left| o^-_{rs} - (o^-_{rs})^{n-1} \right|, \left| \psi^- - (\psi^-)^{n-1} \right| \right\} \]

If \( \xi \leq \eta_0 \), then terminate and go to the output step. If \( \eta_0 \leq \xi < \eta \), where \( \eta_0 \) is a convergence tolerance (e.g., \( 10^{-6} \)) and \( \eta \) is the upper limit (e.g., \( 10^6 \)) of change in dual variables, then set all parameters of the next iteration equal to those of the current iteration, set Inner\((n) = \text{Inner}(n+1)\), and go to step 2. If \( \xi \geq \eta \) (detecting inner loop divergence), then set all parameters of the next iteration equal to those of the current iteration, go to the outer loop divergence step. In the outer loop divergence step, if \( |K|^{\text{outer}(n-1)} = |K|^{\text{outer}(n)} \) and \( \xi^{\text{outer}(n-1)} = \xi^{\text{outer}(n)} \) then terminate and return solution divergence (i.e., no solution can satisfy all the constraints); otherwise, set outer\((n) = \text{outer}(n+1)\), and go to the column generation step.

In the above procedure, we provide only the final form of the adjustment equations for different types of constraint (e.g., observed links, unobserved links, observed intersections, target O-D flows, etc.). The detailed derivations of the adjustment equations can be found in Chen et al. (2009, 2010b), and convergence of the iterative balancing scheme is discussed in details in Bell et al. (1997) and Bell and Iida (1997).
5.3.2 Column generation

The above iterative balancing scheme assumes that a working path set is given. For large networks, it is not practical to enumerate a working path set in advance since the number of possible paths grows exponentially with respect to network size. To circumvent path enumeration, a column (or path) generation procedure can be augmented to the iterative balancing scheme. Basically, the algorithm introduces an outer loop (or iteration) to iteratively generate paths to be added to the working path set as needed to replicate the observed interval constraints (e.g., link counts, turning movement counts, selected prior O-D flows, etc.), and to account for the capacity restraints for the unobserved links as well as the congestion effects, while the iterative balancing scheme iteratively adjusts the primal variables (e.g., path flows, link flows, intersection turning movement flows, O-D flows, etc.) and the dual variables in the inner loop for a given working path set from the outer loop. Note that the working path set is generated by a column generation scheme (or a shortest path algorithm) using the generalized link costs, which is based not only on the link costs but also on the dual variables from the active side constraints. The dual variables force the column generation scheme to generate paths that satisfy the side constraints. For additional discussions on the issue of using the generalized link costs to generate paths, refer to Bell et al. (1997) and Chen et al. (2009, 2010b).

5.3.3 Output generation from path flows

The unique path-flow solution from PFE makes it possible to derive useful information at different spatial levels. Given the path-flow solution, the following information at different spatial levels can be derived:
- Total demand: the sum of all path flows from all O-D pairs gives the total demand utilizing the network.
- Zonal production: the sum of all path flows emanating from a given origin gives the zonal production.
- Zonal attraction: the sum of all path flows terminating at a given destination gives the zonal attraction.
- O-D flow: the sum of all path flows connecting that O-D pair gives the O-D flow.
- Intersection turning movement flow: the sum of all path flows passing through that intersection turning movement gives the intersection turning movement flow.
- Link flow: the sum of all path flows passing through a given link gives the link flow.

Remark: Note that the introduction of error bounds in constraints enhances the flexibility of PFE by allowing the user to incorporate local knowledge about the network conditions to the estimation process. More reliable information (e.g., intersection turning movement flows, link flows, etc.) will constrain the estimation to be within a smaller tolerance, whereas less reliable information will allow for a larger deviation. However, the error bounds should be set up judiciously. Otherwise, it may lead to solution non-existence if the error bounds are specified too tight, or a biased solution (i.e., underestimate the network flow) if the error bounds are specified too loose. Alternatively, the PFE can be reformulated by using the norm approximation method to endogenously determine appropriate error bounds. Chen et al. (2009) adopted three norm approximation criteria (i.e., $L_\infty$-norm, $L_1$-norm, and $L_2$-norm) to formulate three $L_p$-PFE models for estimating consistent path flows and O-D flows that simultaneously minimize the deviation between the estimated and
observed link volumes. See Chen et al. (2009, 2010b) for the full details of the $L_p$-PFE formulations and their applications to the O-D estimation problem with traffic counts. In the current research, we use the PFE formulation given by Bell et al. (1997) and Chen et al. (2005) by allowing the user the flexibility of incorporating local knowledge of network conditions. This means the estimation results must satisfy the error bounds specified by the user if there exists a solution. Otherwise, the user needs to adjust the error bounds to find a solution.

5.3.4 Visual PFE software

Visual PFE is an integrated software suite that combines the Path Flow Estimator (PFE) with other software components to facilitate the base year estimation and future year forecasting, using a user-friendly Graphical User Interfaces (GUI) for inputting data and outputting results in tabular and map formats. It is built to be an open source standalone software using an open source programmable Geographical Information Systems (GIS) software (MapWindow, 2011). MapWindow is a mapping tool, a GIS modeling system, and GIS application programming interface all in one convenient redistributed open source form. MapWindow was developed to address the needs for a GIS programming tool that could be used in engineering research and project software, without requiring end users to purchase a complete GIS system, or become GIS experts. With Visual PFE, users can:

- Run PFE with GUIs
- Convert the estimated O-D tables to Microsoft Excel Files
- Change the colors and zoom levels of the O-D table cells
- Interactively display and query O-D desire lines
- Interactively display and query paths between any pairs of O-D
• Convert the PFE outputs to GIS files
• Create thematic maps of network links and traffic analysis zones
• Generate diagnostic scatter plots
• Link the scatter plots to the network for identification of outliers
• Change network link attributes and export the network back to the PFE format for another estimation
• Create and edit PFE networks as text files
• Compare different scenarios

For additional details, readers are referred to Zhang et al. (2010) for a complete description of the Visual PFE software.

5.4 Case studies

For proof of concept, two case studies were conducted using two communities in northern California. The first case study is a small community in the city of St. Helena, where the city's planning department has neither an existing travel demand model nor the budget for a full model development. The second case study is in the city of Eureka, where there is a four-step travel demand forecasting model developed for the Humboldt County Association of Government that can be used for comparison.

5.4.1 Case study 1: City of St. Helena

The City of St. Helena is located in the wine-producing region of Napa Valley in California, approximately 65 miles north of San Francisco. St. Helena is a full-service city with a population of 6,006 (as of January 1, 2005) within an area of 4.0 square miles. The City's development pattern is relatively compact. Commercial development and wineries
concentrate along Highway 29 (Main Street) corridor and residential developments radiate out from Main Street (see Fig. 5.8).

Fig. 5.8 St. Helena network and locations of intersection turning movement counts

The primary planning goal of the City is to preserve the rural, small town quality and agricultural character. Nevertheless, in the past few years, the City has been faced with pressures to grow as demand for service and commercial activities are rapidly rising with the increased number of tourists to the Napa Valley every year. The pressure for regional growth has caused serious concerns in the community with regard to deteriorating traffic conditions and the small town atmosphere. Continuous growth in most of the City is not expected, because there is a substantial difficulty in expanding the public utility systems.
To cater to the need for development, the City designated an area within the city boundary as a specific plan area to carefully guide and support future development of properties within the area while maintaining the desired town characteristics. The specific plan proposes to relieve congestion on Highway 29 by extending a street (Oak Avenue) that runs parallel to the highway. Oak Avenue is designated in the City's General Plan (City of St. Helena, 1993) as a collector street. The identification of potential traffic impacts and the decision-making for right-of-way preservation hinge upon a reliable forecast of design year traffic volumes on the extended street. However, the City's planning department has neither an existing travel model nor the budget for a full model development. The dilemma calls for innovative modeling approaches that can provide quick and reasonable responses with available resources.

5.4.1.1 Base year analysis

For the base year analysis, traffic counts collected since 2001 were retrieved for application of the proposed procedure. Link volumes collected during the evening peak hour, the time of day when traffic congestion on Highway 29 presents a serious issue, were assembled and a network with 28 TAZs was coded with the observed link volumes. The network contains 113 links and 54 of the links do not have traffic count data. Turn penalties, based on the actual traffic conditions, were also applied in the network such that the shortest paths among TAZ centroids replicate actual travel patterns in the area. The standard Bureau of Public Road (BRP) function is adopted as the link travel time function. Turning movement counts at two intersections (Fulton/Main and Mitchell Dr./Oak Ave.) were considered to assist the estimation (used as constraints in addition to link observation constraints). Fig. 5.8 depicts the St. Helena network and locations of two intersections.
(Fulton/Main and Mitchell Dr./Oak Ave.) of which turning movement counts were considered to assist the estimation. The error bounds for link counts were specified as ±5% to ±15% to ensure solution existence, while the error bounds for intersection turning movement counts were set at ±5%. In other words, we assume the intersection turning movement counts are more reliable than the link counts.

Note that there could be many possible O-D demand patterns that can match the observed link counts and intersection turning movement counts. The key in the base year estimation is to make sure the major flows on Main Street (Route 29) are captured within the model. This was done by using majority of the counts on the Main Street and a through trip survey. The though trips on Route 29 were used as selected target O-D demand constraints within ±5% to constrain the base year PFE formulation. The PFE formulation used for the base year estimation is a special case of the base year PFE formulation with \( \theta = 0.5 \) given in Section 5.2.1. It consists of observed traffic count constraints, capacity constraints for unobserved links, and selected target O-D demand constraints (i.e., through trips on Main street). The detailed derivation of analytical path-flow solution is given in Appendix D.

Fig. 5.9 shows the scatter plots of observed and estimated link flows and turning movement flows, respectively. To measure the accuracy of estimated flows, the \( R^2 \) value is adopted. As can be seen in the figure, both estimated link flows and intersection turning movement flows closely match the observed values. Specifically, the estimated link flows are within ±15% errors of the observed values with an \( R^2 \) value of 0.993, while the intersection turning movement flows are within ±5% errors of the observed values with an \( R^2 \) value of 0.998.
Fig. 5.9 Comparison between observed and estimated flows

Fig. 5.10 shows the estimated results in terms of link flows, production flows, and attraction flows to assess the overall flow pattern in the city.

Fig. 5.10 Base year flow pattern estimates
We observe that most flows are concentrated on Main Street, which agrees with the traffic condition in the city. Overall, the results seem valid as shown by the match between observed and estimated flows in Fig. 5.9 and a reasonable check of the network flow pattern in Fig. 5.10.

5.4.1.2 Future year analysis

Since future traffic counts are not available, both future land use information and the base year O-D trip table are used to predict future network traffic conditions. Future land use data are generally available from the General Plan of each city. For this study, we obtained the land use data from the General Plan 2030 of the City of St. Helena (City of St. Helena, 2010). When land use changes are expected to impact trip production and attraction in the study area of future years, the additional number of trips can be estimated with the ITE trip generation rates to convert various land uses to trip productions and attractions. In general, the land use characteristics in St. Helena can be classified into four categories: residential areas, commercial and mixed use areas, business and industrial areas, and community and natural resource areas. Since no other potential development was expected elsewhere in the city, the additional trips produced from and attracted to TAZs in the specific plan area were added to the base year production and attraction to obtain the future production and attraction flows. Trip balancing was subsequently performed to ensure that the total production equals to the total attraction. Fig. 5.11 graphically depicts the future year zonal production and attraction flows for the St. Helena network.
According to the specific plan proposed by city's general plan, two scenarios were set up as follows.

1. No build scenario

2. Build scenario (extending Oak Avenue)

The St. Helena planning network was revised according to the conceptual street design proposed in the specific plan (i.e., extending Oak Avenue). The error bounds for the future zonal production and attraction flows and target O-D demands were specified at ±5%. The future PFE was performed to forecast traffic volumes based on the land use and the two scenarios considered above. The forecasting process is essentially an O-D estimation constrained on the future trip production and attraction flows and the scaled
future O-D trip table to preserve the spatial distribution patterns exhibited in the base year O-D trip table. The PFE formulation used for the future year prediction is a special case of the future year PFE formulation given in Section 5.2.4. It consists of future zonal production and attraction constraints and future target O-D demand constraints. The formulation, analytical solution, and behavioral explanation are given in Appendix E.

Since there are no observed future traffic conditions, the forecast results can only be assessed based on its reasonableness. Fig. 5.12 shows the forecast results in terms of link flows and volume to capacity (V/C) ratios for the two scenarios. As expected, congestion level on Main Street is slightly higher in scenario 1 (no build scenario) compared to those in scenario 2 (build scenario). It appears that by extending Oak Avenue parallel to Main Street can help to alleviate some of the congestion on Main Street, particularly on the segments between Mitchell Dr. and Mills Lane.

Fig. 5.12 Forecast link flows and V/C ratios for two scenarios
To further assess the build scenario (extending Oak Avenue), a level of service (LOS) analysis is performed as a post analysis to justify the build scenario. Maintaining an acceptable LOS is one of the most important controls in roadway design. According to the Fourth Edition of A policy on Geometric Design of Highways and Streets by the American Association of State Highway and Transportation Officials (AASHTO, 2001), the number of lanes to be provided on urban collector streets with high traffic volumes should be determined from capacity analysis. This analysis should consider both intersections and mid-block locations, when appropriate, in assessing the ability of a proposed design to provide the desired LOS. The standard methodology in the Highway Capacity Manual (HCM, 2010) provided in Table 5.3 was used to determine the LOS on various segments on Main Street. To facilitate the analysis of average travel speed on the design street, forecasts of link volumes were derived from the future year PFE. The existing segments on Main Street can be classified as Class III urban streets with a speed limit of between 30 mph and 35 mph.

Table 5.3: Urban street level of service

<table>
<thead>
<tr>
<th>Urban Street Class</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range of Free Flow Speeds (mph)</td>
<td>45 to 55</td>
<td>35 to 45</td>
<td>30 to 35</td>
<td>25 to 35</td>
</tr>
<tr>
<td>Typical Free Flow Speed (mph)</td>
<td>50</td>
<td>40</td>
<td>33</td>
<td>30</td>
</tr>
<tr>
<td>Level of Service</td>
<td>Average Travel Speed (mph)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>&gt;42</td>
<td>&gt;35</td>
<td>&gt;30</td>
<td>&gt;25</td>
</tr>
<tr>
<td>B</td>
<td>&gt;34</td>
<td>&gt;28</td>
<td>&gt;24</td>
<td>&gt;19</td>
</tr>
<tr>
<td>C</td>
<td>&gt;27</td>
<td>&gt;22</td>
<td>&gt;18</td>
<td>&gt;13</td>
</tr>
<tr>
<td>D</td>
<td>&gt;21</td>
<td>&gt;17</td>
<td>&gt;14</td>
<td>&gt;9</td>
</tr>
<tr>
<td>E</td>
<td>&gt;16</td>
<td>&gt;13</td>
<td>&gt;10</td>
<td>&gt;7</td>
</tr>
<tr>
<td>F</td>
<td>≤16</td>
<td>≤13</td>
<td>≤10</td>
<td>≤7</td>
</tr>
</tbody>
</table>

Source: Highway Capacity Manual 2000
Table 5.4 summarizes the analysis results. For the no build scenario, some segments on Main Street are not acceptable with a LOS of D and E. However, with the build scenario of extending Oak Avenue, all segments become acceptable with a LOS of C. The LOS results can be used to assist the decision-making process of extending Oak Avenue parallel to Main Street to alleviate congestion on Main Street.

Table 5.4: LOS analysis for Main Street

<table>
<thead>
<tr>
<th>Street Segment (Scenario 1)</th>
<th>Direction</th>
<th>Volume (vph)</th>
<th>V/C Ratio</th>
<th>Speed (mph)</th>
<th>LOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vintage Ave - Dowdell Lane</td>
<td>Northbound</td>
<td>1,186</td>
<td>0.70</td>
<td>22</td>
<td>C</td>
</tr>
<tr>
<td>Dowdell Lane - Vintage Ave</td>
<td>Southbound</td>
<td>1,258</td>
<td>0.74</td>
<td>21</td>
<td>C</td>
</tr>
<tr>
<td>Dowdell Lane –Mills Lane</td>
<td>Northbound</td>
<td>1,294</td>
<td>0.76</td>
<td>21</td>
<td>C</td>
</tr>
<tr>
<td>Mills Lane - Dowdell Lane</td>
<td>Southbound</td>
<td>1,512</td>
<td>0.89</td>
<td>18</td>
<td>C</td>
</tr>
<tr>
<td>Mills Lane - Charter Oak Ave</td>
<td>Northbound</td>
<td>1,453</td>
<td>0.97</td>
<td>16</td>
<td>D</td>
</tr>
<tr>
<td>Charter Oak Ave - Mills Lane</td>
<td>Southbound</td>
<td>1,564</td>
<td>1.04</td>
<td>14</td>
<td>D</td>
</tr>
<tr>
<td>Charter Oak Ave - Spring St.</td>
<td>Northbound</td>
<td>1,391</td>
<td>0.93</td>
<td>17</td>
<td>D</td>
</tr>
<tr>
<td>Spring St.- Charter Oak Ave</td>
<td>Southbound</td>
<td>1,590</td>
<td>1.06</td>
<td>13</td>
<td>E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Street Segment (Scenario 2)</th>
<th>Direction</th>
<th>Volume (vph)</th>
<th>V/C Ratio</th>
<th>Speed (mph)</th>
<th>LOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vintage Ave - Dowdell Lane</td>
<td>Northbound</td>
<td>1,199</td>
<td>0.71</td>
<td>22</td>
<td>C</td>
</tr>
<tr>
<td>Dowdell Lane - Vintage Ave</td>
<td>Southbound</td>
<td>1,275</td>
<td>0.75</td>
<td>21</td>
<td>C</td>
</tr>
<tr>
<td>Dowdell Lane –Mills Lane</td>
<td>Northbound</td>
<td>1,308</td>
<td>0.77</td>
<td>21</td>
<td>C</td>
</tr>
<tr>
<td>Mills Lane - Dowdell Lane</td>
<td>Southbound</td>
<td>1,532</td>
<td>0.90</td>
<td>18</td>
<td>C</td>
</tr>
<tr>
<td>Mills Lane - Charter Oak Ave</td>
<td>Northbound</td>
<td>1,241</td>
<td>0.83</td>
<td>20</td>
<td>C</td>
</tr>
<tr>
<td>Charter Oak Ave- Mills Lane</td>
<td>Southbound</td>
<td>1,252</td>
<td>0.83</td>
<td>20</td>
<td>C</td>
</tr>
<tr>
<td>Charter Oak Ave- Spring St.</td>
<td>Northbound</td>
<td>1,142</td>
<td>0.76</td>
<td>21</td>
<td>C</td>
</tr>
<tr>
<td>Spring St.- Charter Oak Ave</td>
<td>Southbound</td>
<td>1,252</td>
<td>0.83</td>
<td>20</td>
<td>C</td>
</tr>
</tbody>
</table>

Volume = forecast of number of vehicles during the PM peak hour; Time = travel time (including delays) in seconds; Speed = segment speed in miles per hour; LOS = level of service for urban streets.
5.4.2 Case study 2: City of Eureka

The City of Eureka is located in the Humboldt County in California, approximately 270 miles north of San Francisco. Eureka serves as a principal city for the County with a population of 26,097 (as of 2007). U.S. Route 101 extends north and south through the City. Harris St., Myrtle Ave, H St., and I St. are the main roads in the City of Eureka (see Fig. 5.13). There is a four-step travel demand forecasting model developed by DKS Associates (2006) for the Humboldt County Association of Government. For the trip generation step, a cross classification model is used with planning data including dwelling unit (single family and multi-family) and employment data for each traffic analysis zone (TAZ). In the trip distribution step, a doubly-constrained gravity model is adopted, while a probit-based stochastic user equilibrium (SUE) assignment model is adopted to conduct the traffic assignment step. Note that the modal split step is not considered in the four-step model. To demonstrate how to apply the base year and future year PFE to the Eureka network, comparisons between the results obtained from the four-step model and the PFE models were performed.

5.4.2.1 Input data preparation

The input data preparation for the Eureka network involves a number of steps: (a) subarea analysis, (b) zonal aggregation, and (c) land use and socio-economic data. The first step was to perform subarea analysis by extracting the City of Eureka from the Humboldt County network using the Greater Eureka Area Travel Model (GEATM) developed by DKS Associates (2006). In the second step, zonal aggregation was performed to eliminate zones with very low demand or no demand. Finally, zonal production and attraction flows were estimated from the land use and socio-economic data. Fig. 5.13 shows the extracted
network in the City of Eureka with locations of traffic counts and TAZ after subarea analysis and zonal aggregation. The extracted network consists of 99 zones, 6,335 nodes, 6,344 links, and 9,330 O-D pairs. The BRP function with link-specific parameters obtained from the GEATM model is adopted as the link travel time function. The number of observed counts in the City of Eureka is 136 or about 2% of the 6,344 links in the extracted network. As can be seen, the number of observed counts is not sufficient to estimate a reliable O-D trip table. Many of the O-D pairs in the extracted network simply contain no information that can be used to estimate the O-D demand.

Fig. 5.13 Traffic counts and aggregated TAZ locations in the Eureka network
Since traffic counts are not sufficient (about 2%) to estimate a reliable O-D trip table for the base year, land use and socio-economic data were used to increase the observability for estimating the base year O-D trip table as well as for predicting future network traffic conditions. Land use and socio-economic data are generally available from the General Plan of each city. For this case study, we obtained the data (i.e., the variables used to develop the trip generation models) from the GEATM model. Fig. 5.14 presents a procedure using ITE trip rates to convert land use data to zonal production and attraction trips for the aggregated TAZs. As shown in Fig. 5.14, the first step is to extract the land use and socio-economic data from the GEATM model.

Fig. 5.14 Procedure for converting land use data to zonal production and attraction trips

Fig. 5.15 shows the distributions of land use and socioeconomic data obtained from GEATM in terms of employment and dwelling units for the base year 2005 and future year 2030. The second step is to use appropriate ITE trip rates to calculate the zonal production and attraction flows. Table 5.5 provides an illustration of this step using the aggregated TAZ 320 (i.e., consists of the original TAZs 319, 320, 321, and 324 in the GEATM). The
final step is to perform trip balancing to ensure the total production flows equal the total attraction flows.

Fig. 5.15 Land use and socioeconomic data in terms of employment and dwelling units for years 2005 and 2030

Fig. 5.16 shows the distributions of the estimated zonal production and attraction flows for the base year 2005 and future year 2030. For this case study, the traffic counts in the base year were specified according to the FHWA guidelines given in Table 5.1 [i.e., freeway (±7%), major arterial (±10%), minor arterial (±15%), and collector (±25%)]. As for the zonal production and attraction flows and target O-D demands, the error bounds for both base year and future year were specified as ±10%. The dispersion parameter is set at 0.5 for the PFE formulation.
Table 5.5: Example of using ITE trip rates to determine zonal production and attraction for aggregated TAZ 320

<table>
<thead>
<tr>
<th>Aggregated TAZ</th>
<th>TAZ</th>
<th>Parcel ID</th>
<th>Unit</th>
<th>Number</th>
<th>Trip Rate</th>
<th>Attraction</th>
<th>Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>319</td>
<td>1</td>
<td>Employment</td>
<td>217</td>
<td>0.45</td>
<td>83</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Dwelling units</td>
<td>11</td>
<td>0.44</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sum</td>
<td></td>
<td></td>
<td>84</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>320</td>
<td>1</td>
<td>Employment</td>
<td>135</td>
<td>0.45</td>
<td>52</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Dwelling units</td>
<td>26</td>
<td>0.44</td>
<td>2</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sum</td>
<td></td>
<td></td>
<td>54</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>320</td>
<td>321</td>
<td>Employment</td>
<td>128</td>
<td>0.45</td>
<td>49</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Dwelling units</td>
<td>0</td>
<td>0.44</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sum</td>
<td></td>
<td></td>
<td>49</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>324</td>
<td>1</td>
<td>Employment</td>
<td>103</td>
<td>0.45</td>
<td>39</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Dwelling units</td>
<td>0</td>
<td>0.44</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sum</td>
<td></td>
<td></td>
<td>39</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sum</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sum</td>
<td></td>
<td></td>
<td>226</td>
<td>53</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5.16 Estimated zonal production and attraction flows for years 2005 and 2030
5.4.2.2 Base year analysis

This section compares the results of the base year PFE to those of the four-step model obtained from the GEATM model: (a) zonal production and attraction flows, (b) O-D flows, and (c) link flows.

Zonal Production and Attraction Flow Comparison for the Base Year

Fig. 5.17 provides a graphical visualization of the estimated zonal production and attraction flows obtained from both models. Overall, both models produce similar trip generation patterns with higher zonal production and attraction flows concentrating on the major highways. The R² values between the results of the two models are 0.87 and 0.84 for production and attraction, respectively. Using the base year PFE typically underestimates the zonal production and attraction flows on the major highways while overestimates some of the internal zonal production and attraction flows within the City compared to those of the four-step model. Since traffic counts are not sufficient to estimate a reliable O-D trip table, the production and attraction flows based on the average ITE trip rate values surveyed from many cities of various sizes in North America were also used as constraints in the base year PFE. Therefore, these values may not be able to capture the specific characteristics of the local study area compared to the trip generation (i.e., cross classification) model developed by GEATM.
In the trip distribution step, a histogram shown in Fig. 5.18 is used to compare the O-D flows estimated from the base year PFE model and the four-step model. In the figure, it depicts the percentage of O-D pairs by the absolute O-D flow difference between the base year PFE model and the four-step model. More than 70% of the O-D pairs are
estimated within ±1 trip, and more than 97% of the O-D pairs are within ±10 trips. The overall match between the two models depicted by the correlation coefficient is 0.83.

![Graph](image)

**Fig. 5.18** Percentage of O-D pairs by the absolute O-D flow difference for the base year

**Link Flow Comparison for the Base Year**

Fig. 5.19(a) depicts the estimated link flow results for the four-step model and the base year PFE model. As can be seen, both models produce similar traffic patterns in terms of link flows and V/C ratios (i.e., higher congestion on the major highways). The four-step model estimated a slightly higher congestion level on U.S. 101 compared to those estimated by the base year PFE model. This discrepancy is due to the side constraints [i.e., both traffic counts on the major highways in Fig. 5.13 and the zonal production and attraction flows in Fig. 5.16(a)] used in the logit-based PFE model as opposed to the probit-based SUE assignment results of the four-step model. Fig. 5.19(b) shows the scatter plots of observed and estimated link flows by roadway functional class. It shows that the base year PFE can replicate the observed link counts better than the four-step model. For the estimated flows
by the base year PFE, the errors are within 20%, while some of the link flows estimated by the four-step model have more than 20% error. The $R^2$ values are 0.849 for the four-step model and 0.987 for the base year PFE model.

Fig. 5.19 Link flow comparison for the base year

5.4.2.3 Future year analysis

Similar to the base year analysis, the forecast results from the future year PFE with those obtained from the GEATM model were compared for year 2030. Predicted zonal
production and attraction flows shown in Fig. 5.16(b) were used with selected (important) O-D pairs from the base year trip table to constrain the relationship of travel impedance and trip interchange between each O-D pair in the future year PFE forecasting process.

Zonal Production and Attraction Flows Comparison for the Future Year

Fig. 5.20(a) and Fig. 5.20(b) provide a graphical visualization of the forecast zonal production and attraction flows for year 2030. Visually both models produce similar trip generation patterns for the future year. Fig. 5.20(c) and Fig. 5.20(d) compare the forecast production and attraction flows of the two models using the scatter plots. Recall that the core inputs to the future PFE model for predicting zonal production and attraction flows are the future planning data (i.e., dwelling unit for trip production and employment for trip attraction) and the scaled base year O-D trip table, and both Fig. 5.20(a) and Fig. 5.20(b) show that the predicted zonal flows by ITE trip rates generally agree with those predicted flows by the four-step model, with a slight advantage for production flows over attraction flows. The discrepancy mainly occurs for the low attraction flows (i.e., fewer than 50 trips) in Fig. 5.20(d), which is understandable given that PFE uses the average ITE trip rates to convert the forecast land use and socioeconomic data to trip productions and trip attractions. The $R^2$ values for the predicted production and attraction flows are 0.97 and 0.93, which indicate a good correlation between the four-step model and the PFE model.
O-D Flow Comparison for the Future Year

Fig. 5.21 shows the percentage of O-D pairs by the absolute O-D flow difference between the future year PFE model and the four-step model. More than 87% of the O-D pairs are predicted within ±1 trip, and more than 95% of the O-D pairs are within ±10 trips.
Compared to the base year, the overall match between the two models is much better with a $R^2$ value of 0.91.

Fig. 5.21 Percentage of O-D pairs by the absolute O-D flow difference for the future year

**Link Flow Comparison for the Future Year**

Fig. 5.22 provides a graphical representation of the forecast link flows and V/C ratios obtained from the four-step model and the future year PFE model. From the figure, we can observe that both models produce similar traffic patterns (i.e., congestion on U.S. 101), indicating that both models are capable of capturing the overall congestion pattern in the Eureka network. However, the congestion magnitudes differ slightly between the two models. The future year PFE model appears to forecast a slightly higher congestion levels than those of the four-step model (e.g., some segments on U.S. 101, Myrtle Ave., H & I St., and Harris St.). This discrepancy is due to a higher demand estimate from future land use and socio-economic data (i.e., the total number of trips in PFE is about 1,400 trips (or about 10% of 15,000 trips) higher than the total number of trips predicted by the four-step
model). Fig. 5.23 shows the percentage of links by the forecast V/C ratios for the two models. The results are similar to those shown in Fig. 5.22 with more congested links (i.e., V/C > 1) from the PFE model.

![Fig. 5.22 Forecast link flows and V/C ratios for the future year](image1)

![Fig. 5.23 Percentage of links by the forecast V/C ratio for the future year](image2)
5.5 Conclusions

In this chapter, we presented a simplified methodology for planning applications in small communities by adapting the path flow estimator (PFE). Two versions of PFE were developed: a base year PFE for estimating the O-D trip table using current traffic counts, target trip table, and trip production and trip attraction as constraints, and a future year PFE for predicting network traffic conditions using future trip production and trip attraction and scaled base year O-D trip table to match future total demand as constraints. To show proof of concept, two case studies were conducted using two small communities in northern California.

Some of the specific findings include:

- Based on the experience gained from the two case studies, the proposed simplified planning tool using Visual PFE is applicable for small communities with limited resources.

- In the absence of travel survey data, the proposed method uses similar data (traffic count data, land use and socio-economic data, and standards from ITE and HCM) as a four-step model for model development. The link flow estimates from Visual PFE generally match the observed data with a more satisfactory error bound than the link flow estimates from the four-step model. The results on the spatial distribution of trip making (i.e., O-D trip table) between the two models are satisfactory, while the results on trip generation from Visual PFE are different for some zones due to the use of ITE trip rates to determine the production and attraction flows. These average ITE values may not able to capture the specific
characteristics of the local study area compared to those determined by the calibrated trip generation models in the four-step model.

- Since both ITE trip generation rates and HCM are utilized in the modeling process, the analysis scope and results are consistent with those of common traffic impact studies and other short-range, localized transportation improvement programs.

Several directions for future research are worth noting. These include improving existing modeling capabilities and adding new features to the Visual PFE tool. One possibility is to enhance the trip generation results by incorporating the cross-classification trip rates to better capture the local characteristics of the study area. Another possibility is to consider advanced route choice models, such as the C-logit model (Cascetta et al., 1996; Zhou et al., 2012), path size logit model (Ben-Akiva and Bierlaire, 1999; Chen et al., 2012), cross-nested logit model (Bekhor and Prashker, 1999), paired combinatorial logit model (Bekhor and Prashker, 1999; Pravinvongvuth and Chen, 2005; Chen et al., 2014), multinomial weibit model (Kitthamkesorn and Chen, 2014), and path-size weibit model (Kitthamkesorn and Chen, 2013), to model the route overlapping and heterogeneous perception variance problems in the PFE formulation. New features include: (a) adding a confidence interval estimation module (Chootinan and Chen, 2011) for assessing the quality of the PFE estimates and measurement errors from field data and planning data, (b) adding a mode choice module to assess the impacts of a transit system and/or a non-motorized bicycle mode, (c) adding a sensor location module (Chootinan et al., 2005b; Chen et al., 2007) to help users with where to collect additional data for a given budget (i.e., how many additional counts is needed to provide a 90% coverage and where to collect these additional counts?), (d) extending the Visual PFE tool to include travel time data
collected from automatic vehicle identification (AVI) readers (Chen et al., 2004, 2010a) or other emerging sensors to improve the PFE estimation, and (e) enhancing the proposed approach such that the impacts of long-range, area-wide growth can be modeled within the same framework. Finally, more case studies should be conducted to further validate the usefulness of the simplified PFE planning tool and its applicability across different test areas.

REFERENCES


CHAPTER 6
CONCLUDING REMARKS

6.1 Summary

This chapter summarizes the main results and contributions of this dissertation. Some directions for future research are also highlighted at the end of this chapter.

Chapter 2 provided a literature review of the path flow estimator (PFE). The review includes the concept of the PFE, its mathematical programming formulation, solution algorithm, and applications for transportation planning and operation purposes. The literature review revealed that the PFE is a flexible tool that has the potential to be further developed as an alternative to the current practice (the four-step model). With further development, the PFE could address the deficiencies in modeling bicycles for non-motorized transportation planning applications, for multiple vehicle types (particularly different types of trucks) sharing the same roadway space for freight planning application, and for a simplified travel demand forecasting procedure for small community planning applications.

Chapter 3 provided a two-stage approach for estimating/predicting bicycle volumes on a transportation network given a bicycle origin-destination (O-D) matrix. The first stage used key criteria (e.g., distance related attributes and safety related attributes) that affect bicyclists’ route choice decisions to generate a set of non-dominated (or efficient) paths. Then, the second stage used these efficient paths to determine the flow allocations to the transportation network. A case study of the city of Winnipeg, Canada yielded numerical results that revealed that the two-stage approach is viable in testing the bicycle traffic...
assignment model in real networks. Extension to multiple user classes and multiple criteria provides further flexibility and increases potential for the approach to be used in practice.

Chapter 4 presented a multi-class traffic assignment model for freight planning applications. Different modeling issues were considered to enhance the realism of the multi-class traffic assignment model. These enhancements include asymmetric interactions among different vehicle types (with emphasis on different truck types), a path-size logit model for accounting random perceptions of network conditions with explicit consideration of route overlapping, and various traffic restrictions imposed either individually (e.g., lane restriction, height restriction, and weight restriction) or together (e.g., link capacity constraint) on multiple vehicle types in a transportation network. Model formulation and customized solution algorithm were developed for freight planning purposes. Numerical results revealed that ignoring asymmetric vehicle interactions, route overlapping, and traffic restraints could have a significant impact on the network flow allocations (particularly with medium and heavy trucks), and that the customized solution algorithm would be a practical approach for solving the multi-class traffic assignment problem with realistic modeling considerations.

Chapter 5 developed a simplified travel demand forecasting framework that exploits the O-D estimation capability of the path flow estimator (PFE) to model and forecast network traffic for planning applications in small communities, where limited planning resources hinder the development and application of a full four-step model. Two versions of the PFE were developed to address the specific transportation planning issues and needs of small communities. The base year PFE was used for estimating the current network traffic conditions using field data and planning data (if available) as constraints.
The future year PFE was used for predicting future network traffic conditions using forecast planning data and the scaled base year origin-destination trip table as constraints. Solution algorithms were also developed to solve the two PFE models and were integrated into a GIS-based software called Visual PFE. To show proof of concept, two case studies were conducted using two small communities in northern California. The results showed that the simplified procedure is applicable for small communities with limited resources.

6.2 Contribution of this dissertation

The contributions of this dissertation are as follows.

- The development of a two-stage approach for the bicycle traffic assignment problem, which not only accounts for multiple user classes by acknowledging that there are different types of cyclists with different levels of biking experience, but also accounts for relevant criteria that may affect each user classes’ behavior in route choice decisions. The two-stage approach can also be used as a stand-alone bicycle traffic assignment procedure given a bicycle network and a bicycle O-D demand matrix.

- The development of an enhanced multi-class traffic assignment model that accounts for asymmetric vehicle interactions, route overlapping, and traffic restraints for freight planning application. In addition, the customized path-based algorithm was effective for solving for solving the multi-class traffic assignment problem with different modeling considerations.

- The development of a simplified travel demand methodology for small community planning to overcome the limited resources for developing a transportation planning model that not only meets the planning requirements from the Federal
regulations, but also addresses the specific transportation planning issues and needs of small communities. The Visual PFE tool, fully integrated with a GIS software, has the potential to help the planning applications of many small communities in the United States.

6.3 Directions for future research

Several directions for future research are worth noting. These include improving existing modeling capabilities and adding new features to the three transportation planning applications developed in this dissertation.

6.3.1 Bicycle network analysis tool for non-motorized transportation planning

Several improvements to the bicycle network analysis tool can be considered as follows:

- Consider other surrogate measures, such as the bicycle compatibility index (Harkey et al., 1998) or the stress indicator (Mekuria et al., 2012), for modeling cyclists’ perception of safety (or risk) on different bicycle facility types, and examine its impact on efficient route generation and flow allocations to the bicycle network.

- Consider other surrogate measures, such as nitrogen oxides (NO$_x$), volatile organic compounds (VOC), and particulate matter (PM) of various sizes and composition, for modeling traffic-related air pollutants linked to health risks on different bicycle facility types (Pankow et al., 2014), and examine its impact on efficient route generation and flow allocations to the bicycle network,
• Consider additional criteria such as route cognition using the concept of space syntax (Raford et al., 2007) from the field of urban planning and design to model bicycle route choice behavior.

• Consider the effect of congestion (i.e., link travel times are dependent of flows) in the bicycle traffic assignment procedure.

• Conduct more numerical tests with different network topologies with different bicycle facilities and cyclist characteristics.

• Develop a two-stage procedure for estimating bicycle O-D demand using observed bicycle count data and other available planning data.

6.3.2 Multi-Class network analysis tool for freight transportation planning

It should be pointed out that the multi-class demand matrices used for the multi-class traffic assignment model in Chapter 4 were assumed available from the Emme software (INRO Consultants, 2013). Particularly, the medium and heavy truck demand matrices were assumed to be disaggregated from the truck demands by 70% and 30%, respectively. In practice, the O-D demand matrices for passenger car and different truck classes need to be estimated simultaneously due to the asymmetric interactions between auto traffic and truck traffic and different vehicle restrictions applied to different vehicle classes. Hence, future research will focus on developing a multi-class origin-destination trip estimation method for multiple vehicle classes with multiple data sources. In addition, different physical and environmental side constraints (e.g., Chen et al., 2011) can be considered to model different physical configurations (e.g., lane drop, merging, nodal capacity, etc.), traffic control policies, and environmental restraints (e.g., emission and noise) to improve the realism of multi-class traffic assignment models.
6.3.3 Simplified PFE planning tool for small community transportation planning

Several enhancements to the simplified travel demand forecasting procedure can be considered as follows.

- Enhance the trip generation results by incorporating the cross-classification trip rates to better capture the local characteristics of the study area.
- Consider advanced route choice models, such as the path-size weibit model (Kitthamkesorn and Chen, 2013), to model the route overlapping and heterogeneous perception variance problems in the PFE formulation.
- Add a mode choice module to assess the impacts of a transit system and/or a non-motorized bicycle mode.
- Add a sensor location module (Chootinan et al., 2005; Chen et al., 2007) to help users with identifying where to collect additional data for a given budget (i.e., how many additional counts is needed to provide a 90% coverage and where to collect these additional counts?).
- Enhance the simplified procedure such that the impacts of long-range, area-wide growth can be modeled within the same framework.
- Conduct more case studies to further validate the usefulness of the simplified PFE planning tool and its applicability across different test areas.

REFERENCES


APPENDIX A

Appendix A provides the derivations of the adjustment equations for the vehicle restriction constraints, link capacity constraints, and O-D demand conservation constraints.

Handling vehicle restriction constraints

Consider the vehicle restriction constraints in Eq. (4.6) and the definitional constraints in Eq. (4.10). An adjustment factor \( \gamma_a^m \) associated with the dual variable \( u_a^m \) is created to ensure vehicle restriction (i.e., zero flow) on restricted link \( a \) for class \( m \). The adjustment factor is determined by substituting the analytical path-flow solution given by Eq. (4.21) into Eq. (4.6) and Eq. (4.10):

\[
\sum_{rs \in R} \sum_{k \in K_{rs}} PS_{km}^r \cdot \exp \left( -\theta^m \left( c_{km}^r + \sum_{a \in A} d_a \delta_{kam}^m + \sum_{a \in A_m} \left( u_a^m + \gamma_a^m \right) \delta_{kam}^m - \delta_{kam}^m \right) \right) \cdot \delta_{kam}^m = 0,
\forall a \in \hat{A}_m, m \in M \quad (A1)
\]

which is equivalent to

\[
\sum_{rs \in R} \sum_{k \in K_{rs}} f_{km}^r \cdot \delta_{kam}^m \exp \left( \theta^m \left( -\gamma_a^m \right) \right) = 0 \quad \Rightarrow \quad \gamma_a^m = -\frac{1}{\theta^m} \ln \left( \frac{0}{x_a^m} \right) = \infty, \quad \forall a \in \hat{A}_m, m \in M \quad (A2)
\]

In Eq. (A2), it should be noted that the adjustment factor \( \gamma_a^m \) is infinity, which implies the dual variable \( u_a^m \) is also infinity, in order to enforce the strict vehicle restriction of class \( m \) on link \( a \). However, for practical implementation, the adjustment factor is set at a sufficiently large value (e.g., \( 10^{10} \)).
Handling link capacity constraints

Consider the link capacity constraints in Eq. (4.5) and definitional constraints in Eq. (4.11). An adjustment factor \( \pi_a \) associated with the dual variables \( d_a \) to restrict the flow on link \( a \) within its link capacity. The adjustment factor is determined by substituting the analytical path-flow solution given by Eq. (4.21) into Eq. (4.5) and Eq. (4.11):

\[
\sum_{m \in M} PCE^m \cdot \sum_{rs \in RS} \sum_{k \in K^m_{rs}} PS_{km}^{rs} \cdot \exp \left( -\theta^m \cdot \left( c_{km}^{rs} + \sum_{a \in \bar{A}} (d_a + \pi_a) \delta_{kam}^{rs} + \sum_{a \in A} u^m_a \delta_{kam}^{rs} + \sigma_{rs}^{m} \right) \right) \cdot \delta_{kam}^{rs} = C_a, \quad \forall \ a \in A \quad (A3)
\]

which is equivalent to:

\[
\sum_{m \in M} PCE^m \cdot \sum_{rs \in RS} \sum_{k \in K^m_{rs}} f_{km}^{rs} \cdot \delta_{ka}^{rs} \cdot \exp \left( \theta^m \left( -\pi_a \right) \right) = C_a \\
\Rightarrow \sum_{m \in M} PCE^m_a \cdot x_a^m \cdot \exp \left( \theta^m \left( -\pi_a \right) \right) = C_a, \quad \forall \ a \in \bar{A} \quad (A4)
\]

To find a suitable \( \pi_a \) (i.e., an aggregate link variable) to adjust the individual vehicle class \( m \) sharing the same roadway space, a numerical method (e.g., Newton method) is adopted since there is no closed-form adjustment factor for computing \( \pi_a \). Note that the adjustment factors \( \pi_a \) is within the following interval if \( x_a \geq C_a \):

\[
-\frac{1}{\max \{ \theta \} \ln \left( \frac{C_a}{x_a} \right)} \leq \pi_a \leq -\frac{1}{\min \{ \theta \} \ln \left( \frac{C_a}{x_a} \right)}.
\]

Newton method

\[\text{While } \left( \left( \pi_a \right)^{n+1} - \left( \pi_a \right)^{n} \leq \varepsilon \right) \]

\[\left( \pi_a \right)^{n+1} = \left( \pi_a \right)^{n} + \frac{C_a - f \left( \pi_a \right)^{n}}{f' \left( \pi_a \right)^{n}}
\]

End while

where \( f \left( \pi_a \right) = \sum_{m \in M} PCE^m_a \cdot x_a^m \cdot \exp \left( \theta^m \left( -\pi_a \right) \right) \) and \( \left( \pi_a \right)^{0} = -\frac{1}{\max \{ \theta \} \ln \left( \frac{C_a}{x_a} \right)} \)
Handling O-D demand conservation constraints

Consider the demand conservation constraints in Eq. (4.9). An adjustment factor \((\psi_{rs}^m)\) associated with the dual variable \((o_{rs}^m)\) is used to preserve the demand of O-D pair \(rs\).

The adjustment factor is determined by substituting the analytical path-flow solution given by Eq. (4.21) into Eq. (4.9):

\[
\sum_{k \in K_{rs}^m} PS_{km}^r \cdot \exp \left( -\theta^m \cdot \left( e_{km}^r + \sum_{a \in A} d_{am}^r \delta_{nam}^r + \sum_{a \in A_m} u_{am}^m \delta_{nam}^r \right) - \left( o_{rs}^m + \psi_{rs}^m \right) \right) = q_{rs}^m, \quad \forall m \in M, \; rs \in R S
\]

which is equivalent to:

\[
\sum_{k \in K_{rs}^m} f_{mk}^r \exp(\theta^m \psi_{rs}^m) = q_{rs}^m \quad \Rightarrow \quad \psi_{rs}^m = \frac{1}{\theta^m} \ln \left( \frac{q_{rs}^m}{\sum_{k \in K_{rs}^m} f_{mk}^r} \right), \quad \forall m \in M, \; rs \in R S
\]
APPENDIX B

Appendix B provides the detailed derivations of the analytical path flow solution for the base year PFE formulation derived as a function of path costs and dual variables associated with the constraints.

Minimize:  
\[ Z = \frac{1}{\theta} \sum_{rs \in RS} \sum_{k \in K_{rs}} f_k^{rs} \left( \ln f_k^{rs} - 1 \right) + \sum_{a \in A} \int_0^{x_a} t_a(\omega) d\omega \]  \hspace{1cm} (B1)

subject to:

\[(1 - \varepsilon_a) \cdot v_a \leq x_a \leq (1 + \varepsilon_a) \cdot v_a, \quad \forall a \in \bar{A} \]  \hspace{1cm} (B2)

\[(1 - \varepsilon_m^i) \cdot g_{a_s}^i \leq t_s^i \leq (1 + \varepsilon_m^i) \cdot g_{a_s}^i, \quad \forall m \in M_i, i \in \bar{T} \]  \hspace{1cm} (B3)

\[(1 - \varepsilon_r) \cdot O_r \leq P_r \leq (1 + \varepsilon_r) \cdot O_r, \quad \forall r \in R \]  \hspace{1cm} (B4)

\[(1 - \varepsilon_s) \cdot D_s \leq A_s \leq (1 + \varepsilon_s) \cdot D_s, \quad \forall s \in \bar{S} \]  \hspace{1cm} (B5)

\[(1 - \varepsilon_{rs}) \cdot z_{rs} \leq q_{rs} \leq (1 + \varepsilon_{rs}) \cdot z_{rs}, \quad \forall rs \in \bar{RS} \]  \hspace{1cm} (B6)

\[(1 - \varepsilon) \cdot F \leq T \leq (1 + \varepsilon) \cdot F \]  \hspace{1cm} (B7)

\[x_a \leq C_a, \quad \forall a \in U \]  \hspace{1cm} (B8)

\[f_k^{rs} \geq 0, \quad \forall k \in K_{rs}, rs \in \bar{RS} \]  \hspace{1cm} (B9)

where

\[x_a = \sum_{rs \in RS} \sum_{k \in K_{rs}} f_k^{rs} \delta_{k_{ba}}, \quad \forall a \in \bar{A} \]  \hspace{1cm} (B10)

\[t_s^i = \sum_{rs \in RS} \sum_{k \in K_{rs}} \sum_{a \in \bar{IN}_i} \sum_{b \in \bar{OUT}_i} f_k^{rs} \delta_{k_{ba}} \delta_{k_{sb}}, \quad \forall m \in M_i, i \in \bar{T} \]  \hspace{1cm} (B11)

\[P_r = \sum_{s \in S} \sum_{k \in K_{rs}} f_k^{rs}, \quad \forall r \in \bar{R} \]  \hspace{1cm} (B12)

\[A_s = \sum_{r \in R} \sum_{k \in K_{rs}} f_k^{rs}, \quad \forall s \in \bar{S} \]  \hspace{1cm} (B13)
\[ q_{rs} = \sum_{k \in K_{rs}} f_{rs}^k, \quad \forall rs \in RS \]  \hfill (B14)

\[ T = \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{rs}^k \]  \hfill (B15)

The Lagrangian function of the above PFE formulation is

\[
L(f, u^+, u^-, \tau^+, \tau^-, \rho^+, \rho^-, \alpha^+, \alpha^-, \psi^+, \psi^-, \psi^+, d) = Z
\]

\[
+ \sum_{a \in A} u_a^- \left( v_a (1 - \varepsilon_a) - \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{rs}^k \delta_{ka} \right) + \sum_{a \in A} u_a^+ \left( v_a (1 + \varepsilon_a) - \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{rs}^k \delta_{ka} \right)
\]

\[
+ \sum_{m \in M} \sum_{r \in R} \tau_m^r \left( g_m (1 + \varepsilon_m^r) - \sum_{rs \in RS} \sum_{k \in K_{rs}} \sum_{a \in A} \sum_{be \in OUT_i} f_{rs}^k \delta_{ka} \delta_{kb} \right)
\]

\[
+ \sum_{m \in M} \sum_{r \in R} \tau_m^r \left( g_m (1 - \varepsilon_m^r) - \sum_{rs \in RS} \sum_{k \in K_{rs}} \sum_{a \in A} \sum_{be \in OUT_i} f_{rs}^k \delta_{ka} \delta_{kb} \right)
\]

\[
+ \sum_{r \in R} \rho_r^- \left( O_r (1 - \varepsilon_r) - \sum_{s \in S} \sum_{k \in K_{rs}} f_{rs}^k \right) + \sum_{r \in R} \rho_r^+ \left( O_r (1 + \varepsilon_r) - \sum_{s \in S} \sum_{k \in K_{rs}} f_{rs}^k \right)
\]

\[
+ \sum_{s \in S} \alpha_s^- \left( D_s (1 - \varepsilon_s) - \sum_{r \in R} \sum_{k \in K_{rs}} f_{rs}^k \right) + \sum_{s \in S} \alpha_s^+ \left( D_s (1 + \varepsilon_s) - \sum_{r \in R} \sum_{k \in K_{rs}} f_{rs}^k \right)
\]

\[
+ \sum_{r \in RS} \sigma_r^- \left( Z_{rs} (1 - \varepsilon_{rs}) - \sum_{k \in K_{rs}} f_{rs}^k \right) + \sum_{r \in RS} \sigma_r^+ \left( Z_{rs} (1 + \varepsilon_{rs}) - \sum_{k \in K_{rs}} f_{rs}^k \right)
\]

\[
+ \psi^- \left( F (1 - \varepsilon) - \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{rs}^k \right) + \psi^- \left( F (1 + \varepsilon) - \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{rs}^k \right)
\]

\[
+ \sum_{a \in U} d_a \cdot \left( C_a - \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{rs}^k \delta_{ka} \right)
\]  \hfill (B16)

where \( u_a^- \), \( u_a^+ \), \( \tau_m^r \), \( \tau_m^r \), \( \rho_r^- \), \( \rho_r^+ \), \( \alpha_s^- \), \( \alpha_s^+ \), \( \sigma_r^- \), \( \sigma_r^+ \), \( \psi^- \), \( \psi^- \) and \( d_a \) are the dual variables of constraints (B2), (B3), (B4), (B5), (B6), (B7) and (B8), respectively.
The first partial derivatives with respect to the path-flow variables can be expressed as follows:

\[
\frac{\partial L}{\partial f_{rs}^k} = 0 \Rightarrow \frac{1}{\theta} \ln f_{rs}^k + \sum_{a \in \Lambda} t_a(x_a) \delta_{ka}^{rs} - \sum_{a \in \Lambda} u_a^+ \delta_{ka}^{rs} - \sum_{a \in \Lambda} u_a^- \delta_{ka}^{rs} -
\]
\[
\sum_{m \in M_k} \sum_{i \in T} \sum_{a \in IN_{i}} \sum_{b \in OUT_{i}} \tau_{m}^{i+} \delta_{ka}^{rs} \delta_{bb}^{rs} - \sum_{m \in M_k} \sum_{i \in T} \sum_{a \in IN_{i}} \sum_{b \in OUT_{i}} \tau_{m}^{i-} \delta_{ka}^{rs} \delta_{bb}^{rs} - \rho_r^{-} - \rho_r^{+} - (B17)
\]
\[
\alpha_s^{+} - \alpha_s^{-} - o_{rs}^{-} - o_{rs}^{+} - \psi^{-} - \psi^{+} - \sum_{a \in U} d_a \delta_{ka}^{rs} = 0, \quad \forall \ k \in K_{rs}, \ rs \in RS
\]

From the (B17), analytical path flow solution can be expressed as follows:

\[
f_{rs}^k = \exp \left( \theta \cdot \left( -\sum_{a \in \Lambda} t_a(x_a) \delta_{ka}^{rs} + \sum_{a \in \Lambda} (u_a^- + u_a^+) \cdot \delta_{ka}^{rs} + \sum_{m \in M_k} \sum_{i \in T} \sum_{a \in IN_{i}} \sum_{b \in OUT_{i}} \left( \tau_{m}^{i-} + \tau_{m}^{i+} \right) \cdot \delta_{ka}^{rs} \delta_{bb}^{rs} + \rho_r^{-} + \rho_r^{+} + \alpha_s^{+} + \alpha_s^{-} + o_{rs}^{-} + o_{rs}^{+} + \psi^{-} + \psi^{+} + \sum_{a \in U} d_a \delta_{ka}^{rs} \right) \right), \quad \forall \ k \in K_{rs}, \ \forall rs \in RS \ (B18)
\]
APPENDIX C

Appendix C provides the detailed derivations of the analytical path flow solution for the future year PFE formulation derived as a function of path costs and dual variables associated with the constraints.

Minimize: $Z = \frac{1}{\theta} \sum_{rs \in RS} \sum_{k \in K_{rs}} f^rs_k (\ln f^rs_k - 1) + \sum_{a \in A} \int_{0}^{r^s_k} t_a(\omega)d\omega$ (C1)

subject to:

$$(1 - \varepsilon_r) \cdot O_r \leq P_r \leq (1 + \varepsilon_r) \cdot O_r, \quad \forall r \in \mathbb{R} \quad (C2)$$

$$(1 - \varepsilon_s) \cdot D_s \leq A_s \leq (1 + \varepsilon_s) \cdot D_s, \quad \forall s \in \mathbb{S} \quad (C3)$$

$$(1 - \varepsilon_{rs}) \cdot z_{rs} \leq q_{rs} \leq (1 + \varepsilon_{rs}) \cdot z_{rs}, \quad \forall rs \in \mathbb{RS} \quad (C4)$$

$$(1 - \varepsilon) \cdot F \leq T \leq (1 + \varepsilon) \cdot F \quad (C5)$$

$$f^rs_k \geq 0, \quad \forall k \in K_{rs}, rs \in \mathbb{RS} \quad (C6)$$

where

$$x_a = \sum_{rs \in RS} \sum_{k \in K_{rs}} f^rs_k \delta^rs_k, \quad \forall a \in A \quad (C7)$$

$$P_r = \sum_{s \in S} \sum_{k \in K_{rs}} f^rs_k, \quad \forall r \in \mathbb{R} \quad (C8)$$

$$A_s = \sum_{r \in R} \sum_{k \in K_{rs}} f^rs_k, \quad \forall s \in \mathbb{S} \quad (C9)$$

$$q_{rs} = \sum_{k \in K_{rs}} f^rs_k, \quad \forall rs \in \mathbb{RS} \quad (C10)$$

$$T = \sum_{rs \in RS} \sum_{k \in K_{rs}} f^rs_k$$
The Lagrangian function of the above PFE formulation is

\[
L(f, p^+, p^-, \alpha^+, \alpha^-, o^+, o^-, \psi^+, \psi^-) = Z
\]
\[+
\sum_{r \in R} \rho_r^+ \left( O_r (1 - \epsilon_r) - \sum_{s \in S} \sum_{k \in K_{rs}} f_{rs}^{rs} \right) + \sum_{r \in R} \rho_r^- \left( O_r (1 + \epsilon_r) - \sum_{s \in S} \sum_{k \in K_{rs}} f_{rs}^{rs} \right)
\]
\[+
\sum_{s \in S} \alpha_s^- \left( D_s (1 - \epsilon_s) - \sum_{r \in R} \sum_{k \in K_{rs}} f_{rs}^{rs} \right) + \sum_{s \in S} \alpha_s^+ \left( D_s (1 + \epsilon_s) - \sum_{r \in R} \sum_{k \in K_{rs}} f_{rs}^{rs} \right)
\]
\[+
\sum_{r \in RS} o_{rs}^- \left( z_{rs} (1 - \epsilon_{rs}) - \sum_{k \in K_{rs}} f_{rs}^{rs} \right) + \sum_{r \in RS} o_{rs}^+ \left( z_{rs} (1 + \epsilon_{rs}) - \sum_{k \in K_{rs}} f_{rs}^{rs} \right)
\]
\[+\psi^- \left( F(1 - \epsilon_r) - \sum_{r \in RS} \sum_{k \in K_{rs}} f_{rs}^{rs} \right) + \psi^+ \left( F(1 + \epsilon_r) - \sum_{r \in RS} \sum_{k \in K_{rs}} f_{rs}^{rs} \right)
\]

(C12)

where \( \rho_r^-, \rho_r^+, \alpha_s^-, \alpha_s^+, o_{rs}^-, o_{rs}^+ \), \( \psi^- \) and \( \psi^+ \) are the dual variables of constraints (C2), (C3), (C4) and (C5), respectively.

The first partial derivatives with respect to the path-flow variables can be expressed as follows:

\[
\frac{\partial L}{\partial f_{rs}^{rs}} = 0 \Rightarrow \frac{1}{\theta} \ln f_{rs}^{rs} + \sum_{a \in A} t_a(x_a) \delta_{ka}^{rs} - \rho_r^- - \rho_r^+ - \alpha_s^- - \alpha_s^+ - o_{rs}^- - o_{rs}^+ - \psi^- - \psi^+ = 0, \quad \forall k \in K_{rs}, rs \in RS
\]

(C13)

From the (C13), analytical path flow solution can be expressed as follows:

\[
f_{rs}^{rs} = \exp \left( \theta \left( - \sum_{a \in A} t_a(x_a) \delta_{ka}^{rs} + \rho_r^- + \rho_r^+ + \alpha_s^- + \alpha_s^+ + o_{rs}^- + o_{rs}^+ + \psi^- + \psi^+ \right) \right), \quad \forall k \in K_{rs}, rs \in RS
\]

(C14)
APPENDIX D

Note that Appendices B and C are for the general formulation that contains all available data used as constraints in Sections 5.2.2 and 5.2.4. However, not all the data are required for the PFE formulations. Appendix D provides the derivations of the analytical path flow solution for a special case of the base year PFE formulation with observed traffic count constraints, capacity constraints for unobserved links, and target O-D demand constraints. This special base year PFE formulation is given as follows:

Minimize: $Z = \frac{1}{\theta} \sum_{rs \in RS} \sum_{k \in K_{rs}} f_k^{rs} \left( \ln f_k^{rs} - 1 \right) + \sum_{a \in A} t_a(w) dw$  \hspace{1cm} (D1)

subject to:

$(1 - \epsilon_a) \cdot v_a \leq x_a \leq (1 + \epsilon_a) \cdot v_a, \quad \forall \ a \in \mathcal{A}$  \hspace{1cm} (D2)

$x_a \leq C_a, \quad \forall \ a \in \mathcal{U}$  \hspace{1cm} (D3)

$(1 - \epsilon_{rs}) \cdot z_{rs} \leq q_{rs} \leq (1 + \epsilon_{rs}) \cdot z_{rs}, \quad \forall \ rs \in \mathcal{RS}$  \hspace{1cm} (D4)

$f_k^{rs} \geq 0, \quad \forall \ k \in K_{rs}, rs \in \mathcal{RS}$  \hspace{1cm} (D5)

where

$x_a = \sum_{rs \in RS} \sum_{k \in K_{rs}} f_k^{rs} \delta_{ka}, \quad \forall \ a \in \mathcal{A}$  \hspace{1cm} (D6)

$q_{rs} = \sum_{k \in K_{rs}} f_k^{rs}, \quad \forall \ rs \in \mathcal{RS}$  \hspace{1cm} (D7)
The Lagrangian function of the above PFE formulation is

\[
L(f, u^+, u^-, d, o^+, o^-) = Z + \sum_{a \in A} u_a^- \cdot \left( v_a (1 - \varepsilon_a) - \sum_{r \in RS, k \in K_n} f^u_k \delta^{rs}_{ka} \right) + \sum_{a \in M} u_a^+ \cdot \left( v_a (1 + \varepsilon_a) - \sum_{r \in RS, k \in K_n} f^u_k \delta^{rs}_{ka} \right) + \sum_{a \in U} d_a \cdot \left( C_a - \sum_{r \in RS, k \in K_n} f^u_k \delta^{rs}_{ka} \right) + \sum_{r \in RS} o^- \cdot \left( z_{rs} (1 - \varepsilon_{rs}) - \sum_{k \in K_n} f^u_k \right) + \sum_{r \in RS} o^+ \cdot \left( z_{rs} (1 + \varepsilon_{rs}) - \sum_{k \in K_n} f^u_k \right)
\]

where \( u_a^-, u_a^+, d_a, o^-_{rs} \) and \( o^+_{rs} \) are the dual variables of constraints (D2), (D3), and (D4), respectively.

The first partial derivatives with respect to the path-flow variables can be expressed as follows:

\[
\frac{\partial L}{\partial f^u_k} = 0 \quad \Rightarrow \quad \frac{1}{\theta} \ln f^u_k + \sum_{a \in A} t_a(x_a) \delta^{rs}_{ka} - \sum_{a \in A} u_a^- \delta^{rs}_{ka} - \sum_{a \in A} u_a^+ \delta^{rs}_{ka} - \sum_{a \in U} d_a \delta^{rs}_{ka} - \sum_{r \in RS} o^- \left( z_{rs} (1 - \varepsilon_{rs}) - \sum_{k \in K_n} f^u_k \right) - \sum_{r \in RS} o^+ \left( z_{rs} (1 + \varepsilon_{rs}) - \sum_{k \in K_n} f^u_k \right) = 0, \quad \forall k \in K_n, rs \in RS
\]

Let \( \tilde{c}^u_k = c^u_k - \sum_{a \in A} (u_a^- + u_a^+) \delta^{rs}_{ka} - \sum_{a \in U} d_a \delta^{rs}_{ka} - (o^+ + o^-_{rs}) \) be the generalized route cost, where

\[
(c^u_k = \sum_{a \in A} t_a(x_a) \delta^{rs}_{ka}). \text{ Substituting equation (D9) with } \tilde{c}^u_k, \text{ we obtain}
\]

\[
f^u_k = \exp \left(-\theta \cdot \tilde{c}^u_k\right), \quad \forall k \in K_n, rs \in RS
\]

Hence, the route choice probability function can be expressed as the logit-based probability.

\[
P^u_k = \frac{f^u_k}{\sum_{k \in K_n} f^u_k} = \frac{\exp \left(-\theta \cdot \tilde{c}^u_k\right)}{\sum_{k \in K_n} \exp \left(-\theta \cdot \tilde{c}^u_k\right)}, \quad \forall k \in K_n, rs \in RS
\]
Appendix E provides the derivations of the analytical path flow solution for a special case of the future year PFE formulation with future zonal production and attraction constraints and future target O-D demand constraints. This special future year PFE formulation is given as follows:

Minimize: \[ \frac{1}{\theta} \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{rs t}^k (\ln f_{rs t}^k - 1) + \sum_{a \in A} \int_0^{\rho_a} \omega a(t_a(\omega) d\omega) \] (E1)

subject to:

\[(1 - \epsilon_r) \cdot O_r \leq P_r \leq (1 + \epsilon_r) \cdot O_r, \quad \forall r \in \bar{R} \] (E2)

\[(1 - \epsilon_s) \cdot D_s \leq A_s \leq (1 + \epsilon_s) \cdot D_s, \quad \forall s \in \bar{S} \] (E3)

\[(1 - \epsilon_{rs}) \cdot z_{rs} \leq q_{rs} \leq (1 + \epsilon_{rs}) \cdot z_{rs}, \quad \forall rs \in \bar{RS} \] (E4)

\[f_{rs t}^k \geq 0, \quad \forall k \in K_{rs}, rs \in RS \] (E5)

where

\[x_a = \sum_{rs \in RS} \sum_{k \in K_{rs}} f_{rs t}^k \delta_{ka}, \quad \forall a \in A \] (E6)

\[P_r = \sum_{s \in S} \sum_{k \in K_{rs}} f_{rs t}^k, \quad \forall r \in \bar{R} \] (E7)

\[A_s = \sum_{r \in R} \sum_{k \in K_{rs}} f_{rs t}^k, \quad \forall s \in \bar{S} \] (E8)

\[q_{rs} = \sum_{k \in K_{rs}} f_{rs t}^k, \quad \forall rs \in \bar{RS} \] (E9)
The Lagrangian function of the above PFE formulation is

\[ L(f, \rho^*, \rho^-, \alpha^+, \alpha^-, o^+, o^-) = Z \]

\[ + \sum_{r \in \mathcal{R}} \rho^+_r \left( O_r (1 - \varepsilon_r) - \sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{K}_n} f^+_ks \right) + \sum_{r \in \mathcal{R}} \rho^-_r \left( O_r (1 + \varepsilon_r) - \sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{K}_n} f^-_ks \right) \]

\[ + \sum_{s \in \mathcal{S}} \alpha^+_s \left( D_s (1 - \varepsilon_s) - \sum_{r \in \mathcal{R}} \sum_{k \in \mathcal{K}_n} f^+_ks \right) + \sum_{s \in \mathcal{S}} \alpha^-_s \left( D_s (1 + \varepsilon_s) - \sum_{r \in \mathcal{R}} \sum_{k \in \mathcal{K}_n} f^-_ks \right) \]

\[ + \sum_{rs \in \mathcal{R}S} o^-_{rs} \left( z_{rs} (1 - \varepsilon_{rs}) - \sum_{k \in \mathcal{K}_n} f^+_ks \right) + \sum_{rs \in \mathcal{R}S} o^+_{rs} \left( z_{rs} (1 + \varepsilon_{rs}) - \sum_{k \in \mathcal{K}_n} f^-_ks \right) \]  

(E10)

where \( \rho^-, \rho^+, \alpha^+, \alpha^- \) and \( o^-, o^+ \) are the dual variables of constraints (E2), (E3) and (E4) respectively.

The first partial derivatives with respect to the path-flow variables can be expressed as follows:

\[ \frac{\partial L}{\partial f^e_{rs}} = 0 \Rightarrow \frac{1}{\theta} \ln f^e_{rs} + \sum_{a \in \Lambda} t_a (x_a) \partial^e_{rs} - \rho^-_r - \rho^+_r - \alpha^-_s - \alpha^+_s \]

\[ a^-_{rs} - a^+_{rs} = 0, \quad \forall k \in \mathcal{K}_n, rs \in \mathcal{R}S \]  

(E11)

From equation (E11), we can obtain the path-flow solution:

\[ f^e_{rs} = \exp \left( -\theta \left( c^e_{rs} - a^-_{rs} + a^+_{rs} \right) \right) \cdot \exp \left( \theta \left( \rho^-_r + \rho^+_r \right) \right) \cdot \exp \left( \theta \left( \alpha^-_s + \alpha^+_s \right) \right), \quad \forall k \in \mathcal{K}_n, rs \in \mathcal{R}S \]  

(E12)

Using equation (E9), the demand of each O-D pair \( rs \) is

\[ q_{rs} = \sum_{k \in \mathcal{K}_n} \exp \left( -\theta \left( c^e_{rs} - a^-_{rs} + a^+_{rs} \right) \right) \cdot \exp \left( \theta \left( \rho^-_r + \rho^+_r \right) \right) \cdot \exp \left( \theta \left( \alpha^-_s + \alpha^+_s \right) \right), \quad \forall rs \in \mathcal{R}S \]  

(E13)

Dividing equation (E12) by equation (E13) gives the logit probability for route choice step.

\[ p^e_{krs} = \frac{\exp \left( -\theta e^e_{krs} \right)}{\sum_{k \in \mathcal{K}_n} \exp \left( -\theta e^e_{kr} \right)}, \quad \forall k \in \mathcal{K}_n, rs \in \mathcal{R}S \]  

(E14)
Using equation (E12) and recognizing

\[ B_r P_r = \exp \left( \theta \left( \rho_r - \rho_r^* \right) \right), \quad B_s A_s = \exp \left( \theta \left( \alpha_s - \alpha_s^* \right) \right), \quad \text{and} \]

\[ \pi_{rs} = -\frac{1}{\theta} \ln \sum_{k \in K_{rs}} \exp \left( -\theta \left( c_{rs}^k - o_{rs}^k - o_{rs}^* \right) \right) \]  

(E15)

where \( B_r \) and \( B_s \) are the balancing factors, \( P_r \) and \( A_s \) are the production and attraction flows, and \( \pi_{rs} \) is the O-D cost between origin \( r \) and destination \( s \). Equation (E13) can be rewritten as the optimal solution of a doubly-constrained gravity model with a negative exponential deterrence function for the trip distribution step as follows:

\[ q_{rs} = B_r P_r B_s A_s \exp \left( -\theta \pi_{rs} \right), \quad \forall rs \in RS \]

(E16)

Note that the impedance parameter typically used in the trip distribution problem is equal to the dispersion parameter used in this logit-based SUE assignment problem with zonal production and attraction constraints and future target O-D demand constraints. This model can be considered as an extension of the combined distribution and assignment (CDA) problem with a deterministic user equilibrium (DUE) principle given by Evans (1976), where the trip distribution follows a doubly constrained gravity model with a negative exponential impedance function and the traffic assignment follows the logit-based SUE model (Lundgren and Patriksson, 1998). From a behavioral perspective, the PFE formulation with zonal production and attraction constraints and future target O-D demand constraints can be interpreted from the efficiency principle (Smith, 1978, 1983). That is, trip makers generally prefer lower travel costs to higher ones. In other words, the PFE formulation finds the most probable flow pattern under the assumption of efficient trip making behavior defined by Smith’s efficiency principle.
REFERENCES


PERMISSION LETTER

Will Recker

Hi Seungkyu,

I have no problem with it.

Will

Michael Zhang

Neither from me, Seungkyu. And congratulations on your successful defense!

Michael Zhang
Professor of Transportation Engineering
Department of Civil and Environmental Engineering
University of California Davis
Davis, California 95616

Seungkyu <ryuseungkyu@gmail.com>

Dear Professor Recker and Zhang

Hello.
My name is Seungkyu Ryu who is studying for Ph.D under Professor Chen in Utah State University.

I did the final dissertation defense, and

I need your permission to include the following paper in which you appear as a coauthor in my Ph. D dissertation. I would appreciate you can reply to this email soon


Thank you
CURRICULUM VITAE

SEUNGKYU RYU

Education:
- Ph.D. Civil Engineering, Utah State University, USA, 2015
- M.S. Transportation Engineering, University of Seoul, Korea, 2006
- B.S. Transportation Engineering, University of Seoul, Korea, 2004

Academic Experience:
- Graduate Research Assistant, Utah State University, 2007-2015.
- Graduate Research Assistant and Teaching Assistant, University of Seoul, 2004-2006.

Honors and Awards:
- Outstanding paper award, 2005 Korea Transportation Society Conference (2005)
- Scholarships from University of Seoul, Korea (2003 - 2005)

Professional Affiliation and Service:
- Student member of Transportation Research Board (TRB)
- Student member of Korea Transportation Association in America (KOTAA)
- Student member of Korean American Scientists and Engineers Association (KSEA)
- Reviewer for Transportation Research Board: Bicycle Transportation Committee (ANF20), 2013
- Reviewer for Journal of Advanced Transportation, 2014
- Reviewer for Transportation Research Part C: Emerging Technologies, 2015
International Visiting Positions (Collaborative Research):

- Ajou University, Suwon, Korea, 2013 and 2014 (TOD-Based Sustainable City Transportation Research Center and Prof. Keechoo Choi)
- Tongji University, Shanghai, P.R. China, Department of Traffic and Transportation Engineering, 2013. (Collaborative research with Prof. Chao Yang)
- Central South University, Changsa, P.R. China, School of Traffic and Transportation Engineering, 2013. (Collaborative research with Dr. Jia Yao)
- Southeast University, Nanjing, P.R. China, School of Transportation, 2013. (Collaborative research with Dr. Xiangdong Xu)
- University of Hong Kong, Hong Kong, Department of Civil Engineering, College of Engineering, 2009. (Collaborative research with Prof. S.C. Wong)

Research Interests:

- Transportation Systems Modeling and Analysis
- Network Modeling and Solution Algorithm Development
- Transportation Network Design under Uncertainty
- Transportation and the Environment
- Non-motorized Transportation Modeling
- Path Flow Estimator and Its Application in Transportation Planning

Research Experience:

- A bicycle network analysis tool for planning application in small communities, funded by the Mountain Plains Consortium – A Federally Funded Regional Transportation Center, May 2013 to Dec. 2014.
- A two-stage approach for estimating a statewide truck trip table, funded by the Mountain Plains Consortium – A Federally Funded Regional Transportation Center, May 2012 to Dec. 2013.
- Visual PFE Courseware Training, funded by the California Department of Transportation, Division of Research and Innovation, through the University of
California at Irvine, April 2010 to Dec. 2011.

- Visual PFE Software Enhancements for Planning Applications in Small-Sized Communities, funded by the California Partners for Advanced Transit and Highways (PATH), June 2008 to Nov. 2010.
- Forecasting Network Traffic for Small Communities in Utah, funded by the Utah Transportation Center (UTC) at USU, Sept. 2009 to Aug. 2010.

**Refereed Journal Publications**


Refereed Conference Papers


**Research Reports**


