https://dx.doi.org/10.1016/j.tra.2016.09.013

A TWO-STAGE APPROACH FOR ESTIMATING A STATEWIDE TRUCK TRIP TABLE

Sarawut Jansuwan [a], Seungkyu Ryu [b], and Anthony Chen [c*]

[a] School of Applied Statistics, The National Institute of Development Administration
Bangkok, Thailand
[b] Department of Transportation Engineering, Ajou University, Suwon, 442-749, Korea
[c] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

ABSTRACT

This research develops a two-stage approach to estimate a statewide truck origin-destination (O-D) trip table. The proposed approach is supported by two sequential stages: the first estimates the *commodity-based* truck O-D trip tables primarily derived from the commodity flow database, and the second refines the O-D estimates using the observed truck counts. The first stage uses national commodity flow data from the Freight Analysis Framework Version 3 (FAF^3) database to develop a commodity-based truck trip table, while the second stage uses the *path flow estimator* (PFE) concept to refine the truck trip table obtained from the first stage using the truck counts from the statewide truck count program. The model allows great flexibility for data incorporation at different spatial levels in terms of estimating the statewide truck O-D trip table. To show proof of concept, a case study is conducted using the Utah statewide freight transportation network to demonstrate how the two-stage approach can be implemented in practice. The results show that the proposed approach is applicable for estimating a statewide truck O-D trip table with limited resources, and can be used to conduct truck corridor analysis to determine congested links and potential bottlenecks in Utah.

Keywords: Truck O-D trip table; path flow estimator; commodity flow; statewide planning model; freight transportation

[*] Corresponding author. Tel.: 852 3400-8327; fax: 852 2334-6389.
Email address: anthony.chen@polyu.edu.hk

1 INTRODUCTION

Statewide models including passenger and freight movements are frequently used for supporting numerous statewide planning activities. Many states use them for traffic impact studies, air quality conformity analysis, freight planning, economic development studies, project prioritization, and many other planning needs (Horowitz, 2006). According to the figures provided by the Federal Highway Administration (FHWA, 2009) and the U.S. Census Bureau (U.S. Census Bureau, 2010, 2012), the United States (U.S.) transportation system shipped a total of 17.6 billion tons of goods in 2011 to serve almost 117 million households and 7.4 million business establishments. The importance of truck demand has increased in the statewide planning process because of its strong influence on state and nation economies. Trucks are the dominant mode of freight transportation; the industry hauls 11.9 billion tons in 2011, equating to approximately two-thirds (i.e., 67%) of all freight transported in the U.S. (FHWA, 2009). Truck transportation will continue to grow over the next decade so long as the U.S. economy maintains its steady growth, international merchandise trade increases, freight sector productivity improves, and as demand for an extensive multimodal transportation network becomes available (Bureau of Transportation Statistics (BTS), 2004). According to the Freight Analysis Framework (FAF) database, trucks share 75% of all *domestic* freight shipments and it will be stable from 2007 to 2040 (FHWA, 2009). However, freight transportation capacity, especially roadway transportation, is expanding too slowly to keep up with demand (Cambridge Systematics, 2005). This imbalance in growth could significantly contribute to congestion at highway segments, interchanges, and highway bottlenecks or chokepoints (i.e., locations physically narrow and/or congested and hence very susceptible to incidents and disruptions). Congestion is also caused by restrictions on freight movement, such as the lack of space for trucks in dense urban areas (FHWA, 2008) as posted on the roadways due to height, length, width, weight limits, incident, or construction.

The truck origin-destination (O-D) trip table is an important component that can be used to help strategic transportation planners, providers, and government agencies to identify the potential bottlenecks in their areas. The subsequent results of a truck trip table obtained from the proposed framework will be beneficial for assisting the state departments of transportation (DOTs) and metropolitan planning organizations (MPOs) on evaluating operational strategies to address the consequent impacts due to truck traffic including congestion, infrastructure deterioration, safety, and environment. The demand forecast can further support the long-term strategies for the infrastructure management and investment decisions.

The current practices in estimating a statewide truck origin-destination (O-D) trip table use the truck trip rates estimated either in the Quick Response Freight Manual (QRFM) developed by Cambridge Systematics (2007) or in a commercial freight database (i.e., TRANSEARCH developed by IHS Global Insight, Inc.). However, because of the nature of the shared databases, the state DOT has to spend tremendous efforts to improve the accuracy of the estimations to match the local observations (e.g., truck counts, vehicle-miles of travel (VMT), etc.). The calibration process is usually a lengthy process and requires specialized technical staff to operate. In addition, commercial freight databases (e.g., TRANSEARCH by Global Insight, Inc.) are typically proprietary, not available for public access. Many small states usually do not have sufficient resources to conduct freight surveys or employ technical staff to develop the freight demand model. Many existing models thus overlook this component or just simply make assumptions that freight trips follow some behavioral mechanism similar to passenger trips (that is, truck traffic is estimated as a function of passenger-car traffic) (Ogden, 1992). This could be a potential weakness of truck demand modeling in the statewide model, where truck flow characteristics have been determined by other contributing factors such as location factors (i.e., places of production and market), physical factors (i.e., method by which goods can be transported: in bulk, tank, flat bed, or refrigerated container), geographical factors (the location and density of population may influence the distribution of end products), and so on (Ortuzar and Willumsen, 2002).

Holguín-Veras and Thorson (2000) summarized different methods of modeling freight transportation demand, and divided them into two major modeling approaches: trip-based modeling and commodity-based modeling. For the trip-based modeling approach, the model has three major components: trip generation, trip distribution, and traffic assignment. The trip-based model does not need a modal split step as it assumes mode choice has already been selected. List *et al*. (2002), for instance, used the trip-based modeling method to estimate a truck O-D trip table from partial and fragmentary truck observations in the New York region. The main advantage of the trip-based modeling method is that it typically requires less data (i.e., only truck traffic counts) to reproduce an O-D matrix. However, the trip-based modeling method tends to overlook the behavioral characteristics of commodity flows. The commodity-based modeling method, on the other hand, uses the commodity flows to estimate truck flows produced and attracted by each zone in the study area. Sorratini and Smith (2000), for example, developed a statewide truck trip model using commodity flow data obtained from the commodity flow survey (CFS) and improved the estimation using the input-output (I-O) economic data. Although the commodity-based models can capture the fundamental economic mechanisms of freight movements more accurately than the trip-based models, the truck O-D trip tables estimated from commodity-based models often overlook the *non-freight* truck trips (e.g., light commercial truck or empty truck trips).

To fill this modeling gap, this research proposes a two-stage approach to estimate a statewide truck O-D trip table. The proposed approach is supported by two sequential stages: Stage 1 estimates the commodity-based truck O-D trip table primarily derived from the commodity flow database, and Stage 2 adopts the concept of path flow estimator (PFE) to refine the commodity-based truck O-D trip table using the observed truck counts. The proposed approach uses the secondary data sources available for public and research access such as the Freight Analysis Framework (FAF) database, statewide traffic counts, and socioeconomic and land use data to estimate statewide network truck traffic. A case study using the Utah statewide freight transportation network is conducted to demonstrate the application of the proposed method. This paper is divided into five sections. Section 2 provides an overview and review of methods for estimating truck O-D trip table including commodity-based and trip-based models. Section 3 explains the approach for estimating the statewide truck O-D trip table. Section 4 presents the analysis and findings in the Utah statewide freight transportation network. In Section 5, we conclude and discuss the findings and future research direction.

2 LITERATURE REVIEW ON TRUCK O-D ESTIMATION

Holguín-Veras and Thorson (2000) summarized different ways that could be used for modeling freight demand and divided them into two major modeling platforms: (1) trip-based modeling and (2) commodity-based modeling. Figure 1 depicts the modeling of these two approaches. This section provides a literature review based on these two modeling approaches.

2.1 Trip-based modeling approach

For the trip-based modeling approach, the model has three major components: trip generation, trip distribution, and traffic assignment. The trip-based model begins with trip generation. In this step, regression models for trip production and trip attraction are estimated in conjunction with land use and socio-economic characteristics for each traffic analysis zone (TAZ). The next step is trip distribution, which is accomplished through a spatial interaction model (i.e., gravity model or growth factor method). The last step is to assign the trip table from the trip distribution step to the network. This trip-based modeling approach is also known as a three-step model as the mode choice has been already made in the truck freight model.

The current practice in estimating truck trip table is through the use of the truck trip rates estimated in the Quick Response Freight Manual (QRFM) II developed by Cambridge Systematics (2007). The QRFM provides truck trip generation rates based on the survey data collected from Phoenix, Arizona. Using trip rates to reflect the trip-making propensity based on land use configurations is a common practice and is an economical and reasonable method when planning resources are limited.

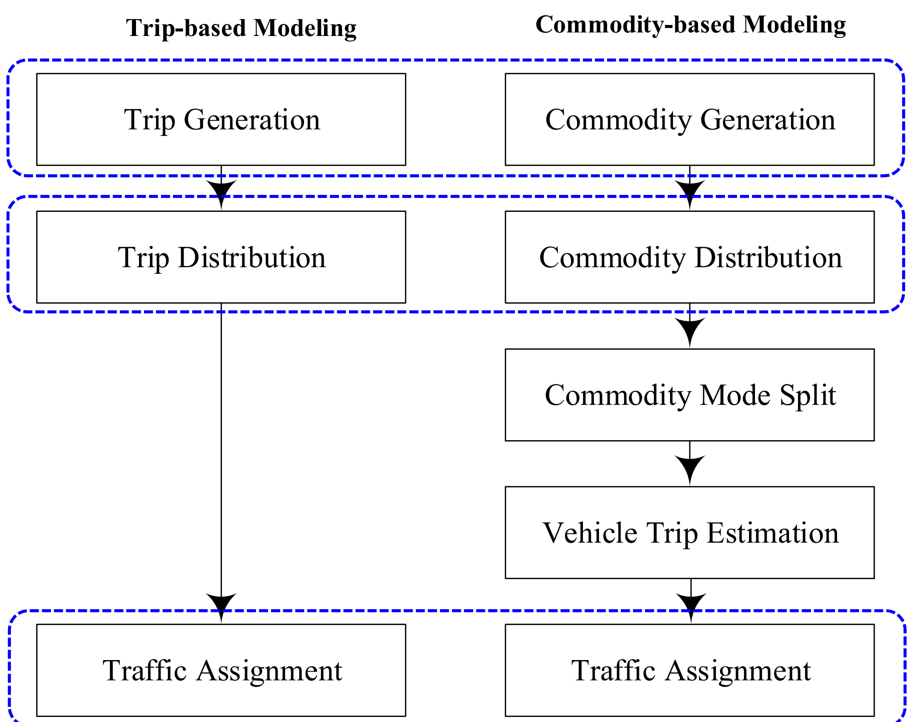


Figure 1 Trip-based and commodity-based approaches

(modified from Holguín-Veras and Thorson, 2001)

Many researchers have also demonstrated that the estimation of truck O-D trip table could be achieved using secondary data sources based on the trip-based modeling approach. Tamin and Willumsen (1989) introduced a three-step model to estimate freight demand from observed traffic count data. They used two types of gravity models in the trip distribution step including the gravity model and the gravity-opportunity model. They proposed the nonlinear least square and maximum likelihood estimation methods to ensure that the models estimate link flows as close as possible to the observed data. List and Turnquist (1994) developed a linear programming (LP) method to synthesize the truck flow pattern from the observed truck counts on some links and cordon lines. This LP method minimizes the weighted sum of the residual between the estimated and observed values using fixed link-use coefficients for each O-D pair from a probabilistic path assignment procedure.

Later, List *et al*. (2002) used a similar technique to estimate a large-scale truck O-D trip table in the New York region. The model was implemented in a two-step process: the first step estimates the trip production and trip attraction at each traffic analysis zone (TAZ), while the second step uses the link-use coefficients based on a multi-path traffic assignment procedure to estimate the truck O-D trip table. Crainic *et al*. (2001) used a bi-level optimization program to adjust the target freight demand matrix such that the differences between the observed and assigned truck flows in the upper level are minimized. The lower level for this bi-level program is a system optimum (SO) traffic assignment procedure. They implemented the bi-level programming method in the Strategic Planning of Freight transportation (STAN) software, an interactive-graphic transportation planning package for multimodal multiproduct freight transportation. The main advantage of the trip-based modeling approach is that it typically requires less data (i.e., only truck traffic counts) with some existing planning data (e.g., trip production, trip attraction, partial or full size of target trip table) to estimate an O-D matrix. However, the main disadvantage of the trip-based modeling approach is that it tends to overlook the behavioral characteristics of commodity flows in the urban and regional models. Holguín-Veras *et al*. (2001) noted that trip-based models have a limited range of applicability to account for major changes of the study areas such as changes in land use and it could be difficult to model multimodal systems using this modeling approach.

2.2 Commodity-based modeling approach

The commodity-based modeling approach, on the other hand, uses the commodity flows to estimate truck flows produced and attracted by each TAZ. In the United States (US), the FAF estimates commodity flows over the national highway networks, waterways, and rail systems among the states and regions. The current version of the FAF commodity O-D database (FAF version 3) provides estimates of commodity flows by origin, destination, and by mode for the base year 2007 and the forecast years from 2010 to 2040 with a five-year interval. Note that the FAF commodity O-D database was developed using the 2007 Commodity Flow Survey (CFS) and other public data sources. To estimate the truck demand from the CFS data, the commodity flows in tonnage have to be disaggregated from the state to the finer zonal level such as TAZ by county and then convert them to truck trips using the truck payload equivalent factor (TPEF).

Because the CFS database is based on survey data established through a shipper-based survey, the commodity-based models can better capture the fundamental behavioral characteristics of commodity flows. Sorratini and Smith (2000), for example, developed a statewide truck trip model using the commodity flow data obtained from the CFS database and improved the estimation using the input-output (I-O) economic data. A similar technique was also adopted by Fischer *et al*. (2005) for estimating the heavy-duty truck O-D trip table for the Southern California Association of Government (SCAG) region. The commodity-based modeling approach is often used in the statewide and regional practices. Zhang *et al*. (2003), for instance, estimated the intermodal freight flow patterns of highway, railway, and waterway networks for the state of Mississippi using the CFS database and public domain data. They further developed a simulation model to assess the freight operations and the modal shift effect (i.e., from truck to intermodal barge/truck). Liedtke (2006) and Wisetjindawat *et al*. (2006) used microsimulation models to replicate the commodity movements and to assess different scenarios of urban freight distribution. Although this approach provides a much finer resolution of truck traffic flows over time, this technique is usually data demanding, computationally expensive, and may be more suitable for assessing truck operations of urban freight traffic instead of regional freight traffic for strategic planning. Al-Battaineh and Kaysi (2005) used a genetic algorithm (GA) procedure to find the best O-D matrix that gives the minimum deviation between observed and estimated data when the O-D matrix is assigned to the network. Trip production and trip attraction derived from the trip generation step were also used to preserve the spatial distribution of the commodity flow pattern.

However, it is known that GA cannot guarantee to find the global optimum. Stefan et al. (2005) noted that it may be difficult to obtain the I-O data for certain regional and urban areas. While the commodity-based models have more advantages than the trip-based models (they can capture the fundamental economic mechanisms of freight movements more accurately), a truck O-D trip table estimated from the commodity-based method often overlooks the non-freight truck trips (e.g., commercial truck or empty truck trips). Hybrid models have also been developed to bridge the modeling gap of trip-based and commodity-based models. Holguín-Veras and Patil (2008) developed a multi-commodity O-D estimation model that combines two submodels: (1) a commodity-based model, and (2) a complementary model of empty truck trips. The findings of this study highlighted the significant benefits of considering an empty truck trip model in the estimation process as it can improve the ability to replicate the observed traffic counts. The hybrid approach was also adopted in the SCAG's truck demand model. Hybrid models forecast the internal-internal (I-I) truck trips through the use of a trip-based model and forecast the external truck trips through the use of a commodity flow survey. Ma *et al*. (2012) proposed a new concept of origin destination tuple (ODT) that can bridge between the macro transportation planning model and the micro activity-based or trip chain model to estimate travel demand. Data from multiple sources including the travel survey of individual trip chains, traffic flow data from loop detectors, and cameras were utilized for constructing and calibrating the model. Sánchez-Díaz *et al*. (2015) developed the time-dependent freight tours synthesis (TD-FTS) model to replicate freight activities in urban areas. Specifically, a time-dependent entropy maximization program was developed based on three major components (e.g., tour, freight demand, and link count traffic) to find aggregate tour flows. In summary, some of the truck freight demand modeling approaches including trip-based, commodity-based, and hybrid models are provided in Table 1.

Table 1 Freight demand modeling approaches, methods, and data sources

Authors	Modeling approaches		Methods	Data sources
	Trip-based	Commodity-based		
List and Turnquist (1994)	●		Linear programming model	Observed truck counts for some links and cordon lines
Sorratini and Smith (2000)		●	I-O model	CFS, TRANSEARCH
List et al. (2002)	●		Linear programming model	Observed truck counts for some links and cordon lines
Zhang et al. (2003)		●	Planning and simulation models	CFS, TRANSEARCH, intermodal databases
Al-Battaineh and Kaysi (2005)		●	I-O model, Genetic Algorithm	Commodity flows, observed truck counts
Liedtke (2006), Wisetjindawat et al. (2006), de Jong and Ben-Akiva (2007); Ruan *et al.* (2011)		●	Micro-simulation, Tour-based	Commodity flow surveys, trip chain surveys.
Fischer et al. (2005)	■	■	Hybrid model	Shipper and receiver surveys (for internal trips), commodity flow surveys (for external trips)
Holguín-Veras and Patil (2008)	■	■	Hybrid model, least square method	Multi-commodity flows, estimated empty truck trips, observed truck counts
Ma et al. (2012)	■	■	Activity-based model with Bayesian Network	Travel survey of individual trip chain, traffic flow of loop detectors, and cameras
Sánchez-Díaz et al. (2015)	■	■	Time dependent entropy maximization program (two objectives)	Freight demand and partial link count traffic

Note: ■ represents a hybrid model

3 A TWO-STAGE APPROACH FRAMEWORK

Our approach divides the process into two stages: 1) development of a commodity-based truck trip table from the recent developed FAF^3 database, and 2) usage of the PFE concept to refine the truck trip table obtained from the first stage. Figure 2 depicts an overall framework of this approach. The estimation is accomplished through the observed truck counts collected from the permanent count stations within the state and state borders as a part of the statewide truck count programs. The commodity-based matrix will help to guide the estimating process in the second stage as it preserves the spatial distribution of the O-D demand pattern. It is important to note that the commodity flow derived from the FAF database is in unit of tonnage per year. Conversion factors are required to convert the annual commodity flows to daily truck trips. Additionally, the commercial and empty trucks are considered in the estimation process as it can improve the ability to replicate the observed traffic counts. Details of these two stages are described in the following subsections.

3.1 Stage 1: Development of a commodity-based truck O-D trip table

A simplified procedure shown in Figure 2 was developed in the first stage to estimate truck O-D trip table from commodity flows. This method accounts for all types of truck flows including intrastate trips (within state), interstate trips (trips originating from the state and trips destined to the state), and through trips. It includes four steps: (1) Extraction of truck flows by weight from FAF database, (2) Distribution of truck flows to internal and external state zones, (3) Disaggregation of truck flows to the county level, and (4) Conversion of truck flows to truck trips. The steps are briefly explained as follows:

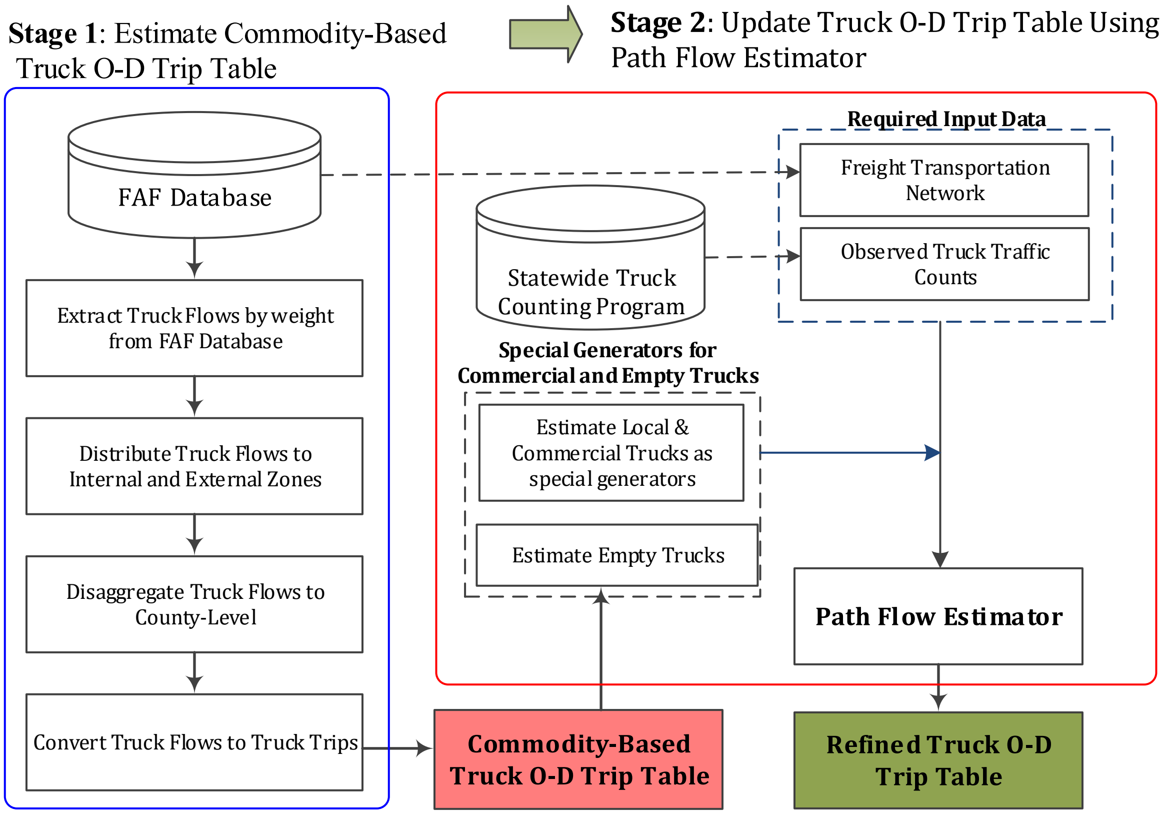


Figure 2 Conceptual framework of the two-stage approach

1) Extract truck flows by weight from FAF database

The first step is to extract truck flows from the FAF commodity flow database. It should be noted that the FAF[3] commodity database can be publicly accessed from the Freight Management and Operations Database website[1] (developed by Oak Ridge National Laboratory, 2011). The FAF data, developed from the Commodity Flow Survey (CFS) and the international trade data from the Census Bureau, represent the freight movements among states (state-to-state flows) of 43 commodity types, 123 domestic regions, and 8 international regions. Most of the major sectors (e.g., agriculture, coal, fuel oils, gasoline, electronics, pharmaceuticals, etc.) are included in the FAF database. It provides an approximate volume and sources of traffic passing through different jurisdictions at the corridor level. Additionally, it informs states about their major trading partners either domestic or international trades, the approximate volume and sources of traffic passing through jurisdictions at the corridor level (FHWA, 2012). It consists of three major

[1] Available at: http://ops.fhwa.dot.gov/freight/freight_analysis/faf/index.htm

databases: (1) DOM database: the commodity flows between domestic origins and domestic destinations, (2) BRD database: the commodity flows by land from Canada and Mexico to domestic destinations via ports of entry on the U.S. border and vice versa, and (3) SEA database: the commodity flows by water from overseas origins via ports of entry to domestic destinations and vice versa. The commodity flows are classified based on the Standard Classification of Transported Goods (SCTG). The measurement units of the commodity flow database are in thousands of tons (KT) and millions of dollars (MDOL). In our study, the DOM truck flows were extracted from the FAF database and the outputs of this step are truck flows by weight in unit of thousand tons (KT).

2) Distribute Truck Flows to Internal and External State Zones

This step requires quantifying four types of truck flows including: (1) truck flows within a state (Internal-Internal, I-I), (2) truck flows from a given state to other states (Internal-External, I-E), (3) truck flows from other states to a given state (External-Internal, E-I), and (4) through truck flows (External-External, E-E). It should be noted that the FAF database does not provide enough information to estimate the through truck flows (E-E). In order to estimate the through truck flows, the *subarea analysis technique* using the user equilibrium (UE) assignment in the planning software (i.e., CUBE by Citilabs) was used in this step. CUBE automatically identifies the external stations that enter and exit to/from a given state. Note that the subarea analysis technique is also available in other planning software packages such as TransCAD and EMME4.

3) Disaggregate Truck Flows to the County Level

This step disaggregates the truck flows from the state level to the county level using disaggregation factors. The disaggregation factors were developed from the information of population and employment of each county. Note that the employment and population are the most common disaggregation factors and they can be obtained from the state government organizations (e.g., Utah Governor's Office of Planning and Budget (GOPB) for population and Utah Department of Work Force Services for employment in this study (Utah Department of Workforce Services, 2008). To perform the disaggregation, the activity variables of a specific state were collected. The activity variables including population, income, employment, Gross Domestic Product (GDP), warehouse space, etc., are available at different regions. The most common variables that are widely used for disaggregation include employment and population due to its availability and accessibility from various geographical scales. Population can be an accurate reflection of the majority consumption of commodities, while the employment factor can be correlated to the tonnage output (production) in some industries. Based on the FAF database, the correlation of the top five commodities by volume in Utah (i.e., coal, nonmetallic mineral products, gravel, waste/scrap) ranges between 60%-86%. Therefore, we used employment as the disaggregate factor for truck trip production, and population as the disaggregate factor for truck trip attraction. We believe population and employment can serve as a surrogate factor for Utah due to their simplicity, availability, and accessibility. However, the production and consumption of some industrial commodities may not correlate well with population and employment. The disaggregation factors by population and employment should be implemented with caution. One should follow the Guidebook for Developing Subnational Commodity Flow Data (2013) to overcome this issue.

4) Convert Truck Flows to Truck Trips

This step converts the truck flows from Step 3 to truck trips. The truck payload equivalent factor (TPEF) derived from the Federal Vehicle Inventory and User Survey (VIUS) data was employed. It should be noted that the truck payload data for all states are provided in the VIUS database. The average payload data could be replaced by observed data from the statewide truck weigh stations and weigh-in-motion (WIM) stations. For Utah, the average payload is 41,196 lbs/vehicle or 20.6 tons/vehicle. This number is within a reasonable range compared to studies in other states (e.g., 16.07 tons/vehicle for Ohio (Cambridge Systematics, 2002), 24.00 tons/vehicle for Wisconsin (Wisconsin Department of Transportation,1995), and 25.77 tons/vehicle for Texas (Cambridge Systematics, 2004)). After converting truck flows to truck trips, the unit is the number of truck trips per year or annual truck trips. These annual truck trips must be converted to daily truck trips using the average number of the working days per year for trucks. For practical purposes, we adopted 300 days per year as an average annual truck workdays as suggested by the Highway Capacity Manual (Transportation Research Board, 2000). This number was obtained based on 5 working weekdays plus 44% capacity on weekends, yielding 306 workdays and subtracting 6 federal holidays. The results of this final process are the estimated daily truck flows at the county-level.

It should be noted that estimating truck O-D trip table from the commodity flows often underestimates the local truck trips such as the light commercial and empty truck trips. Additionally, the FAF does not provide local detail or temporal (seasonal, daily, or hourly) variation in freight flows that are typically necessary to support project planning. It estimates of truck tonnage and number of trucks on the network, particularly in regions with multiple routes or significant local traffic between major centers of freight activities, should be supplemented with local data to support local applications. Thus, in this study, we estimate the commercial trucks for each TAZ using the commercial truck trip generation model. The commercial truck trip generation model is expressed as:

$$O_r^{comm} = \lambda^{agriculture} x_r^{agriculture} + \lambda^{basic} x_r^{basic} + \lambda^{retail} x_r^{retail} + \lambda^{office} x_r^{office} + \lambda^{household} x_r^{household} \quad (1)$$

where O_r^{comm} is the commercial truck trip production of origin *r*; $x_r^{agriculture}$, x_r^{basic}, x_r^{retail}, x_r^{office} are the employment rates for agriculture, basic (e.g., manufacturing, transportation, wholesale and utilities), retail, and office, respectively; and $x_r^{household}$ is the number of households in origin *r*. The calibrated coefficients ($\lambda^{agriculture}$, λ^{basic}, λ^{retail}, λ^{office}, $\lambda^{household}$) were borrowed from the Utah Statewide Travel Model (Wilbur Smith Associates and Resource Systems Group, 2009) (i.e., (0.166, 0.141, 0.133, 0.065, 0.038) for the urban area, and (0.050, 0.222, 0.133, 0.065, 0.038) for the rural area). The commercial truck trip attraction of destination *s* (D_s^{comm}) is assumed to be the same as the commercial truck trip production. Note that the commercial truck trip production and attraction are not necessary the same. When they are not the same, it is important to balance the total production to match with the total attraction. Deciding whether to use total production or total attraction as the control total depends on which one is more accurate. This process should be done prior to Stage 2 as these estimates (e.g., O_r and D_s) serve as inputs to the PFE formulation. The commercial truck production and attraction flows (O_r^{comm}, D_s^{comm}) in Eq. (1) are the total zonal production of origin *r* and total zonal attraction of destination *s* (i.e., $O_r^{comm} + O_r^{loaded} = O_r$ and $D_s^{comm} + D_s^{loaded} = D_s$). Note that O_r composes of the commodity-based and commercial truck trip production of origin *r,* while D_s composes of the commodity-based and commercial truck trip attraction of destination *s*.

The empty truck trips were estimated using the Holguín-Veras and Thorson (HV-T) model developed by Holguín-Veras *et al*. (2010). The empty truck trip model was developed using the destination choice probability function expressed as a function of trip distance and the magnitude of opposing commodity flows. In our study, we focus more on the long haul truck freight transportation so the empty truck trips are estimated using the HV-T model III with zero order trip chains. Specifically, the empty truck trips were estimated based on the logit probability function as follows:

$$z_{sr}^{empty} = \frac{\exp(\omega_0 - \omega_1 d_{sr})}{\sum_{l \in R} \exp(\omega_0 - \omega_1 d_{sr})} z_{rs}^{loaded}, \forall rs \in RS \tag{2}$$

where z_{sr}^{empty} are the empty truck trips returning from destination *s* to origin *r* (*s*, *r*), z_{rs}^{loaded} are the loaded truck trips between (*r*, *s*), d_{sr} is the returning distance between (*s*, *r*), and ω_0 and ω_1 are the coefficients empirically calibrated in the same study for large trucks (i.e., ω_0 =0.689, ω_1 =3.452). The parameters in this model should be further calibrated using the statewide observed empty truck trip data. Note that the logit formulation implies that the longer distance trucks (e.g., through truck traffic) would have lower probabilities of returning to their origins. The empty truck trips estimated from Eq. (2) were used as supplements to truck movements on the statewide highway network. In our study, the analysis of the empty truck trips indicates a high number of empty or returning truck trips (i.e., 46.81% of total truck trips are empty truck trips) in Utah. Hence, ignoring the empty truck could significantly underestimate the statewide truck O-D demand. The commercial truck production, attraction, and empty truck trips derived above were finally added to commodity-based production, attraction, and O-D flows.

3.2 Stage 2: Update truck O-D trip table using PFE

This stage uses the optimization approach to refine the commodity-based truck O-D trip table obtained from the first stage. The basic idea is to use the concept of Path Flow Estimator (PFE) to estimate path flows that can reproduce the observed link counts and flows on other spatial levels. PFE is capable of estimating path flows and path travel times using only traffic counts from a subset of network links. PFE was originally developed by Bell and Shields (1995) and further enhanced Chen *et al*. (2005, 2009, 2010). The core component of PFE is a logit-based path choice model in which the perception errors of path travel times are assumed to be independently and identically Gumbel variates. The logit model interacts with link cost functions to produce a stochastic user equilibrium (SUE) traffic pattern. It should be noted that the SUE traffic assignment procedure was also implemented to estimate the freight flows in the FAF version 3 (please refer to Chapter 5 of FAF[3] report (FHWA, 2009)). This subsection includes the PFE formulation, optimality conditions, and solution procedure. Hereafter, the following notation in Table 2 is considered.

Table 2 Notation for the PFE model

Notation	Description
Set of Variables	
M	: Set of network links with truck counts
U	: Set of network links without truck counts
A	: Set of all network links $A = M \cup U$
R	: Set of origins
S	: Set of destinations
RS	: Set of O-D pairs
K_{rs}	: Set of paths connecting origin *r* and destination *s*
$\overline{R}$	: Set of origins with commodity-based data
$\overline{S}$	: Set of destinations with commodity-based data
$\overline{RS}$	: Set of target (or prior) O-D pairs
Decision Variables	
f_k^{rs}	: Flow on path *k* connecting O-D pair *rs*
Input Variables and Parameters	
v_a	: Observed truck counts on link *a*
C_a	: Capacity of link *a*
O_r	: Daily truck trip production of origin *r*
D_s	: Daily truck trip attraction of destination *s*
z_{rs}	: Daily O-D truck trips between origin *r* and destination *s*
F	: Total daily truck trips
ε_a	: Error bound allowed for truck count on link *a*
ε_r	: Error bound allowed for truck trip production of origin *r*
ε_s	: Error bound allowed for truck trip attraction of destination *s*
ε_{rs}	: Error bound allowed for the O-D truck trips between origin *r* and destination *s*
ε	: Error bound allowed for the total demand
θ	: Dispersion parameter in the logit model
$t_a(\cdot)$	: Truck travel time on link *a*
δ_{ka}^{rs}	: Path-link indicator, 1 if link a is on path *k* between O-D pair *rs* and 0 otherwise
x_a	: Estimated truck traffic volume on link *a*
P_r	: Estimated truck trip production of origin *r*
A_s	: Estimated truck trip attraction of destination s
q_{rs}	: Estimated truck O-D flows between origin *r* and destination *s*
α_a, β_a	: Parameters for BPR link cost function

The aim of this stage is to adapt the PFE to take not only truck traffic counts but also the available freight planning data (i.e., truck production and attraction flows) to update the commodity-based truck O-D trip table. PFE requires traffic count data to estimate the statewide truck O-D trip table while the planning data is an optional input in this process. However, the commodity-based truck O-D trip table obtained from the first stage can enhance the observability of the O-D estimation problem as well as preserving the spatial commodity flow pattern in the study area.

$$\text{Min } Z = \sum_{a\in \mathrm{A}} \int_0^{x_a} t_a(\omega)dw + \frac{1}{\theta} \sum_{rs\in \mathrm{RS}} \sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\left(\ln f_k^{rs} - 1\right) \tag{3}$$

s.t.

$$(1-\varepsilon_a)\cdot v_a \le x_a \le (1+\varepsilon_a)\cdot v_a, \forall a \in \mathrm{M}, \tag{4}$$

$$x_a \le C_a, \forall a \in \mathrm{U}, \tag{5}$$

$$(1-\varepsilon_{rs})\cdot z_{rs} \le q_{rs} \le (1+\varepsilon_{rs})\cdot z_{rs}, \forall rs \in \overline{\mathrm{RS}}, \tag{6}$$

$$(1-\varepsilon_r)\cdot O_r \le P_r \le (1+\varepsilon_r)\cdot O_r, \forall r \in \overline{\mathrm{R}}, \tag{7}$$

$$(1-\varepsilon_s)\cdot D_s \le A_s \le (1+\varepsilon_s)\cdot D_s, \forall s \in \overline{\mathrm{S}}, \tag{8}$$

$$(1\text{-}\varepsilon\ )\cdot F\ \le T\ \le (1+\varepsilon\ )\cdot F, \tag{9}$$

$$f_k^{rs} \ge 0, \qquad \forall k \in \mathrm{K}_{rs}, rs \in \mathrm{RS}, \tag{10}$$

where

$$x_a = \sum_{rs\in \mathrm{RS}} \sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\delta_{ka}^{rs}, \forall a \in \mathrm{A}, \tag{11}$$

$$q_{rs} = \sum_{k\in \mathrm{K}_{rs}} f_k^{rs}, \forall rs \in \mathrm{RS}, \tag{12}$$

$$P_r = \sum_{s\in \mathrm{S}} \sum_{k\in \mathrm{K}_{rs}} f_k^{rs}, \forall r \in \mathrm{R}, \tag{13}$$

$$A_s = \sum_{r\in \mathrm{R}} \sum_{k\in \mathrm{K}_{rs}} f_k^{rs}, \forall s \in \mathrm{S}, \tag{14}$$

$$T = \sum_{rs \in \mathrm{RS}} \sum_{k \in \mathrm{K}_{rs}} f_k^{rs}, \tag{15}$$

The objective function in Eq. (3) has two terms: a user equilibrium term and an entropy term and. 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, PFE finds path flows that minimize the SUE objective function while simultaneously reproducing truck traffic counts on all observed links in Eq. (4), commodity-based demands of certain O-D pairs in Eq. (6), truck production and attraction of certain origins and destinations in Eqs. (7) and (8), and total demand in Eq. (9) within some predefined error bounds. Note that z_{rs} (i.e., $z_{rs} = z_{rs}^{loaded} + z_{rs}^{empty}$ including both loaded and empty truck trips) is the daily truck trips between O-D pair rs, O_r and D_s (i.e., $O_r^{comm} + O_r^{loaded} = O_r$ and $D_s^{comm} + D_s^{loaded} = D_s$ including both commodity-based and commercial truck trips) are the daily truck trip production of origin r and attraction of destination s, and F is the total daily truck trips for the whole network estimated from the first stage. These values can be treated as target values for the second stage PFE procedure, similar to those often used in O-D estimation from traffic counts to increase observability. The introduction of the error bounds in the side constraints (i.e., Eq. (4) and Eqs. (6)-(9)) enhances the flexibility of PFE by allowing the analyst to incorporate local knowledge about the network conditions (i.e., field data) and socio-economic and land use data (i.e., planning data) at different spatial levels to the estimation procedure. More reliable information will constrain the estimation to be within a smaller tolerance, whereas less reliable information will allow for a larger deviation (Bell et al., 1997; Chen et al., 2005). However, note that the error bounds should be set up judiciously. Otherwise, improper error bound specifications 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. In this paper, we adopted the guidelines provided by the FHWA (1990) and Cambridge Systematics (2010) according to the roadway functional class as follows: Freeway (+/- 7%), major arterial (+/- 10%), minor arterial (+/- 15%), and collector (+/- 25%). If there exists a solution, then the estimation results must satisfy all the constraints including those with error bounds specified by the analyst. Otherwise, the analyst needs to adjust the error bounds (i.e., fine-tuning the error bounds) until a solution exists. To resolve the error bound specification issue, one can also adopt the norm approximation method to endogenously determine appropriate error bounds as suggested by Chen et al. (2009, 2010). For the unobserved links, the estimated flows cannot exceed their respective capacities as indicated by Eq. (5). Eq. (10) constrains the path flows to be non-negativity, while Eqs. (11)-(15) are definitional constraints that sum up the estimated path flows to obtain the link flows, O-D flows, zonal production flows, zonal attraction flows, and total demand, respectively.

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.

$$\begin{aligned}
L(\mathbf{f},\mathbf{u}^{+},\mathbf{u}^{-},\mathbf{d},\mathbf{o}^{+},\mathbf{o}^{-},\boldsymbol{\rho}^{-},\boldsymbol{\rho}^{+},\boldsymbol{\eta}^{-},\boldsymbol{\eta}^{+},\boldsymbol{\psi}^{-},\boldsymbol{\psi}^{+}) = Z + \sum_{a\in \mathrm{M}} u_a^{-}\cdot\left(v_a\left(1-\varepsilon_a\right)-\sum_{rs\in \mathrm{RS}}\sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\delta_{ka}^{rs}\right)+ \\
\sum_{a\in \mathrm{M}} u_a^{+}\cdot\left(v_a\left(1+\varepsilon_a\right)-\sum_{rs\in \mathrm{RS}}\sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\delta_{ka}^{rs}\right)+\sum_{a\in \mathrm{U}} d_a\cdot\left(C_a-\sum_{rs\in \mathrm{RS}}\sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\delta_{ka}^{rs}\right)+ \\
\sum_{rs\in \mathrm{RS}} o_{rs}^{-}\cdot\left(z_{rs}\left(1-\varepsilon_{rs}\right)-\sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\right)+\sum_{rs\in \mathrm{RS}} o_{rs}^{+}\cdot\left(z_{rs}\left(1+\varepsilon_{rs}\right)-\sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\right)+ \\
\sum_{r\in \mathrm{R}} \rho_r^{-}\cdot\left(P_r\left(1-\varepsilon_r\right)-\sum_{s\in \mathrm{S}}\sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\right)+\sum_{r\in \mathrm{R}} \rho_r^{+}\cdot\left(P_r\left(1+\varepsilon_r\right)-\sum_{s\in \mathrm{S}}\sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\right)+ \\
\sum_{s\in \mathrm{S}} \eta_s^{-}\cdot\left(A_s\left(1-\varepsilon_s\right)-\sum_{r\in \mathrm{R}}\sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\right)+\sum_{s\in \mathrm{S}} \eta_s^{+}\cdot\left(A_s\left(1+\varepsilon_s\right)-\sum_{r\in \mathrm{R}}\sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\right)+ \\
\psi^{-}\left(T\left(1-\varepsilon\right)-\sum_{rs\in \mathrm{RS}}\sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\right)+\psi^{+}\left(T\left(1+\varepsilon\right)-\sum_{rs\in \mathrm{RS}}\sum_{k\in \mathrm{K}_{rs}} f_k^{rs}\right)
\end{aligned} \tag{16}$$

where u_a^{-}, u_a^{+}, d_a, o_{rs}^{-}, o_{rs}^{+}, ρ_r^{-}, ρ_r^{+}, η_s^{-}, η_s^{+}, ψ^{-}, and ψ^{+} are the dual variables of constraints (4), (5), (6), (7), (8), and (9) respectively. The values of u_a^{+}, d_a, o_{rs}^{+}, ρ_r^{+}, η_s^{+}, and ψ^{+} are restricted to be non-positive, while the value of u_a^{-}, o_{rs}^{-}, ρ_r^{-}, η_s^{-}, and ψ^{-} 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}^{-}, o_{rs}^{+}, ρ_r^{-}, ρ_r^{+}, η_s^{-}, η_s^{+}, ψ^{-}, and ψ^{+} can be interpreted as corrections to the O-D travel times, zonal production attractiveness, zonal attraction attractiveness, and total demand attractiveness, respectively, that can be used to steer the estimated path flow pattern to within the interval constraints specified by Eqs. (6), (7), (8), and (9). These dual variables are zero if the estimated link flows, O-D flows, zonal production flows, zonal attraction flows, and total demand are within an acceptable range defined by the error bounds, and non-zero if they are binding at one of the limits. d_a is related to the link queuing delay when the estimated link flow reaches its capacity (Bell and Iida 1997). 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 (4)-(9) as follows.

$$f_k^{rs} = \exp\left(\theta \cdot \left(\begin{array}{l} -c_k^{rs} + \sum_{a \in \mathrm{M}} \left(u_a^- + u_a^+\right)\delta_{ka}^{rs} + \sum_{a \in \mathrm{U}} d_a \delta_{ka}^{rs} + \\ \left(o_{rs}^+ + o_{rs}^-\right) + \left(\rho_r^- + \rho_r^+\right) + \left(\eta_s^- + \eta_s^+\right) + \left(\psi^- + \psi^+\right) \end{array}\right)\right) \forall\, k \in \mathrm{K}_{rs}, rs \in \mathrm{RS} \tag{17}$$

The solution procedure for solving PFE 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 (e.g., link flows, O-D flows, production flows, attraction flows, and total demand) using the path-flow solution from PFE. For details of the solution procedure, please refer to Bell and Shield (1995), Chen *et al*. (2005, 2009, 2010). The overall solution procedure for solving the PFE formulation is provided in Figure 3. Using the path flow solution, various outputs from path flows are derived as follows.

- 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 paths flows connecting that O-D pair gives the O-D flow.
- Link flow: the sum of all path flows passing through a given link gives the link flow.

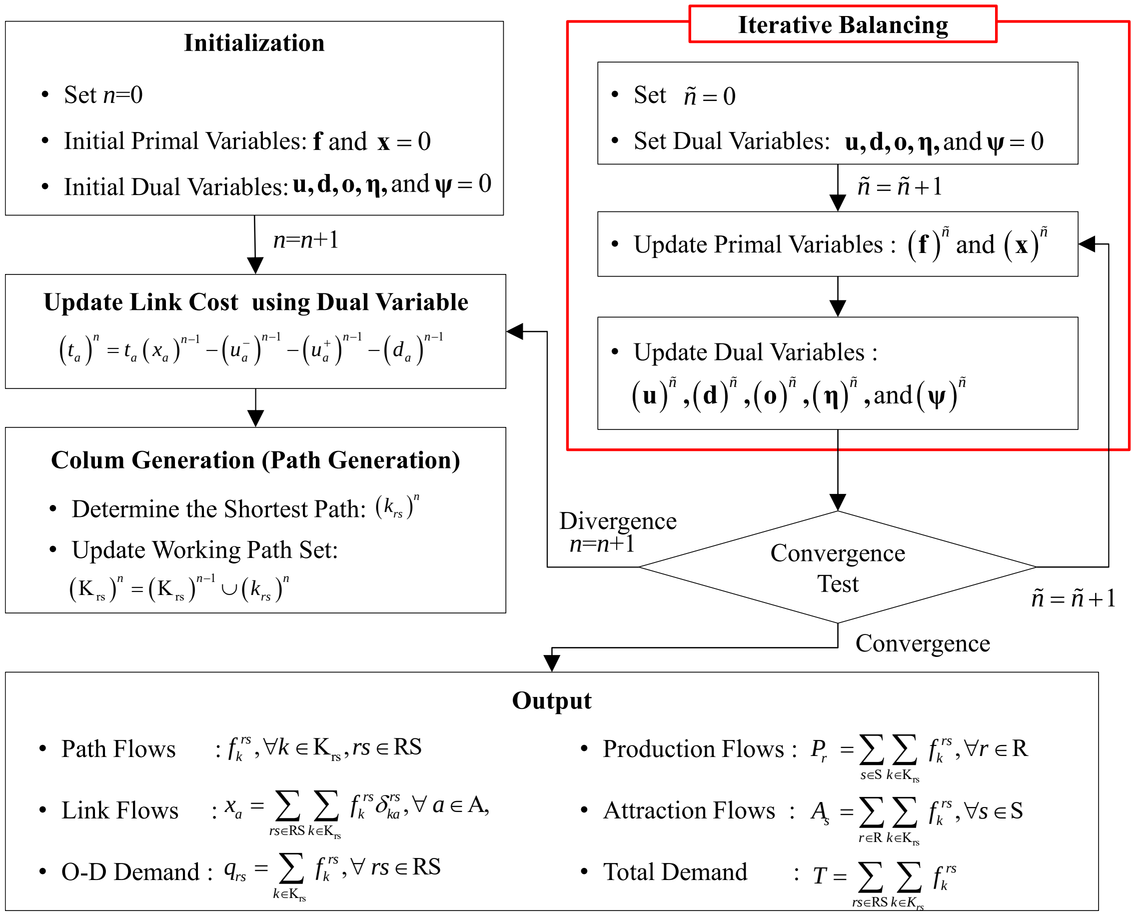


Figure 3 PFE solution procedure

4 CASE STUDY: UTAH STATEWIDE FREIGHT TRANSPORTATION NETWORK

This section presents numerical results to demonstrate the features of the proposed approach as well as its applications to the Utah statewide freight transportation network. The freight transportation network of Utah depicted in Figure 4 was extracted from the FAF[3] network. The network consists of 385 nodes, 944 links, and 2,256 O-D pairs. The study area consists of 29 counties and 19 external stations (i.e., entry and exit points around the state borders). The Wasatch Front Regional Council (WFRC), the major truck generation area in the state highlighted in Figure 4, consists of three major counties: Salt Lake, Weber, and Davis counties.

Truck traffic counts from 222 locations (about 23% of network links) were collected from the Utah Department of Transportation (UDOT) traffic map (UDOT traffic maps, 2013). The observations are mainly located on the major interstate freeways of Utah, such as I-15, I-70, I-80, and I-84 (see the interstate freeways in Figure 4). These major interstate freeways are the major truck routes for Utah, especially I-15, which runs north-south and passes through Salt Lake City and many other cities. Note that the freight demand derived from the FAF^3 database was based on the average annual daily truck traffic (AADTT), so link capacity values were required to replicate the daily equivalent capacity for a given link. To do so, we adopted the daily capacity conversion factors based on the functional class of the roadways. The capacity was then expanded by dividing the hourly capacity by the conversion factor and used for subsequent steps.

4.1 Results of Commodity-based truck O-D trip table

The estimation procedure described in Figure 2 was applied for the base year (2007) FAF commodity flow database. There are 29 internal zones or counties within Utah and 27 external stations. However, we found that only 19 external stations are used for truck traffic, hence the size of the trip table is 48×48 (i.e., 29 internal zones or counties within Utah and 19 external stations). Note that four types of statewide truck flows are estimated at this stage: (i) truck flows within Utah (Internal-Internal, I-I), (ii) truck flows from Utah to other states (Internal-External, I-E) or production flows, (iii) truck flows from other states to Utah (External-Internal, E-I) or attraction flows, and (iv) through truck flows (E-E), respectively.

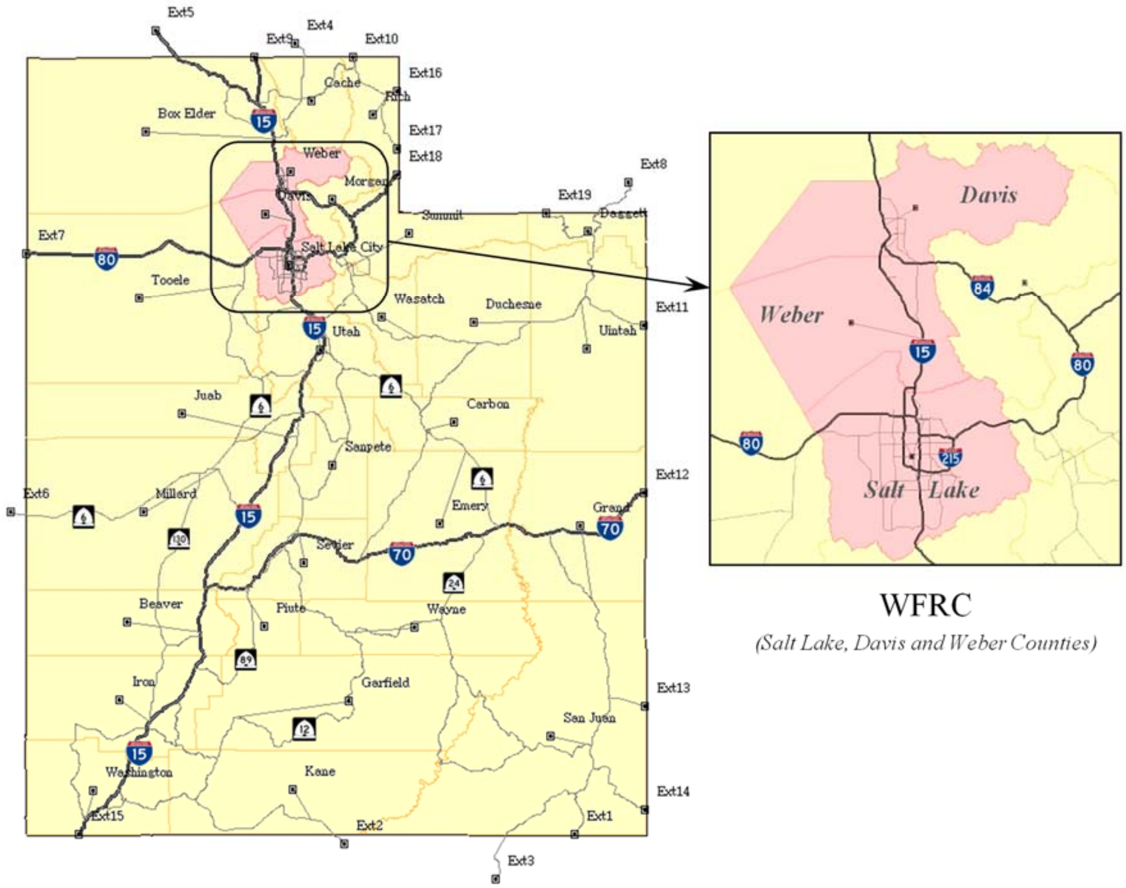


Figure 4 Utah statewide freight transportation network

Figure 5 depicts the truck O-D trip table for Utah (i.e., from all origins to all destinations). The total daily truck trips obtained from the first stage was 25,508 truck trips/day. Specifically, this total consists of 45.6% within Utah (I-I), 9.8% from Utah to other states (I-E), 11.0% from other states to Utah (E-I), and 33.6% through truck flows (E-E). We can observe the through truck flows (E-E) between I-80E and I-15N and between I-80E to I-80W are quite heavy. The highlighted bar series (in dark green) represent the production flows from the major counties along the Wasatch Front area such as Salt Lake, Cache, Weber, Davis, and Utah. The number of commercial truck trips estimated using the USTM commercial trips derived from Eq. (1), and the number of empty truck trips estimated from Eq. (2) are estimated and included to the commodity-based truck O-D trip table. At this stage, we found that the daily empty truck trips account approximately 40% of the commodity flows. If this component is not considered in the estimation, the total truck traffic and congestion in the study area could be significantly underestimated.

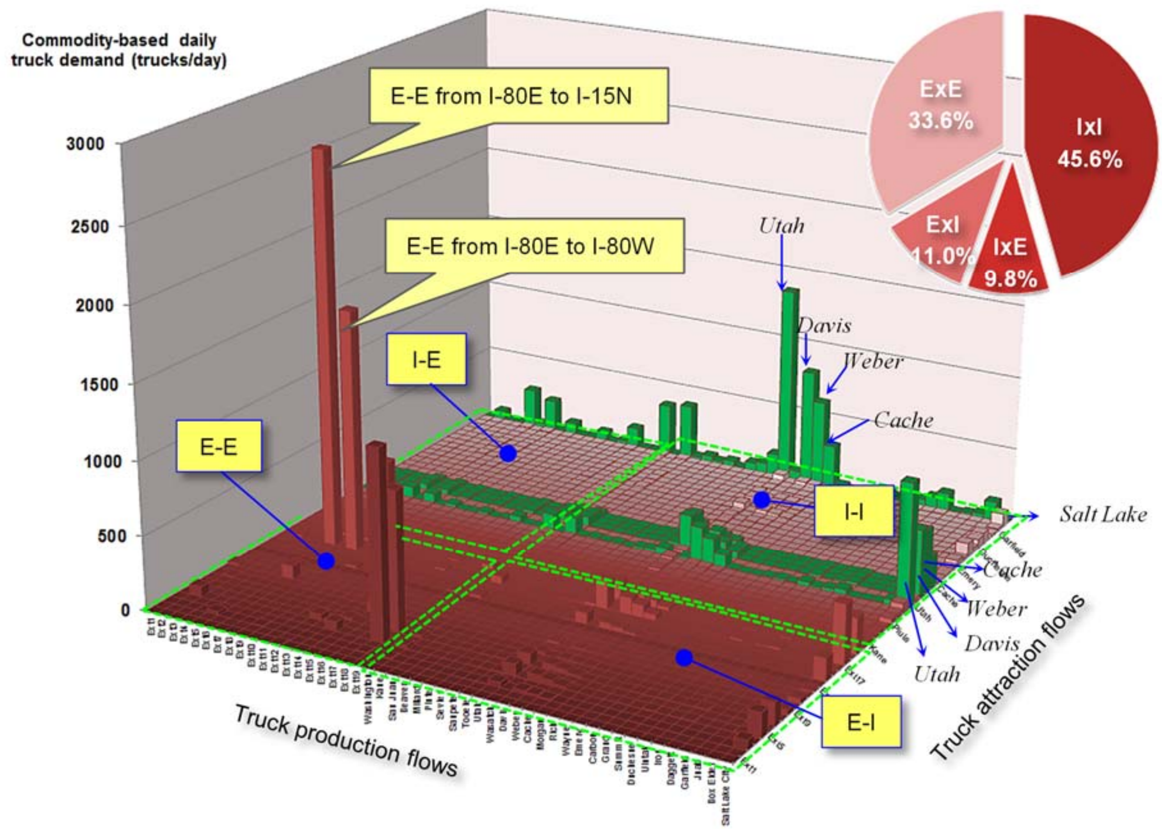
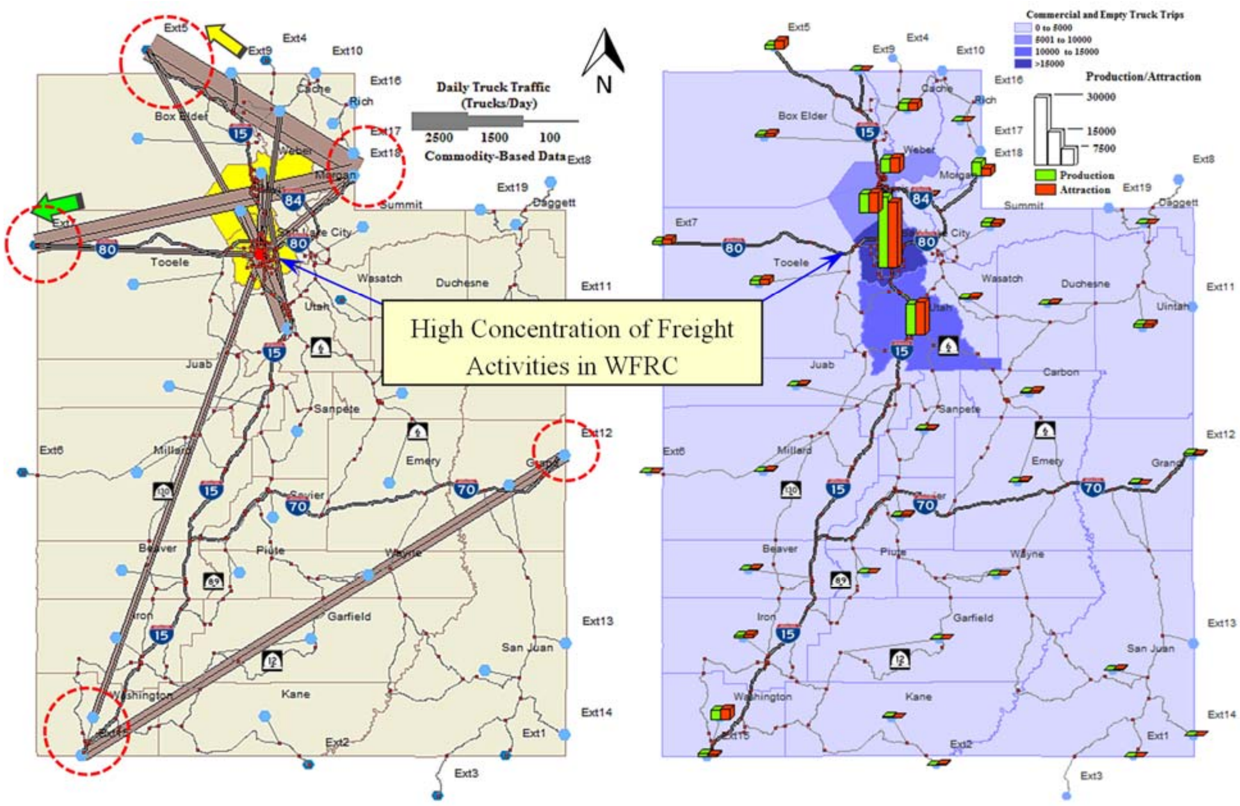


Figure 5 Commodity-based truck O-D trip table

Furthermore, we used the desire lines to highlight selected O-D pairs with high truck flows (i.e., greater than 500 trucks/day) in Figure 6(a). The circles in the figure show the entering and exiting truck flows at major external stations along the interstate freeways. We can observe high entering and exiting freight flows at the external stations: between I-15 South and I-70, I-80 East and I-15 North, I-80 West and I-80 East via I-15 near Salt Lake City and so on. These are the important interstate truck routes in Utah and are used for connecting the through trips from/to other states. The O-D flows were then aggregated to show the truck trip production and attraction flows at the county level as well as the external stations shown in Figure 6(b). As can be seen, truck trip production and attraction flows derived from the first stage are relatively concentrated around the WFRC area compared to other counties. Figure 6(b) reveals most commercial and empty truck trips are concentrated in the WFRC area and Utah County (shaded areas). This is to be expected because the major freight activities in Utah are mainly generated from these counties where warehousing and distribution centers are located. Overall, the truck flows within and through Utah are two major demand components as they account for almost 80% of total commodity flows transported in Utah.



(a) Commodity truck flows (desire lines)

(b) Total production/attraction flows

Figure 6 Estimated statewide commodity-based truck flows

4.2 Results of updated truck O-D trip table using PFE

The second stage used the path flow estimator (PFE) to refine the truck trip table obtained from the first stage using the truck counts from the statewide truck count program. Three different types of information, including truck counts, partial set of target O-D flows, target production and attraction flows (commodity and empty truck flows), were used to update the truck O-D trip table. Figure 7 depicts the scatter plots of observed and estimated link flows obtained from the two stage approach and compares to the one stage approach (i.e., the commodity-based truck O-D trip table from stage one). Note that the one stage approach assigns the commodity-based truck trips using an all-or-nothing (AON) traffic assignment procedure, which is a typical method used to preload trucks to the statewide network. Truck flows obtained from this method are usually assigned based on the shortest distance or travel time, and there is no consideration of congestion. As can be seen, the results obtained from this method are underestimated, especially in the WFRC area and the high freight activity locations in the study area (i.e., areas around Salt Lake City International Airport and perimeters of the Salt Lake County). Consequently, such issues often hinder the statewide planning as it is incapable of capturing the freight movements under congestion and of explaining the truck traffic variations in the urban areas mentioned above. On the other hand, the two stage approach using PFE to refine the commodity-based truck O-D trip table with truck counts, target O-D flows, and target production and attraction flows can provide a more reasonable match between the observed and estimated values. The majority of the observations are within an acceptable tolerance with a few points outside of the error bound. Figure 8(a) shows the complete truck flow pattern on the statewide network based on the two-stage approach. The figure reveals a high concentration of truck traffic on I-15 around the Salt Lake County, and I-80W in the Summit County and Salt Lake County. To highlight the congested links, Figure 8(b) shows the volume to capacity (V/C) ratios. As can be seen, many interstate and state routes (e.g., I-15, interchange to I-215, SR 130, SR 12, and SR 6) are quite congested.

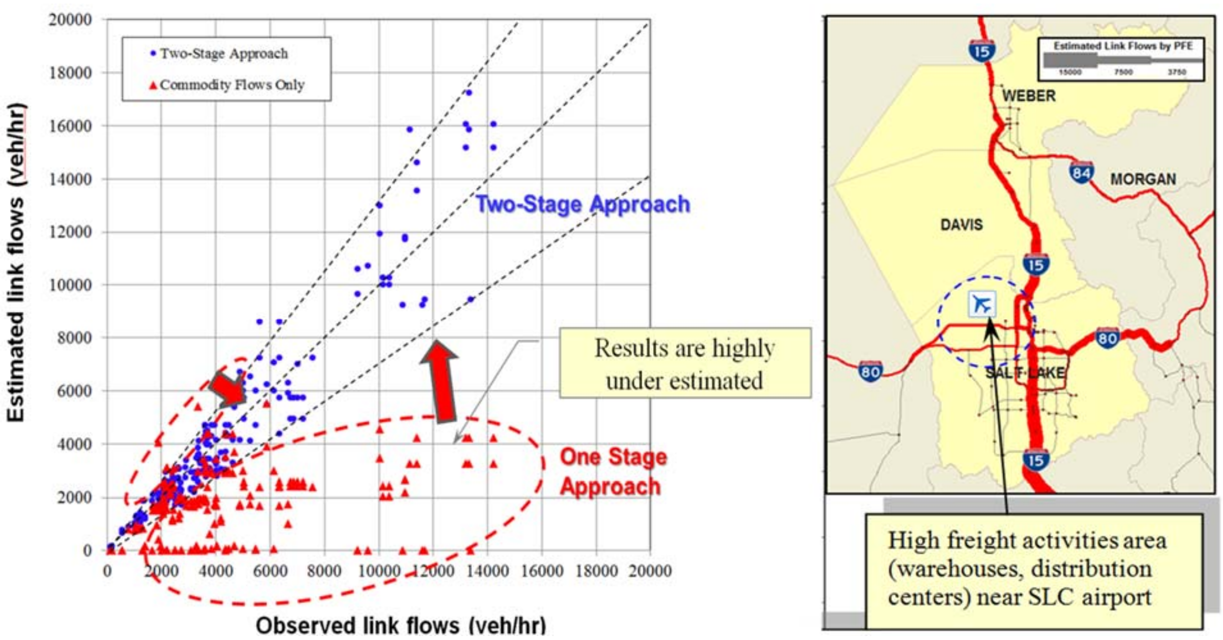
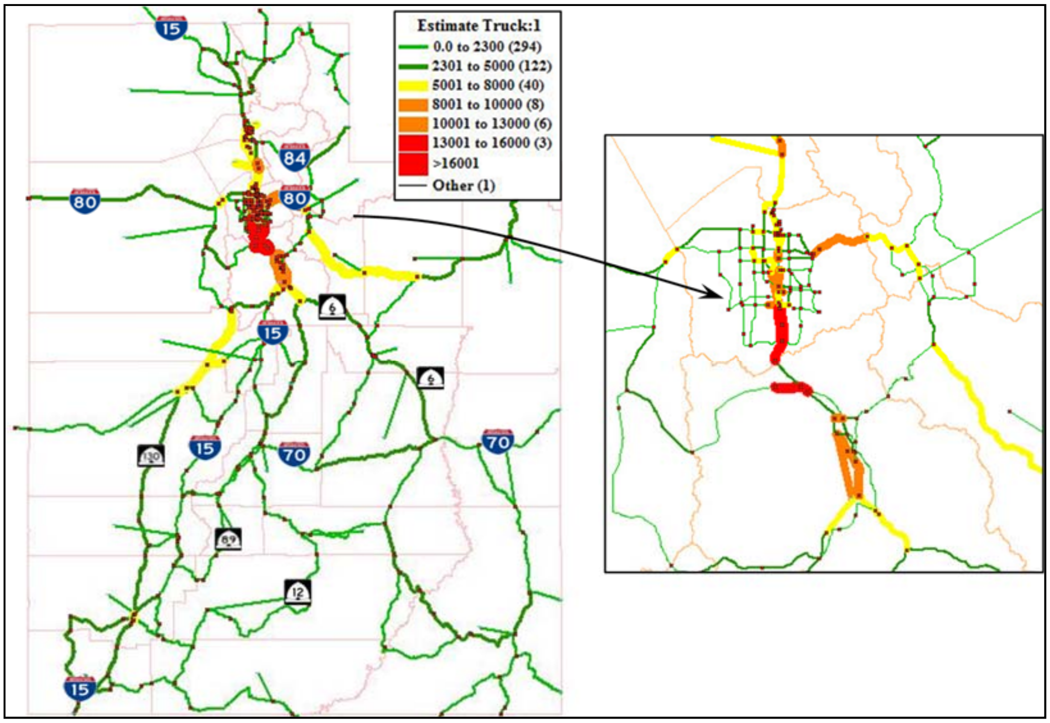
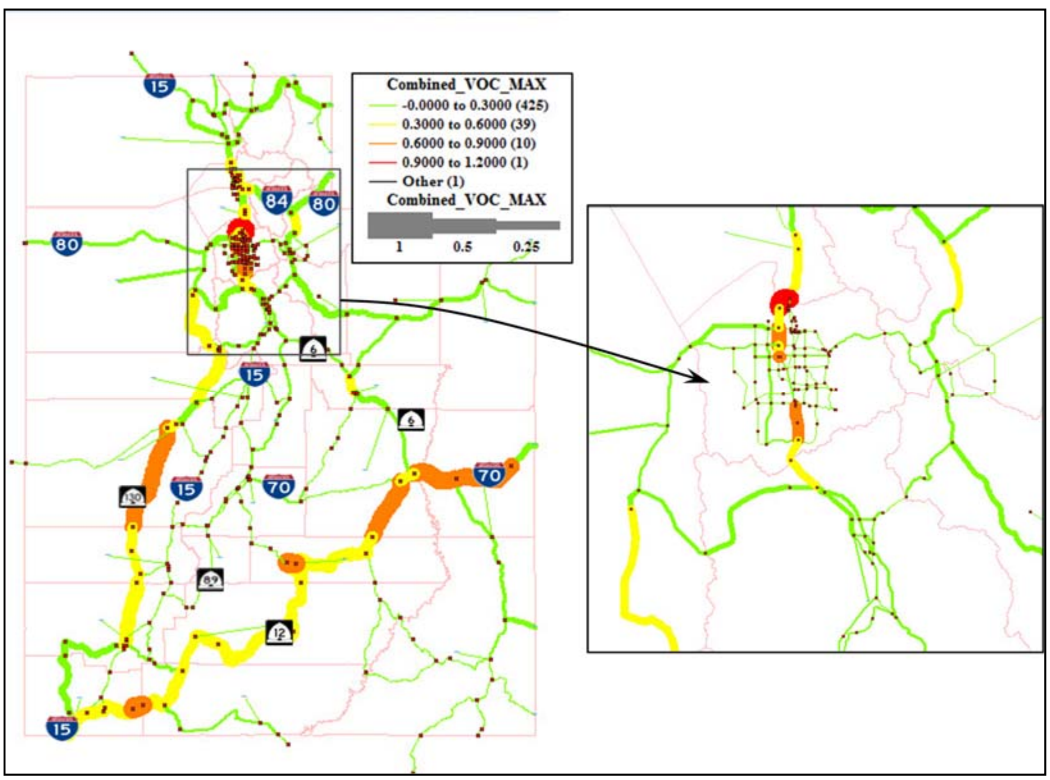


Figure 7 Comparison between one stage and two stage approaches



(a) Statewide truck traffic (AADTT)



(b) VC ratios

Figure 8 Statewide truck traffic and volume/capacity analysis

4.3 Effect of spatial constraints

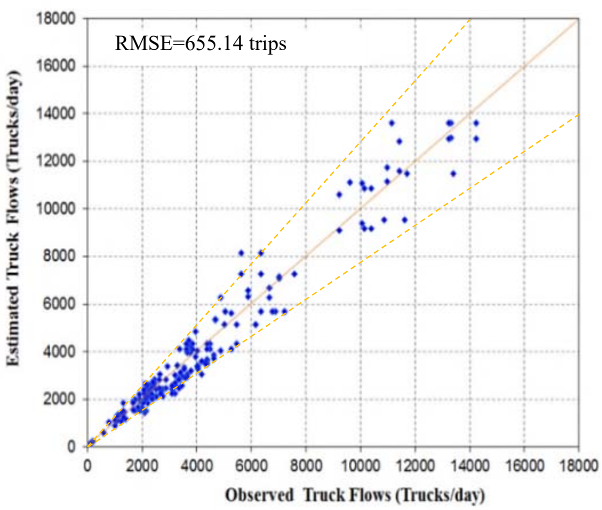
In this section, two cases are considered for assessing the effect of including spatial constraints in the PFE:

- Case 1: PFE with truck counts only, and
- Case 2: PFE with truck counts with zonal production and attraction flow constraints derived from the first stage.

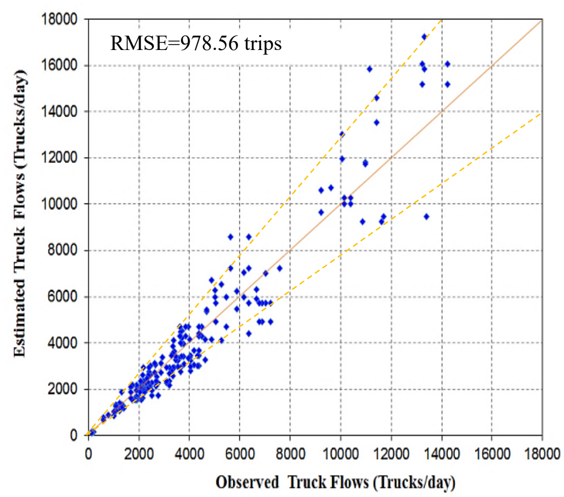
Accuracy of the estimates can be measured by the root mean square error (RMSE) as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}\left(x_{est}^{n} - x_{obs}^{n}\right)^{2}} \tag{18}$$

where N is the number of observations, x_{est}^{n} and x_{obs}^{n} are the estimated and observed truck flows, respectively. Figure 9(a) a and Figure 9(b) depict the scatter plots of observed and estimated link flows and estimated trip production for these two cases.



(a) Truck count only (case 1)



(b) Truck count only + commodity based data (case 2)

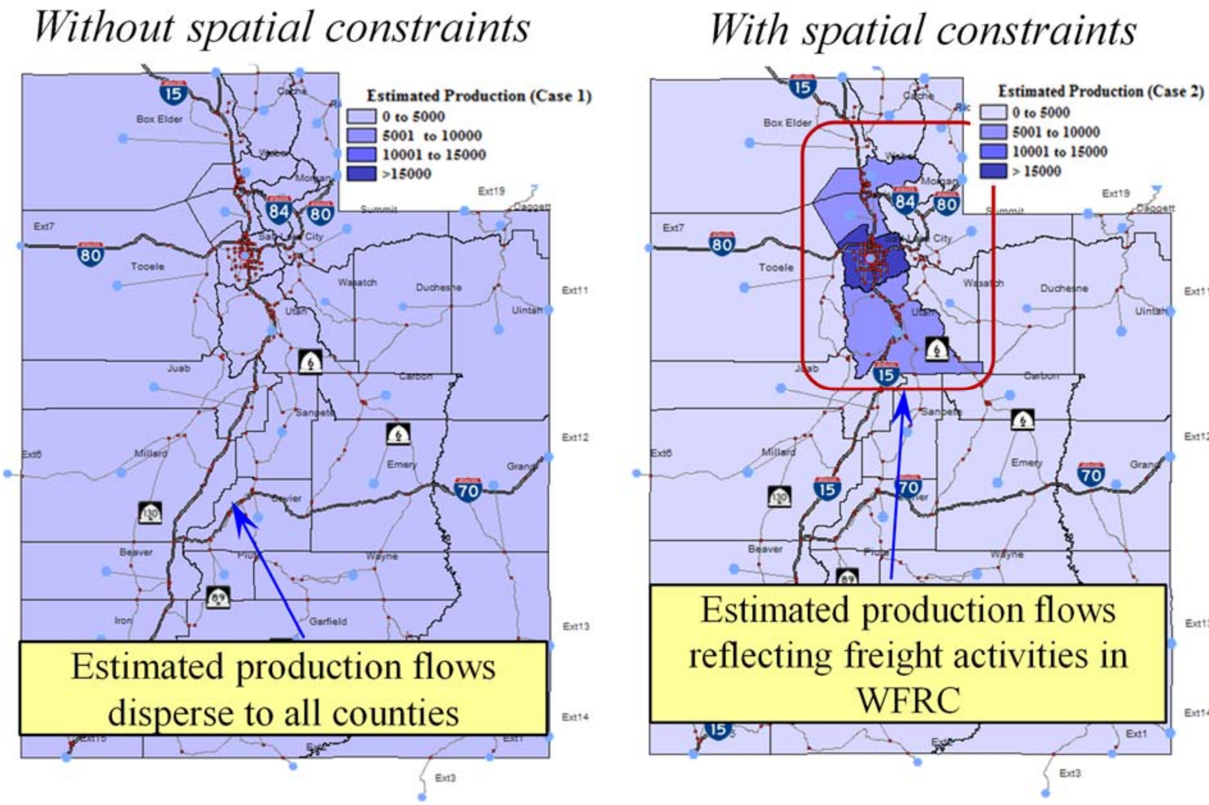
Figure 9 Comparisons of observed and estimated statewide truck flows

The results show the truck trip table estimated by PFE produces a fairly good match for both cases (i.e., case 1: RMSE= 655.14 trucks/day, R^2=0.9562, MAPE=13.06%; case 2: RMSE=

978.56 trucks/day, R^2=0.9228, MAPE=14.62%). It should be noted that the RMSE indicates the aggregated quality of O-D estimates. A smaller value indicates a higher quality of the estimation process. Between the two cases, including spatial constraints into the estimation slightly deteriorates the matching of truck counts as indicated by the higher RMSE. This is compensated by the better estimates of zonal production and attraction flows. The estimated total demand of Case 1 is approximately 38% less than the total demand estimated from the first stage. This highlights the importance of including the spatial constraints into the PFE model, which can better capture the total demand in Case 2 (i.e., slightly over 6%). However, we still observe that Case 2 underestimates some link flows, especially those links with high truck flows on I-15 near Salt Lake City.

This is because those links are located closed to areas with a higher level of freight activities near the Salt Lake City International Airport. This is the concentrated area with high truck traffic accessing to/from the shipping companies and intermodal facilities such as rail-truck and air-truck modes. Special generators of truck trips from surveys of high freight density areas such as warehouses and freight distribution centers were added to resolve this issue. From the modelling point of view, these special generators can be implemented in the PFE framework as they are handled by the zonal production and attraction constraints (in Eqs. (7), and (8)) similar to the commercial and empty truck trips. Further, one of the possible ways to overcome the shortfalls is to further develop the tour-based distribution particularly for the finer spatial resolution like the urban area. A larger data set (e.g., truck trip chain surveys) is needed for modeling the vehicle tours. More factors such as transshipment nodes, commodity types, truck types, touring distances, and time window constraints should be considered in the future to overcome the behavioral limitations of the proposed model.

Figure 10(a) and Figure 10(b) depict the truck production flows for Case 1 and Case 2, respectively. From these two figures, we can observe the trip productions in Case 2 are more distributed when the spatial constraints are considered in the estimation process. By adding zonal production and attraction flows as constraints in Case 2, it can improve the observability of the trip generation pattern. Thus, this emphasizes the importance of using a two-stage approach to capture both the commodity flows and truck counts in the field, so that the statewide truck flow pattern can better reflect the reality.



(a) Estimated trip production (case 1) (b) Estimated trip production (case 2)

Figure 10 Comparisons of estimated production flows

4.4 Truck Corridor Analysis

This section further provides the truck corridor analysis. In Utah, I-15 is a primary corridor for both passenger and freight movements. The truck corridor serves as a backbone route for truck movements of agricultural and energy (i.e., oil, gas, and coal) products in the southern Utah region and onward to major cities in the state such as Provo, Salt Lake City, and Ogden. Additionaly, the I-15 corridor also helps to connect the through truck traffic as part of the CANAMEX corridor. Figure 11(a) depicts the daily truck traffic flows on the I-15 corridor. Figure 11(b) and Figure 11(c) show additional details of the truck flow profile starting from the northern border (from Idaho) to the southen border (to Arizona) and the corresponding daily truck V/C ratios.

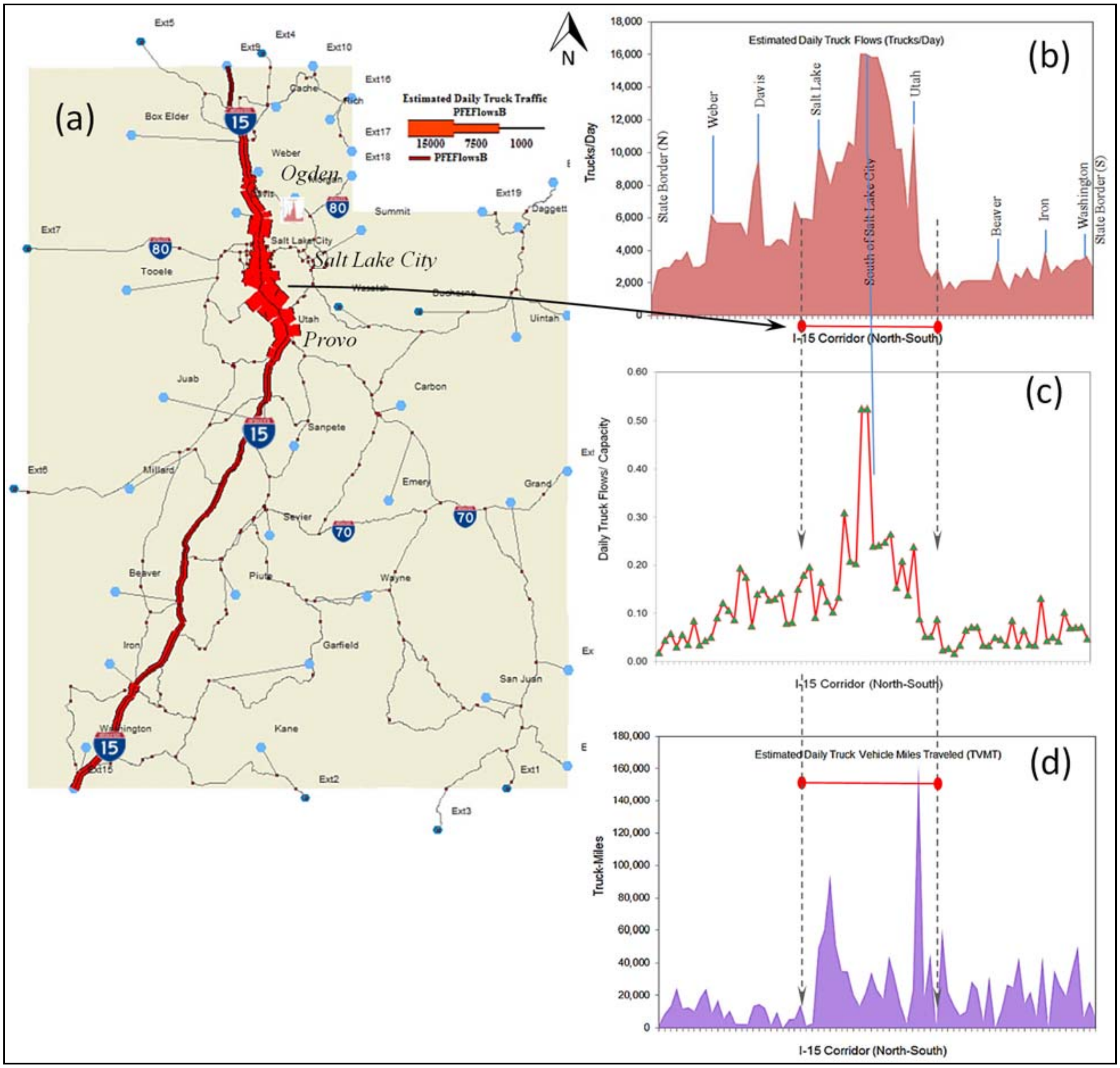


Figure 11 Estimated truck flows and truck vehicle miles traveled on I-15 corridor, Utah

As expected, the heavily used truck links are in the WFRC area, especially the links near Salt Lake City and its pheripheral urbanized areas such as Weber County, Davis County, and Utah County. The most congested link carries a daily truck traffic of 16,058 trucks/day with an AADT of 34,634 passenger cars/day or about 30% of this segment are truck traffic. Additionally, in this area, the daily truck flow to capacity ratios range between 0.3 to 0.5. The most congested link is about 0.52 indicates that truck traffic highly contributes to the congestion on this particular link in the urban areas. Figure 11(d) depicts the daily truck vehicle mile traveled (TVMT) for this corridor.

The daily TVMT is calculated based on the truck travel distance and the daily truck flows estimated from the two-stage approach. As can be seen, the TVMT in Salt Lake county is lower than those of Davis and Utah counties. The major reason is that higher truck flows can travel a longer distance in those counties, while a similar amount of truck flows can travel a shorter distance within the Salt Lake county. This suggests that these links could have higher congestion, which could lead to stop and go traffic conditions around this area.

5 CONCLUSIONS

This study has developed a two-stage approach for estimating truck O-D trip table using both commodity flows and truck counts data. The model is supported by two sequential stages: Stage 1 estimates the commodity-based truck O-D trip tables primarily derived from the commodity flow database, while Stage 2 uses the path flow estimator (PFE) to refine the truck trip table to better match the observed truck counts.

In the first stage, we have developed a simplified procedure to estimate a commodity-based truck trip table using the commodity flows from the newly released FAF^3. The FAF^3 provides commodity flow estimates based on tonnage and value by commodity type, mode, origin, and destination for 2007, and forecasts through 2040. It is publicly accessible from the Freight Management and Operations Database from the FHWA website. This stage considers intrastate, interstate, and through truck flows. To accomplish this, it involves four tasks: i) extract state-specific commodity flows from FAF^3, ii) conduct subarea analysis to estimate through truck flows iii) disaggregate the state-specific to county-specific commodity flows, and iv) convert the commodity flows into truck trips.

The second stage adopts the path flow estimator (PFE) to refine the truck trip table obtained from the first stage by using the up-to-date truck counts. The basic idea is to find a set of path flows that can reproduce the observed truck counts collected from permanent count station locations within the state and state borders from the statewide truck count program, while preserving the spatial distribution of the O-D commodity flow pattern obtained from the first stage. To enhance the observability of the truck O-D trip table, additional planning data such as production and attraction flows are included in the PFE estimation. Validation of the results of PFE is assessed by the accuracy of the assignment estimates measured by the root mean square error (RMSE), mean absolute percentage error (MAPE), and R^2 between the estimated and observed truck counts.

The flexibility of aggregating path flows at different spatial levels in PFE allows us to makes use of various existing data (e.g., truck counts, production and attraction commodity flows, truck VMT at the state level, etc.) and commodity-based data with commercial and empty truck trips for estimating the statewide truck trip table. The proposed approach can be also used to conduct the truck corridor analysis to determine the congested links and potential bottlenecks. Although the results using Utah as a case study are satisfactory, accurate and consistent truck counts are required in the PFE to produce reliable results. Extending the PFE to handle inconsistent traffic counts at the statewide level should be explored (see Chen *et al*., 2009, 2010). Constraints such as trip length frequency distribution is needed to model different types of statewide truck traffic (i.e., short-haul, long-haul, and empty truck trips) in PFE. Hence, further work should consider multiclass and multimode (e.g., commercial, single-, multiple-unit trucks, and passenger cars) (e.g., Yang and Huang, 2004; Marcotte and Wynter, 2004; Wong *et al*., 2005), so that it can better reflect the actual congestion of the statewide network. Truck surveys at freight companies and distribution centers for each county and state border (e.g., Weigh-in-motion (WIM), Port of Entry (POE) stations) should be conducted to understand the freight movements in the statewide network. The current truck O-D trip table is estimated from the commodity flow data from FAF and truck counts collected by the Utah Department of Transportation (UDOT). It should be updated using the newly developed Utah Statewide Travel Model (USTM) to improve the accuracy and quality of the truck O-D trip table. Finally, the two-stage method developed in this paper, particularly the PFE formulation in the second stage, only estimates the current O-D truck trip table based on both commodity flows and truck counts of the base year. To forecast the truck O-D trip table would require developing a future PFE formulation that can accept forecast commodity flow data (i.e., available from FAF up to 2045), future socio-economic and land use data, and the estimated/calibrated base year truck O-D trip table as constraints since observed truck counts are not available in the future. The idea is similar to those developed in PFE for planning applications in small communities (Ryu et al., 2014) and for small metropolitan planning organizations (Jansuwan et al., 2012).

ACKNOWLEDGMENT

The authors are grateful to two anonymous references and three guest editors for their constructive comments and suggestions to improve the quality and clarity of the paper. This research was supported by the Mountain-Plains Consortium (MPC), the Utah Transportation Center at Utah State University, the Hong Kong Polytechnic University, and the Chang Jiang Chair Professorship funded by the Ministry of Education in China. The contents of this paper reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein and do not necessarily reflect the views of our sponsor.

REFERENCES

Al-Battaineh, O., Kaysi, I., A., 2005. Commodity-based truck origin-destination matrix estimation using Input-Output data and genetic algorithms. In Transportation Research Record: Journal of the Transportation Research Board 1923, 37-45.

Bell, M.G.H., Iida, Y., 1997. Transportation Network Analysis. John Wiley & Sons.

Bell, M.G.H., Shield, C.M., 1995. A log-linear model for path flow estimation. In Proceedings of the 4th International Conference on the Applications of Advanced Technologies in Transportation Engineering, Carpi, Italy, 695-699.

Bell, M. G. H., Shield, C. M., Busch, F., Kruse, G., 1997. A stochastic user equilibrium path flow estimator. *Transportation Research* Part C 5 (3-4), 197-210.

Bureau of Transportation Statistics (BTS), 2004. Freight shipments in America. U.S. Department of Transportation. Available via: http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/freight_shipments_in_america/pdf/entire.pdf Accessed February 5, 2012.

Cambridge Systematics, 2002. Freight impacts on Ohio's roadway system. Report prepared for Ohio Department of Transportation (OH-2002/026).

Cambridge Systematics, 2005. An initial assessment of freight bottlenecks on highway. Report prepared for Office of Transportation Policy Studies, Federal Highway Administration (FHWA).

Cambridge Systematics, 2007. Quick Response Freight Manual II. Report prepared for Federal Highway Administration (FHWA).

Cambridge Systematics, 2010. Travel Model Validation and Reasonableness Checking Manual.

Cambridge Systematics, 2013. TRB's National Cooperative Freight Research Program (NCFRP) Report 26: Guidebook for Developing Subnational Commodity Flow Data. The National Academies Press. Washington D.C.

Cambridge Systematics and Eastern Research Group and Alliance-Transportation Group Inc., 2004. TxLED VMT Estimation Project. Report prepared for the Texas Commission on Environmental Quality.

Chen, A., Chootinan, P., Recker, W., 2005. Examining the quality of synthetic origin-destination trip table estimated by path flow estimator. Journal of Transportation Engineering 131(7), 506-513.

Chen, A., Chootinan, P., Recker, W., 2009. Norm approximation method for handling traffic count inconsistencies in path flow estimator. Transportation Research Part B 43(8), 852-872.

Chen, A., Ryu, S., Chootinan, P., 2010. $L\infty$-norm path flow estimator for handling traffic count inconsistencies: Formulation and solution algorithm. ASCE Journal of Transportation Engineering136(6), 565-575.

Crainic, T.G., Dufour, G., Florian, M., Larin, D., Leve, Z., 2001. Demand matrix adjustment for multimodal freight networks. In Transportation Research Record: Journal of the Transportation Research Board 1771, 140-147.

de Dios Ortuzar, J., Willumsen, L.G., 2002. Modeling transport, 3rd Edition, Wiley, New York.

de Jong, G., Ben-Akiva, M., 2007. A Micro-simulation model of shipment size and transport chain choice. Transportation Research Part B 41(9), 950–965.

Federal Highway Administration (FHWA), 1990. Calibration and Adjustment of System Planning Models.

Federal Highway Administration (FHWA), 2008. Freight Story: A National Perspective on Enhancing Freight Transportation. Available via: http://ops.fhwa.dot.gov/freight/publications/fhwaop03004/freight.pdf Accessed February 5, 2013.

Federal Highway Administration. Freight Analysis Framework (FAF), 2009. Freight Management and Operations. U.S. Department of Transportation. Federal publication. Available via: http:// www.ops.fhwa.dot.gov/freight/freight_analysis/faf/ Accessed February 5, 2013.

Fischer, M.J., Outwater, M.L., Cheng, L.L., Ahanotu, D.N., Calix R., Cambridge Systematics, Inc., 2005. An innovative framework for modeling freight transportation in Los Angeles County. Report prepared for Los Angeles County Metropolitan Transportation Authority.

Holguín-Veras, J., List, G., Meyburg, A., H., Paaswell, R., E., Teng, H., Yahalom, S., 2001. An assessment of methodological alternatives for a regional freight model in the NYMTC Region. Report prepared for New York Metropolitan Transportation Council (NYMTC).

Holguín-Veras, Patil, G., 2008. A multicommodity integrated freight origin-destination synthesis model. Networks and Spatial Economics 8(2-3), 309-326.

Holguin-Veras, J., Thorson, E., Zorrilla, J.C., 2010. Commercial vehicle empty trip models with variable zero empty trip probabilities. Networks and Spatial Economic, 10(2), 241-259.

Holguín-Veras, J., Thorson, E., 2000. Trip length distributions in commodity-based and trip-based freight demand modeling. In Transportation Research Record: Journal of the Transportation Research Board 1707, 37-48.

Horowitz, A., 2006. Statewide travel forecasting models. National Cooperative Highway Research Program Synthesis 358, Transportation Research Board, Washington, DC.

Jansuwan, S., Chen, A., Ryu, S., 2012. An alternative planning tool for a small metropolitan planning organization in Utah. *Transportation Research Record* 2307, 68-79.

Liedtke, G, 2006. An actor-based approach to commodity transport modelling. PhD thesis, Institute for Economic Policy Research (IWW), University of Karlsruhe, Germany.

List, G., Konieczny, L., Durnford, C., Papayanoulis, V., 2002. Best-practice truck-flow estimation model for the New York City Region. In Transportation Research Record: Journal of the Transportation Research Board 1790, 97-103.

List, G.F., Turnquist, M.A., 1994. Estimating truck travel patterns in urban areas. In Transportation Research Record: Journal of the Transportation Research Board 1430, 1-9.

Ma, Y., Kuik, R., van Zuylen, H., 2013. Day-to-day origin-destination tuple estimation and prediction with hierarchical Bayesian networks using multiple data sources. *Transportation Research Record: Journal of the Transportation Research Board* 2343(1), 51-61.

Marcotte, P., Wynter, L., 2004. A new look at the multiclass network equilibrium problem. Transportation Science 38(3), 282-292.

Oak Ridge National Laboratory, 2011. FAF Freight Traffic Analysis. Technical Report. Available at: http://faf.ornl.gov/fafweb/Data/Freight_Traffic_Analysis/faf_fta.pdf Accessed January 5, 2013

Ogden, K.W., 1992. Urban goods movement, a guide to policy and planning. Ashgate, Surrey, U.K.

Ruan, M., J. Lin, K. Kawamura, 2012. Modeling urban commercial vehicle daily tour chaining. Transportation Research Part E 48, 1169-1184.

Ryu, S., Chen, A., Zhang, H.M., Recker, W., 2014. Path flow estimator for planning applications of small communities. *Transportation Research Part A* 69, 212-242.

Sánchez-Díaz, I., Holguín-Veras, J., Ban, X. 2015. A time-dependent freight tour synthesis model. *Transportation Research Part B* 78(1), 144-168.

Sorratini, J., Smith, R., 2000. Development of a statewide truck trip forecasting model based on commodity flows and input-output coefficients. In Transportation Research: Journal of Transportation Research Board Record 1707, 49-55.

Stefan, K.J., McMillan, J.D.P., Hunt, J.D., 2005. Urban commercial vehicle movement model for Calgary, Alberta, Canada. In Transportation Research Record: Journal of the Transportation Research Board 1921, 1-10.

Tamin, O.Z., Willumsen, L.G., 1989. Transport demand model estimation from traffic count. Transportation 16(1), 3-26.

Transportation Research Board, 2000. Highway Capacity Manual. Special Report No. 209.

U.S. Bureau of Economic Analysis, Utah Department of Work Force Services, 2000. State of Utah employment projections by county and multi-county district 2000-2030.

U.S. Census Bureau, 2010. Households and families. Available via: www.census.gov/prod/cen2010/briefs/c2010br-14.pdf. Accessed Jul. 15. 2013.

U.S. Census Bureau, 2012. Census of governments: organization component preliminary estimates. Available via:
www.census.gov/govs/cog2012 Accessed Jul. 15. 2013.

Utah Department of Transportation (UDOT) Traffic maps for 2008. www.udot.utah.gov. Accessed Jan. 15, 2013.

Utah Department of Workforce Services, 2008. Utah nonagricultural employment. Available via https://jobs.utah.gov. Accessed Jul. 15. 2013.

Wilbur Smith Associates and Resource Systems Group, Inc., 2009. Utah statewide travel model, final model validation report. Report prepared for Utah Department of Transportation.

Wisconsin Department of Transportation, 1995. Translinks 21 Technical Report Series: Multimodal Freight Forecasts for Wisconsin, Draft No. 2.

Wisetjindawat, W., Sano, K., and Matsumoto, S. Commodity distribution model incorporating spatial interactions for urban freight movement. Proceeding of the 85th Annual Meeting of the Transportation Research Board. CD-ROM Transportation Research Board, 2006.

Wong, S.C., Tong, C., Wong, K., Lam, W., Lo, H. Yang, H., Lo, H., 2005. Estimation of multiclass origin-destination matrices from traffic counts. Journal of Urban Planning and Development 131(1), 19-29.

Wong, S.C., Yang, H., 1997. Reserve capacity of a signal-controlled road network. Transportation Research Part B 31(5), 397-402.

Yang, H., Huang, H., 2004. The multi-class, multi-criteria traffic network equilibrium and systems optimum problem. Transportation Research Part B 38(1), 1-15.

Zhang, Y., Bowden, R.O., Allen, A.J., 2003. Intermodal freight transportation planning using commodity flow data. Report prepared for the National Center of Intermodal Transportation (NCIT), the Mississippi Department of Transportation (MDOT).