# Weight-Categorized Truck Flow Estimation: A Data-fusion Approach and a Florida Case Study

Dongfang Zhao[a], Suryaprasanna Kumar Balusu[a], Parvathy Vinod Sheela[a], Xiaopeng Li[a,1],
Abdul Rawoof Pinjari[b], Naveen Eluru[c]

a. Department of Civil and Environmental Engineering
University of South Florida, FL 33620

b. Department of Civil Engineering, Center for Infrastructure, Sustainable Transportation and Urban Planning (CiSTUP)
Indian Institute of Science (IISc), Bangalore, India 560012

c. Department of Civil, Environmental and Construction Engineering
University of Central Florida, FL 32816

**Abstract**

Knowledge of spatial distributions of weight-categorized truck flows in a region is critical to the understanding of movements of empty or partially-loaded trucks and devising appropriate strategies to reduce empty or partially-loaded truck flows and improve truck utilization efficiency in the region. However, such disaggregated information cannot be directly obtained from existing data sources and models. In this paper, we propose a compact model for estimating weight-categorized truck origin-destination (OD) flows and link-level truck counts by fusing several freight datasets. The proposed model minimizes the squared errors between the estimated and observed truck OD flows and link volumes considering the flow conservation of trucks and commodity weights. To illustrate a real-world application of this model, a case study is conducted to estimate the spatial distribution of empty or partially-loaded truck flows into, within, and out of the State of Florida. With the case study results, high production- and attraction-zones of empty or partially-loaded truck trips are also identified. Such results can potentially inform freight planning and policy decisions to learn spatial patterns of empty or partially-loaded backhauling truck flows and devise countermeasures to reduce such flows and improve freight transportation efficiency. This is particularly relevant to a terminal state such as Florida that has large volumes of back-hauling truck flow.

---

[1] Corresponding author. Tel: +1-813-974-0778; Email: xiaopengli@usf.edu.

## 1. Introduction

Trucks backhauling from areas with a significant imbalance in the consumption and production of goods comprises a notable proportion of empty or partially-loaded trucks. According to the National Private Truck Council, approximately 25% of the vehicle miles traveled by the private trucks in the United States correspond to empty truck movements (White paper on backhaul networking, 2007). The empty or partially loaded backhauling trucks cause significant loss to the trucking industry in terms of wasted fuel, workforce, time and other resources. Matching loads for at least 5% of the empty trucks, which constitute more than 100,000 non-revenue generating trucks, could save at least 25 million gallons of diesel fuel (Sheckler et al., 2009). Therefore, the estimation of empty or partially loaded truck counts can help inform strategies to reduce empty or partially-loaded backhauls through balancing trade activities across the region, e.g., development of production centers and attraction of imports to the region's sea ports.

In the context of estimating the spatial distribution of truck counts, a number of efforts have been made to model estimate commodity flows and truck trips. As summarized in Hautzinger (1984), there are mainly three models until 1984 including the naïve proportionality model, the formulation by Noortman and vanEs (1978), and the one presented in his paper. In the naïve proportionality model, truck trips are modeled as a direct function of the commodity flows. However, as discussed in Hautzinger (1984), the model is limited especially when there are inconsistencies between commodity flows and truck trips. Noortman and van Es (1978) estimate the number of empty trips between OD pairs as a function of the commodity flow in the opposing direction, which however only considers two types of vehicles (empty and fully-loaded) instead of more detailed multiple weight categories. Both models only consider empty trips back to the home base instead of complex trip chains.

In order to formally characterize commercial vehicle trip chain models, Holguin-Veras and Thorson (2003a) discuss some practical insights to the challenges involved in modeling the empty trips using traditional freight demand models. Although capable of estimating both the empty and loaded truck trips, vehicle trip-based models cannot account for the economic characteristics of the cargo, which are critical in the decision making, as discussed in Holguin-Veras and Jara-Diaz (1999), Roorda et al. (2010), McFadden et al. (1986), Bernardin et al. (2015) and Holguin-Veras (2002). On

the other hand, commodity- based models cannot estimate the empty truck counts accurately since they ignore the fact that the logistic decisions cannot be directly explained using commodity flows when empty trips are prevailing (Holguin-Veras and Thorson, 2003a). Other studies (Holguin-Veras and Thorson, 2003a; Holguín-Veras and Thorson, 2003b; Holguín-Veras et al., 2010; Jansuwan et al., 2017) have attempted to use statistical models where the empty flows were modeled as a function of loaded truck counts. Holguín-Veras and Patil (2008), integrate a commodity-based demand model based on a gravity model with a statistical model estimating empty trips to develop a freight origin–destination synthesis that includes both loaded and empty truck trips. Studies based on such statistical models would however require extensive data collection efforts such as OD surveys (Mesa-Arango et al., 2013), which may not always be feasible in practice.

These issues can be addressed by integrating distinct datasets by formulating a hybrid approach using optimization techniques (Jansuwan et al., 2017). Some of the early works in this area are done by Crainic et al. (1993) and Crainic and Laporte (1997). Later, Mesa-Arango et al. (2013) formulate an optimization function to minimize the overall system cost while ensuring the truck flow conservation for both loaded and empty trips. Additionally, Guelat et al. (1990) and Chow et al. (2014) propose a nonlinear inverse optimization technique for the freight assignment at different network equilibrium conditions. Besides, four-step models (McNally (2000), Giuliano et al. (2010), Agrawal et al. (2018)) are frequently used to estimate freight flows or trip distribution over a transportation network. However, they have not considered the weight categories of trucks and detailed truck trip-chains.

There is a considerable array of freight modelling works estimating the truck flows for two weight categories, i.e. empty and full truck loads (Kulpa, T., 2014; Middela et al., 2018). This study adds to the existing freight literature by exploring into the estimation of truck flows for more than two weight categories, such as empty, partially filled and full truck loads. Classifying trucks into multiple weight categories provides more detailed and quantitative information of both vehicle distributions and commodity distributions over the time space. It potentially enables more flexible and versatile strategies of managing and consolidating truck commodities (e.g., consolidating several partially filled trucks into fewer truckloads to save empty trips and associated energy consumption and costs). In the model, we propose an optimization procedure to estimate origin-destination (OD) matrices of truck counts (OD-level truck counts) and link-level truck counts in several truck weight categories. In doing so, this paper proposes and demonstrates an easy-to-use method for integrating

various available freight datasets for a region into useful freight data products. The fused datasets include link-level truck counts from traffic monitoring and measurement programs such as Department of Transportation, OD level truck counts from transportation research institutes, commodity flows between OD pairs typically available at a county-level (or similar) resolution in the US and other countries, and the path flows for truck counts from the assignment stage in a four-step freight demand model. The proposed data fusion approach is used to derive OD-level and link-level truck counts in weight categories at a county-level spatial resolution in the State of Florida. Furthermore, the study utilizes a spatial disaggregation procedure proposed by Holguín-Veras and Patil (2008) to disaggregate the estimated empty truck counts between OD pairs at the county level into a finer spatial resolution.

In the remainder of this paper, Section 2 describes the proposed optimization model for fusing alternative freight data to estimate OD-level and link-level truck counts in different weight categories. This section also presents a simple method for disaggregating the estimated county-level empty truck counts between OD pairs to a finer TAZ (Traffic Analysis Zone) level. Section 3 applies the model and the associated disaggregation method to a case study for Florida and presents a comparison of the results in different scenarios. Then, we provide the validation of the results against the observed data and discuss trends of the estimated empty truck counts in Florida. The final section summarizes and concludes the study.

## 2. Methodology

This section aims to fuse the observed truck flow data from multiple sources (including commodity mass and truck counts from the sampled links and all relevant OD pairs) to produce the best estimation of weight-categorized truck flows at different resolutions over the studied region. We propose a convex optimization model to estimate the weight-categorized truck counts for the sample links and OD pairs that best match the observations from all these sources. The objective function of this model is set to minimize the summation of the squared errors between the estimated and the observed truck flows for both weight-categorized truck counts and associated commodity masses. Flow conservation constraints are applied to ensure the estimated OD flows are consistent with the estimated link flows. Proper weight factors are multiplied to each error term to balance the effects of the different data sizes and error magnitudes from these multiple data sources. In section 2.1, we fuse

truck travel data from multiple sources to estimate the weight-categorized OD-level and link-level truck counts in different truck weight categories. Then, in Section 2.2, we propose a disaggregation approach that breaks county-level empty truck counts between OD pairs into relatively finer TAZs level.

## 2.1 Estimation of truck counts in different weight categories

The proposed model aims to estimate link-level and OD-level truck counts in several weight categories. Required data sets and the proposed non-linear optimization model with its objective function and constraints are introduced in this section. For the convenience of readers, the variables and parameters in the model are listed in Table 1.

Table 1 Notation used in model formulation

| Notation | Description |
|---|---|
| $A^{WS}$ | set of links for which truck counts by weight categories are available |
| $A^T$ | set of links for which only total truck counts are available without weight information |
| $A$ | set of all links, $A = A^{WS} \cup A^T$ |
| $W^C$ | set of OD pairs for which commodity flows by weight are available |
| $W^T$ | set of OD pairs for which truck counts are available |
| $W$ | set of all OD pairs, $W = W^C \cup W^T$ |
| $A'_w$ | set of links used by trucks traveling between OD pair $w \in W$ |
| $W_a$ | set of OD pairs contributing to the truck count on link $a \in A$ |
| $C_1, C_2, C_3, C_4, C_5, C_6$ | optimization weightage factors for different error terms |
| $\mathcal{L}$ | set of weight categories for trucks, $\mathcal{L} = \{1, 2, \dots, l, \dots, L\}$ |
| $n_{la}$ | number of category $l$ trucks passing through link $a \in A^{WS}$, $l \in \mathcal{L}$ |
| $\hat{n}_{la}$ | estimated number of category $l$ trucks passing through link $a \in A^{WS}$, $l \in \mathcal{L}$ |
| $\bar{n}_l^{WS}$ | average number of category $l$ trucks passing through all links in $A^{WS}, l \in \mathcal{L}$ |

| | |
|---|---|
| $n_a$ | truck count on link $a \in A^T$ |
| $\hat{n}_a$ | estimated truck count on link $a \in A^T$ |
| $\bar{n}^A$ | average of all truck counts on all links in $A^T$ |
| $m_{la}$ | commodity weight of category $l$ trucks passing through link $a \in A^{WS}, l \in \mathcal{L}$ |
| $\widehat{m}_{la}$ | estimated commodity weight of category $l$ trucks passing through link $a \in A^{WS}, l \in \mathcal{L}$ |
| $\bar{m}_l^{WS}$ | average commodity weight of category $l$ trucks passing through all links in $A^{WS}, l \in \mathcal{L}$ |
| $m_w$ | commodity weight between OD pair $w \in W^C$ |
| $\widehat{m}_w$ | estimated commodity weight between OD pair $w \in W^C$ |
| $\bar{m}^C$ | average observed commodity weight across all OD pairs in $W^C$ |
| $n_w$ | truck count between OD pair $w \in W^T$ |
| $\hat{n}_w$ | estimated truck count between OD pair $w \in W^T$ |
| $\bar{n}^W$ | average truck OD count across all OD pairs in $W^T$ |
| $P_{wa}$ | percentage of truck count between OD pair $w \in W$ passing through the link $a \in A_w$ |
| $v_1$ | weight of an empty truck |
| $v_l$ | average commodity weight carried by a category $l$ truck, $l \in \mathcal{L}$, excluding empty truck weight |
| $v_{lg}$ | average commodity weight of a truck in category $l$, $v_{lg} = v_l + v_1$, $l \in \mathcal{L}$ |
| $x_{la}$ | estimated number of weight-category $l$ trucks passing through link $a \in A, l \in \mathcal{L}$ |
| $y_{lw}$ | estimated number of weight-category $l$ trucks between OD pair $w \in W, l \in \mathcal{L}$ |
| $\varepsilon_{la}$ | error term for weight-category $l$ trucks passing through link $a \in A, l \in \mathcal{L}$ |

The proposed model considers a freight truck transportation network in a region specified by a set of OD pairs $W$ and a set of links $A$. All trucks are classified into a set of weight categories, denoted by $\mathcal{L}$. To serve the purpose of estimating empty truck flows, the simplest $\mathcal{L}$ comprises only two categories, i.e., an empty-truck category and a loaded truck category. Nonetheless, this study allows a greater number of categories in $\mathcal{L}$ and a finer weight categorization to increase the accuracy of estimation. This study assumes all empty trucks have the same weight, and let $v_1$ denote the empty truck weight. For each weight category $l \in \mathcal{L}$, let $v_{lg}$ and $v_l$ denote the average commodity weight (not including empty truck weight) and the average truck weight (the summation of commodity weight $v_{lg}$ and empty truck weight $v_1$), respectively. This model aims to estimate (i) weight-categorized OD-level truck count $y_{lw}$ in weight category $l \in \mathcal{L}$ for OD pair $w \in W$ in the study region and (ii) weight-categorized link-level truck count $x_{la}$ in weight category $l \in \mathcal{L}$ and link $a \in A$ in the study region. Estimating variables $x_{la}$ and variables $y_{lw}$ in each truck weight categories $l \in \mathcal{L}$ in each link $a \in A$ and OD pair $w \in W$ will help planners better design commodity delivery operations and help them obtain more information about the accuracy of weight in motion stations. To estimate these variables, a variety of observed data are used as inputs to the model. The input data is described in the following.

a. <u>Weight-categorized link-level truck count data</u>: Let $A^{WS}$ denote the set of links on which truck counts are measured for trucks traveling though these links; for example, through weigh-in-motion stations (WIM). Let $n_{la}$ denote the truck count for category-$l$ trucks on link $a \in A^{WS}$ and define $\bar{n}_l^{WS} = \sum_{a \in A^{WS}} n_{la} / |A^{WS}|$ as the average observed category-$l$ truck count across all links in $A^{WS}$. Further, we are able to obtain $\hat{n}_{la} = x_{la}$ as the estimated truck count of category $l$ trucks on link $a \in A^{WS}$. Such dataset $A^{WS}$ are typically available in many regions through the weigh-in-motion stations located on a number of highways.

b. <u>Aggregated (or non-weight-categorized) link-level truck count data:</u> Let $A^T$ denote the set of links for which aggregated truck counts are observed (without any weight information). Let $n_a$ denote the total truck count on link $a \in A^T$ and define $\bar{n}^A = \sum_{a \in A^T} n_a / |A^T|$ as the average observed aggregated truck count across all links in $A^T$. Further, we are able to obtain $\hat{n}_a = \sum_{l \in \mathcal{L}} x_{la}$ as the estimated aggregated truck count on link $a \in A^T$. Such dataset $A^T$ is typically available in many regions through Telemetered Traffic Monitoring Sites (TTMS) or other means of collection traffic counts by vehicle class.

c. <u>Weight-categorized link-level commodity weight data</u>: This data is associated to the same link set $A^{WS}$ as above. Let $m_{la}$ denote the observed commodity weight of all category $l$ trucks on link $a \in A^{WS}$ and define $\bar{m}_l^{WS} = \sum_{a \in A^{WS}} m_{la} / |A^{WS}|$ as the average observed commodity weight across all links in $A^{WS}$. Further, we calculate $\hat{m}_{la} = x_{la} v_{lg}$ as estimated commodity weight of category $l$ trucks on link $a \in A^{WS}$. Such dataset $A^{WS}$ can be obtained from the same sources as the previous weight-categorized link-level truck count data.

d. <u>OD commodity weight data:</u> Let $W^C$ denote the set of OD pairs for which commodity weight (i.e., the total tonnage of commodity weight from origin to destination by truck) is observed. Let $m_w$ denote the observed commodity weight between OD pair $w \in W^C$ and define $\bar{m}^C = \sum_{w \in W^C} m_w / |W^C|$ as the average observed commodity weight across all OD pairs in $W^C$. Further, we are able to obtain $\hat{m}_w = \sum_{l \in \mathcal{L}} y_{lw} v_l$ as estimated commodity weight between OD pair $w \in W^C$. Such dataset $W^C$ can be extracted from commodity flow databases provided by both public and proprietary agencies (e.g., the Transearch data from IHS Global Insight, Inc.).

e. <u>OD-level truck count data</u>: Let $W^T$ denote the set of OD pairs for which OD-level truck counts are observed. Let $n_w$ denote the truck count between OD pair $w \in W^T$ and define $\bar{n}^W = \sum_{w \in W^T} n_w / |W^T|$ as the average observed truck count across all OD pairs in $W^T$. Further, we can obtain $\hat{n}_w = \sum_{l \in \mathcal{L}} y_{lw}$ as the estimated truck count between OD pair $w \in W^T$. Such dataset $W^T$ can be obtained after the trip distribution and mode choice step in a regional four-step freight travel demand.

f. <u>OD-link assignment percentage data (i.e., path flow data):</u> It is worth noting that a link in set $A$ is passed by certain paths between certain OD pairs in set $W$. With an appropriate OD-to-link assignment method, we are able to obtain the percentage of an OD truck count passing through a link. We further define $A'_w$ as the set of links used by truck counts between OD pair $w \in W$. Then we denote the percentage of the trucks traveling between OD pair $w \in W$ going through link $a \in A_w$ as $P_{wa}$. We assume that truck counts of all weight categories between OD pair $w$ have the same assignment percentage $P_{wa}$ on link $a$.

It is important to note that these datasets contain noise and each data set may just contain incomplete information. Therefore, to accurately estimate truck counts in multiple weight categories on links and between OD pairs, it is necessary to build an estimation model to determine the most likely values of the truck counts over the network by fusing all the data sets. Of course, one can still use the method even if only a subset of these datasets is available. We will first formulate the

constraints that associate estimated weight-categorized link-level truck counts $x := \{x_{la}\}_{l \in \mathcal{L}, a \in A}$ with OD- level truck counts $y := \{y_{lw}\}_{l \in \mathcal{L}, w \in W}$ based on allocation percentages $\{P_{wa}\}$ as follows.

$$\sum_{w \in W} y_{lw} P_{wa} = x_{la} + \varepsilon_{la}, \forall\, l \in \mathcal{L}, a \in A'_w \tag{1}$$

where $\varepsilon_{la}$ ($\varepsilon_{la} \in \{\varepsilon_{la}\}_{l \in \mathcal{L}, a \in A}$.) is an error variable to recognize inconsistencies between the two sides of equations due to estimation errors and inconsistencies between OD-level truck counts and link-level truck counts. Without the error variable, the optimization problem can be solved but it is not able to capture the inconsistency. Models with or without the error variable are tested in case study later.

Now we formulate the optimization objective as the sum squared errors between the observed and estimated values of five different terms as well as the sum of squares of the error variable $\varepsilon_{la}$ as the sixth term to minimize the inconsistencies between link-level and OD-level truck counts. Specifically, the objective function is shown in Equation (2) below:

$$\min_{x,y,\varepsilon} \quad \underbrace{\left[ \sum_{l \in \mathcal{L}} \sum_{a \in A^{WS}} C_1 (n_{la} - x_{la})^2 \right]}_{Conservation\ of\ link-level, weight-categorized\ truck\ counts} \tag{2}$$

$$+ \quad \underbrace{\left[ \sum_{a \in A^T} C_2 \left( n_a - \sum_{l \in \mathcal{L}} x_{la} \right)^2 \right]}_{Conservation\ of\ link-level\ truck\ counts}$$

$$+ \quad \underbrace{\left[ \sum_{l \in \mathcal{L}} \sum_{a \in A^{WS}} C_3 \left( m_{la} - x_{la} v_{lg} \right)^2 \right]}_{Conservation\ of\ link-level\ commodity\ weight\ flow}$$

$$+ \quad \underbrace{\left[ \sum_{w \in W^c} C_4 \left( m_w - \sum_{l \in \mathcal{L}} y_{lw} v_l \right)^2 \right]}_{Conservation\ of\ OD-level\ commodity\ weight\ flow}$$

$$+ \quad \underbrace{\left[ \sum_{w \in W^T} C_5 \left( n_w - \sum_{l \in \mathcal{L}} y_{lw} \right)^2 \right]}_{Conservation\ of\ OD-level\ truck\ counts}$$

$$+ \underbrace{\left[ \sum_{w \in W^T} C_6 \, \varepsilon_{la}^2 \right]}_{Inconsistencies\ between\ link-level\ and\ OD-level\ truck\ counts}$$

In the above objective function, the first term represents the sum of squared errors between observed and estimated weight-categorized link-level truck counts on link $a \in A^{WS}$. The second term represents the sum of squared errors between observed and estimated total truck counts without weight categorization on link $a \in A^T$. The third term corresponds to a sum of squared differences between observed and estimated total commodity weight of all trucks in weight categories passing through link $a \in A^{WS}$. The fourth term represents the sum of squared errors between OD-level observed and estimated commodity weight. The fifth term represents sum of squared errors between OD-level observed and estimated truck counts without weight categorization. Finally, the sixth term captures inconsistencies between estimated link-level trucks and OD-level truck counts.

The objective function is subject to constraints (1) and non-negativity constraints on all weight categorized truck counts: $x_{la}, y_{lw} \geq 0, \forall\, l \in \mathcal{L}, a \in A, w \in W$. Note that the above six terms include weights $C_1$ through $C_6$ to weigh the squared error magnitudes. These terms are determined in the following equations to normalize the impact of sample size and data dispersion:

$$C_1 = c_1 / \sum_{l \in \mathcal{L}} \sum_{a \in A^{WS}} (n_{la} - \bar{n}_l^{WS})^2 \tag{3}$$

$$C_2 = c_2 / \sum_{a \in A^T} (n_a - \bar{n}^A)^2 \tag{4}$$

$$C_3 = c_3 / \sum_{l \in \mathcal{L}} \sum_{a \in A^{WS}} (m_{la} - \bar{m}_l^{WS})^2 \tag{5}$$

$$C_4 = c_4 / \sum_{w \in W^C} (m_w - \bar{m}^C)^2 \tag{6}$$

$$C_5 = c_5 / \sum_{w \in W^T} (n_w - \bar{n}^W)^2 \tag{7}$$

$$C_6 = c_6 / \left\{ \sum_{l \in \mathcal{L}} \sum_{a \in A^{WS}} (n_{la} - \bar{n}_l^{WS})^2 + \sum_{a \in A^T} (n_a - \bar{n}^A)^2 \right\} \tag{8}$$

The proposed optimization model is a convex nonlinear optimization model that can be efficiently solved by commercial solvers including Gurobi.

## 2.2 Disaggregation of estimated truck counts

More often than not, the estimated empty truck counts from the above analysis might be relatively aggregated and further processes might be needed to disaggregate them into smaller geographic areas to suit the needs of specific applications. This section proposes the use of a simple disaggregation method based on a model proposed by Holguín-Veras and Patil (2008) for the purpose of disaggregating the OD-level empty truck counts from a coarse spatial resolution (e.g., county-level) to a finer resolution (e.g., TAZ-level).

In the method, each original OD pair $w \in W^T$ are now divided into a number of finer pairs $k$, where $k \in K_w$ and $K_w$ is set of finer OD pairs in OD pair $w \in W$. Further, we set $b_k$ as estimated empty truck counts between OD pair $k \in K_w$ and $a_k$ as the observed empty truck count between OD pair $k \in K_w$. We define operator $\overline{\phantom{w}}$ as the opposite direction of an OD pair such that $\overline{w}$ is an OD pair in the opposite direction of OD pair $w \in W^T$. In Section 2.1, we are able to obtain $y_{1\overline{w}}$ as the empty truck count between OD pair $\overline{w} \in W^T$. According to the proposed method by Holguín-Veras and Patil (2008), we can assume that there is a proportion parameter $p_w$ that:

$$b_{\overline{k}} = p_w * a_k, \forall k \in K_w, \forall w \in W^T \qquad (9)$$

Then the sum of estimated empty truck counts $\sum_{\overline{k} \in \overline{w}} b_{\overline{k}}$ should be equal to the estimated empty truck counts between OD pair $\overline{w}$ , $\boldsymbol{y_{1\overline{w}}}$. Therefore, we have

$$\sum_{\overline{k} \in K_{\overline{w}}} b_{\overline{k}} = y_{1\overline{w}}, \forall w \in W^T \qquad (10)$$

Thus, the estimated empty truck counts in finer OD pairs can be obtained by using Equations (9) and (10) and eliminating parameter $p_w$:

$$b_{\overline{k}} = y_{1\overline{w}} * \frac{a_k}{\sum_{k \in K_w} a_k}, \forall k \in K_w, \forall w \in W^T \qquad (11)$$

## 3. Florida Case Study

In this section, we apply the proposed optimization model to a case study for Florida. In Section 3.1, the observed truck counts and associated commodity weights obtained from Florida are added to parameters in the model. Two sets of truck weight categories and four scenarios with

different values of optimization weight coefficients are tested in the case study. Afterwards, in Section 3.1, we analyzed the goodness of fit of the results for four scenarios and picked the scenarios with the best goodness of fit. Further, we showed and analyzed the 45-degree results of the estimated data versus the observed data. It is necessary to note that we kept aside the observed data for a number of links and OD pairs for validation. The 45-degree results of estimated values versus validation links are shown in this section as well. Further, we showed the total empty truck counts from Florida and the total empty truck counts to Florida at the county and TAZ levels. The satisfied results for the Florida case study provides a numerical example in detail and illustrates a potential application of the model in the other states or countries.

## 3.1 Data description

In this section, to apply the model to the case study for Florida, we provide the data description of Florida truck counts data and commodity data applied to parameters in the proposed optimization model. Settings of truck weight categories and optimization weight coefficients for four scenarios are also listed in this section.

a. Link-level truck counts in weight categories ($m_{la}$ and $n_{la}$): Weigh-In-Motion (WIM) data for the 2011 year was obtained from the Florida Department of Transportation (FDOT). It contains 24.50 million truck records within Florida. 29 WIM stations were operational in 2011, and some of the stations had the capability to measure the truck weight in both the traffic directions. This made up to 53 links corresponding to WIM stations which are available for the model estimation and validation. We kept aside 11 links for validation and thus we have 42 links in set $A^{WS}$ with commodity weight measured.

*Figure 1 Truck counts on links in WIM sites*

b.  Aggregated link-level truck count data ($n_a$): TTMS for the year 2010 provides 353 links where truck counts are available for model estimation and validation. In the dataset, we kept aside 14 links for validation and 339 links in link set $A^T$ for model estimation.
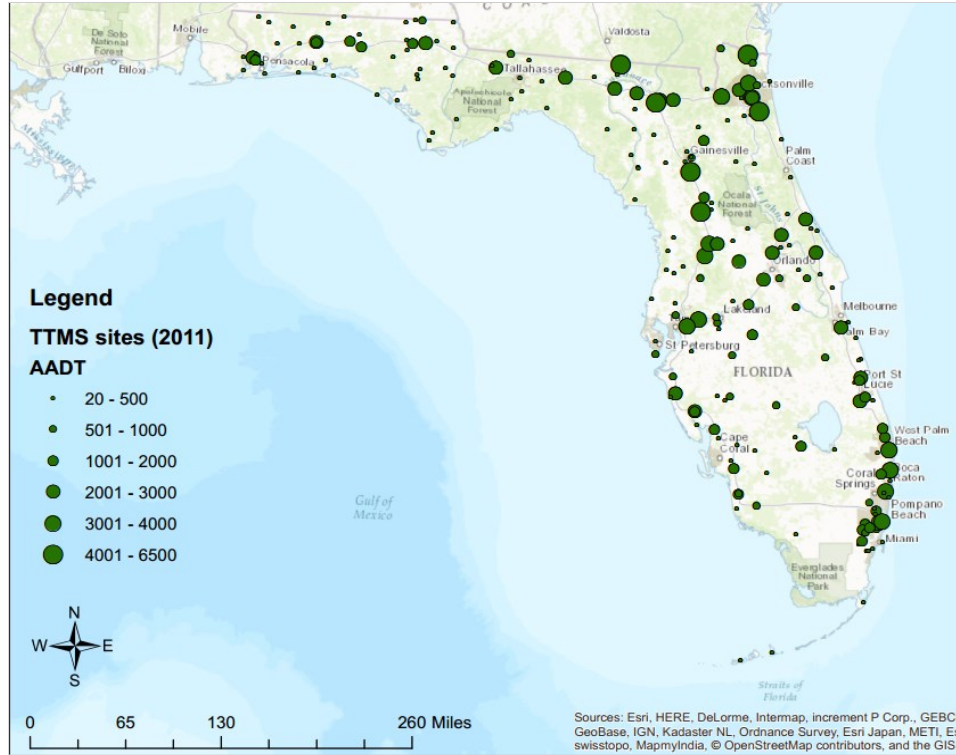
*Figure 2 Truck counts on links in TTMS sites*

c.  OD commodity weight data ($m_w$): OD commodity weight data $m_w$ is obtained from Transearch, developed by IHS Global Insight, Inc for the year 2011. The database provides Florida-centric data on the commodity weight between 379 zones inside the country with commodity flow at the county-level resolution in Florida. This dataset has $|W^C| = 17700$ OD pairs with observed commodity weights.

d.  OD-level truck count data ($n_w$): Zanjani et al. used the GPS data for the trucks in the year 2010 jointly provided by American Transportation Research Institute and Federal Highway Administration (FHWA) and the counts from TTMS sites to estimate Florida centric OD-level truck counts at both county and statewide TAZ levels (Zanjani et al., 2015). Therefore, we have the set of OD pairs for which truck counts are observed $W^T$ with 11087 OD pairs.

e.  OD-link assignment percentage data ($P_{wa}$): The percentage of truck counts on the links in WIM sites and TTMS sites $P_{wa}$ are extracted from the OD-level truck counts estimated from ATRI data for the year 2010. The path flows are obtained using the traffic assignment step using the Cube software, as described in the OD-level truck counts estimation study by Zanjani et al. (2015).

f.  Empty truck weight: From the WIM data corresponding to the Florida, it was observed that the heavy-duty trucks (class 8 and above according to the FHWA trucks classification) constitute 80% of the total truck counts. The individual empty weight of tractor and trailer varies depending on the manufacturer. According to a survey conducted in 2014 by American Transportation Research Institute (ATRI), the majority of fleets operated truck-tractors, and the most prevalent trailer types were 53-foot and 28-foot trailers respectively. Using this proportion and the information from the manufacturers on the range of weights for truck-tractor units, the weight of an empty truck can range from 21 kips to 37 kips. So different empty weight within the given range were tested, and the optimum value for the empty truck $v_1 = 35 kips$ was chosen to obtain the best prediction.

g.  Optimization parameter settings: In the optimization procedure, we have tried two sets of truck-weight categories (set 1 and set 2) and four scenarios with different values of optimization weight coefficients $(c_1, c_2, c_3, c_4, c_5, c_6)$, as listed in Table 2 and Table 3, respectively. Since truck counts on links and OD pairs are forced to match each other as we have Constraint (1) as a connection between truck counts on links and OD pairs, different ratio between c1 and c2, c3 or ratio between c4 and c5 will not change the results much, which we have tested in the case study. Since the number of OD pairs is much larger than the number of links, we set different weightage for OD pairs and links to balance the magnitude difference. Since the purpose of assigning different weights is to balance the data noises and disproportion between the link data and the OD data instead of those within the link data alone or within the OD data alone, c1, c2 and c3 are set at the same values in each scenario, and so are c4 and c5..

Table 2 Truck-weight categories

| Category no. | Weight range in kips (kilo pounds) | |
| --- | --- | --- |
| | Set 1 | Set 2 |
| 1 | ≤ 35 | ≤ 35 |
| 2 | 35 - 60 | 35 - 40 |
| 3 | > 60 | 40 - 45 |
| 4 | -- | 45 - 50 |
| 5 | -- | 50 - 55 |
| 6 | -- | 55 - 60 |
| 7 | -- | > 60 |

Table 3 Scenarios with different optimization weights

| Weightage | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 |
|---|---|---|---|---|
| $c_1$ | 1 | 10 | 100 | 1 |
| $c_2$ | 1 | 10 | 100 | 1 |
| $c_3$ | 1 | 10 | 100 | 1 |
| $c_4$ | 1 | 1 | 1 | 10 |
| $c_5$ | 1 | 1 | 1 | 10 |
| $c_6$ | 1 | 10 | 100 | 1 |

In the optimization process, three weight categories are used. They are 0-35 kips, 35-60 kips, and 60kips or above. Again, the categorization is based on the weight ranges considering the typically empty, partially loaded, and fully loaded trucks. In addition to these broad weight categories, the results analyzed for finer categories with 5kip intervals are also considered for the optimization procedure, for a better quality of fitting.

## 3.2 Results

This section presents the results from the optimization procedure, in which the truck counts with multiple weight categories between the OD pairs are estimated at the county level resolution for the state of Florida. The average solution time of the model is 25 sec for three weight categories and 75sec for seven weight categories.

In this study, two sets of truck-weight categories in 'Set 1' and 'Set 2' as given in Table 2 are analyzed for the four scenarios given in Table 3. Different empty weight values within the range of 21 kips to 37 kips are tested in all the 4 scenarios for two truck-weight categories, and an optimal value for the empty truck is chosen as $v_1 = 28 \, kips$ which provides the best prediction. The different sets of 'c' values as shown in Table 3 are then used for the analysis.

After we apply the Florida data to the proposed model in Section 2, estimated link-level truck count $x_{la}$ and estimated OD-level truck count $y_{lw}$ are obtained. Thus, we can calculate the estimated value of each parameter as explained in Section 2.

To compare the results of the four scenarios with two truck-weight categories, we calculate the mean absolute error (MAE) to test the difference between the estimated and observed values of a parameter in the optimization model. To compare the magnitude of the errors across parameters, we further divide the MAE by the mean value of the observed value of the parameter and name it as

mean absolute error to mean (MAEM). For example, the MAEM of truck link count $n_a$, $\forall a \in A^T$, is

$$\frac{\sum_{a \in A^T} |\hat{n}_a - n_a| \Big/ |A^T|}{\bar{n}^A}.$$

Figure 3 shows the MAEM values in all four scenarios for both sets of truck-weight categories with or without considering error term $\varepsilon_{la}$. Figure 3 (a) and (b) show the MAEM values of parameters $m_{la}, n_{la}, n_a, m_w, n_w$ in the four scenarios for both sets of truck-weight categories with considering error term $\varepsilon_{la}$ in Equation (1) and Objective function (2). Figure 3 (c) and (d) show the MAEM values of parameters $m_{la}, n_{la}, n_a, m_w, n_w$ without considering error term $\varepsilon_{la}$ in Equations (1) and (2). According to the MAEM values, the model in categories in 'Set 1' performes better than the model with categories in 'Set 2'. It indicates that the model considering error term $\varepsilon_{la}$ helps improve the model performance. The results show that when the weights associated with links are higher, we obtain lower total MAEM despite higher OD-associated MAEM. As the empty truck counts between the OD pairs are important for planners, we select categories in 'Set 1' and scenario 4 with considering error term $\varepsilon_{la}$ for further analysis.

Narrowing down the results to each of the parameters, Figure 4 shows the 45-degree result of estimated data vs. observed data. It shows the following four comparisons: (1) estimated truck count $\hat{n}_a$ vs. observed truck count $n_a$ at each link $a \in A^T$ in TTM sites; (2) estimated truck weight $\hat{m}_{la}$ vs. observed truck weight $m_{la}$ at each link $a \in A^{WS}$ in each category $l \in L = \{1,2,3\}$ in WIM sites; (3) estimated truck count $\hat{n}_w$ vs. observed truck count $n_w$ in each OD pair $w \in W^T$; (4) estimated commodity weight $\hat{m}_w$ vs. observed commodity weight $m_w$ for each OD pair $w \in W^C$. We see that overall, most observed and estimated values are distributed around the 45-degree line within 25% error bounds, verifying that the estimates from the proposed model are reasonable. Nonetheless, the estimated and observed data do not match each other perfectly, indicating noise and inconsistency between different datasets, which again highlights the needs for fusing different data sources to reduce noise. Part of the reason for the mismatched data is that truck weight $m_{la}$ and truck count $n_{la}$ in category $l \in L$ at link $a \in A^{WS}$ are obtained from databases in year 2011 while other datasets are extracted from databases in year 2010.

Figure 5 uses color coding to differentiate between the WIM data used for optimization and that kept aside for validation. In each panel of this figure, the estimated and observed truck counts at are presented separately for the WIM sites whose data is used in the optimization and data kept aside for validation. Three of these panels (a), (b) and (c) are for the comparison of estimated truck count

$\hat{n}_{la}$ and observed truck count $n_{la}$ for $a \in A^{WS}$ at WIM sites for each category in 'Set 1'. The fourth panel (d) makes such comparison for total truck counts in three categories $\sum_{l=1,2,3} n_{la}$ on link $a \in A^{WS}$ at the WIM sites. It is evident from all panels in the figure that the estimated truck counts in all three weight categories are close to the observed values (or at least within 25% error) for the validation sites. This highlights the efficacy of the optimization procedure used to estimate the truck counts by weight category.
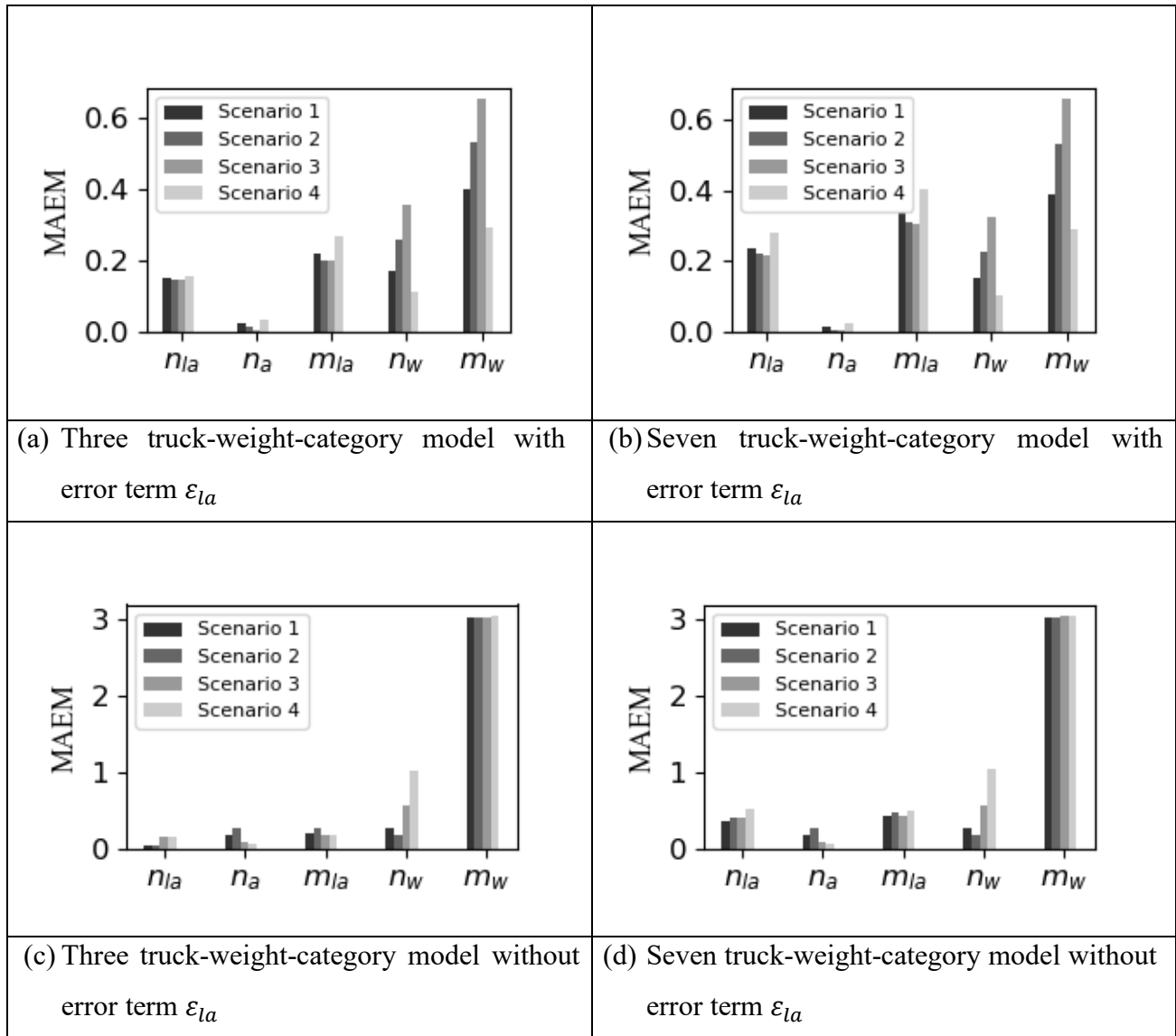
(a) Three truck-weight-category model with error term $\varepsilon_{la}$

(b) Seven truck-weight-category model with error term $\varepsilon_{la}$

(c) Three truck-weight-category model without error term $\varepsilon_{la}$

(d) Seven truck-weight-category model without error term $\varepsilon_{la}$

*Figure 3 MAEM of each type of category for 4 scenarios of optimization weightages with error term or without error term*

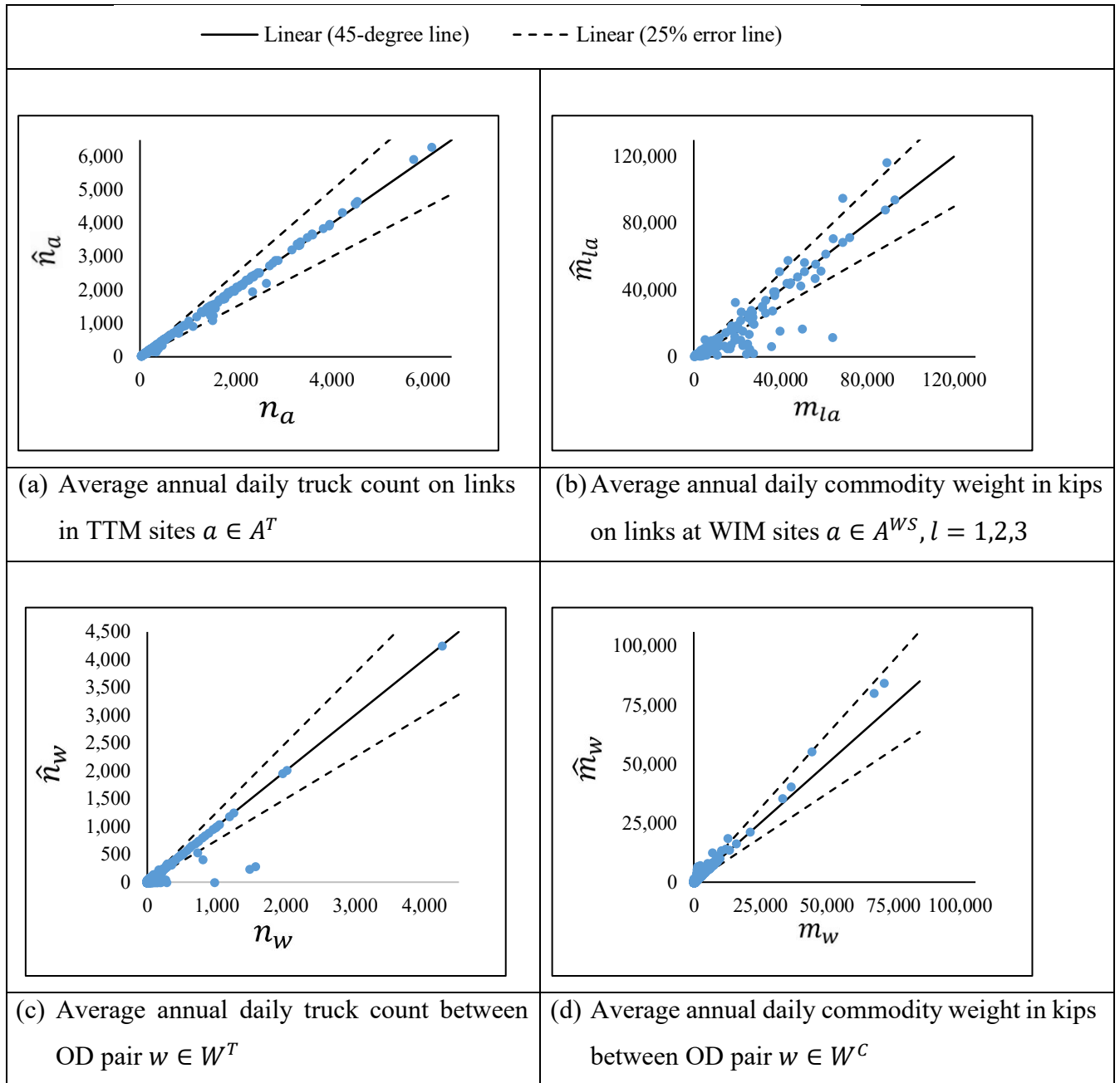| | |
|---|---|
| (a) Average annual daily truck count on links in TTM sites $a \in A^T$ | (b) Average annual daily commodity weight in kips on links at WIM sites $a \in A^{WS}, l = 1,2,3$ |
| (c) Average annual daily truck count between OD pair $w \in W^T$ | (d) Average annual daily commodity weight in kips between OD pair $w \in W^C$ |

*Figure 4 Observed versus estimated link-level truck counts, OD-level truck counts and OD-level commodity weights per day*

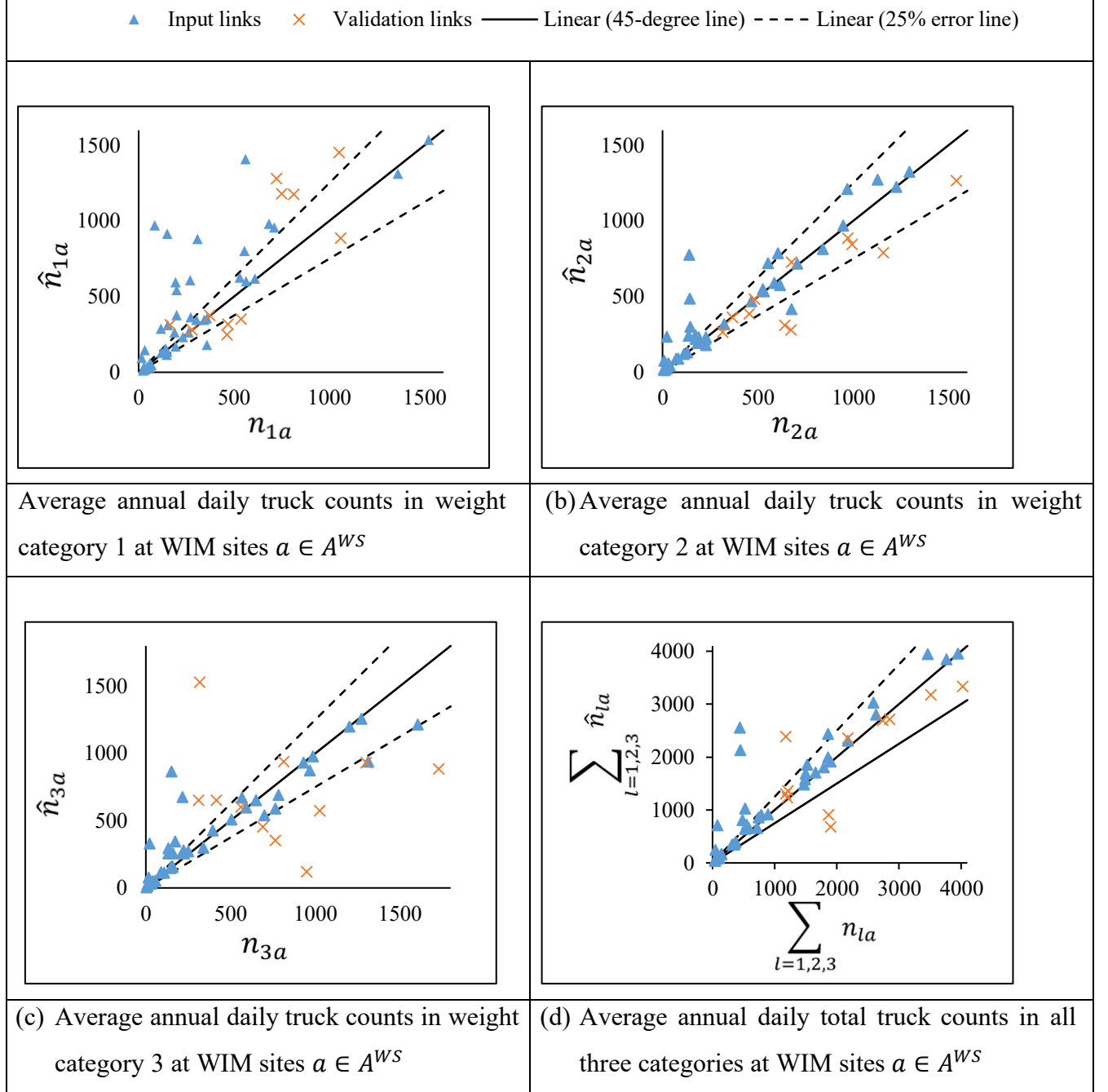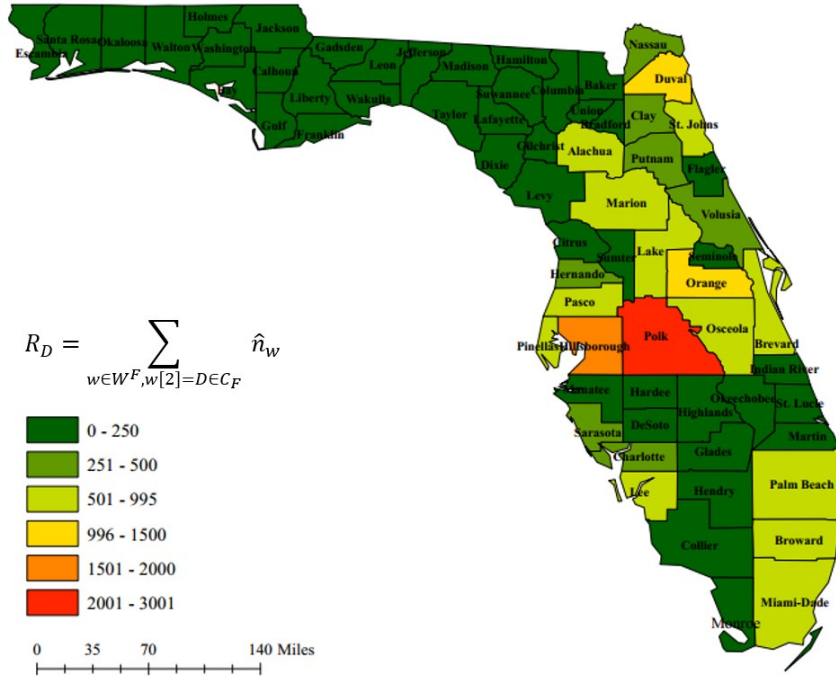| | |
|---|---|
| Average annual daily truck counts in weight category 1 at WIM sites $a \in A^{WS}$ | (b) Average annual daily truck counts in weight category 2 at WIM sites $a \in A^{WS}$ |
| (c) Average annual daily truck counts in weight category 3 at WIM sites $a \in A^{WS}$ | (d) Average annual daily total truck counts in all three categories at WIM sites $a \in A^{WS}$ |

*Figure 5 Observed versus estimated average annual daily link-level truck counts at WIM sites*
$a \in A^{WS}$

For further analysis, we define $w[1], w[2]$ as the origin and destination of an OD pair $w \in W$ separately. Next, we denote the total truck counts from origin O to all destination as trip production $R_O$, and the total truck counts from all origin to destination D as attraction $R_D$. Therefore, by the definition of the attraction and production, we are able to obtain $R_O = \sum_{w \in W, w[1]=O} n_w$, $R_D =$
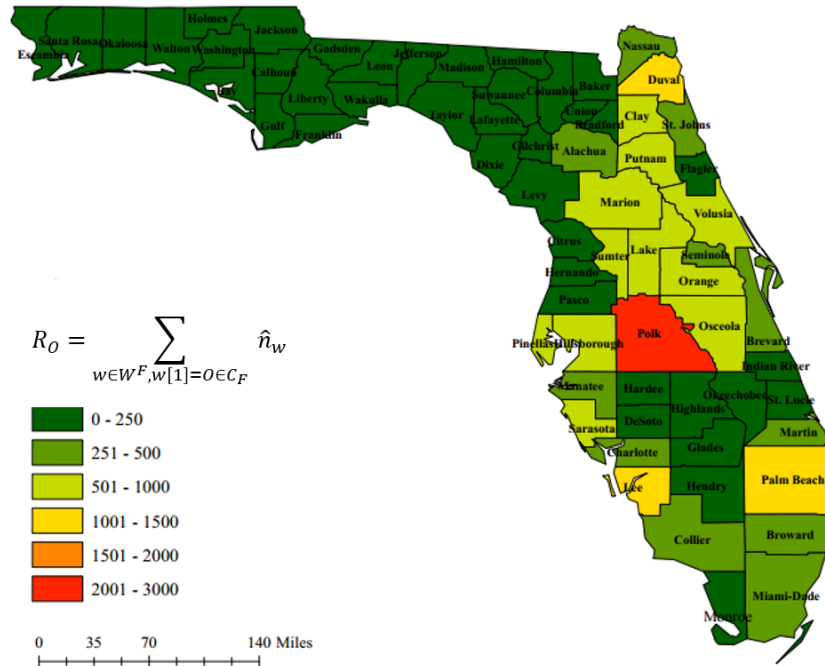
$\sum_{w \in W, w[2]=D} n_w$, $\forall w \in W$. Further, we define the set of OD pairs inside Florida as $W^F$ and the counties inside Florida as $C_F$ so that we can classify attraction and production of each county inside Florida. Figure 6 shows the county level trip productions and attractions (excluding intra county movements) for trucks moving within Florida and in weight category 1 (truck load $\leq 35$ kips), most of which are empty trucks. Similarly, Figure 7 shows the county level trip attractions and production of category 1 truck counts between counties in Florida and other states in USA. One can use such results to identify the areas with high productions and attraction of empty truck counts and design appropriate policies to reduce the empty truck flows.

Figure 8 shows the spatial distribution of empty truck counts (truck weight $\leq 35$ kips) from the state of Florida to other states in the United States. It is important to know that the link data in TTM sites and WIM sites used in the modeling are only in the Florida, thus the truck counts between Florida and nearby states are much reliable as compared to the flows between Florida and far away states. From Figure 8, it can be observed that a considerable proportion of empty trucks from Florida are destined to Alabama and Georgia. A possible explanation could be that the trucks delivering goods in Florida and leaving empty while returning may go to Alabama and Georgia to get loads. Traffic operators can use such results to identify the specific OD pairs with high empty truck counts, so that appropriate strategies may be used to reduce the empty back-hauls.

Using the methodology described in section 2.2 and information on loaded truck counts at Statewide TAZ (SWTAZ) level within Florida from the Transearch data, the estimated OD matrix (67×67) of empty truck counts within Florida at Transearch county level are disaggregated into the OD matrix of 8518 × 8518 at SWTAZ level. Figure 9 shows the SWTAZ level attractions and productions of empty truck trips within Florida. Such data can help stakeholders visualize empty flow patterns in a higher resolution to support decisions made at a finer geographic scale.
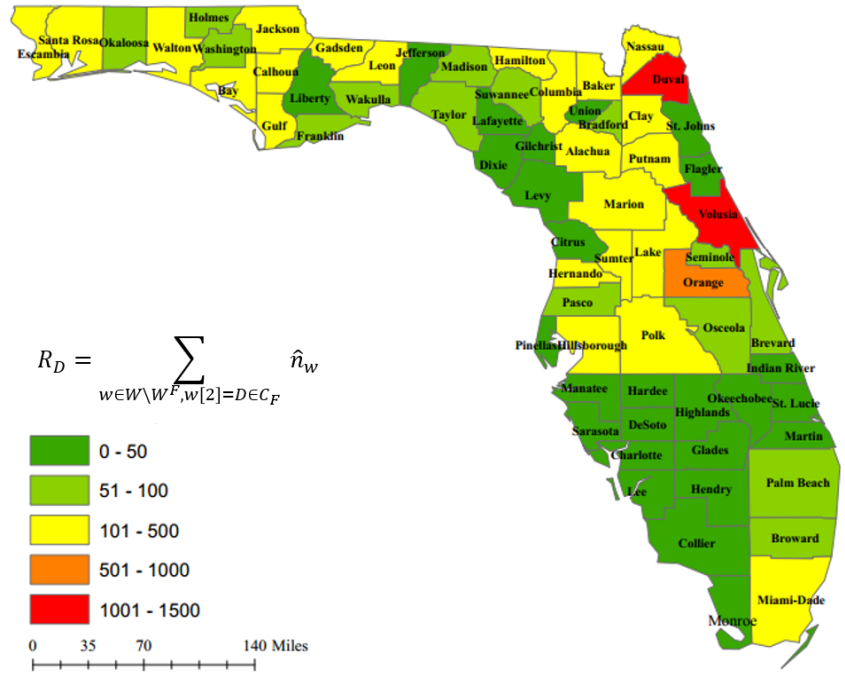
$$R_D = \sum_{w \in W^F, w[2]=D \in C_F} \hat{n}_w$$

- ■ 0 - 250
- ■ 251 - 500
- ■ 501 - 995
- ■ 996 - 1500
- ■ 1501 - 2000
- ■ 2001 - 3001

0    35    70         140 Miles

(a) County level trip attractions $R_D$ for $OD$ pair



$$R_O = \sum_{w \in W^F, w[1]=O \in C_F} \hat{n}_w$$

- ■ 0 - 250
- ■ 251 - 500
- ■ 501 - 1000
- ■ 1001 - 1500
- ■ 1501 - 2000
- ■ 2001 - 3000

0    35    70         140 Miles
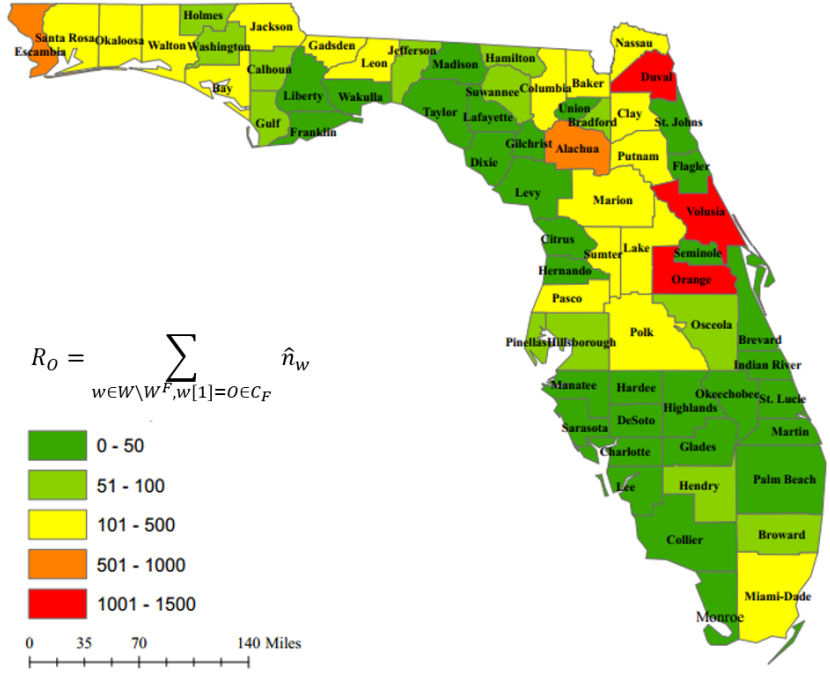
(b) County level trip productions $R_O$ for $OD$ pair

*Figure 6 Estimated county level in-state trip attractions and productions for empty trucks*

23

$$R_D = \sum_{w \in W \setminus W^F, w[2]=D \in C_F} \hat{n}_w$$

| | |
|---|---|
| ■ (dark green) | 0 - 50 |
| ■ (light green) | 51 - 100 |
| ■ (yellow) | 101 - 500 |
| ■ (orange) | 501 - 1000 |
| ■ (red) | 1001 - 1500 |

0   35   70            140 Miles

(a) County level trip attractions $R_D$ of each county $D$



$$R_O = \sum_{w \in W \setminus W^F, w[1]=O \in C_F} \hat{n}_w$$

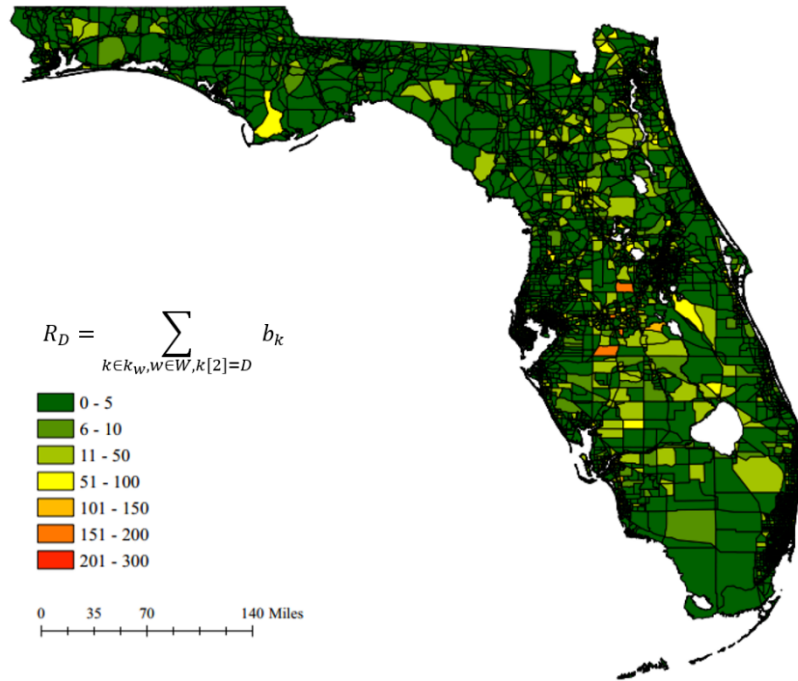| | |
|---|---|
| ■ (dark green) | 0 - 50 |
| ■ (light green) | 51 - 100 |
| ■ (yellow) | 101 - 500 |
| ■ (orange) | 501 - 1000 |
| ■ (red) | 1001 - 1500 |

0   35   70            140 Miles

(b) County level trip productions $R_O$ of each county $O$

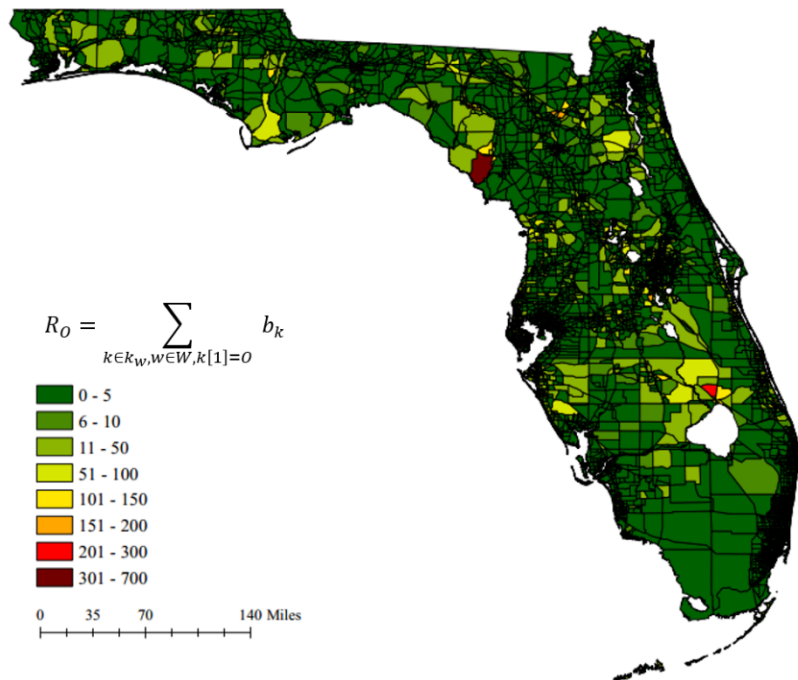*Figure 7 Estimated county level out-of-state trip attractions and productions for empty trucks*
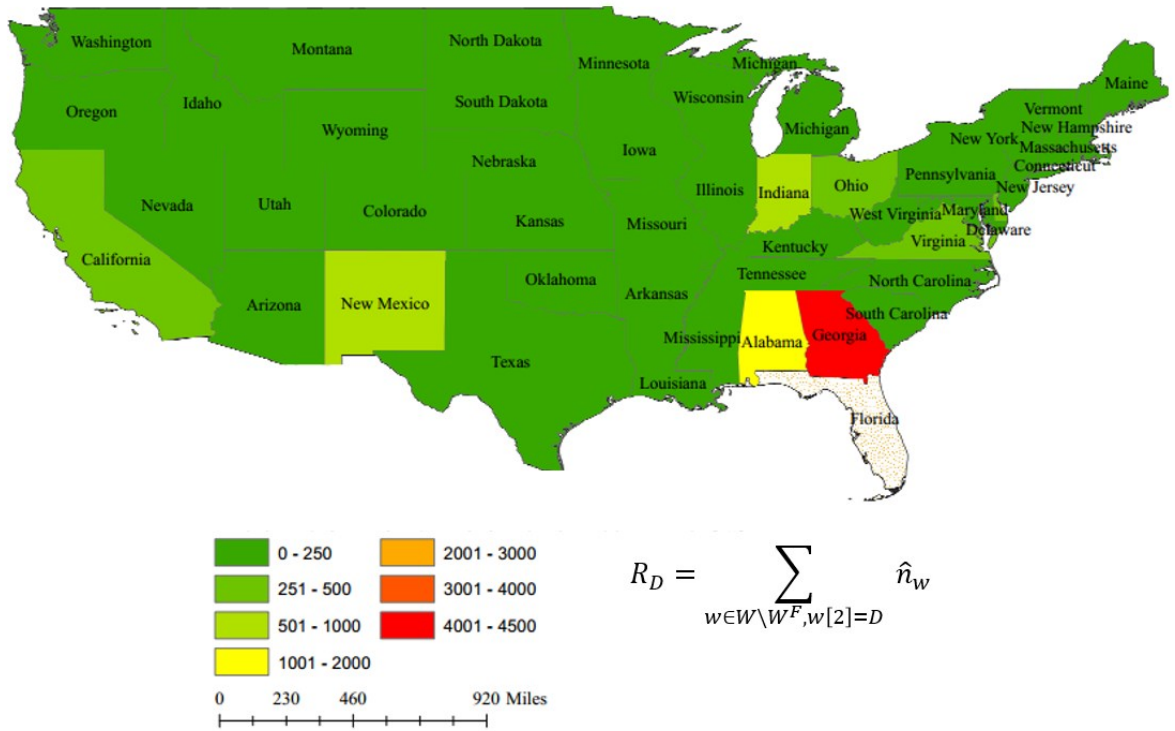
$$R_D = \sum_{w \in W \setminus W^F, w[2]=D} \hat{n}_w$$

*Figure 8 Empty truck counts from Florida to other states of United States*

$$R_D = \sum_{k \in k_w, w \in W, k[2]=D} b_k$$

- 0 - 5
- 6 - 10
- 11 - 50
- 51 - 100
- 101 - 150
- 151 - 200
- 201 - 300

0   35   70         140 Miles

(a) SWTAZ level trip attractions $R_D$ of each region $D$ in SWTAZ level



$$R_O = \sum_{k \in k_w, w \in W, k[1]=O} b_k$$

- 0 - 5
- 6 - 10
- 11 - 50
- 51 - 100
- 101 - 150
- 151 - 200
- 201 - 300
- 301 - 700

0   35   70         140 Miles

(b) SWTAZ level trip productions $R_O$ of each region $O$ in SWTAZ level

*Figure 9 Estimated SWTAZ level trip attractions and productions for empty trucks*

**4. Summary and conclusions**

This paper proposes a nonlinear optimization model to estimate truck counts in different weight categories. The optimization model minimizes an objective function with sum of squared errors to estimate truck counts in multiple truck-weight categories. And the estimated empty truck counts are disaggregated into finer granularity to get better understanding about the empty truck flows. The proposed approach is then applied to estimate truck counts on links and OD pairs in different weight categories for the State of Florida. A variety of different scenarios are considered to derive appropriate weightages for different datasets used in the optimization program. For the final set of truck-weight categories and weightage scheme used in the study, a validation exercise was undertaken to compare the estimated truck counts and observed truck counts by weight at selected locations in the network. The validation results were satisfactory and highlighted the efficacy of the proposed method. The estimated OD-level truck counts and link-level truck counts in different categories can be used for understanding the spatial distribution of empty truck counts and for formulating policies targeting the trade imbalance in the region.

Recently, a large scale study was conducted by Holguin-Veras et al., (2017) and Holguín-Veras et al., (2011) to use the commodity flow survey microdata and other establishment data to develop freight production models that can be applied to employment data to estimate – at the level of establishments and zip codes in the US – the amount of freight produced and the corresponding freight trips and related service trips generated. Such an approach to estimating freight production and related trip generation can be viewed as a bottom-up approach to estimating freight flows, where the total freight production is estimated based on economic relationships and the corresponding freight trip generation and service trip generation is estimated based on logistics relationships between freight production and trip generation. The resulting trips produced from such models need to be further taken through additional modeling procedures to estimate modal and spatial distribution of the freight trips. Further, additional data, understanding and models will be needed to estimate spatial distribution of freight truck trips by different weight categories, including empty truck trips. Further research is necessary to advance the bottom-up approach to use it for estimating the spatial distribution of freight truck trips by weight categories. On the other hand, our proposed approach is top-down in that it utilizes already estimated spatial distribution of freight flows (e.g., from Transearch data) and other data to estimate freight flows by different weight categories (including empty freight truck flows). While the bottom-up approach is better than the top-down approach due

to the insights it might offer on the factors that influence spatial distribution of truck flows by different weight categories, the top-down approach works well (to estimate weight-categorized truck flows) in the absence of detailed microdata for gaining such insights. Further, the top-down approach is useful to identify current imbalances in empty truck flows that can be used to inform short-term strategies to address such imbalances. However, the top-down approach may not be suitable for long-term forecasting of truck flows by weight categories.

While the proposed study provides methods and insights into weight-category-based truck flow estimation, it can be improved in several directions for even better accuracy of estimation. Firstly, in Section 3.1, OD commodity weight data ($m_w$) is obtained from Transearch for the year 2011 while the other datasets are collected from the year 2010. This inconsistency might bring errors to the results and influence the accuracy. When data in the same year is available, it may further improve the estimation results. Next, some datasets may have certain errors and missing records during the data collection process. Better data collection and correction techniques may be helpful. Besides, although we have tried multiple combinations of terms in the objective function, a systematic procedure to determine the optimal combination of the weights is yet to be proposed. A possible solution to obtain an optimal combination could be using a bi-level optimization model with the upper level determining the optimal combination of weights and the lower level estimating truck counts in different truck weight categories. In conclusion, the results can be improved in several ways including the use of data on the observed truck counts in neighboring states, extracting data parameters from dataset in the same year, omitting bad data in datasets, improvements to the optimization weightage factors for different error terms and the inclusion of path flows using observed route choice patterns by using GPS data.

**Acknowledgements**

# References

Agrawal, A., Udmale, S.S. and Sambhe, V.K., (2018). Extended Four-Step Travel Demand Forecasting Model for Urban Planning. In Information and Communication Technology for Sustainable Development (pp. 191-198). Springer, Singapore.

Bernardin Jr, V.L., Trevino, S. and Short, J., (2015). Expanding truck GPS-based passive origin-destination data in Iowa and Tennessee (No. 15-4687).

Chow, J. Y. J., Ritchie, S. G. and Jeong, K., (2014). "Nonlinear inverse optimization for parameter estimation of commodity-vehicle-decoupled freight assignment." Transportation Research Part E: Logistics and Transportation Review, Vol. 67, pp. 71-91.

Crainic, T. G., and Laporte, G., (1997). "Planning models for freight transportation." European journal of operational research, Vol. 97, No. 3, pp. 409-438.

Crainic, T. G., Gendreau, M. and Dejax, P., (1993). "Dynamic and stochastic models for the allocation of empty containers." Operations research, Vol. 41, No. 1, pp. 102-126.

Giuliano, G., Gordon, P., Pan, Q., Park, J. and Wang, L., (2010). Estimating freight flows for metropolitan area highway networks using secondary data sources. Networks and Spatial Economics, 10(1), pp.73-91.

Guelat, J., Florian, M. and Crainic, T. G., (1990). "A multimode multi product network assignment model for strategic planning of freight flows." Transportation Science, Vol. 24, No. 1, pp. 25-39.

Florida department of Transportation., (2018). "Truck empty backhaul." http://www.fdot.gov/statistics/docs/FDOT_Truck%20Empty%20Back%20Haul_Final%20Report_2018.pdf.

Florida Department of Transportation, Florida Chamber Foundation, (2010). "Florida trade and logistics study." http://www.flchamber.com/wp-content/uploads/2016/08/FloridaTradeandLogisticsStudy_December2010.pdf.

Hautzinger, H., (1984). The prediction of interregional goods vehicle flows: some new modeling concepts, Ninth International Symposium on Transportation and Traffic Theory, VNU Science Press, pp. 375–396

Holguín-Veras, J., (2002). "Revealed Preference Analysis of the Commercial Vehicle Choice Process." Journal of Transportation Engineering, Vol. 128, No. 4.

Holguín-Veras, J., Jaller, M., Destro, L., Ban, X., Lawson, C. and Levinson, H.S., (2011). Freight generation, freight trip generation, and perils of using constant trip rates. Transportation Research Record, 2224(1), pp.68-81.

Holguín-Veras, J., Lawson, C., Wang, C., Jaller, M., González-Calderón, C., Campbell, S., Kalahashti, L., Wojtowicz, J. and Ramirez-Ríos, D., (2017). Using commodity flow survey microdata and other establishment data to estimate the generation of freight, freight trips, and service trips: guidebook (No. Project NCFRP-25 (01)).

Holguin-Veras, J., and Thorson, E., (2003a). "Practical implications of modeling commercial vehicle empty trips." Transportation Research Record: Journal of the Transportation Research Board, No. 1833. pp. 87-94.

Holguín-Veras, J., and Thorson, E., (2003b). "Modeling commercial vehicle empty trips with a first order trip chain model." Transportation Research Part B: Methodological, Vol. 37, No. 2, pp. 129-148.

Holguín-Veras, J., Thorson, E. and Zorrilla, J. C., (2010). "Commercial vehicle empty trip models with variable zero order empty trip probabilities." Networks and Spatial Economics, Vol. 10, No. 2, pp. 24 1.

Holguín-Veras, J., and Patil, G. R., (2008). "A multicommodity integrated freight origin–destination synthesis model." Networks and Spatial Economics, Vol. 8, No. 2, pp. 309-326.1-259.

Holguín-Veras, J., and Jara-Díaz, S., (1999). "Optimal Space Allocation and Pricing for Priority Service at Container Ports." Transportation Research, Part B, Vol. 33, No. 2, 1999, pp. 81–106.

Jansuwan, S., Ryu, S. and Chen, A., (2017). "A two-stage approach for estimating a statewide truck trip table." Transportation Research Part A: Policy and Practice, Vol. 102, pp. 274-292.

Kulpa, T., (2014). "Freight truck trip generation modelling at regional level." Procedia-Social and Behavioral Sciences, 111, 197-202.

McFadden, D., C. Winston, and Boersch-Supan, A., (1986). "Joint Estimation of Freight Transportation Decisions under Nonrandom Sampling." Analytical Studies in Transport Economics. Cambridge University Press, Cambridge, England.

McNally, M. G. (2000). The four step model.

Mesa-Arango, R., Ukkusuri, S., and Sarmiento, I., (2013). "Network flow methodology for estimating empty trips in freight transportation models." Transportation Research Record: Journal of the Transportation Research Board, No. 2378, pp. 110-119.

Middela, M. S., Pulipati, S. B., and Prasad, C. S. R. K., (2018). "Modelling Freight Generation and Distribution for Nationwide Interstate Freight Movement." Transportation in Developing Economies, 4(1), 6.

National Private Truck Council., (2007). "Whitepaper on backhaul networking." http://www.nptc.org/network/learn-more/download-center/general-miscellaneous/318-whitepaper-on-backhaul-networking-1.

Noortman, H.J., van Es, J., (1978). Traffic Model. Manuscript for the Dutch Freight Transport Model

Roorda, M. J., Hain, M., Amirjamshidi, G., Cavalcante, R., Abdulhai, B. and Woudsma, C. 2010. Exclusive truck facilities in Toronto, Ontario, Canada: analysis of truck and automobile demand. Transportation Research Record, 2168(1), pp.114-128.

Sheckler, R., Maynus, L.W. and Calmar Telematics, L.L.C., (2009). Empty backhaul, an opportunity to avoid fuel expended on the road (No. C-08-31). New York (State). Dept. of Transportation.

Zanjani, A.B., Pinjari, A., Kamali, M., Thakur, A., Short, J., Mysore, V. and Tabatabaee, S.F., (2015). Estimation of statewide origin–destination truck flows from large streams of GPS data: Application for Florida statewide model (No. 15-5463).