Analysis and Multi-Level Modeling of Truck Freight Demand

Huili Wang$^{1,2}$, Kitae Jang$^1$, Ching-Yao Chan$^{1,3}$

1. California PATH, University of California at Berkeley, Richmond, CA, USA, 94804
2. Northwestern Polytechnical University, Xi’an, Shaanxi, P.R.China, 710072
3. Corresponding Author: TEL: 510-665-3621, E-mail: cychan@path.berkeley.edu

Abstract: Trucks play a significant role of cargo shipping through the ground transportation systems in U.S.. From the perspectives of logistics planning as well as infrastructure management, it is important to be able to predict and estimate the amount of freight in a highway network. In this paper, we start with an analysis of truck weight and volume data obtained from weight-in-motion stations to examine temporal and seasonal patterns. We then propose a hierarchical model by coupling the truck traffic and weight data with socio-economic variables to establish a freeway-level statistical model for freight demand estimation and prediction. The model reveals that the truck freight demand can be estimated by the average daily truck traffic on freeway segments along with the population, number of firms and median income at the county level. The model can serve as a useful for practitioners to predict the growth of freight demand in a freeway network, thus allowing the identification of potential freight transport bottlenecks as well as achieving design for optimal flows of cargo.

Key words: Freight demand, Multi-level model, Hierarchical model, Temporal variation, Seasonal variation

I. Introduction

Trucks are one of the most widely used transportation modes to carry freight. Other modes, such as airplanes, ships, and trains are restricted to specific routes and ports. Trucks can transport goods directly “door-to-door” and other modes sometimes require the use of trucks for deliveries to their final destinations. According to a 2007 commodity flow survey, in the United States, 8.78 billion tons of shipments were transported annually by trucks for an average of 206 miles per carry [1]. Given the importance of truck shipping, it will be extremely beneficial to be able to predict the amount of freight moving in segments of a highway network for both planning and operational purposes so as to address issues and concerns of infrastructure deterioration, truck-involved safety, etc.

Previous studies in this field mostly attempted to model large-scale freight demands. Sorratini and Smith developed a model, which used commodity flow data, a private freight database and input-output coefficients to formulate a statewide truck travel demand model [2]. In another approach presented by Fite et al., researchers attempted to predict future freight volumes using several economic indices [3]. Researchers Fernandez et al developed a different kind of model to forecast intercity freight transportation, which considered supply-demand equilibrium between shippers and carriers [4]. Garrido and Mahmassani used a multinomial probit model with spatially and temporally correlated error structure to predict freight demand [5]. Since the modeling techniques are developed with a large-scale focus, these earlier studies were not applicable for the investigation of freight demand at the freeway level, which is critical for improving operational efficiency.

In this study, a multi-level statistical model for freight demand was developed using economic data from the U.S. Census Bureau and California freight data from the Performance Measurement Systems (PeMS) database [6], which includes both truck volumes and weight data that were obtained from Weight-in-Motion (WIM) stations. WIM is a technology for determining the weight of a commercial vehicle without requiring it to stop on a scale. It can capture and record axle weights and gross vehicle weights as vehicles drive over a measurement site at the normal traffic speeds. Weight and volume data from 66 WIM
stations in California were used to examine the patterns of temporal and seasonal variations and to estimate parameters of statistical model for freight demand.

The remaining portion of this paper is organized as follows: in Section II, the PeMS freight data was analyzed in hourly, daily, and monthly intervals to observe the temporal and seasonal variations of freight movement. In Section III, a multi-level model was developed, which linked truck traffic volume and the region’s economic conditions to the amount of freight transported through various segments in a freeway network. In Section IV, concluding remarks are given.

II. Temporal & Seasonal Variations

Due to the nature of freight movement, the temporal and seasonal variations of freight do not follow the same patterns as general traffic in a freeway network. Trucks generally operate at all hours of the day, whereas passenger vehicles are typically more concentrated in peak hours.

After analyzing the data from 66 WIM sites, it was found that all exhibit similar patterns and characteristics. The WIM data, including truck traffic volume, average truck weight, and total freight weigh from one study site are analyzed and presented below for illustration purposes. These data are taken from a WIM station in Caltrans District 1 in Humboldt County on route US-101 at CA post-mile 64.287 in the northbound direction from March 2008 to February 2009. The following figures show clear temporal patterns of freight movement variations in different time intervals from hours of a day to months of a year.

- Hourly variation

![Figure 1. Hourly Average Truck Weight](image1)
![Figure 2. Hourly Truck Volumes](image2)
![Figure 3. Hourly Total Freight Weight](image3)

![Figure 4. Daily Average Truck Weight](image4)
![Figure 5. Daily Truck Volumes](image5)
![Figure 6. Daily Total Freight Weight](image6)

![Figure 7. Monthly Average Truck Weight](image7)
![Figure 8. Monthly Truck Volumes](image8)
![Figure 9. Monthly Total Freight Weight](image9)
It is noticeable that there is a disproportional amount of heavy truck movements during the night hours. The truck traffic also shows a single peak in the middle of the day, in contrast to the typical double peaks of general traffic patterns.

- **Daily variation**
  From the daily data, it can be seen that all have a similar trend in that the truck traffic is heavier during the weekdays than on weekends, indicating the freight movements still conform to the business activities.

- **Monthly variation**
  The freight demand varies throughout the year because some goods movements and economic activities are seasonal. The monthly truck volumes and total freight weight are higher in the summer months than in the winter months. On the other hand, the average truck weight is fairly constant over the year.

### III. Truck Freight Demand Modeling

Trucks transport goods from the origin to the final destination through freeways. Since there can be substantial amount of influential factors between origins and destinations, and this would significantly affect the freight demand. Such unobserved (latent) variables may result in more correlated observations within groups than between groups. In the present study, those groups are: i) counties in which same socio-economies affect the freight demand; and ii) freeways that connect comparable sets of origins and destinations. In this section, therefore, a model is developed to relate freight demand and truck AADT under consideration of potential correlated structure of data.

#### A. Multi-Level Model Specification

The model in the present study was developed to explain the relationship between Truck AADT, socio-economic factors and freight demand. To consider variability between counties and freeways, random intercepts were specified. The conceptual hierarchy of the data is that freeways transverse multiple counties and, within each county, there can be many WIM stations as shown in Figure 10.

![Figure 10. Conceptual Hierarchy of Data](image)

In the model, thus, random intercepts were specified for two higher levels–freeways and counties. It is explained as the following:

\[
\begin{align*}
    y_{ijk} &= \beta_i + \beta_k \cdot \text{truck AADT}_k + \epsilon_{ijk} & \text{(Level-1 Equation)} \\
    \pi_{ij} &= \gamma_{ij} \cdot x_{ij} + u_{ij} & \text{(Level-2 Equation)} \\
    \pi_i &= u_i & \text{(Level-3 Equation)}
\end{align*}
\]

Then, the reduced form of these equations becomes:

\[
y_{ijk} = \beta_i + \gamma_{ij} \cdot x_{ij} + \beta_k \cdot \text{truck AADT}_k + u_{ij} + u_i + \epsilon_{ijk}
\]  

(1)
Where the index $i$ is for the freeway, $j$ for the county and the $k$ for the WIM station. $y_{ijk}$ denote the freight, truck AADT$_k$ is the volume detected by WIM $k$, and $x_{ij}$ is the socio-economic variable on freeway $i$ in county $j$.

Therefore, the coefficients, $\beta_i$ and $\beta_k$, are fixed intercept and the coefficient for the Truck AADT respectively, and $\gamma_{ij}$ are the county-level coefficients for socio-economic variables. $u_{ij}$ and $u_i$ are the random intercepts, that follow a normal distribution with mean zero and variances $\psi^{(2)}$ and $\psi^{(3)}$, and $\epsilon_{ijk}$ is error term of the model with mean zero and variance $\theta$. As defined and described for the hierarchical model, the Level-2 Equation signifies that the random intercept varies between counties depending on socio-economic variables and the Level-3 Equation indicates that the random intercept varies between freeways.

**B. Model Estimation**

By analyzing the data from the PeMS web and the economic data, it was found that Truck AADT was a key predictor for the freight demand with statistical significance. Among many other socio-economic variables, three variables were entered in the model as regional socio-economic indices. Population can represent the user scale of the county and, thus, correlated with the level of goods needed for consumptions. Median income is an index of wealth of a county that can be interpreted as population groups’ consumption power or the characteristics of the regional landscape. To capture the impact of commercial activities, the number of firms was entered to represent the characteristics of a region. Table 1 summarizes the estimates of coefficients and the log-likelihood test against linear regression, which indicates the hierarchical structure of the model performs better significantly. It can be seen clearly that truck traffic volume is a dominant variable. One unit (vehicle count) increase in truck AADT is associated with about 12 units (tons) increase in freight demand. Also listed in table 1, for median income, population and firms, the result is significant at 10%. As an example of presenting outcome of such a freight demand model, the actual freight demand data for the California WIM stations in different counties and freeways from March 2008 to February 2009 is given in Figure 11.

<table>
<thead>
<tr>
<th>Table 1. Estimate of Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
</tr>
<tr>
<td><strong>Fixed Part</strong></td>
</tr>
<tr>
<td>Constant ($\beta_1$)</td>
</tr>
<tr>
<td>Level-1</td>
</tr>
<tr>
<td>Truck AADT ($\beta_k$)</td>
</tr>
<tr>
<td>Level-2</td>
</tr>
<tr>
<td>Median Income</td>
</tr>
<tr>
<td>Population</td>
</tr>
<tr>
<td>Firms</td>
</tr>
<tr>
<td><strong>Random Part</strong></td>
</tr>
<tr>
<td>Level-2</td>
</tr>
<tr>
<td>$\sqrt{\psi^{(2)}}$</td>
</tr>
<tr>
<td>Level-3</td>
</tr>
<tr>
<td>$\sqrt{\psi^{(3)}}$</td>
</tr>
<tr>
<td>Residual</td>
</tr>
<tr>
<td>Log likelihood at convergence</td>
</tr>
</tbody>
</table>

*: significant at 10%, **: significant at 1%

---

**Figure 11. Actual Freight Demand (2008.3-2009.2)**
C. Intraclass Correlation
Intraclass Correlation (IC) describes how strongly observations in the same group correlate each other. In other words, IC for each level indicates the level of compaction within the group. Thus, higher IC means the structure of the model is valid and vice versa.

For the same freeway $i$ but different counties $j$ and $j'$, we obtained

$$
\rho_1 = \frac{\psi^{(3)}}{\psi^{(2)} + \psi^{(3)} + \theta} = 0.64
$$

(2)

Whereas for the same freeway $i$ and same county $j$, we got

$$
\rho_2 = \frac{\psi^{(2)} + \psi^{(3)}}{\psi^{(2)} + \psi^{(3)} + \theta} = 0.77
$$

(3)

The value of the intraclass correlation coefficient is defined between 0 and 1. There are many factors that can affect freight demand, however just for three factors in the model, the correlation for level 2 and 3 are greater than 0.6. From the calculations above, freight demand is correlated for the same freeway and same county. The result verifies the reasonable structure of the model.

IV. Discussions and Conclusion
In this paper, the truck data obtained from California WIM stations were first analyzed in hourly, daily and monthly time windows to display the temporal and seasonal variation of the freight movement. It was found that at most sites, heavier trucks operated during the daytime and there was a “plateau” peak of truck traffic in the middle of the day. Truck volumes were typically greater during the weekdays than on the weekend. And lastly, both truck volumes and the monthly average weight were greater during the summer months than during the winter months demonstrating the seasonal supply and demand of goods.

A multi-level model is developed to provide a tool for estimating freight demand on different freeways and in different counties. The truck data from a number of California WIM stations were combined with socio-economic factors obtained from US Census Bureau to construct a three-level model. The random effect was considered in this model to consider correlated structure of data that may be caused by unobserved (latent) variables.

As expected, freeway segments with higher average truck volumes and regions with greater population yield a higher freight demand. The median income was found to have a negative relationship with freight demand. The reason may be that high-income neighborhoods are less likely to have a larger freeway network running through it or to have large volume of truck traffic. Similarly, the number of firms was also found to be inversely related with freight demand. It can be hypothesized that this reflects the pattern of freight flow in these regions as the goods are transported away from the central business district where a large number of firms tend to be located.

Reference: