A BEHAVIORAL MODELING IN MICRO-SIMULATION FOR URBAN FREIGHT TRANSPORTATION

Wisinee WISETJINDAWAT
Graduate Student
Transportation Engineering Program
Asian Institute of Technology
P.O. Box 4, Klong Luang, Pathumthani,
12120 Thailand
Fax: +66-2524-5509
E-mail: st019991@ait.ac.th

Kazushi SANO
Associate Professor
Transportation Engineering Program
Asian Institute of Technology
P.O. Box 4, Klong Luang, Pathumthani,
12120 Thailand
Fax: +66-2524-5509
E-mail: sano@nagaokaut.ac.jp

Abstract: This paper proposes and applies a micro-simulation model for the modeling of urban freight transportation that describes freight transportation at the behavioral level. The proposed model is a modification of the traditional four-step approach considering the behavior of each freight decision maker individually. The model structure consists of three stages: commodity production and consumption, commodity distribution, and conversion of commodity flow to vehicle movement. The proposed model has been applied to estimate truck movement in the Tokyo Metropolitan Area for model validation. This paper discusses the issues involved in model development and validation including conceptual framework, mathematical formulation, and estimated results.

Key Words: urban freight transportation, micro-simulation, behavioral model, disaggregate model

1. INTRODUCTION

Freight Transportation is a major factor contributing to economic growth and development. However, freight transportation issues should be considered carefully in both modeling and policy making because they involve with a significant amount of energy consumption and pollution. As a result, a proper freight transportation-modeling tool is needed for transportation planners in order to assess to the effect of policy decision on freight transportation system.

The current direction of demand modeling is to model at the micro level for both freight demand modeling and passenger demand modeling. It is widely accepted that demand modeling at the micro level results in a more realistic and policy-sensitive model. In passenger demand models, the activity-based approach is now receiving more attention than in the past. The activity-based approach describes an individual's daily activities more realistically. It is possible to views freight transportation at the behavioral level as well as in passenger transportation. Freight transportation characteristics in particular are more complex than those of passenger transportation. Freight transportation involves a very complex linkage among many freight decision makers (such as shippers, customers, and carriers), while in passenger transportation, there is only one decision maker: the passenger. In addition, freight transportation deals with commodities that vary in volume, weight, and shape, whereas relevant passenger characteristics are less variable. The objective of this paper is therefore to develop a model that focuses on the behavioral level of freight transportation. The proposed model attempts to incorporate the behavior of freight decision makers interacting in a supply chain. The research utilizes micro-simulation as an approach to the modeling, considering the behavior of each firm individually. The research scope covers three of the four steps in the traditional four-step approach: commodity generation, commodity distribution, and conversion of commodity OD to freight OD. It does not cover vehicle trip assignment. The model has been tested through application urban freight movement in the Tokyo Metropolitan Area.

The first section of the paper provides the background information on freight transportation modeling in which the earlier freight transportation models are reviewed. The key behavior of freight decision makers is also described. The next section discusses the model framework and details on model development. In addition, the mathematical formulation used in the model is presented. Finally, the last section contains the results of model application to a real

transportation system and recommendation on the proposed model.

2. BACKGROUD

2.1 Freight Transportation Modeling

Earlier freight transportation models can be categorized into two groups: trip-based and commodity-based approaches (Veras and Thorson, 2000). Trip-based approach deals with truck trips that are generated directly from the size indicators for trip generation such as the number of employees and floor area without concerning the amount of commodity production or consumption. The generated trips are then distributed from generation points to attraction points resulting in trip origins and destinations, which are assigned to a network via traffic assignment. An example of this approach is the model of truck estimation with limited origin-destination survey data by Park and Smith (1997). The model generates truck trips using trip rate technique and utilized the gravity model, which was calibrated from the available OD survey data, for distributing the generated trips. Matt and Visser (1996) utilized the Geographic Information System (GIS) to incorporate the trip chain characteristic of freight transportation. After trip generation and attraction were performed, the GIS was used for route simulation resulting in the sequence of picking up or delivery at the optimal distance, time, or cost. Nevertheless, this approach has limitations, as freight trips are directly generated; it is therefore not applicable to evaluate logistics policies that affect the changes in the characteristics of commodity movement.

Commodity-based approach attempts to overcome the drawback of the former approach. It starts from the movement of commodities. At the stage of trip generation, instead, the approach generates the level of commodity production and consumption. Consequently, it requires the additional stage to convert the commodity flows to vehicle tours. Sivakumar and Baht (2002) developed the fractional split distribution model that focused on the stages of commodity generation and distribution. The consumption level was determined by commodity generation and commodity distribution was estimated by the fractional split distribution model, which was similar to a multinomial logit model, instead of generally using a gravity model. Boerkamps et al (2000) proposed the behavioral model of freight movement in a supply chain. The model predicted the consumption demand based on the end user in a supply chain. The demand was then linked to the suppliers according to the spatial distribution and the market share. The commodity flows were then allocated to vehicle tours and modes and assigned to an infrastructure network. The good point of this model is the logistical concept incorporated in the model that reflects the actual behavior of freight transportation. The proposed model in this paper also belongs to this category.

2.2 Micro-Simulation

Micro-simulation is an analysis approach to modeling the behavior of an individual, in other word, it is an extension of the disaggregate model. The disaggregate model needs enormous data that contains all the components, such as, in case of freight transportation, firm's location, delivery frequencies, shipment size, cost and so on. From the data viewpoint, it is impossible to gather such data from all the individuals. Therefore, the application of the disaggregate model is limited to a small area. In contrast, the micro-simulation allows expanding the survey data to all the individuals from their distribution; micro-simulation is consequently applicable to the wide area. The research of Vovsh et al (2002) is an example of the application of micro-simulation to transportation demand model. The authors discussed the advantages of micro-simulation that was successfully applied to modeling the passenger trips in New York/New Jersey/Connecticut metropolitan area. They stated that micro-simulation had ability to handle the complex linkages across multiple trips. It was because micro-simulation considered the various chain decisions and time-space constraints on individual travel behavior. In addition, it could incorporate variability of travel demand, as the distribution of traffic volume was more important than the average volume due to the design based on the critical maximum volume. The micro-simulation was able to estimate the range of likely traffic volume rather than the average daily or hourly volume. All the above is the results while applying micro-simulation to passenger transportation demand. Likewise, the advantage of micro-simulation is also suitable for the problem of urban freight transportation that involved in the complex trip chains and systems.

2.3 Freight Decision Makers in a Supply Chain

The complexity of freight transportation is caused from the complicated interactions among the freight decision makers, which include shippers, customers, carriers, and administrators at the city level (Taniguchi *et al.*, 2001). Shipper is the person who sells the products to customers. Manufacturer, wholesaler, and retailer can behave as a shipper or a customer at the same time. The carriers are the person who distributes the products to the customers. The shippers can select either asking for the carriers or delivering directly by themselves. Both shippers and carriers attempt to maximize their profit through minimizing the costs of pickup or delivery the products to the customers, while customers need the products to be pickup or delivered on time. Administrator at the city level is the person who makes a decision on freight transportation in order to enhance the environmental condition, alleviate the traffic congestion, and any other enhancement to the society toward the logistics initiatives.

Each freight decision maker has the different role on freight movement. Boerkamps (2000) summarized the key behavior of freight decision makers in the freight movement market. Shippers have the important role in product availability, transport mode and carrier choice, and location. Customers strongly influence on the demand of products and location. Demand involving the amount of products, delivery frequency, and lot size are depends on customers. Carriers have a major role in delivering the product to customers. The task includes transport mode choice, services, and route choice. Lastly, administrators control the freight movement in an indirect way. The policy measures, regarding freight movement system such as truck regulation and construction of infrastructure, significantly affect the behavior of the former decision makers. Understanding the role of all decision makers is necessary to capture the complex mechanism of freight transportation in the modeling.

3. MODEL STRUCTURE

This research develops a micro-simulation model that is a modified version of the traditional four-step approach. The proposed model is a commodity-based model that considers the behavior of all firms individually. In order to estimate truck trip OD matrices, the model comprises three parts: a commodity production and consumption model, a commodity distribution model, and a conversion of commodity flows to vehicle tours. This section discusses the concept and mathematical formulations of the developed model.

3.1 Commodity Production and Consumption Model

The production and consumption model determines the total monthly amount of commodity produced and consumed by each firm. Production and consumption amounts are estimated from the firm characteristics such as number of employees, floor area, and other indicators. This model utilizes regression techniques to estimate production and consumption amounts. Commodity production and consumption are estimated by type of commodity. The following equation shows the amount of commodity *k* produced by firm *i*, represented by G_i^k, and the amount of commodity *k* consumed by firm *j*, represented by A_j^k.

$$G_i^k = f(x_1, x_2, ..., x_n) \tag{1}$$

$$A_j^k = f(x_1, x_2, ..., x_n) \tag{2}$$

3.2 Commodity Distribution Model

Consumption demand fundamentally determines the commodity flows. Originating from the demand side, this model makes a connection from demand to supply according to the attractiveness of suppliers and distribution channels. The attractiveness of suppliers is derived from supplier location and the amount of commodity produced by that supplier. Distribution channels can be defined as the paths connecting the customers and shippers of a commodity. In a supply chain, there are a number of possible distribution channels. In a simple example, as shown in Figure 1, retailers can purchase a commodity from wholesalers or the higher levels.

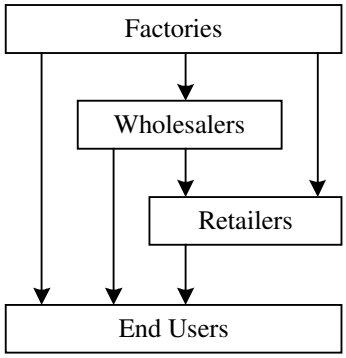

Figure 1. Distribution Channels

In that case, the distribution model consists of three parts: probability of distribution channel, probability of shipper location, and probability of selecting each shipper. The product of all three parts yields the probability of a given shipper being selected. The following shows the mathematical form of the model.

$$P^k(i) = P(C^k) \cdot P(z|C^k) \cdot P(i|C^k, z), \quad i \in C^k, z \tag{3}$$

where,

$P^k(i)$ is the probability of shipper i being selected for commodity k

$P^k(C^k)$ is the probability of distribution channel C being used for commodity k

$P^k(z|C^k)$ is the probability of zone z being selected, given distribution channel C

$P^k(i|C^k, z)$ is the probability of shipper i is selected, given distribution channel C and zone z

The first part, the distribution channel probability, $P(C^k)$ is calculated directly from the empirical data. The location probability, on the other hand, can be viewed as a choice selection problem with multiple alternatives. The research utilizes a multinomial logit model and assumes a utility function for zone attractiveness, which is represented by firm spatial distribution and attractive potential, such as the number of firms in a zone and the total amount of commodity generated within a zone. A customer is assumed to select the zone that maximizes their utility function among the zone alternatives. The conditional probability that zone z is chosen, given distribution channel C^k, is expressed below:

$$P(z|C^k) = \frac{\exp(\mu^z V_z)}{\sum_{z' \in z} \exp(\mu^z V_{z'})} \tag{4}$$

where,

μ^z is the scale parameter for zone z

V_z is the utility function of zone z and can be written as:

$$V_z = f(D_{IJ}, N_I^{C,k}, G_I^k) \tag{5}$$

D_{IJ} is the distance between zone I and zone J

$N_I^{C,k}$ is the number of firm type C that produce commodity k in zone I

G_I^k is the total production amount of commodity k from zone I

The third part, the shipper probability is necessary to identify the shippers from which a customer purchases. A multinomial logit model is applied in this part as well. However, due to the limitation of the survey data that does not identify the exact purchased shipper, we then assume that the shipper probability can be derived directly from the proportion of the size

indicators of each shipper; in this research, the production amount is utilized. The conditional probability that shipper *i* is chosen, given C^k and *z*, is as shown below:

$$P(i|C^k, z) = \frac{\exp(G_i^k)}{\sum_{i' \in i} \exp(G_{i'}^k)} \tag{6}$$

where,

G_i^k is the production amount of commodity *k* from shipper *i*

The number of shippers from which a customer purchases, is estimated by regression techniques from the indicators for customer size. This research utilizes number of employees as an indicator as expressed below:

$$Ns_j = f(E_j) \tag{7}$$

where,

Ns_j is the number of shippers from that a customer purchases

E_j is the number of employees in firm *j*

Once the shipper probability and the number of shippers are determined, random numbers are generated and used with the shipper probability to indicate the selected shippers and determine the share for each selected shipper. Commodity flows are then derived from the product of shipper share and consumption amount for each firm as expresses in equation (8). In simulation, shipper probability will be recalculated every iterations in order that the commodity flows will be satisfied both constraints of the total production and consumption amounts.

$$Q_{ij}^k = P^k(i) \cdot A_j^k \tag{8}$$

where,

Q_{ij}^k is monthly commodity flow between firm *i* to firm *j* for commodity *k*

$P^k(i)$ is probability of shipper *i* is selected for commodity *k*

A_j^k is monthly consumption amount consumed by firm *j* for commodity *k*

3.3 Conversion of Commodity Flows to Trip Chain

Commodity flows will be assigned to vehicle movement according to the following three steps: delivery lot size and frequency, carrier and vehicle choice selection, and vehicle routing.

3.3.1 Delivery Lot Size and Frequency

Delivery lot size is the amount of commodities delivered to a customer at a time. Likewise, delivery frequency is defined as the number of deliveries to a customer during a certain time period. The determination of delivery lot size and frequency is a necessary step in converting monthly commodity flows to commodity flow in one time. Assuming that lot sizes are the same at every delivery for each customer, thus, delivery lot size and frequency have the following relationship:

$$Q_{ij}^k = L_{ij}^k \cdot F_{ij}^k \tag{9}$$

where,

Q_{ij}^k is the monthly commodity flow of commodity *k* between firm *i* and firm *j*

L_{ij}^k is the lot size of commodity *k* that shipper *i* delivers to customer *j*

F_{ij}^k is the delivery frequency of commodity *k* that shipper *i* delivers to customer *j*

Delivery lot size and frequency depend on many factors including inventory cost, transportation cost, demand, shipper availability, and other factors. Delivery frequency and lot size are selected in the way that minimizes the total costs. This research calculates the total cost as a function of inventory cost and transportation cost. Inventory cost comprises three types of cost: purchasing cost, ordering cost, and holding cost. Purchasing cost is proportional to monthly purchased amount, while ordering cost is proportional to delivery frequency, and, holding cost is proportional to delivery lot size. Inventory cost function can therefore be expressed as follows:

$$IC_{ij}^{k} = \alpha^{k} \cdot Q_{ij}^{k} + \beta^{k} \cdot F_{ij}^{k} + \gamma^{k} \cdot L_{ij}^{k} \tag{10}$$

where,

α^{k} is a parameter for the purchase cost of commodity k

β^{k} is a parameter for the ordering cost of commodity k

γ^{k} is a parameter for the holding cost of commodity k

Transportation cost is proportional to delivery frequency and travel distance. Transportation costs can be calculated as follows:

$$TC_{ij}^{k} = \delta^{k} \cdot F_{ij}^{k} \cdot D_{ij}^{k} \tag{11}$$

where,

δ^{k} is a parameter for transportation cost of commodity k

Total cost is therefore the summation of inventory cost and transportation cost as expressed below:

$$C_{ij}^{k} = \alpha^{k} \cdot Q_{ij}^{k} + \beta^{k} \cdot F_{ij}^{k} + \gamma^{k} L_{ij}^{k} + \delta^{k} \cdot F_{ij}^{k} \cdot D_{ij}^{k} \tag{12}$$

Based on minimization of the total cost, the first derivative of total cost with respect to delivery frequency provides the optimal delivery lot size and frequency. The following expresses the relationship between the optimal delivery lot size and frequency.

$$\frac{1}{\gamma^{k}}(\beta^{k} + \delta^{k} \cdot D_{ij}^{k}) = \frac{Q_{ij}^{k}}{(F_{ij}^{k})^{2}} \tag{13}$$

or,

$$a^{k} + b^{k} \cdot D_{ij}^{k} = \frac{L_{ij}^{k}}{F_{ij}^{k}} \tag{14}$$

Parameters a^{k} and b^{k} can be calibrated directly from survey data; therefore, delivery lot size and frequency are determined from the following equation:

$$F_{ij}^{k} = \sqrt{\frac{Q_{ij}^{k}}{a^{k} + b^{k} \cdot D_{ij}^{k}}} \tag{15}$$

3.3.2 Carrier and Vehicle Choice Selection

Once commodity movements are determined, carrier and vehicle choices are decided for each commodity flow. Shippers play a major role in the selection of carrier and vehicle choice. The characteristics of the shippers, customers, transported commodities, and firm spatial distribution strongly influences the decision. The characteristics of shippers and customers can be represented by type of firms (retailer, wholesaler, or manufacturer), number of employees, and other characteristics. Attributes of transported commodities include commodity type, delivery lot size and frequency, and other characteristics related to the commodities.

Discrete choice model is generally applied to a mode choice problem. This research utilizes a

nested logit model to describe the choice decision process. The model is structured on two levels (see Figure 2): carrier choice and vehicle choice. At the first level, the choice is carrier choice between private truck and business truck. Vehicle choice at the second level indicates truck size. This research categorizes truck size into two types: small truck and large truck. The small truck category covers the trucks that have maximum carrying weight less than five tons, including light truck, small truck, and pickup truck. In contrast, the large truck category covers the trucks that have more than five-ton maximum carrying weight including large truck and special truck.

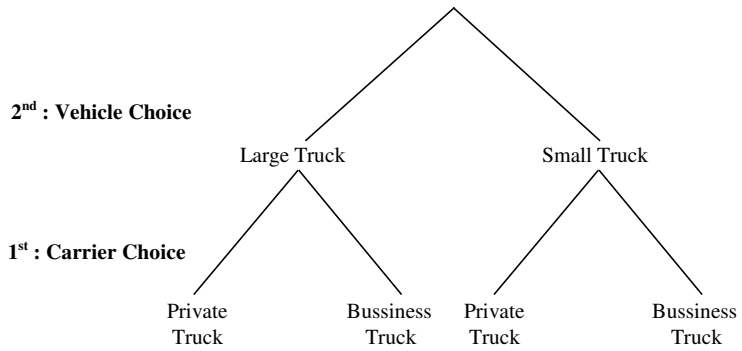

Figure 2. Model Structure for Carrier and Vehicle Choice

First level: Carrier choice between private and business truck

$$P_{ij}(private) = \frac{\exp(V_{prv})}{\exp(V_{bss}) + \exp(V_{prv})} \tag{16}$$

where,

V_{prv} is the utility function for a private truck. It can be written as:

$$V_{prv} = f(Nc_i, E_i, L_{ij}, F_{ij}, TT_{ij}) \tag{17}$$

V_{bss} is the utility function for a business truck, in this case, this value is zero

Nc_i is the number of customers for shipper *i*

E_i is the number of employees for shipper *i*

L_{ij} is the lot size delivered from shipper *i* to customer *j*

F_{ij} is the delivery frequency from shipper *i* to customer *j*

TT_{ij} is the travel time between shipper *i* and customer *j*

Second level: Vehicle choice between small and large Truck

$$P_{ij}(Large) = \frac{\exp(V_{lrg} + V_p^{'} \cdot \mu_p)}{\exp(V_{sml}) + \exp(V_{lrg} + V_p^{'} \cdot \mu_p)} \tag{18}$$

where,

V_{lrg} is the utility function for a large truck, which can be expressed as:

$$V_{lrg} = f(Nc_i, E_i, L_{ij}, F_{ij}, TT_{ij}) \tag{19}$$

V_{sml} is the utility function for a small truck, in this case, this value is zero

Nc_i is the number of customers of shipper *i*

E_i is the number of employees of shipper *i*

L_{ij} is the lot size delivered from shipper *i* to customer *j*

F_{ij} is the delivery frequency from shipper *i* to customer *j*

TT_{ij} is travel time from shipper *i* to customer *j*

μ_p is the scale parameter for private truck

$V_p^{'}$ is the log-sum variable of the nested logit model which can be written as:

$$V_p^{'} = \ln\left[\exp(V_{prv}) + \exp(V_{bss})\right] \tag{20}$$

Therefore, carrier and vehicle choice probability for selecting large private truck can be expressed as:

$$P_{ij}(LP) = \frac{\exp(V_{prv})}{1+\exp(V_{bss})} \cdot \frac{\exp(V_{lrg} + V_p^{'} \cdot \mu_p)}{1+\exp(V_{lrg} + V_p^{'} \cdot \mu_p)} \tag{21}$$

3.3.3 Vehicle Routing

After assigning lot size, frequency, and vehicle type to each commodity flow between a shipper and a customer, vehicle routing simulation will provide an order of visiting customers for each shipper in delivery, resulting in the delivery route or saying that vehicle tours are constructed. Vehicle routing is simulated based on the commodity flow characteristics including lot size, carrier and vehicle choice. The delivery route is therefore decided in the way that minimizes the route travel time, which is constrained by the maximum working hours of a driver and limited carrying weight of a truck. The formulation can be written as:

Min $$TC_i = \sum_{m=1}^{n}\sum_{l=1}^{n}(TT_{l,m} \cdot x_{l,m}) \tag{22}$$

Subject to: $$\sum_{m=1}^{n} x_{l,m} = 1 \tag{23}$$

$$\sum_{l=1}^{n} x_{l,m} = 1 \tag{24}$$

$$\sum_{m=1}^{n}\sum_{l=1}^{n}(TT_{l,m} \cdot x_{l,m}) \le HR_{\max} \tag{25}$$

$$\sum_{m=1}^{n}\sum_{l=1}^{n}(L_{l,m} \cdot x_{l,m}) \le WT_{\max} \tag{26}$$

$$x_{l,m} \in \{0,1\} \tag{27}$$

where,

TC_i is the total travel time of a delivery route for shipper *i*

$TT_{l,m}$ is the travel time between customer *l* and customer *m*

n is the number of customers

$x_{l,m}$ is 1, when the link between customer *l* and customer *m* exists

0, otherwise

$HR_{\max}$ is the maximum working hours of a driver

$WT_{\max}$ is the maximum carrying weight of a truck

4. ESTIMATED RESULTS

4.1 Data set

The Tokyo Metropolitan Goods Movement Survey (TMGMS) data used in the model calibration were collected from firms all over the Tokyo Metropolitan Area by the City and the Regional Development Bureau of the Japanese Ministry of Construction in 1982. The data consist of records on commodity movement and truck movement of each firm. The data were from approximately 46,000 firms, corresponding to three percent of all firms in the study area. Each record provides information about firm characteristics, commodity movement, and truck movement. The information of firm characteristics includes industry type, number of employees, floor area, and other related information. In the same way, the information on commodity movement and vehicle movement includes commodity type, weight carried in and out, delivery frequency, truck type, carrier type, and other related information.

The Establishment and Enterprise Census data (EEC) were collected by the Statistics Bureau in the Japanese Ministry of Public Management, Home Affairs, Posts and Telecommunication in 1999. This data covers all the enterprises in Japan. The data provide the general information about firms, for example, industry type, location, number of employees, and other related issues. The 174,892 virtual firms and their attributes are generated using the Monte-Carlo simulation according to the distribution of the attributes in the EEC data. The number of virtual firms utilized in this research corresponds to 10 percent of all firms in the study area. In model validation, the virtual firms are input to the model to estimate truck origin and destination volumes, which will be compared with the actual volumes from the survey data, which were collected by the Road Traffic Census in 1999.

4.2 General Information

4.2.1 Industries and Commodities

This research categorizes firms into 13 industry types, and commodities into 8 types based on the TMGMS classification. The categories of industry types and commodity types are presented in Tables 1 and 2 respectively.

Table 1 Industry Type Classification

Industry Type	Description
1	Agriculture, Forestry, and Fishery
2	Mining
3	Construction
4	Chemical Manufacturer
5	Metal Manufacturer
6	Machinery Manufacturer
7	Other Manufacturer
8	Material Wholesaler
9	Product Wholesaler
10	Retailer
11	Warehouse
12	Electricity, Gas and Water Supplier
13	Service and Government Work

Table 2 Commodity Type Classification

Commodity Type	Description
1	Agricultural Products
2	Forestry Products
3	Mineral Products
4	Metal and Machinery Products
5	Chemical Products
6	Light Industry Products
7	Other Products
8	Wastes and Scraps

4.2.2 Zoning

This research utilizes A-zone classification of the Tokyo Metropolitan Goods Movement Survey zone system. There are 56 zones, comprising 52 zones within the study area (Tokyo Prefecture, Kanagawa Prefecture, Chiba Prefecture, Saitama Prefecture, and the southern part of Ibaraki Prefecture) and 4 zones for the prefectures near the study area for analysis of the external trips.

4.3 Estimation Results of Model Calibration

4.3.1 Commodity Production and Consumption

Commodity production and consumption models are developed separately for each industry type and each commodity type. The probability of commodity type produced and consumed by each industry type is directly calculated from the TMGMS data and summarized in Table 3. The commodity type produced or consumed by fewer than 5 percent of the firms will be neglected. The monthly amount of production and consumption are estimated from the number of employees and floor area. It should be noted that we are indirectly considering the end users as their purchasing amount can be indicated by the consumption amount of retailers. The production and consumption functions are assumed linear relationship, which can be written as follow:

$$G_i^k = \alpha^k + \beta^k \cdot Emp_i + \gamma^k \cdot Flr_i \tag{28}$$

$$A_j^k = \alpha^k + \beta^k \cdot Emp_j + \gamma^k \cdot Flr_j \tag{29}$$

Table 3. Probability of Firms Producing and Consuming Commodities, Percentage

Industry Type	Commodity type															
	1		2		3		4		5		6		7		8	
1	**64.5**	**(28.8)**	0.8	(1.2)	0.2	(0.7)	**8.9**	(0.0)	0.5	**(43.1)**	**3.8**	(2.7)	0.5	(1.7)	**22.8**	**(21.0)**
2	1.6	(0.9)	0.0	(0.0)	**84.7**	**(60.5)**	3.9	(4.3)	4.4	**(22.7)**	0.4	(4.1)	1.9	(4.6)	4.3	(4.9)
3	0.3	(1.3)	**22.4**	**(24.2)**	**13.0**	**(13.2)**	**38.6**	**(47.0)**	**22.0**	**(24.1)**	2.0	(1.3)	**9.5**	**(5.9)**	**14.8**	**(7.0)**
4	0.5	(1.7)	0.5	(1.1)	**8.4**	**(28.0)**	**8.0**	**(18.3)**	**80.8**	**(76.1)**	2.9	**(11.7)**	**9.2**	**(13.6)**	4.7	(1.6)
5	0.1	(0.6)	0.4	(2.4)	0.6	**(8.5)**	**91.6**	**(81.5)**	3.0	**(18.3)**	0.8	**(5.4)**	3.6	**(5.9)**	**7.2**	**(8.3)**
6	0.1	(0.5)	0.3	(0.5)	**5.8**	(1.1)	**86.1**	**(92.3)**	**6.5**	**(11.4)**	0.4	**(5.7)**	**5.8**	**(7.2)**	2.4	(1.0)
7	**5.2**	**(6.5)**	2.6	**(6.2)**	0.5	(1.6)	**5.3**	**(8.0)**	**7.9**	**(12.6)**	**30.9**	**(39.6)**	**52.1**	**(40.3)**	2.5	(0.5)
8	0.5	(0.2)	**32.9**	**(32.6)**	**12.0**	**(11.6)**	**30.4**	**(35.0)**	**35.5**	**(33.2)**	1.4	(2.5)	**8.6**	**(10.0)**	**5.9**	(1.5)
9	**13.5**	**(17.5)**	0.1	(0.4)	1.0	(0.6)	**25.6**	**(22.2)**	**16.3**	**(18.1)**	**20.4**	**(22.0)**	**26.6**	**(30.4)**	**11.7**	(2.3)
10	**17.4**	**(43.9)**	1.1	(1.0)	0.3	(0.1)	**10.6**	**(9.4)**	**10.9**	**(11.4)**	**32.6**	**(36.1)**	**14.5**	**(18.8)**	**24.2**	(0.4)
11	**24.8**	**(19.2)**	3.3	(2.6)	2.5	(2.0)	**35.4**	**(34.1)**	**28.2**	**(27.3)**	**44.2**	**(41.8)**	**30.7**	**(35.3)**	3.7	(1.0)
12	0.0	(4.7)	0.3	(0.3)	1.4	**(9.5)**	**47.3**	**(48.6)**	**8.4**	**(31.2)**	4.5	**(13.0)**	**20.6**	**(25.9)**	**38.5**	(4.4)
13	2.1	**(23.7)**	0.2	(0.4)	0.1	(1.3)	**9.0**	**(9.8)**	**3.2**	**(21.3)**	**4.4**	**(20.0)**	**32.6**	**(41.8)**	**54.5**	(2.8)

Note: Consumption is in parentheses

Figure 3 shows the average value of correlation coefficient (R) of regression models. Although the value of R is quite low, the amount of production and consumption are significantly related to the number of employees and floor area, as implied by their t-statistic. The average values for t-statistic of coefficients α^k, β^k, and γ^k are 3.5, 5.6, and 7.7 respectively, with an average number of samples of 426. The reasons the regression model is not very accurate in terms of R-value might be the variation in commodity weight for each commodity type. As the results implies, commodity type 3 (Mineral Products), which has the least variation in unit weight, has the highest R-value as well.

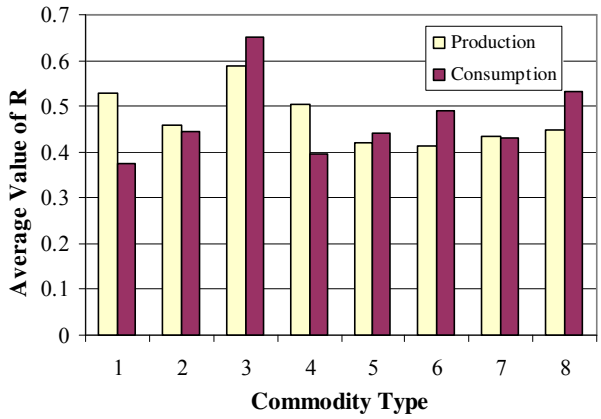

Figure 3. Average Value of Coefficient of Correlation (R) by Commodity Type

In simulation, random numbers are generated according to the proportion in Table 3 in order to indicate the commodity types that a firm produces and consumes. After estimating the amount of commodity production and consumption by each commodity type from the regression model, the production and consumption amounts were adjusted to satisfy the total production and consumption, which are expressed in monetary terms and estimated by regression techniques from the indicators of firm size. This research estimates the average value of each commodity type from the TMGMS data.

4.3.2 Commodity Distribution

In the commodity distribution model, the commodity consumptions generated from the former stage are linked to the production points. As mentioned in the model development section, the shipper probability describes the fraction of commodity that a customer selects to purchase from a shipper. The shipper probability ($P^k(i)$) is a product of three parts: distribution channel probability ($P(C^k)$), shipper location probability ($P(z|C^k)$), and shipper selection probability ($P(i|C^k, z)$). The first part, distribution channel probability ($P(C^k)$), is determined directly from the empirical data. The probability varies by industry type of customers, industry type of shippers, and type of purchased commodity. The probability is therefore calculated separately for each commodity type that is purchased by each industry type of customer from each industry type of shipper.

The second part, shipper location probability ($P(z|C^k)$), is estimated by a multinomial logit model. This model is also developed separately for each industry type of customer, each industry type of shipper, and each commodity type. The variables used in the model include zone size variables (total zone production ($G_z^{k,C}$) and total number of firms in a zone (N_z^C)) and an impedance variable (the distance between the customer and shipper zones ($D_{I,J}$)). This research assumes non-linear utility functions as follows:

$$\text{Type 1: } V_Z = \alpha^{k,C} \cdot \sqrt{G_Z^{k,C}} + \beta^{k,C} \cdot N_Z^C + \gamma^{k,C} \cdot D_{I,J} \tag{30}$$

$$\text{Type 2: } V_Z = \alpha^{k,C} \cdot \sqrt{G_Z^{k,C}} + \beta^{k,C} \cdot \sqrt{N_Z^C} + \gamma^{k,C} \cdot D_{I,J} \tag{31}$$

$$\text{Type 3: } V_Z = \alpha^{k,C} \cdot \sqrt{G_Z^{k,C}} + \beta^{k,C} \cdot N_Z^C + \gamma^{k,C} \cdot \log(D_{I,J}) \tag{32}$$

Among three types of the utility function, the utility type that provides the best fitness in term

of ρ^2 is selected for each model. For the estimated results, the value of t-statistic of all the variables is statistically significant, as the t-statistics of coefficients $\alpha^{k,C}$, $\beta^{k,C}$, and $\gamma^{k,C}$ are 133.5, 4.6, and -45.1 respectively. The sign of the variables indicates the customer's preference for the proportion of commodity being purchased. The total zone production and the total number of firms have a positive sign indicating that a bigger purchase amount comes from zone producing a higher amount of commodity and having a greater number of firms. Conversely, distance has a negative sign, which implies that customers prefer purchasing larger amounts from nearer zones. The measure of fit of the models, represented by ρ^2, ranges from 0.058 to 0.544. The models that deal with sub items are apt to have low value of ρ^2.

The third part, shipper selection probability ($P(i|C^k, z)$), is calculated directly using the proportion of production amount of shippers by commodity type, because the survey data do not provide information about the shippers that each customer selects. Lastly, the numbers of shippers is estimated by regression techniques the number of employees of a customer firm. The number of shippers for customer industry types 1, 2, 4, 7, 8, and 9 are assumed proportional to the square root of the number of employees, while the numbers of shippers for other industries are assumed proportional to the number of employees.

4.3.3 Delivery Lot size and Frequency

Delivery lot size and frequency are determined using the concept of cost minimization from the viewpoint of customers. The models are developed separately by industry type of customers and by purchased commodity type. Parameters a^k and b^k of equation (14) are calibrated by regression analysis from the TMGMS data. The average coefficients of correlation (R) of all models are 0.48. The average values of t-statistic of parameters a^k and b^k are 3.3 and 10.5 respectively, when the average sample size is 350. The inaccuracy in the estimated results is due to this model's utilization of direct distance between customer and shipper, which means that the model considers cost for each trip; in actual, delivery lot size and frequency will not necessarily the same for the same distance customers. For example, a shipper might deliver a large lot size to the biggest customer and a smaller lot size to the smaller customers located in the same region. Future development of this model will consider the total cost for each delivery tour.

4.3.4 Carrier and Vehicle Choice Selection

Nested logit models are developed separately for each industry type of shipper. The model consists of two levels on which the first level is choice between private and business truck and the second level is choice between large and small truck. However, there is an exception for the models of industry types 1, 3, and 8, which are vise versa. The estimated results are presented in Table 4.

For the models of industries apart from industry types 1, 3, and 8, the first level of model is the carrier choice between private and business truck, with business truck as the reference alternative The t-statistic of the most variables in all models is statistically significant. Among all industries, travel time variables have negative sign, which indicates that a private truck is generally used in short distance delivery. Constant for large and small private trucks have positive and negative signs, respectively, for most industries. This can be interpreted as a large truck being preferable for private trucks in most industries. The coefficient for the number of employees has a negative sign for most industries, which implies that large firms usually use business trucks rather than private trucks except in industry type 5 and 11. For industry type 2, this variable is less significant. Estimated results of the second level are the choice between large and small trucks, with small truck as the reference alternative. Most variables are significant, as their t-statistic value is rather high. The sign of the variables indicates the preference of the shippers. The positive sign proves that, as per intuition, large trucks are usually used to deliver large lot sizes. The positive sign of travel time indicates that, for long distance delivery, large truck tends to be more preferable. This is reasonable because the usage of large trucks can minimize the delivery costs for long distance delivery, as they can deliver to more customers.

Table 4. Carrier and Vehicle Choice Model Estimated Coefficients (t-statistics in parenthesis)

Industry Type	Choice level	Const-PL Const-LP*		Const-PS Const-LB*		Const-L Const-P*		CUS		EMP		TT		LOT		FRQ		μ_L μ_p*		N ρ^2 Hit Ratio
1*	1st	-14.11	*(-2.5)*	-23.85	(-10.1)			1.35	*(21.6)*							8.33	*(14.2)*			181
	2nd					38.30	*(12.3)*	-3.61	*(5.8)*	-0.0106	*(7.2)*	-22.28	*(15.3)*					0.37	*(14.1)*	0.574
																				0.696
2	1st	9.53	*(3.5)*	-0.99	*(-1.5)*			-0.0277	*(-3.2)*			-0.0156	*(-2.0)*							1,181
	2nd					-132	*(-1.6)*	0.3777	*(1.6)*			0.2010	*(2.3)*	0.0000028	*(2.9)*			0.07	*(1.1)*	0.624
																				0.798
3*	1st	3.76	*(1.7)*	-4.53	*(-2.9)*											0.0379	*(1.9)*			4,964
	2nd					-12.9	*(-14.4)*	0.1585	*(22.1)*	-0.00252	*(18.7)*					-0.1210	*(8.7)*	0.25	*(11.2)*	0.623
																				0.783
4	1st	6.27	*(33.3)*	-2.13	*(-19.7)*			-0.0159	*(-19.8)*	-0.00310	*(-19.3)*	-0.0125	*(-26.3)*	-0.0000014	*(-7.8)*	-0.0118	*(-2.0)*			24,367
	2nd					-9.44	*(-21.8)*	0.0060	*(9.0)*	0.00137	*(9.5)*	0.0056	*(10.2)*	0.0000012	*(8.2)*	0.0282	*(2.8)*	0.55	*(8.9)*	0.654
																				0.807
5	1st	6.40	*(18.1)*	-1.91	*(-10.0)*			-0.0188	*(-5.5)*	0.00019	*(3.0)*	-0.0134	*(-9.9)*			0.0093	*(2.4)*			4,158
	2nd					-7.24	*(-6.3)*					0.0046	*(2.9)*	0.000551	*(12.8)*			0.71	*(1.8)*	0.617
																				0.764
6	1st	5.33	*(43.2)*	-1.75	*(-23.6)*			-0.0100	*(-7.9)*	-0.00026	*(-8.4)*	-0.0101	*(-24.2)*	-0.0000449	*(-12.7)*					20,328
	2nd					-9.78	*(-16.9)*	-0.0053	*(-3.0)*	0.00050	*(7.7)*	0.0063	*(10.7)*	0.000008	*(5.5)*	0.0124	*(6.6)*	0.46	*(8.9)*	0.578
																				0.770
7	1st	5.64	*(50.1)*	-2.23	*(-31.7)*			-0.0045	*(-9.4)*	-0.00041	*(-12.1)*	-0.0107	*(-30.1)*	-0.000138	*(-16.1)*					26,965
	2nd					-9.13	*(-17.8)*	0.0039	*(4.1)*			0.0067	*(12.2)*	0.0000998	*(5.5)*	0.0218	*(14.0)*	0.53	*(8.7)*	0.559
																				0.743
8*	1st	4.79	*(1.8)*	-5.67	*(-2.1)*											0.0259	*(1.9)*			23,965
	2nd					-2.39	*(17.4)*	-0.0266	*(11.1)*			-0.0181	*(3.2)*			-0.0356	*(9.7)*	0.59	*(4.5)*	0.725
																				0.872
9	1st	6.34	*(37.8)*	-2.49	*(-23.4)*					-0.00185	*(-10.3)*	-0.0087	*(-16.4)*	-0.0000082	*(-6.6)*	-0.0254	*(-3.9)*			28,189
	2nd					-22.74	*(-7.0)*			0.00331	*(6.0)*	0.0160	*(5.7)*	0.0000183	*(5.1)*	0.1220	*(4.7)*	0.26	*(5.3)*	0.579
																				0.752
10	1st	4.23	*(35.3)*	-2.84	*(-22.6)*			-0.0009	*(-7.2)*	-0.00246	*(-20.3)*	-0.0095	*(-9.2)*							32,351
	2nd					-6.34	*(-15.2)*	-0.0013	*(-5.5)*	0.00048	*(6.6)*					0.0337	*(3.8)*	0.54	*(8.4)*	0.814
																				0.913
11	1st	4.14	*(20.5)*	-4.52	*(-23.1)*			0.0027	*(1.9)*	0.00048	*(2.0)*	-0.0050	*(-6.5)*							7,448
	2nd					-33.83	*(-2.3)*	-0.0248	*(-1.7)*			0.0284	*(2.3)*	0.000492	*(16.9)*			0.12	*(1.9)*	0.621
																				0.777
12	1st	4.14	*(6.4)*	-5.74	*(-5.8)*			0.1830	*(5.7)*	-0.00485	*(-4.3)*	-0.0150	*(-3.0)*	0.0000654	*(1.4)*	0.0943	*(3.2)*			431
	2nd					-16.30	*(-1.5)*	-0.7412	*(-1.8)*	0.01825	*(2.0)*	0.0515	*(1.8)*			-0.4723	*(-1.5)*	0.22	*(1.2)*	0.623
																				0.768
13	1st	6.34	*(27.5)*	-0.55	*(-7.9)*			-0.0203	*(-10.8)*	-0.00033	*(-4.8)*	-0.0084	*(-11.6)*	-0.0000241	*(-6.3)*					7,205
	2nd					-27.14	*(-4.9)*	0.0816	*(5.2)*	0.00104	*(2.6)*	0.0197	*(4.8)*	0.000103	*(4.4)*	-0.0184	*(-6.8)*	0.21	*(4.0)*	0.430
																				0.680

Note: * indicates that the model structure is changed as “vehicle choice being the first level and carrier choice being the second level”.

For industry types 1, 3, and 8, the first level of model is choice between large and small truck, with small truck as the reference alternative. Constant for private and business large trucks have positive and negative signs respectively, which imply that private truck is preferable for large truck, except for industry type 1. The second level of model is carrier choice between private and business truck, with business truck as the reference alternative. The negative sign of number of employee implies that business truck is preferable for large company, which confirms the fact that the cost of using business truck for large company is cheaper than its cost of using private truck. In addition, as the scale parameter of all models is not equal to one, it indicates that a nested logit model is better suited for this data than would be a simple multinomial logit model.

The measures of fitness of the model, represented by ρ^2 and hit-ratio, which are shown in the right side of Table 4, are rather high for all models. The values of hit-ratio of most models are more than 70 percent, which indicates that the proposed nested logit models for carrier and vehicle choices are sufficient to predict the choice behavior.

4.3.5 Vehicle Routing

Commodity flows are assigned to vehicle tours after deciding carrier and vehicle choices. Before assigning a truck to customers, the customers having the same delivery frequency are grouped by delivery time. Customers having the same delivery frequency are then grouped again by location in order to minimize the delivery cost. Vehicle routing simulation provides a sequence for visiting customers that minimizes the total cost considering the maximum carrying weight of truck and maximum working hour of the driver. The maximum carrying weight utilized in this research is derived from the 80th percentile of the actual carrying weight of trucks from the TMGMS data. This value is computed by industry type and by truck type. For the constraint of the total travel time, we assume that the maximum working hours of a driver are 10 hours. In simulation, the summation of total travel time to visit all customers and total staying time spent at customer locations should be less than or equal to the maximum number of working hours. The staying time at customers, this research utilizes an average value by industry.

4.4 Model Validation

The 174,892 virtual firms generated by the Monte-Carlo simulation from the EEC data, were used for model validation. The virtual firms were input through the three stages of the model (commodity production and consumption, commodity distribution, and conversion of commodity flow to vehicle movement) resulting in the estimated truck origin and destination volumes by truck type. The model sequence can be summarized as follows:

- Identify the commodity types for each virtual firm consuming and producing, using the probability of firms producing and consuming commodities.
- Generate the amount of commodity production and consumption by each firm using the commodity production and consumption model.
- Make the connections between shippers and customers according to their relationship in supply chains using the commodity distribution model and resulting in commodity flows between firms.
- Assign the vehicle trip characteristics to each commodity flow through three stages: delivery lot size and frequency, carrier and vehicle choice selection, and vehicle routing.
- Truck movement resulted from the former stage then can be summed up for truck trip origin and destination matrices.

The estimated volumes by truck type were compared with the actual volumes from the Road Traffic Census survey. Figures 4 and 5 show the results between the actual volume and the estimated volume for small trucks and large trucks, respectively. The estimated volume for large trucks is similar to the actual volume while the volume for small truck is significantly overestimated. For small truck, the simulation volume for the zones within the center of Tokyo (shown in triangle labels) is particularly distinct from the other zones because the characteristics of freight movement in the center of Tokyo are much difference from those in other locations. The estimation error may be due to data obsolescence; some characteristics may have changed over the 20 years. Although the estimated results are not very accuracy now, the results suggest that further model development under this concept could be very beneficial.

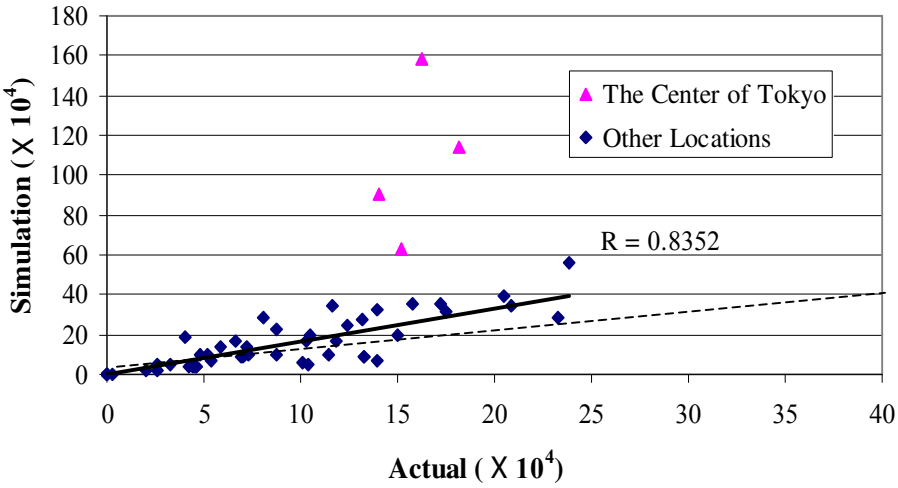

Figure 4. Generated OD from Simulation Compared with Actual OD for Small Trucks

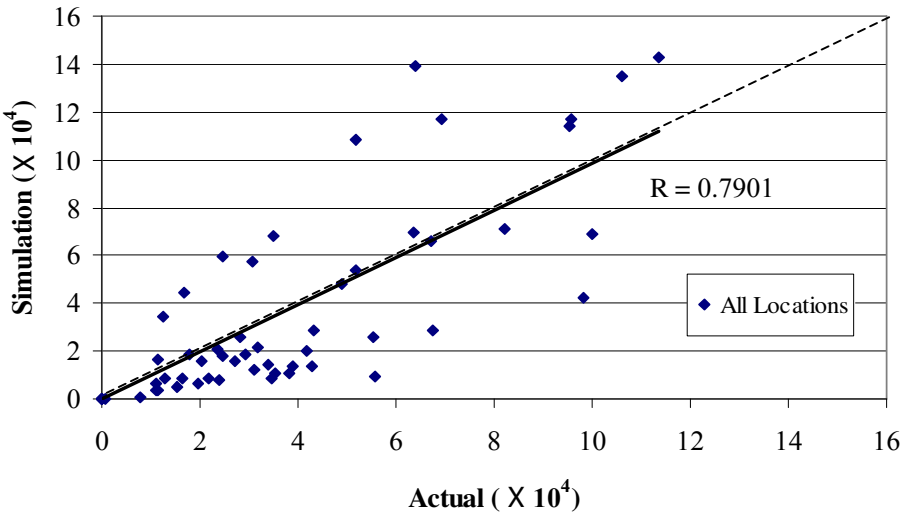

Figure 5. Generated OD from Simulation Compared with Actual OD for Large Trucks

5. CONCLUSION

The micro-simulation model proposed in this paper is a comprehensive approach to modeling freight transportation in a way that systematically reflects the individual behavior of freight decision makers. The model is a modification to the traditional four-step approach that follows commodity production and consumption, commodity distribution, and conversion of commodity flows to vehicle movement. The model incorporates individual behaviors considering the largest influence of each freight decision maker at each stage. For example, in the commodity distribution model, customers have the most significant influence on the decision of the fraction of commodity being purchased. As a result, the proposed model structure can incorporate the complex relationship among freight decision makers and their

interactions within the model. As the model is developed at the micro level, the model is capable of analyzing the wide range of logistics initiatives, which includes not only the effect of large-scale changes (such as new transportation infrastructure), but also the effect of small-scale changes (such as changes in logistical structure).

As presented in the model application part in this paper, the proposed model has been tested by estimating truck trip origin and destination volumes in the Tokyo Metropolitan Area. Model calibration was performed using the data sets of the Tokyo Metropolitan Goods Movement Survey. For model validation, the model was applied to the virtual firms generated by the Monte-Carlo simulation from a data set of the Establishment and Enterprise Census, which provides information of firms operating in the Tokyo Metropolitan Area and in nearby prefectures. The estimated truck volumes are compared with the actual volumes obtained from the Road Traffic Census survey. The results of model validation indicate that the proposed model provides promising results.

Freight transportation is, in nature, a very complex system; the proposed model, attempted to capture the mechanism of the system, however, could not include all aspects of the system. The supply chain considered in this paper is rather a simple case; in fact, there are many more actors involved in freight distribution: agents, brokers, and etc. As of recent, the involved actors are increasing as the growth in concerning of the quality of services. For example, third party logistics (3PL), Just-In-Time distribution (JIT), and other innovations in freight distribution adopted currently have much effect to the modeling structure. All these should be taken into account in the future improvement.

REFERENCES

a) Books and Books chapters

Taniguchi E., Thompson R.G., Yamada T., and Duin R.V., (2001) **City Logistics: Network Modeling and Intelligent Transport Systems**. Pergamon, Netherlands.

b) Journal papers

Boerkamps J.H.K., Binsbergen A.J.V., and Bovy P.H.L. (2000) Modeling Behavioral Aspects of Urban Freight Movement in Supply Chains, **Transportation Research Record 1725**, 17-25.

Park M.B., and Smith R.L., (1997) Development of Statewide Truck-Travel Demand Model with Limited Origin-Destination Survey Data, **Transportation Research Record 1602**, 14-21

Veras J.H., and Thorson E., (2000) Trip Length Distributions in Commodity-Based and Trip-Based Freight Demand Modeling, **Transportation Research Record 1707**, 37-48.

c) Papers presented to conferences

Boerkamps J., and Binsbergen A.V., (1999) GoodTrip-A New Approach for Modeling and Evaluation of Urban Goods Distribution. **Proceedings 1st International Conference on City Logistics: City Logistics I**, Australia.

Maat K., and Visser J., (1996) A Simulation Model for Urban Freight Transport. **Proceedings 2nd Joint European Conference & Exhibition on Geographic Information**, Spain.

Sivakumar A., and Bhat C., (2002) A Fractional Split Distribution Model for Statewide Commodity Flow Analysis. **The 81st Annual Meeting of the Transportation Research Board**, Washington D. C.

Vovsha P., Petersan E., and Donnelly R., (2002) Micro-Simulation in Travel Demand Modeling: Lessons from the New York "Best Practice" Model. **The 81st Annual Meeting of the Transportation Research Board**, Washington D. C.